Measure Theory

November 17, 2020

Contents

Part I
Class Notes

1 Introduction

Measure theory is a central subject in analysis. Let us consider four problems which helped lead to the development of measure theory.

1) When we studied inner-products in Hilbert spaces, we considered the inner-product space $(C[a,b], \langle \cdot, \cdot \rangle_1)$, where $C[a,b]$ is the $\mathbb{C}$-vector space of all continuous functions defined on the interval $[a,b]$ and $\langle \cdot, \cdot \rangle_1$ is the inner-product defined by

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\mathrm{d}x \tag{1}$$

for all $f, g \in C[a,b]$. One of the problems that we ran into when studying this inner-product space is that it is *not* a Hilbert space. In other words, $C[a,b]$ is not complete with respect to topology induced by this inner-product. It turns out however that just as how $\mathbb{R}$ can be viewed as the "completion" of $\mathbb{Q}$ with respect to the usual topology on $\mathbb{Q}$ induced by the usual absolute value $|\cdot|$, there is a similar space which can be viewed as the "completion" of $C[a,b]$ with respect to the topology induced by the inner-product $\langle \cdot, \cdot \rangle_1$. Measure theory will give us a nice interpretation of what this completed space looks like.

2) Now consider the normed linear space $(C[a,b], \|\cdot\|_\infty)$, where $\|\cdot\|_\infty$ is the supremum norm. This time $C[a,b]$ *is* complete with respect to the topology induced by this norm. In other words, $(C[a,b], \|\cdot\|_\infty)$ is a Banach space. In this case, we'd like to know what the dual space looks like. Recall that the dual of a normed linear space $(\mathcal{X}, \|\cdot\|)$ is defined to be the space $(\mathcal{X}^*, \|\cdot\|)$ where

$$\mathcal{X}^* := \{\ell \colon \mathcal{X} \to \mathbb{C} \mid \ell \text{ is a bounded linear functional}\}$$

and where the norm, denoted $\|\cdot\|$ again, is defined by

$$\|\ell\| = \sup\{|\ell(x)| \mid \|x\| \leq 1\}$$

for all $\ell \in \mathcal{X}^*$. For instance, here are three examples of a bounded linear functionals on $C[a,b]$ (with respect to the sup norm)

$$\begin{aligned}
\ell_1(f) &= \int_a^b f(x)\mathrm{d}x \\
\ell_2(f) &= f(a) \\
\ell_3(f) &= f(a) + f(b).
\end{aligned}$$

Here again we will find that measure theory will give us a natural interpretation of what the dual space of $C[a,b]$ looks like. We shall that see that, in a certain sense, the bounded linear functionals on $C[a,b]$ will correspond to different ways you can integrate on $[a,b]$.

3) The third problem we consider goes as follows: let $T \colon \mathcal{H} \to \mathcal{H}$ be a bounded linear map from a Hilbert space to itself. If T is compact and self-adjoint, then the Spectral Theorem tells us that there exists an orthonormal basis of eigenvectors (e_n) with eigenvalues (λ_n) such that

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n$$

for all $x \in \mathcal{H}$. Can we still get a spectral theorem if we remove the compactness condition? It turns out that there is a way to do this. We may not be able to completely show this in this class (perhaps in a functional analysis class), but we will make progress using measure theory.

4) The foundations of probability theory is based on measure theory. We will better understand the connection between probability theory and measure theory.

Let's go back to $C[a,b]$ equipped with the inner-product $\langle\cdot,\cdot\rangle_1$ described above. As we said, the main problem with $(C[a,b],\langle\cdot,\cdot\rangle_1)$ is that this space is not complete. Ideally, we would like to be able to measure "the length" of any subset of $\mathbb{R}$ such that

1. the length of an interval (c,d) is $d-c$,
2. if we translate a set, its length should stay the same,
3. if (E_n) is a sequence of pairwise disjoint subsets of $\mathbb{R}$ with lengths $\mu(E_n)$, then the length of their union satisfy
$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right)=\sum_{n=1}^{\infty}\mu(E_n).$$

2 Intervals, Algebras, and Measures

Throughout the rest of this section, we consider the normed linear space $(C[a,b],\|\cdot\|)$[1] where, unless otherwise specified, $\|\cdot\|$ is the L_1 norm:
$$\|f\|:=\int_a^b |f(x)|\mathrm{d}x$$
for all $f\in C[a,b]$ where the integral is understood to be the Riemann integral. It is easy to prove that $C[a,b]$ equipped with the L_1 norm is a normed linear space, but it is not complete. As shown in the Appendix, we know that we can construct the completion of $C[a,b]$ using Cauchy sequences, but this space is a little too abstract for our desires. We would like to describe the completion of $C[a,b]$ in a more concrete way. More specifically, we'd like to realize the completion of $C[a,b]$ as a quotient of a certain space of functions (rather than as a quotient of a space of Cauchy sequences).

Indeed, if $\mathcal{Y}$ is a dense subspace of a normed linear space $\mathcal{X}$, then every bounded linear functional $\ell\colon\mathcal{Y}\to\mathbb{C}$ can be extended in a unique way to a bounded linear functional $\widetilde{\ell}\colon\mathcal{X}\to\mathbb{C}$ with the same norm. That is, $\widetilde{\ell}|_{\mathcal{Y}}=\ell$ and $\|\widetilde{\ell}\|=\|\ell\|$. The linear functional $\ell\colon C[a,b]\to\mathbb{C}$ defined by
$$\ell(f)=\int_a^b f(x)\mathrm{d}x$$
for all $f\in C[a.b]$ is bounded with $\|\ell\|=1$. So if we can define the completion of $C[a,b]$ as a quotient of a space of functions, then we will be able to compute $\widetilde{\ell}(f)$ for every f. In turn, this can be naturally viewed as computing $\int_a^b f(x)\mathrm{d}x$. This new integral will be called the **Lebesgue integral** and the representative functions in this space will be called **Lebesgue measurable functions**. Let us denote the completion of $C[a,b]$ with respect to $\|\cdot\|$ by $(\mathscr{C}[a,b],\|\cdot\|)$, or more simply $\mathscr{C}[a,b]$.

2.1 Intervals

Unless otherwise specified, by a a **subinterval of** $[a,b]$, we shall mean a subinterval of $[a,b]$ of the following types: (c,d), $(c,b]$, $[c,b)$, and $[c,d]$, where
$$-\infty<a\leq c\leq d\leq b<\infty.$$
Note that $(c,c)=\varnothing$ and $[c,c]=\{c\}$ are considered intervals as well. If I is a subinterval of $[a,b]$, then the **indicator function** $1_I\colon[a,b]\to\{0,1\}$ with respect to I is defined by
$$1_I(x)=\begin{cases}1 & \text{if } x\in I\\ 0 & \text{if } x\notin I\end{cases}$$
for all $x\in[a,b]$. If I is a subinterval of $[a,b]$, then it closure in $[a,b]$ has the form
$$\overline{I}=[c,d]$$

[1] Unless otherwise specified, we will often simply write $C[a,b]$ instead of $(C[a,b],\|\cdot\|)$ and view $C[a,b]$ as a normed linear space equipped with the L_1 norm.

for some $a \le c \le d \le b$. In this case, we define the **length** of I to be

$$\text{length}(I) = d - c.$$

We denote by $\mathcal{I}_{[a,b]}$ (or more simply just $\mathcal{I}$ if context is clear) to be the collection of all finite unions of subintervals of $[a, b]$.

2.1.1 Collection of all Subintervals of $[a, b]$ Forms a Semialgebra of Sets

Definition 2.1. A nonempty collection $\mathcal{E}$ of subsets of X is said to be a **semialgebra** of sets if it satisfies the following properties:

1. $\emptyset \in \mathcal{E}$;
2. If $E, F \in \mathcal{E}$, then $E \cap F \in \mathcal{E}$;
3. If $E \in \mathcal{E}$, then E^c is a finite disjoint union of sets from $\mathcal{E}$.

Proposition 2.1. *The collection of all subintervals of $[a, b]$ forms a semialgebra of sets.*

Proof. Let $\mathcal{I}$ denote the collection of all subintervals of $[a, b]$. We have $\emptyset \in \mathcal{I}$ since $\emptyset = (c, c)$ for any $c \in [a, b]$. Now we show $\mathcal{I}$ is closed under finite intersections. Let I_1 and I_2 be subintervals of $[a, b]$. Taking the closure of I_1 and I_2 gives us closed intervals, say

$$\overline{I}_1 = [c_1, d_1] \quad \text{and} \quad \overline{I}_2 = [c_2, d_2].$$

Assume without loss of generality that $c_1 \le c_2$. If $d_1 < c_2$, then $I_1 \cap I_2 = \emptyset$, so assume that $d_1 \ge c_2$. If $d_1 \ge d_2$, then $I_1 \cap I_2 = I_2$, so assume that $d_1 < d_2$. If $c_1 = c_2$, then $I_1 \cap I_2 = I_1$, so assume that $c_2 > c_1$. So we have reduced the case to where

$$c_1 < c_2 \le d_1 < d_2.$$

With these assumptions in mind, we now consider four cases:

Case 1: If $I_1 = [c_1, d_1]$ or $I_1 = (c_1, d_1]$ and $I_2 = [c_2, d_2]$ or $I_2 = [c_2, d_2)$, then $I_1 \cap I_2 = [c_2, d_1]$.

Case 2: If $I_1 = [c_1, d_1)$ or $I_1 = (c_1, d_1)$ and $I_2 = [c_2, d_2]$ or $I_2 = [c_2, d_2)$, then $I_1 \cap I_2 = [c_2, d_1)$.

Case 3: If $I_1 = [c_1, d_1]$ or $I_1 = (c_1, d_1)$ and $I_2 = (c_2, d_2]$ or $I_2 = (c_2, d_2)$, then $I_1 \cap I_2 = (c_2, d_1)$.

Case 4: If $I_1 = [c_1, d_1)$ or $I_1 = (c_1, d_1)$ and $I_2 = (c_2, d_2]$ or $I_2 = (c_2, d_2)$, then $I_1 \cap I_2 = (c_2, d_1)$.

In all cases, we see that $I_1 \cap I_2$ is a subinterval of $[a, b]$.

Now we show that compliments can be expressed as finite disjoint unions. Let I be a subinterval of $[a, b]$ and write $\overline{I} = [c, d]$. We consider four cases:

Case 1: If $I = [c, d]$, then $I^c = [a, c) \cup (d, b]$.

Case 2: If $I = (c, d]$, then $I^c = [a, c] \cup (d, b]$.

Case 3: If $I = [c, d)$, then $I^c = [a, c) \cup [d, b]$.

Case 4: If $I = (c, d)$, then $I^c = [a, c] \cup [d, b]$.

Thus in all cases, we can express I^c as a disjoint union of intervals since $a \le c \le d \le b$. □

2.1.2 Indicator Function 1_I Represents an Element in $\mathscr{C}[a, b]$

The next proposition tells us that 1_I should represent an element in $\mathscr{C}[a, b]$ for all subintervals I of $[a, b]$ and that $\|1_I\| = \text{length}(I)$.

Proposition 2.2. *Let I be a subinterval of $[a,b]$. Then there exists a Cauchy sequence (f_n) in $(C[a,b], \|\cdot\|)$ such that (f_n) converges pointwise to 1_I on $[a,b]$. Moreover we have*

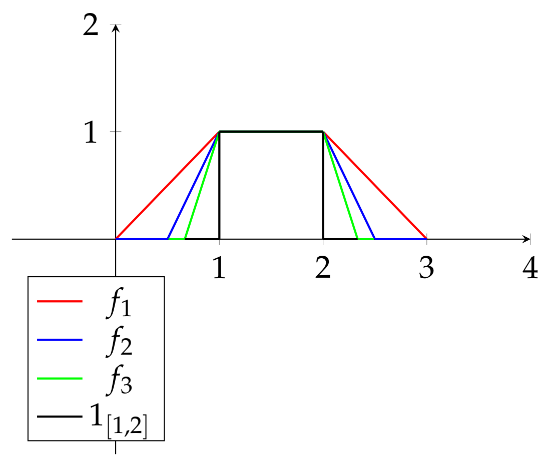
$$\lim_{n\to\infty} \|f_n\| = \text{length}(I).$$

Proof. If $I = \emptyset$, then we take $f_n = 0$ for all $n \in \mathbb{N}$. Thus assume I is a nonempty subinterval of $[a,b]$. We consider two cases; namely $I = (c,d)$ and $I = [c,d]$. The other cases ($I = (c,d]$ and $I = [c,d)$) will easily be seen to be a mixture of these two cases.

Case 1: Suppose $I = [c,d]$. For each $n \in \mathbb{N}$, define

$$f_n(x) = \begin{cases} 0 & \text{if } a \leq x < c - \left(\frac{c-a}{n}\right) \\ \frac{n}{c-a}(x-c)+1 & \text{if } c - \left(\frac{c-a}{n}\right) \leq x < c \\ 1 & \text{if } c \leq x \leq d \\ \frac{n}{d-b}(x-d)+1 & \text{if } d < x \leq d + \left(\frac{b-d}{n}\right) \\ 0 & \text{if } d + \left(\frac{b-d}{n}\right) < x \leq b \end{cases}$$

The image below gives the graphs for f_1, f_2, and f_3 in the case where $[a,b] = [0,3]$ and $[c,d] = [1,2]$.



For each $n \in \mathbb{N}$, the function f_n is continuous since each of its segments is continuous and are equal on their boundaries.

Let us check that (f_n) converges pointwise to 1_I: If $x \in [a,c)$, then we choose $N \in \mathbb{N}$ such that

$$x \leq c - \left(\frac{c-a}{N}\right).$$

Then $f_n(x) = 0$ for all $n \geq N$. Thus

$$\lim_{n\to\infty} f_n(x) = 0 = 1_I(x).$$

Similarly, if $x \in (d,b]$, then we choose $N \in \mathbb{N}$ such that

$$x \geq d + \left(\frac{b-d}{N}\right).$$

Then $f_n(x) = 0$ for all $n \geq N$. Thus

$$\lim_{n\to\infty} f_n(x) = 0 = 1_I(x).$$

Finally, if $x \in [c,d]$, then $f_n(x) = 1$ for all $n \in \mathbb{N}$ by definition and thus

$$\lim_{n\to\infty} f_n(x) = 0 = 1_I(x).$$

Let us check that (f_n) is Cauchy in $(C[a,b], \|\cdot\|_1)$. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that

$$\frac{c-a+b-d}{n} < \varepsilon$$

for all $n \geq N$. Then $n \geq m \geq N$ implies

$$\begin{aligned}
\|f_n - f_m\|_1 &= \int_a^b |f_n(x) - f_m(x)| \mathrm{d}x \\
&= \int_a^b (f_n(x) - f_m(x)) \mathrm{d}x \\
&= \int_{c-\left(\frac{c-a}{m}\right)}^{c} (f_n(x) - f_m(x)) \mathrm{d}x + \int_d^{d+\left(\frac{b-d}{m}\right)} (f_n(x) - f_m(x)) \mathrm{d}x \\
&\leq \int_{c-\left(\frac{c-a}{m}\right)}^{c} \mathrm{d}x + \int_d^{d+\left(\frac{b-d}{m}\right)} \mathrm{d}x \\
&= \frac{c-a}{m} + \frac{b-d}{m} \\
&= \frac{c-a+b-d}{m} \\
&< \varepsilon.
\end{aligned}$$

Thus the sequence (f_n) is Cauchy in $(C[a,b], \|\cdot\|_1)$.

Finally, we check that $\|f_n\|_1 \to \text{length}(I)$ as $n \to \infty$. We have

$$\begin{aligned}
d - c &\leq \|f_n\|_1 \\
&= \int_a^b |f_n(x)| \mathrm{d}x \\
&= \int_a^b f_n(x) \mathrm{d}x \\
&= \int_{c-\left(\frac{c-a}{n}\right)}^{c} f_n(x) \mathrm{d}x + \int_c^d \mathrm{d}x + \int_d^{d+\left(\frac{b-d}{n}\right)} f_n(x) \mathrm{d}x \\
&\leq \int_{c-\left(\frac{c-a}{n}\right)}^{c} \mathrm{d}x + \int_c^d \mathrm{d}x + \int_d^{d+\left(\frac{b-d}{n}\right)} \mathrm{d}x \\
&= \frac{c-a}{n} + d - c + \frac{b-d}{n} \\
&\to d - c.
\end{aligned}$$

Thus for each $n \in \mathbb{N}$, we have

$$d - c \leq \|f_n\|_1 \leq d - c + \frac{c-a+b-d}{n}. \tag{2}$$

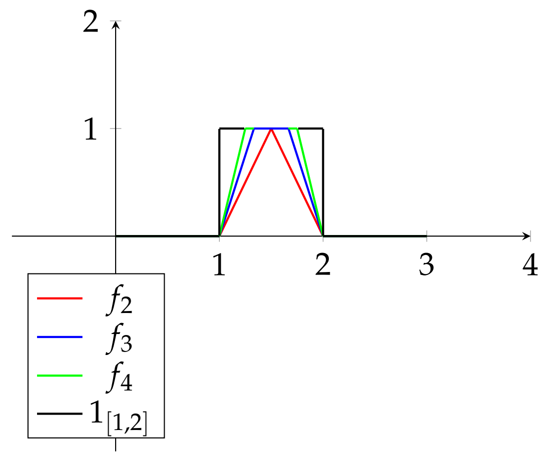
By taking $n \to \infty$ in (2), we see that

$$\begin{aligned}
\lim_{n \to \infty} \|f_n\|_1 &= d - c \\
&= \text{length}(I).
\end{aligned}$$

Case 2: Suppose $I = (c, d)$. For each $n \geq 2$, define

$$f_n(x) = \begin{cases} 0 & \text{if } a \leq x \leq c \\ \frac{n}{d-c}(x-c) & \text{if } c < x \leq c + \left(\frac{d-c}{n}\right) \\ 1 & \text{if } c + \left(\frac{d-c}{n}\right) \leq x \leq d - \left(\frac{d-c}{n}\right) \\ \frac{n}{c-d}(x-d) & \text{if } d - \left(\frac{d-c}{n}\right) \leq x \leq d \\ 0 & \text{if } d \leq x \leq b \end{cases}$$

The image below gives the graphs for f_2, f_3, and f_4 in the case where $[a, b] = [0, 3]$ and $(c, d) = (1, 2)$.



That (f_n) is a Cauchy sequence of continuous funtions in $(C[a,b], \|\cdot\|_1)$ which converges pointwise to 1_I and $\|f_n\|_1 \to \text{length}(I)$ as $n \to \infty$ follows from similar arguments used in case 1. $\square$

Remark 1. Note that $1_{[c,c]}$ is not the zero function, even though $\|1_{[c,c]}\| = 0$. Thus we need to identify the function $1_{[c,c]}$ and the constant zero function in $\mathscr{C}[a,b]$. In other words, $1_{[c,c]}$ and the constant zero function should represent the same element in $\mathscr{C}[a,b]$. Similarly, we have

$$\begin{aligned}\|1_{(c,d]} - 1_{(c,d)}\| &= \|1_{[d,d]}\| \\ &= 0,\end{aligned}$$

and so the functions $1_{(c,d]}$ and $1_{(c,d)}$ should represent the same element in $\mathscr{C}[a,b]$.

2.1.3 Finite Sums of Indicator Functions of Intervals Represents Elements in $\mathscr{C}[a,b]$

Let E be any finite union of intervals. Since the collection of all intervals forms a semialgebra, we can express E as a finite union of disjoint intervals, say

$$E = \bigcup_{i=1}^{k} I_i$$

where $I_i \cap I_j = \emptyset$ whenever $i \neq j$. The next proposition tells us that 1_E should represent an element in $\mathscr{C}[a,b]$ and that

$$\|1_E\| = \sum_{i=1}^{k} \text{length}(I_i)$$

.

Proposition 2.3. *With the notation as above, there exists a Cauchy sequence (f_n) in $(C[a,b], \|\cdot\|)$ such that (f_n) converges pointwise to 1_E on $[a,b]$. Moreover we have*

$$\lim_{n\to\infty} \|f_n\| = \sum_{i=1}^{k} \text{length}(I_i). \tag{3}$$

Proof. For each $1 \le i \le k$, choose a Cauchy sequence $(f_n^i)_{n\in\mathbb{N}}$ such that (f_n^i) converges pointwise to 1_{I_i} on $[a,b]$ and satisfies

$$\lim_{n\to\infty} \|f_n^i\| = \text{length}(I_i).$$

For each $n \in \mathbb{N}$, set $f_n = \sum_{i=1}^{k} f_n^i$. We claim that (f_n) is a Cauchy sequence that converges pointwise to 1_E and satisfies (3). Indeed, it is Cauchy since a sum of Cauchy sequences is Cauchy. Let us check that it converges pointwise to 1_E. Let $x \in E$ and let $\varepsilon > 0$. Then $x \in I_i$ for some $1 \le i \le k$, and so there exists an $N \in \mathbb{N}$ such that $n \ge N$ implies

$$|f_n^i(x) - 1_{I^i}(x)| < \frac{\varepsilon}{k}$$

for all $1 \leq i \leq n$. Choose such an $N \in \mathbb{N}$. Then $n \geq N$ implies

$$\begin{aligned}|f_n(x) - 1_E(x)| &= \left|\sum_{i=1}^{k} f_n^i(x) - \sum_{i=1}^{k} 1_{I_i}(x)\right| \\ &\leq \sum_{i=1}^{k} |f_n^i(x) - 1_{I_i}(x)| \\ &< \sum_{i=1}^{k} \frac{\varepsilon}{k} \\ &= \varepsilon.\end{aligned}$$

It follows that f_n converges pointwise to 1_E. Finally, let us check that (3) holds. Let $\varepsilon > 0$. Without loss of generality, we may assume that $\|f_n^i\| \geq \text{length}(I_i)$ and $\|f_n\| \geq \text{length}(E)$ for all $1 \leq i \leq k$ and for all $n \in \mathbb{N}$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\|f_n^i\| - \text{length}(I_i) < \frac{\varepsilon}{k}$$

for all $i = 1, \ldots, k$. Then $n \geq N$ implies

$$\begin{aligned}\|f_n\| - \text{length}(E) &= \left\|\sum_{i=1}^{k} f_n^i\right\| - \text{length}\left(\bigcup_{i=1}^{k} I_i\right) \\ &= \left\|\sum_{i=1}^{k} f_n^i\right\| - \sum_{i=1}^{k} \text{length}(I_i) \\ &\leq \sum_{i=1}^{k} \|f_n^i\| - \sum_{i=1}^{k} \text{length}(I_i) \\ &< \varepsilon.\end{aligned}$$

It follows that (3) holds. □

2.1.4 Well-Definedness of Length

Given the result above, it seems logical that the indiciator function of any set E which can be expressed as a finite disjoint union of intervals should represent an element in $\mathscr{C}[a,b]$. Moreover, the norm of 1_E (or the measure of E) ought to be the sums of the norms of 1_{I_i} (or the sum of the lengths of I_i). We need to be careful though! We may have two different ways of expressing E as a finite disjoint union, say

$$E = \bigcup_{i=1}^{k} I_i \quad \text{and} \quad E = \bigcup_{i'=1}^{k'} I'_{i'}.$$

We must check that if this is the case, then

$$\sum_{i=1}^{k} \text{length}(I_i) = \sum_{i'=1}^{k'} \text{length}(I'_{i'}).$$

It turns out that this is indeed true, but we will leave the proof of this to the reader.

2.2 Algebras

Definition 2.2. Let $\mathcal{A}$ be a nonempty collection of subsets of X. We make the following definitions.

1. We say $\mathcal{A}$ is an **algebra** if it is closed under finite intersections and is closed under complements. To be closed under finite intersections means that if $A, B \in \mathcal{A}$, the $A \cap B \in \mathcal{A}$. To be closed under complements means that if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$. Note that in this case, we automatically have $\emptyset \in \mathcal{A}$. Indeed, since $\mathcal{A}$ is nonempty, we can choose $A \in \mathcal{A}$. Then $\emptyset = A \cap A^c \in \mathcal{A}$. Note also that $\mathcal{A}$ is closed under finite unions. This means that if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$. Indeed, we have

$$\begin{aligned} A \cup B &= ((A \cup B)^c)^c \\ &= (A^c \cap B^c)^c \\ &\in \mathcal{A}. \end{aligned}$$

2. We say $\mathcal{A}$ is a **semialgebra** if it contains the emptyset, is closed under finite intersections, and complements can be expressed by finite disjoint unions of members of $\mathcal{A}$. The last part means if $A \in \mathcal{A}$, then there exists a pairwise disjoint sequence $A_1, \dots, A_n \in \mathcal{A}$ such that $A^c = \bigcup_{i=1}^n A_i$. Here we need include $\emptyset \in \mathcal{A}$ as part of the definition, since we don't necessarily get this from the other two axioms. However something that we do get from the other two aximos is that *relative* complements can be expressed by finite disjoint unions of members of $\mathcal{A}$. Indeed, let $A, B \in \mathcal{A}$. Choose a pairwise disjoint sequence $B_1, \dots, B_n \in \mathcal{A}$ such that $B^c = \bigcup_{i=1}^n B_i$. Then we have

$$\begin{aligned} A \backslash B &= A \cap B^c \\ &= A \cap \left(\bigcup_{i=1}^n B_i \right) \\ &= \bigcup_{i=1}^n A \cap B_i, \end{aligned}$$

   where $A \cap B_1, \dots A \cap B_n$ is a pairwise disjoint sequence of members of $\mathcal{A}$. Clearly every algebra is a semialgebra.

3. We say $\mathcal{A}$ is a σ**-algebra** if it is closed under *countable* intersections and is closed under complements. To be closed under countable intersections means that if (A_n) is a sequence of members of $\mathcal{A}$, the $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$. Clearly every σ-algebra is an algebra.

Remark 2. We typically use $\mathcal{E}$ to denote a semialgebra, $\mathcal{A}$ to denote an algebra, and $\mathcal{M}$ to denote a σ-algebra.

2.2.1 Obtaining an Algebra from a Semialgebra

Proposition 2.4. *Let $\mathcal{E}$ be a semialgebra of subsets of X. Then the collection $\mathcal{A}$ consisting of all sets which are finite disjoint union of sets in $\mathcal{E}$ forms an algebra of sets.*

Proof. Let us show that $\mathcal{A}$ is closed under finite intersections. Let $A, A' \in \mathcal{A}$ and express A and A' as a finite disjoint union of members of $\mathcal{E}$, say

$$A = E_1 \cup \cdots \cup E_n \quad \text{and} \quad A' = E_1' \cup \cdots \cup E_{n'}'.$$

Then we have

$$\begin{aligned}
A \cap A' &= \left(\bigcup_{i=1}^{n} E_i \right) \cap \left(\bigcup_{i'=1}^{n'} E'_{i'} \right) \\
&= \bigcup_{i'=1}^{n'} \left(\left(\bigcup_{i=1}^{n} E_i \right) \cap E'_{i'} \right) \\
&= \bigcup_{i'=1}^{n'} \left(\bigcup_{i=1}^{n} E_i \cap E'_{i'} \right) \\
&= \bigcup_{\substack{1 \leq i \leq n \\ 1 \leq i' \leq n'}} E_i \cap E'_{i'}
\end{aligned}$$

where the union is disjoint since the E_i and $E'_{i'}$ are disjoint from one another whenever $i \neq i'$. Thus $A \cap A' \in \mathcal{A}$, and hence $\mathcal{A}$ is closed under finite intersections.

Now let us show that $\mathcal{A}$ is closed under complements. Let $A \in \mathcal{A}$ and express A as a finite disjoint union of members of $\mathcal{E}$, say

$$A = \bigcup_{i=1}^{n} E_i.$$

For each E_i, express E_i^c as a finite disjoint union of members of E, say

$$E_i^c = \bigcup_{j_i=1}^{n_i} E_{i,j_i}.$$

Then we have

$$\begin{aligned}
A^c &= \left(\bigcup_{i=1}^{n} E_i \right)^c \\
&= \bigcap_{i=1}^{n} E_i^c \\
&= \bigcap_{i=1}^{n} \bigcup_{j_i=1}^{n_i} E_{i,j_i} \\
&= \bigcup_{j_i=1}^{n_i} \bigcap_{i=1}^{n} E_{i,j_i}
\end{aligned}$$

where we were allowed to commute the union with the intersection from the third line to the fourth line since these are finite unions and finite intersections. Also, the union is disjoint since the E_{i,j_i} and E_{i,j'_i} are disjoint from one another whenever $j_i \neq j'_i$. Thus $A^c \in \mathcal{A}$, and hence $\mathcal{A}$ is closed complements. □

2.2.2 Ascendification, Contractifcation, and Disjointification

Let $\mathcal{A}$ be an algebra of subsets of X and let (A_n) be a sequence in $\mathcal{A}$. There are several operations we can perform on (A_n) to obtain a new sequence in $\mathcal{A}$. They are called **ascendification**, **contractification**, and **disjointification**.

Definition 2.3. Let $\mathcal{A}$ be an algebra of subsets of X and let (A_n) be a sequence in $\mathcal{A}$.

1. Define the sequence (B_n) as follows: for all $n \in \mathbb{N}$, we set $B_n = \bigcup_{i=1}^n A_i$. Since $\mathcal{A}$ is closed under finite unions, it is clear that (B_n) is a sequence in $\mathcal{A}$. The (B_n) is an **ascending sequence**, which means $B_n \subseteq B_{n+1}$ for all $n \in \mathbb{N}$. We call (B_n) the **ascendification** of the sequence (A_n). Note that
$$\bigcup_{n=1}^N A_n = \bigcup_{n=1}^N B_n$$
for all $N \in \mathbb{N}$, and thus in particular, we have
$$\bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty B_n.$$

2. Define the sequence (C_n) as follows: for all $n \in \mathbb{N}$, we set $C_n = \bigcap_{i=1}^n A_i$. Since $\mathcal{A}$ is closed under finite intersections, it is clear that (C_n) is a sequence in $\mathcal{A}$. The (C_n) is a **contracting sequence**, which means $C_n \supseteq C_{n+1}$ for all $n \in \mathbb{N}$. We call (C_n) the **ascendification** of the sequence (A_n). Note that
$$\bigcap_{n=1}^N A_n = \bigcap_{n=1}^N C_n$$
for all $N \in \mathbb{N}$, and thus in particular, we have
$$\bigcap_{n=1}^\infty A_n = \bigcap_{n=1}^\infty C_n.$$

3. Define the sequence (D_n) as follows: for all $n \in \mathbb{N}$, we set $D_n = A_n \setminus \left(\bigcup_{i=1}^{n-1} A_i\right)$. Since $\mathcal{A}$ is closed under finite unions and is closed under relative complements, it is clear that (D_n) is a sequence in $\mathcal{A}$. The (D_n) is a **pairwise disjoint sequence**, which means $D_m \cap D_n = \emptyset$ whenever $m \neq n$. . We call (D_n) the **disjointification** of the sequence (A_n). Note that
$$\bigcup_{n=1}^N A_n = \bigcup_{n=1}^N D_n$$
for all $N \in \mathbb{N}$, and thus in particular, we have
$$\bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty D_n.$$

2.2.3 Equivalent Definitions For σ-Algebra

Proposition 2.5. *Let $\mathcal{A}$ be an algebra of subsets of X. Then $\mathcal{A}$ is a σ-algebra if and only if $\mathcal{A}$ is closed under ascending unions: if (A_n) is an ascending sequence of members of $\mathcal{A}$, then $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$.*

Proof. Clearly, if $\mathcal{A}$ is a σ-algebra, then it is closed under ascending unions since it is closed under countable unions. Conversely, suppose $\mathcal{A}$ is an algebra and that is is closed under ascending unions. Let (A_n) be a sequence in $\mathcal{A}$. Let (B_n) be the ascendification of the sequence (A_n). Then we have
$$\bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty B_n \in \mathcal{A}.$$
Thus $\mathcal{A}$ is closed under countable unions. It follows that $\mathcal{A}$ is a σ-algebra. □

2.2.4 Generating σ-Algebra from a Collection of Subsets

Proposition 2.6. *Let X be a set and let $\mathcal{C}$ be a nonempty collection of subsets of X. Then there exists a smallest σ-algebra which contains $\mathcal{C}$. It is called the **σ-algebra generated by** $\mathcal{C}$ and is denoted by $\sigma(\mathcal{C})$.*

Proof. Let $\mathscr{F}$ be the family of all σ-algebras $\mathcal{F}$ such that $\mathcal{F}$ contains $\mathcal{C}$. The family $\mathscr{F}$ is nonempty since the power set $\mathcal{P}(X)$ of X is a σ-algebra which contains $\mathcal{C}$. Define

$$\sigma(\mathcal{C}) := \bigcap_{\mathcal{F} \in \mathscr{F}} \mathcal{F}.$$

We claim that $\sigma(\mathcal{C})$ is the smallest σ-algebra which contains $\mathcal{C}$.

Let us first show that $\sigma(\mathcal{C})$ is a σ-algebra. Let (A_n) be a sequence of sets in $\sigma(\mathcal{C})$ and let A be a set in $\sigma(\mathcal{C})$. Then (A_n) is a sequence of sets in $\mathcal{F}$ and A is a set in $\mathcal{F}$ for all $\mathcal{F} \in \mathscr{F}$. Therefore $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ and $X \backslash A \in \mathcal{F}$ for all $\mathcal{F} \in \mathscr{F}$ (as each $\mathcal{F}$ is a σ-algebra). Therefore $\bigcup_{n=1}^{\infty} A_n \in \sigma(\mathcal{C})$ and $X \backslash A \in \sigma(\mathcal{C})$.

Now we will show that $\sigma(\mathcal{C})$ is the smallest algebra which contains $\mathcal{C}$. Suppose Σ' is a another σ-algebra which contains $\mathcal{C}$. Then $\Sigma' \in \mathscr{F}$, hence

$$\sigma(\mathcal{C}) \subseteq \bigcap_{\mathcal{F} \in \mathscr{F}} \mathcal{F} \subseteq \Sigma'.$$

□

Definition 2.4. The smallest σ-algebra containing $\mathcal{I}$ is called the **Borel σ-algebra** and the elements of this σ-algebra are called **Borel sets**.

2.3 Premeasures and Measures

Definition 2.5. Let $\mathcal{A}$ be an algebra of subsets of X and let $\mu \colon \mathcal{A} \to [0, \infty]$. We make the following definitions:

1. The function μ is **monotone** if $\mu(A) \leq \mu(B)$ whenever $A, B \in \mathcal{A}$ such that $A \subseteq B$. In this case, we say μ is **finite** if $\mu(X) < \infty$. Thus if μ is finite, we have $\mu(A) < \infty$ for all $A \in \mathcal{A}$ by monotonicity of μ.

2. The function μ is **finitely additive** if $\mu(\varnothing) = 0$ and

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

for all $A, B \in \mathcal{A}$ such that $A \cap B = \varnothing$. In this case, μ is automatically monotone. Indeed, assume that $A, B \in \mathcal{A}$ such that $A \subseteq B$. Then

$$\begin{aligned} \mu(B) &= \mu((B \backslash A) \cup A) \\ &= \mu(B \backslash A) + \mu(A). \end{aligned}$$

Since $\mu(B \backslash A) \in [0, \infty]$, we conclude that $\mu(A) \leq \mu(B)$. If moreover $\mu(A) < \infty$, then we may subtract $\mu(A)$ from both sides to obtain

$$\mu(B \backslash A) = \mu(B) - \mu(A).$$

3. The function μ is **countably subadditive** if $\mu(\varnothing) = 0$ and

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

for all sequences (A_n) in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

4. The function μ is **countably additive** if $\mu(\varnothing) = 0$ and

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

for all pairwise disjoint sequences (A_n) in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. In this case, we call the function μ a **premeasure** and we call the triple $(X, \mathcal{A}, \mu)$ a **premeasure space**. If $\mathcal{A}$ is a σ-algebra, then we call μ a **measure** and we call the triple $(X, \mathcal{A}, \mu)$ a **measure space**.

2.3.1 Equivalent Definitions for Premeasure

Proposition 2.7. *Let $\mathcal{A}$ be an algebra of subsets of X and let $\mu\colon \mathcal{A} \to [0, \infty]$. The following statements are equivalent.*

1. *μ is a premeasure;*
2. *μ is finitely additive and countably subadditive;*
3. *μ is finitely additive and satisfies*
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n). \tag{4}$$
for all ascending sequences (A_n) in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Proof. We first show 1 implies 2. Suppose that μ is a premeasure. Then $\mu(\emptyset) = 0$ by definition. Let $A, B \in \mathcal{A}$ such that $A \cap B = \emptyset$. Then set $A_1 = A$, $A_2 = B$, and $A_n = \emptyset$ for all $n \geq 3$. Then

$$\begin{aligned}
\mu(A \cup B) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) \\
&= \sum_{n=1}^{\infty} \mu(A_n) \\
&= \mu(A_1) + \mu(A_2).
\end{aligned}$$

Thus μ is finitely additive. It is also countably subadditive. Indeed, let (A_n) be any sequence in $\mathcal{A}$. Disjointify the sequence (A_n) to the sequence (D_n): set $D_1 = A_1$ and $D_n = A_n \setminus \bigcup_{m=1}^{n-1} A_m$ for all $n > 1$. Then

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} D_n\right) \\
&= \sum_{n=1}^{\infty} \mu(D_n) \\
&\leq \sum_{n=1}^{\infty} \mu(A_n).
\end{aligned}$$

Thus μ is countably subadditive.

Now we show 2 implies 1. Suppose that μ is finitely additive and countably subadditive. Then $\mu(\emptyset) = 0$ by definition. Let (A_n) be a sequence of pairwise disjoint sets in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Countable subadditivity of μ gives us

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

For the reverse inequality, observe that for each $N \in \mathbb{N}$, finite additivity of μ gives us

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{\infty} A_n\right) &\geq \mu\left(\bigcup_{n=1}^{N} A_n\right) \\
&= \sum_{n=1}^{N} \mu(A_n).
\end{aligned}$$

Taking $N \to \infty$, we find that

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

Thus μ is a premeasure.

Now we will show 1 implies 3. Suppose μ is a premeasure. Let $A, B \in \mathcal{A}$ such that $A \cap B = \emptyset$. Then set $A_1 = A$, $A_2 = B$, and $A_n = \emptyset$ for all $n \geq 3$. Then

$$\begin{aligned} \mu(A \cup B) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) \\ &= \sum_{n=1}^{\infty} \mu(A_n) \\ &= \mu(A_1) + \mu(A_2). \end{aligned}$$

Thus μ is finitely additive. Next let (A_n) be an ascending sequence in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Disjointify (A_n) into the sequence (D_n): let $D_1 = A_1$ and $D_n = A_n \backslash \bigcup_{m=1}^{n-1} A_m$ for all $n \in \mathbb{N}$. Then we have

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} D_n\right) \\ &= \sum_{n=1}^{\infty} \mu(D_n) \\ &= \lim_{m\to\infty} \sum_{m=1}^{n} \mu(D_m) \\ &= \lim_{m\to\infty} \mu\left(\bigcup_{m=1}^{n} D_m\right) \\ &= \lim_{n\to\infty} \mu(A_n). \end{aligned}$$

Therefore μ satisfies (4).

Now we will show 3 implies 1. Suppose μ is finitely additive and satisfied (4). Let (A_n) be a sequence of pairwise disjoint sets in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Construct an ascending sequence (B_n) in $\mathcal{A}$ as follows: we set $B_1 = A_1$ and $B_n = \bigcup_{m=1}^{n} A_n$ for all $n \in \mathbb{N}$. Clearly $B_n \in \mathcal{A}$ for all $n \in \mathbb{N}$, and their union is equal to $\bigcup_{n=1}^{\infty} A_n$. Thus

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} B_n\right) \\ &= \lim_{n\to\infty} \mu(B_n) \\ &= \lim_{n\to\infty} \mu\left(\bigcup_{m=1}^{n} A_m\right) \\ &= \lim_{n\to\infty} \sum_{m=1}^{n} \mu(A_m) \\ &= \sum_{m=1}^{\infty} \mu(A_m). \end{aligned}$$

Thus μ is a premeasure. □

2.3.2 Measure of descending sequences may not commute with limits

It seems reasonable to expect that, if (X, Σ, μ) is a measure space and (A_n) is a descending sequence of sets in Σ, then

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

Unfortunately, this assertion can fail to be true. For example, consider the case where $X = \mathbb{Z}_{\geq 0}$, $\Sigma = \mathcal{P}(\mathbb{Z}_{\geq 0})$, and μ is the counting measure. Define $A_m := \{n \in \mathbb{Z}_{\geq 0} \mid n \geq m\}$ for all $m \in \mathbb{N}$. Then

$$\begin{aligned}
\lim_{m\to\infty} \mu(A_m) &= \lim_{m\to\infty} \infty \\
&= \infty \\
&\neq 0 \\
&= \mu(\emptyset) \\
&= \mu\left(\bigcap_{m=1}^{\infty} A_m\right).
\end{aligned}$$

On the other hand, there is a positive statement we can make about descending sequences. This is done in the following proposition.

Proposition 2.8. *Let (X, Σ, μ) be a measure space and let (A_n) be a descending sequence of sets in Σ such that $\mu(A_1) < \infty$. Then*

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

Proof. The sequence $(A_1 \backslash A_n)$ is an ascending sequence, hence

$$\begin{aligned}
\mu(A_1) - \lim_{n\to\infty} \mu(A_n) &= \lim_{n\to\infty} \mu(A_1) - \lim_{n\to\infty} \mu(A_n) \\
&= \lim_{n\to\infty} (\mu(A_1) - \mu(A_n)) \\
&= \lim_{n\to\infty} \mu(A_1 \backslash A_n) \\
&= \mu\left(\bigcup_{n=1}^{\infty} (A_1 \backslash A_n)\right) \\
&= \mu\left(A_1 \backslash \left(\bigcap_{n=1}^{\infty} A_n\right)\right) \\
&= \mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right),
\end{aligned}$$

where we used the fact that $\mu(A_1) < \infty$ to get from line 2 to line 3. Also since $\mu(A_1) < \infty$, we can subtract $\mu(A_1)$ from both sides to obtain

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

□

2.3.3 Outer Measure

Definition 2.6. Let $\mu \colon \mathcal{P}(X) \to [0, \infty]$. We say μ is an **outer measure** if μ is monotone and countably subadditive.

Proposition 2.9. *Let $(X, \mathcal{A}, \mu)$ be a premeasure space. Define $\mu^* \colon \mathcal{P}(X) \to [0, \infty]$ by*

$$\mu^*(S) = \inf\left\{\sum_{n=1}^{\infty} \mu(A_n) \mid \{A_n\} \subseteq \mathcal{A} \textit{ covers } S\textit{, that is, } S \subseteq \bigcup_{n=1}^{\infty} A_n\right\}.$$

Then μ^* is an outer measure.

Proof. We first show μ^* is monotone. Let $S, T \in \mathcal{P}(X)$ such that $S \subseteq T$. Suppose that $\{A_n\} \subseteq \mathcal{A}$ covers T. Then $\{A_n\} \subseteq \mathcal{A}$ covers S too since $S \subseteq T$. Since the covering $\{A_n\}$ was arbitrary, we see that

$$\begin{aligned}
\mu^*(S) &\leq \inf\left\{\sum_{n=1}^{\infty} \mu(A_n) \mid \{A_n\} \subseteq \mathcal{A} \text{ covers } T\right\} \\
&= \mu^*(T).
\end{aligned}$$

Now we will show that μ^* is countably subadditive. First observe that $\{\emptyset\}$ is a covering of $\emptyset$. Thus

$$\begin{aligned} 0 &\leq \mu^*(\emptyset) \\ &\leq \mu(\emptyset) \\ &= 0 \end{aligned}$$

implies $\mu^*(\emptyset) = 0$. Now let (S_n) be a sequence in $\mathcal{P}(X)$ and let $\varepsilon > 0$. For each $n \in \mathbb{N}$ choose a covering $\{A_{n,k}\}_{k \in \mathbb{N}} \subseteq \mathcal{A}$ of S_n such that

$$\sum_{k=1}^{\infty} \mu(A_{n,k}) \leq \mu^*(S_n) + \frac{\varepsilon}{2^n}.$$

Then observe that $\{A_{n,k}\}_{k,n \in \mathbb{N}} \subseteq \mathcal{A}$ is a covering of $\bigcup_{n=1}^{\infty} S_n$, and so we have

$$\begin{aligned} \mu^*\left(\bigcup_{n=1}^{\infty} S_n\right) &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(A_{n,k}) \\ &\leq \sum_{n=1}^{\infty} \left(\mu^*(S_n) + \frac{\varepsilon}{2^n}\right) \\ &= \sum_{n=1}^{\infty} \mu^*(S_n) + \varepsilon. \end{aligned}$$

Taking $\varepsilon \to 0$ gives us our desired result. □

3 Extending Finite Premeasures

Throughout this section, let $(X, \mathcal{A}, \mu)$ be a finite premeasure space. Our goal in this section is to show that μ can be *uniquely* extended to a measure on $\sigma(\mathcal{A})$. In other words, there exists a measure $\widetilde{\mu}$ on $\supset(\mathcal{A})$ such that $\widetilde{\mu}|_{\mathcal{A}} = \mu$. Furthermore, if ν is any other measure on $\sigma(\mathcal{A})$ such that $\nu|_{\mathcal{A}} = \mu$, then $\widetilde{\mu} = \nu$. Let us state this as a theorem up front.

Theorem 3.1. *(Caratheodory, Hahn, Kolmogorov) Let $(X, \mathcal{A}, \mu)$ be a finite premeasure space. Then μ has a unique extension to a measure on $\sigma(\mathcal{A})$.*

Our proof of Theorem (3.1) will involve Topology, where much is known about unique extensions of continuous functions. In particular, we will realize $\mathcal{A}$ as a topological space and we will realize $\mu \colon \mathcal{A} \to [0, \infty)$ as a continuous function defined on $\mathcal{A}$. In this setting, it will be easy to show that μ can be uniquely extended to a continuous function on the larger space $\sigma(\mathcal{A})$.

3.1 Defining a Pseudometric on $\mathcal{P}(X)$

The first step is to construct a pseudometric on $\mathcal{P}(X)$.[2] In particular, we define $\mathrm{d}_\mu \colon \mathcal{P}(X) \times \mathcal{P}(X) \to [0, \infty]$ by

$$\mathrm{d}_\mu(A, B) = \mu^*(A \Delta B)$$

for all $A, B \in \mathcal{P}(X)$.

Proposition 3.1. *d_μ is a pseudometric on $\mathcal{P}(X)$.*

Proof. We first check reflexivity of d_μ. Let $A \in \mathcal{P}(X)$, then have

$$\begin{aligned} \mathrm{d}_\mu(A, A) &= \mu^*(A \Delta A) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

Next we check symmetry of d_μ. Let $A, B \in \mathcal{P}(X)$. Then we have

$$\begin{aligned} \mathrm{d}_\mu(A, B) &= \mu^*(A \Delta B) \\ &= \mu^*(B \Delta A) \\ &= \mathrm{d}_\mu(B, A). \end{aligned}$$

[2]See the Appendix for more details on pseudometric spaces.

Finally, we check triangle inequality. Let $A, B, C \in \mathcal{P}(X)$. Then we have

$$\begin{aligned} \mathrm{d}_\mu(A, C) &= \mu^*(A \Delta C) \\ &= \mu^*(A \Delta B \Delta B \Delta C) \\ &\leq \mu^*((A \Delta B) \cup (B \Delta C)) \\ &\leq \mu^*(A \Delta B) + \mu^*(B \Delta C) \\ &= \mathrm{d}_\mu(A, B) + \mathrm{d}_\mu(B, C), \end{aligned}$$

where we obtained the third line from the second line by monotonicity of μ^*, and where we obtained the fourth line from the third line by finite subadditivity of μ^*. □

3.1.1 Metric Space Induced by Pseudometric Space

Proposition (3.1) tells us that $(\mathcal{P}(X), \mathrm{d}_\mu)$ is a pseudometric space. The reason d_μ is a pseudometric and not a metric is because we not have identity of indiscernibles: we may have $\mu^*(A \Delta B) = 0$ with $A \neq B$. All is not lost however as every pseudometric space induces a metric space in a natural way. Let us briefly describe the metric space induced by the pseudometric space $(\mathcal{P}(X), \mathrm{d}_\mu)$. More details can be found in the Appendix. We introduce an equivalence relation $\sim$ on $\mathcal{P}(X)$ as follows: let $A, B \in \mathcal{P}(X)$. Then

$$A \sim B \text{ if and only if } \mathrm{d}_\mu(A, B) = 0.$$

One checks that $\sim$ is an equivalence relation on $\mathcal{P}(X)$ and so we may consider quotient space

$$[\mathcal{P}(X)] := \mathcal{P}(X)/\sim.$$

We shall use the notation $[A]$ to denote a coset in $[A]$ with $A \in \mathcal{P}(X)$ as a particular representative. We define a metric $[\mathrm{d}_\mu]$ on $[\mathcal{P}(X)]$ by

$$[\mathrm{d}_\mu]([A], [B]) = \mathrm{d}_\mu(A, B) \tag{5}$$

One checks that (5) is well-defined and satisfies all of the properties required for it to be a metric. Furthermore, one shows that the quotient topology on $[\mathcal{P}(X)]$ is the same as the topology induced by the metric $[\mathrm{d}_\mu]$. In particular, the projection map

$$\pi\colon \mathcal{P}(X) \to [\mathcal{P}(X)]$$

is continuous, and for any topological space Y (such as $[0, \infty]$!) we have a bijection

$$\left\{ \begin{array}{c} \text{continuous functions from } \mathcal{P}(X) \text{ to } Y \\ \text{which are constant on equivalence classes} \end{array} \right\} \cong \{\text{continuous functions from } [\mathcal{P}(X)] \text{ to } Y\}.$$

Indeed, if $\nu\colon [\mathcal{P}(X)] \to Y$ is continuous, then the function $\nu \circ \pi\colon \mathcal{P}(X) \to Y$ is continuous since it is a composition of continuous functions and it is constant on equivalence classes: if $A \sim B$, then

$$\begin{aligned} (\nu \circ \pi)(A) &= \nu(\pi(A)) \\ &= \nu(\pi(B)) \\ &= (\nu \circ \pi)(B). \end{aligned}$$

Convsersely, if $\eta\colon \mathcal{P}(X) \to Y$ is continuous and constant on equivalence classes, then it induces a unique continuous function $\nu\colon [\mathcal{P}(X)] \to Y$ such that $\nu \circ \pi = \eta$.

There many other properties which are both shared by $(\mathcal{P}(X), \mathrm{d}_\mu)$ and $([\mathcal{P}(X)], [\mathrm{d}_\mu])$. For instance, $(\mathcal{P}(X), \mathrm{d}_\mu)$ is complete if and only if $([\mathcal{P}(X)], [\mathrm{d}_\mu])$ complete. For this and many other reasons, we choose to work in the pseudometric space $(\mathcal{P}(X), \mathrm{d}_\mu)$ rather than the metric space $([\mathcal{P}(X)], [\mathrm{d}_\mu])$.

3.1.2 Complement Map is Isometry

Proposition 3.2. *The complement map $-^c\colon \mathcal{P}(X) \to \mathcal{P}(X)$ given by*

$$-^c(A) = A^c$$

for all $A \in \mathcal{P}(X)$ is an isometry on $[\mathcal{P}(X)]$.

Proof. We first check that $-^c$ is constant on equivalence classes. Suppose $A, A' \in \mathcal{P}(X)$ with $A \sim A'$ (so $A \Delta A' = \emptyset$). Then

$$\begin{aligned} A^c \Delta A'^c &= A \Delta A \\ &= \emptyset. \end{aligned}$$

Thus $A^c \sim A'^c$, and so the complement map is constant on equivalence classes. Now we check that it is an isometry. Let $A, B \in \mathcal{P}(X)$. Then

$$\begin{aligned} \mathrm{d}_\mu(A, B) &= \mu^*(A \Delta B) \\ &= \mu^*(A^c \Delta B^c) \\ &= \mathrm{d}_\mu(A^c, B^c). \end{aligned}$$

Thus $-^c$ is an isometry. □

3.1.3 Continuity of Finite Unions and Finite Intersections

Proposition 3.3. *The union map* $\cup \colon \mathcal{P}(X) \times \mathcal{P}(X) \to \mathcal{P}(X)$, *defined by*

$$\cup(A, B) = A \cup B$$

for all $(A, B) \in \mathcal{P}(X)$, *is continuous on* $[\mathcal{P}(X)] \times [\mathcal{P}(X)]$.

Proof. We first check that the union map is constant on equivalence classes. Suppose $A \sim A'$ and $B \sim B'$ where $A, A', B, B' \in \mathcal{P}(X)$. Thus $A \Delta A' = 0$ and $B \Delta B' = 0$. Then

$$\begin{aligned} (A \cup B) \Delta (A' \cup B') &\subseteq (A \Delta A') \cup (B \Delta B') \\ &= \emptyset \cup \emptyset \\ &= \emptyset. \end{aligned}$$

It follows that $A \cup B \sim A' \cup B'$. Now we will show that the union map is continuous. Suppose $A_n \to A$ and $B_n \to B$ in $(\mathcal{P}(X), \mathrm{d}_\mu)$. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that

$$\mathrm{d}_\mu(A_n, A) < \frac{\varepsilon}{2} \quad \text{and} \quad \mathrm{d}_\mu(B_n, B) < \frac{\varepsilon}{2}.$$

Then $n \geq N$ implies

$$\begin{aligned} \mathrm{d}_\mu((A_n \cup B_n), (A \cup B)) &= \mu^*((A_n \cup B_n) \Delta (A \cup B)) \\ &= \mu^*((A_n \cup B_n) \Delta (A \cup B)) \\ &\leq \mu^*((A_n \Delta A) \cup (B_n \Delta B)) \\ &\leq \mu^*(A_n \Delta A) + \mu^*(B_n \Delta B) \\ &< \mathrm{d}_\mu(A_n, A) + \mathrm{d}_\mu(B_n, B) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

It follows that the union map is continuous on $[\mathcal{P}(X)] \times [\mathcal{P}(X)]$. □

Proposition 3.4. *The intersection map* $\cap \colon \mathcal{P}(X) \times \mathcal{P}(X) \to \mathcal{P}(X)$, *defined by*

$$\cap(A, B) = A \cap B$$

for all $(A, B) \in \mathcal{P}(X)$, *is continuous on* $[\mathcal{P}(X)] \times [\mathcal{P}(X)]$.

Proof. The intersection is a composition of the union map with the complement map. Thus it is a composition of continuous functions, and hence must be continuous. □

3.1.4 Uniform Continuity of μ^*

Proposition 3.5. *The function $\mu^* \colon \mathcal{P}(X) \to [0,\infty]$ is Lipschitz continuous on $[\mathcal{P}(X)]$.*

Proof. Let $A, B \in \mathcal{P}(X)$. Then

$$\begin{aligned} |\mu^*(A) - \mu^*(B)| &\leq \max\{\mu^*(A\backslash B), \mu^*(B\backslash A)\} \\ &\leq \mu^*((A\backslash B) \cup (B\backslash A)) \\ &= \mu^*(A\Delta B) \\ &= \mathrm{d}_\mu(A,B). \end{aligned}$$

It follows that μ^* is Lipschitz continuous on $\mathcal{P}(X)$. To see that it is Lipschitz continuous on $[\mathcal{P}(X)]$, we just need to check that it is constant on equivalence classes. Let $A, A' \in \mathcal{P}(X)$ such that $A \sim A'$ (so $\mu^*(A\Delta A') = 0$). Then

$$\begin{aligned} \mu^*(A') &= \mu^*((A\Delta A)\Delta A') \\ &= \mu^*(A\Delta(A\Delta A')) \\ &\leq \mu^*(A \cup (A\Delta A')) \\ &\leq \mu^*(A) + \mu^*(A\Delta A') \\ &= \mu^*(A). \end{aligned}$$

By a similar argument, we also have $\mu^*(A) \geq \mu^*(A')$. Thus μ^* is constant on equivalence classes. □

Remark 3. Since $\mu^*|_{\mathcal{A}} = \mu$, we have also shown that μ is continuous on $[\mathcal{A}]$ and that μ^* is a continuous extension of μ.

3.2 Completion of $(\mathcal{A}, \mathrm{d}_\mu)$

The pseudometric d_μ on $\mathcal{P}(X)$ restricts to a pseudometric on $\sigma(\mathcal{A})$ and it also restricts to a pseudometric on $\mathcal{A}$. Formally, we should denote these restrictions by $\mathrm{d}_\mu|_{\sigma(\mathcal{A})\times\sigma(\mathcal{A})}$ and $\mathrm{d}_\mu|_{\mathcal{A}\times\mathcal{A}}$ respectively, however in order to clean notation, we will simply denote these restricts by d_μ. Thus we have the following inclusions of pseudometric spaces

$$(\mathcal{A}, \mathrm{d}_\mu) \subseteq (\sigma(\mathcal{A}), \mathrm{d}_\mu) \subseteq (\mathcal{P}(X), \mathrm{d}_\mu).$$

Our goal in this subsection is to show that $(\mathcal{A}, \mathrm{d}_\mu)$ is dense in $(\sigma(\mathcal{A}), \mathrm{d}_\mu)$ and that $(\sigma(\mathcal{A}), \mathrm{d}_\mu)$ is complete. Thus $(\sigma(\mathcal{A}), \mathrm{d}_\mu)$ is the completion of $(\mathcal{A}, \mathrm{d}_\mu)$. Before doing this, we introduce the limsup and liminf of a sequence of sets.

3.2.1 Limit Supremum and Limit Infimum

Recall that if (a_n) is a sequence of real numbers, we define its limit supremum by the formula

$$\limsup a_n = \inf_{N\geq 1} \sup_{n\geq N} a_n.$$

Similarly, we define its limit infimum by the formula

$$\liminf a_n = \sup_{N\geq 1} \inf_{n\geq N} a_n.$$

There is an analagous notion of limit supremum and limit infimum of a sequence of sets.

Definition 3.1. Let (A_n) be a sequence of sets. The **limit supremum** of (A_n) is defined to be

$$\limsup A_n = \bigcap_{N\geq 1} \bigcup_{n\geq N} A_n.$$

The **limit infimum** of (A_n) is defined by

$$\liminf A_n = \bigcup_{N\geq 1} \bigcap_{n\geq N} A_n.$$

3.2.2 $(\sigma(\mathcal{A}), \mathrm{d}_\mu)$ is complete

We now want to show that $(\sigma(\mathcal{A}), \mathrm{d}_\mu)$ is complete. In fact, we will do better than this. We will show that (Σ, d_μ) is complete, where Σ is any σ-algebra which contains $\mathcal{A}$. Recall that this means that every Cauchy sequence in (Σ, d_μ) converges to a limit in (Σ, d_μ) . Before we prove this, we need to establish two lemmas.

Lemma 3.2. *Let (X, d) be a pseudometric space and let (x_n) be a Cauchy sequence in X. Suppose there exists a subsequence $(x_{\pi(n)})$ of the sequence (x_n) such that $x_{\pi(n)} \to x$ for some $x \in X$. Then $x_n \to x$.*

Proof. Let $\varepsilon > 0$. Since $(x_{\pi(n)})$ is convergent, there exists an $N \in \mathbb{N}$ such that $\pi(n) \geq N$ implies

$$\mathrm{d}(x_{\pi(n)}, x) < \frac{\varepsilon}{2}.$$

Since (x_n) is Cauchy, there exists $M \in \mathbb{N}$ such that $m, n \geq M$ implies

$$\mathrm{d}(x_m, x_n) < \frac{\varepsilon}{2}.$$

Choose such M and N and assume without loss of generality that $N \geq M$. Then $\pi(n) \geq n \geq N$ implies

$$\begin{aligned} \mathrm{d}(x_n, x) &\leq \mathrm{d}(x_{\pi(n)}, x_n) + \mathrm{d}(x_{\pi(n)}, x) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

It follows that $x_n \to x$. □

Lemma 3.3. *Let (A_n) be a sequence in $\mathcal{P}(X)$. Then*

$$\bigcup_{n=1}^{\infty} (A_n \Delta A_{n+1}) = \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m.$$

Proof. Suppose $x \in \bigcup_{n=1}^{\infty} (A_n \Delta A_{n+1})$. Choose $n \in \mathbb{N}$ such that $x \in A_n \Delta A_{n+1}$. Thus either $x \in A_n \backslash A_{n+1}$ or $x \in A_{n+1} \backslash A_n$. Without loss of generality, say $x \in A_n \backslash A_{n+1}$. Then since $x \in A_n$, we see that $x \in \bigcup_{n=1}^{\infty} A_n$ and since $x \notin A_{n+1}$, we see that $x \notin \bigcap_{n=1}^{\infty} A_n$. Therefore $x \in \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m$. This implies

$$\bigcup_{n=1}^{\infty} (A_n \Delta A_{n+1}) \subseteq \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m.$$

Conversely, suppose $x \in \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m$. Since $x \in \bigcup_{n=1}^{\infty} A_n$, there exists some $n \in \mathbb{N}$ such that $x \in A_n$. Since $x \notin \bigcap_{m=1}^{\infty} A_m$, there exists some $k \in \mathbb{N}$ such that $x \notin A_k$. Assume without loss of generality that $k < n$. Choose m to be the least natural number number such that $x \in A_m$, $x \notin A_{m-1}$, and $k < m \leq n$. Clearly this number exists since $x \notin A_k$ and $x \in A_n$. Then $x \in A_m \Delta A_{m-1}$, which implies $x \in \bigcup_{n=1}^{\infty} (A_n \Delta A_{n+1})$. Thus

$$\bigcup_{n=1}^{\infty} (A_n \Delta A_{n+1}) \supseteq \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m.$$

□

Theorem 3.4. *Let Σ be any σ-algebra which contains $\mathcal{A}$. Then (Σ, d_μ) is complete.*

Proof. Let (A_n) be a Cauchy sequence in (Σ, d_μ) . To show that (A_n) is convergent, it suffices to show that a subsequence of (A_n) is convergent, by Lemma (12.1). We construct such a subsequence as follows: Since (A_n) is Cauchy, for each $n \geq 1$ there exists a $\pi(n) \geq n$ such that $k, m \geq \pi(n)$ implies

$$\mathrm{d}_\mu(A_k, A_m) < \frac{1}{2^n}.$$

Choose such $\pi(n) \geq n$ such that $m < n$ implies $\pi(m) < \pi(n)$. In particular, for each $n \geq 1$, we have

$$\mathrm{d}_\mu(A_{\pi(n)}, A_{\pi(n+1)}) < \frac{1}{2^n}. \tag{6}$$

By passing to the subsequence $(A_{\pi(n)})$ of (A_n) if necessary, we may as well assume that

$$\mathrm{d}_\mu(A_n, A_{n+1}) < \frac{1}{2^n}$$

for all $n \geq 1$. Now set $A = \limsup A_n$. We claim that $A_n \to A$. Indeed, let $\varepsilon > 0$ and choose $N \geq 1$ such that $2^{2-N} < \varepsilon$. Then $n \geq N$ implies

$$\begin{aligned}
\mathrm{d}_\mu(A, A_n) &= \mu^*(A \Delta A_n) \\
&\leq \mu^*(A \backslash A_n) + \mu^*(A_n \backslash A) && \text{finite subadditivity of } \mu^* \\
&\leq \mu^*\left(\bigcup_{m \geq N} A_m \backslash \bigcap_{m \geq N} A_m\right) + \mu^*\left(\bigcup_{m \geq N} A_m \backslash \bigcap_{m \geq N} A_m\right) && \text{monotonicity of } \mu^* \\
&= 2\mu^*\left(\bigcup_{m \geq N} A_m \backslash \bigcap_{m \geq N} A_m\right) \\
&= 2\mu^*\left(\bigcup_{m \geq N} (A_m \Delta A_{m+1})\right) \\
&\leq 2 \sum_{m \geq N} \mu^*(A_m \Delta A_{m+1}) \\
&= 2 \sum_{m \geq N} \mathrm{d}_\mu(A_m \Delta A_{m+1}) \\
&< 2 \sum_{m \geq N} \frac{1}{2^m} \\
&= 2^{2-N} \\
&< \varepsilon.
\end{aligned}$$

It follows that (A_n) converges to A. Finally, note that

$$\begin{aligned}
A &= \limsup A_n \\
&= \bigcap_{N \geq 1} \bigcup_{n \geq N} A_n \\
&\in \Sigma.
\end{aligned}$$

Thus every Cauchy sequence of sets in (Σ, d_μ) converges to a set in (Σ, d_μ). Therefore (Σ, d_μ) is complete. □

3.2.3 $(\mathcal{A}, \mathrm{d}_\mu)$ is dense in $(\sigma(\mathcal{A}), \mathrm{d}_\mu)$

Now we want to show that $(\mathcal{A}, \mathrm{d}_\mu)$ is dense subset in $(\sigma(\mathcal{A}), \mathrm{d}_\mu)$. In other words, we want to show that $\overline{\mathcal{A}} = \sigma(\mathcal{A})$, where $\overline{\mathcal{A}}$ denotes the closure of $\mathcal{A}$ in the pseudometric space $(\mathcal{P}(X), \mathrm{d}_\mu)$. Recall that $\overline{\mathcal{A}}$ is the *smallest* closed set which contains $\mathcal{A}$. Since $\sigma(\mathcal{A})$ is complete, it is certainly closed in $(\mathcal{P}(X), \mathrm{d}_\mu)$, and so we have

$$\overline{\mathcal{A}} \subseteq \sigma(\mathcal{A}).$$

On the other hand, $\sigma(\mathcal{A})$ is the *smallest* σ-algebra which contains $\mathcal{A}$. Thus if we can show that $\overline{\mathcal{A}}$ is a σ-algebra, then we will have the reverse inclusion.

Proposition 3.6. *$\overline{\mathcal{A}}$ is a σ-algebra.*

Proof. The proof consists of three steps.

Step 1: We first show that $\overline{\mathcal{A}}$ is an algebra. We have $\emptyset \in \mathcal{A} \subseteq \overline{\mathcal{A}}$. Next, let $A \in \overline{\mathcal{A}}$. Choose a sequence (A_n) in $\mathcal{A}$ such that $A_n \to A$. Then since taking complements is continuous, we have $A_n^c \to A^c$, which implies $A^c \in \overline{\mathcal{A}}$. Finally, let $A, B \in \overline{\mathcal{A}}$. Choose sequences (A_n) and (B_n) in $\mathcal{A}$ such that $A_n \to A$ and $B_n \to B$. Then since taking unions is continuous, we have $A_n \cup B_n \to A \cup B$, which implies $A \cup B \in \overline{\mathcal{A}}$.

Step 2: We show that μ^* restricts to a measure on $\overline{\mathcal{A}}$. It suffices to show that μ^* is finitely additive on $\overline{\mathcal{A}}$ since we already know it is countably subadditive. Let A and B be two disjoint members of $\overline{\mathcal{A}}$ and choose sequences

(A_n) and (B_n) in $\mathcal{A}$ such that $A_n \to A$ and $B_n \to B$. Then $A_n \cup B_n \to A \cup B$, and so it follows by continuity of μ^* that

$$\begin{aligned}
\mu^*(A \cup B) &= \lim_{n\to\infty} \mu^*(A_n \cup B_n) \\
&= \lim_{n\to\infty} \mu(A_n \cup B_n) \\
&= \lim_{n\to\infty} (\mu(A_n) + \mu(B_n) - \mu(A_n \cap B_n)) \\
&= \lim_{n\to\infty} (\mu^*(A_n) + \mu^*(B_n) - \mu^*(A_n \cap B_n)) \\
&= \mu^*(A) + \mu^*(B) - \mu^*(A \cap B) \\
&= \mu^*(A) + \mu^*(B) - \mu^*(\emptyset) \\
&= \mu^*(A) + \mu^*(B).
\end{aligned}$$

where we needed to use the fact that μ is a finite measure in order to get the third line from the second line.

Step 3: We now show that $\overline{\mathcal{A}}$ is a σ-algebra. Let (A_n) be a sequence of sets in $\overline{\mathcal{A}}$. We want to show $\bigcup_{n=1}^\infty A_n \in \overline{\mathcal{A}}$. To do this, we will shows that $\bigcup_{n=1}^N A_n \to \bigcup_{n=1}^\infty A_n$ as $N \to \infty$. Then since $\bigcup_{n=1}^N A_n \in \overline{\mathcal{A}}$ since $\overline{\mathcal{A}}$ is an algebra by step 2, it will then follow that $\bigcup_{n=1}^\infty A_n \in \overline{\mathcal{A}}$.

Disjointify the sequence (A_n) to the sequence (B_n): set $B_1 = A_1$ and $B_n = A_n \backslash \bigcup_{i=1}^{n-1} A_i$ for all $n > 1$. Then (B_n) is a sequence of pairwise disjoint sets in $\overline{\mathcal{A}}$ since $\overline{\mathcal{A}}$ is an algebra by step 2. Moreover, for each $N \in \mathbb{N}$, we have

$$\begin{aligned}
\mathrm{d}_\mu \left(\bigcup_{n=1}^\infty A_n, \bigcup_{n=1}^N A_n \right) &= \mathrm{d}_\mu \left(\bigcup_{n=1}^\infty B_n, \bigcup_{n=1}^N B_n \right) \\
&= \mu^* \left(\left(\bigcup_{n=1}^\infty B_n \right) \Delta \left(\bigcup_{n=1}^N B_n \right) \right) \\
&= \mu^* \left(\bigcup_{n=N+1}^\infty B_n \right) \\
&= \sum_{n=N+1}^\infty \mu^*(B_n),
\end{aligned}$$

where the last term tends to 0 as $N \to \infty$ since it is the tail of a convergent series:

$$\begin{aligned}
\sum_{n=1}^\infty \mu^*(B_n) &= \mu^* \left(\bigcup_{n=1}^\infty B_n \right) \\
&\leq \mu^*(X) \\
&< \infty.
\end{aligned}$$

Therefore $\bigcup_{n=1}^N A_n \to \bigcup_{n=1}^\infty A_n$ as $N \to \infty$. □

3.3 Uniqueness of Measure

Recall that in the proof of Proposition (3.6) we showed that μ^* restricts to a measure on $\overline{\mathcal{A}} = \sigma(\mathcal{A})$. In particular, since $\mu^*|_{\mathcal{A}} = \mu$, we see that the measure $\mu^*|_{\sigma(\mathcal{A})}$ is an extension of μ to all of $\sigma(\mathcal{A})$. This gives us precisely one example of a measure which extends μ to all of $\sigma(\mathcal{A})$. It is not at all clear however that this is the *only* such extension. If μ is not finite, then there may be more than one extension. Indeed, see problem 6 in homework 2 for an example of this. In our case however where μ is finite, we shall see that this extension is in fact unique. We will use ideas from topology to show this; all of the topological work we just did is about to payoff!

3.3.1 Uniqueness Extensions of Continuous Functions

We first begin with a result from topology.

Proposition 3.7. *Let X be a topological space, let A be a dense subspace of X, let Y be a Hausdorff space, and let $f\colon A \to Y$ be a continuous function. If there exists a continuous extension of f to all of X, then it must be unique, that is, if $\widetilde{f}_1\colon X \to Y$ and $\widetilde{f}_2\colon X \to Y$ are two continuous functions such that*

$$\widetilde{f}_1|_A = f = \widetilde{f}_2|_A,$$

then $\widetilde{f}_1 = \widetilde{f}_2$.

Proof. Assume for a contradiction that $\widetilde{f}_1\colon X \to Y$ and $\widetilde{f}_2\colon X \to Y$ are two continuous extensions of f such that $\widetilde{f}_1 \neq \widetilde{f}_2$. Choose $x \in X$ such that $\widetilde{f}_1(x) \neq \widetilde{f}_2(x)$. Since Y is Hausdorff, we may choose open neighborhoods V_1 and V_2 of $\widetilde{f}_1(x)$ and $\widetilde{f}_2(x)$ respectively such that $V_1 \cap V_2 = \emptyset$. Then $\widetilde{f}_1^{-1}(V_1) \cap \widetilde{f}_2^{-1}(V_2)$ is an open neighborhood of x, and so it must have a nonempty intersection with A. Choose $a \in A \cap \widetilde{f}_1^{-1}(V_1) \cap \widetilde{f}_2^{-1}(V_2)$. Then

$$f(a) = \widetilde{f}_1(a) \in V_1.$$

Similarly,

$$f(a) = \widetilde{f}_2(a) \in V_2.$$

Thus $f(a) \in V_1 \cap V_2$, which is a contradiction since V_1 and V_2 were chosen to disjoint from one another. □

3.3.2 Continuity of Finite Measure

Lemma 3.5. *Let $\mathcal{A}$ be an algebra and let μ be a measure on $\sigma(\mathcal{A})$. Then*

$$(\mu|_{\mathcal{A}})^*(A) \geq \mu(A)$$

for all $A \in \sigma(\mathcal{A})$.

Proof. Let $A \in \sigma(\mathcal{A})$. Then

$$\begin{aligned}
(\mu|_{\mathcal{A}})^*(A) &= \inf\left\{ \sum_{n=1}^{\infty} (\mu|_{\mathcal{A}})(E_n) \mid (E_n) \text{ is a sequence in } \mathcal{A} \text{ such that } A \subseteq \bigcup_{n=1}^{\infty} E_n \right\} \\
&= \inf\left\{ \sum_{n=1}^{\infty} \mu(E_n) \mid (E_n) \text{ is a sequence in } \mathcal{A} \text{ such that } A \subseteq \bigcup_{n=1}^{\infty} E_n \right\} \\
&\geq \inf\left\{ \mu\left(\bigcup_{n=1}^{\infty} E_n \right) \mid (E_n) \text{ is a sequence in } \mathcal{A} \text{ such that } A \subseteq \bigcup_{n=1}^{\infty} E_n \right\} \\
&\geq \mu(A),
\end{aligned}$$

where we used countable subadditivity of μ to get from the second line to the third line, and where we used monotonicity of μ to get from the third line to the fourth line. □

Proposition 3.8. *Let $\mathcal{A}$ be an algebra and let μ be a finite measure on $\sigma(\mathcal{A})$. Then μ is Lipschitz continuous with respect to $\mathrm{d}_{\mu|_{\mathcal{A}}}$.*

Proof. Let $A, B \in \sigma(\mathcal{A})$. Assume without loss of generality that $\mu(A) \geq \mu(B)$. Then

$$\begin{aligned}
\mu(A) - \mu(B) &\leq \mu(A \backslash B) \\
&\leq \mu((A \backslash B) \cup (B \backslash A)) \\
&= \mu(A \Delta B) \\
&\leq (\mu|_{\mathcal{A}})^*(A \Delta B) \\
&= \mathrm{d}_{\mu|_{\mathcal{A}}}(A, B),
\end{aligned}$$

where we used the fact that μ is finite in the first line. □

3.3.3 Uniqueness of Extension for Finite Measures

Theorem 3.6. *Let μ and ν be two finite measures defined on $\sigma(\mathcal{A})$ which coincide on $\mathcal{A}$. Then $\mu = \nu$.*

Proof. We first note that $\mathrm{d}_{\mu|_{\mathcal{A}}} = \mathrm{d}_{\nu|_{\mathcal{A}}}$ since μ and ν agree on $\mathcal{A}$. Indeed, let $A, B \in \sigma(\mathcal{A})$. Then we have

$$\begin{aligned}
\mathrm{d}_{\mu|_{\mathcal{A}}}(A,B) &= (\mu|_{\mathcal{A}})^*(A\Delta B) \\
&= \inf\left\{\sum_{n=1}^{\infty}(\mu|_{\mathcal{A}})(E_n) \mid (E_n) \text{ is a sequence in } \mathcal{A} \text{ such that } A\Delta B \subseteq \bigcup_{n=1}^{\infty} E_n\right\} \\
&= \inf\left\{\sum_{n=1}^{\infty}(\nu|_{\mathcal{A}})(E_n) \mid (E_n) \text{ is a sequence in } \mathcal{A} \text{ such that } A\Delta B \subseteq \bigcup_{n=1}^{\infty} E_n\right\} \\
&= (\nu|_{\mathcal{A}})^*(A\Delta B) \\
&= \mathrm{d}_{\nu|_{\mathcal{A}}}(A,B).
\end{aligned}$$

Therefore $\mathrm{d}_{\mu|_{\mathcal{A}}}$ and $\mathrm{d}_{\nu|_{\mathcal{A}}}$ induce a common topology on $\sigma(\mathcal{A})$. Both $\mu\colon \sigma(\mathcal{A}) \to [0,\infty]$ and $\nu\colon \sigma(\mathcal{A}) \to [0,\infty]$ are continuous extensions of $\mu|_{\mathcal{A}} = \nu|_{\mathcal{A}}$ with respect to this common topology by Proposition (3.8). Since $[0,\infty]$ is Hausdorff and since $\mathcal{A}$ is dense in $\sigma(\mathcal{A})$ with respect to this common topology, it follows from Proposition (3.7) that $\mu = \nu$. □

3.4 Measurability

We've just shown that any finite premeasure space $(X, \mathcal{A}, \mu)$ can be uniquely extended to the finite measure space $(X, \sigma(\mathcal{A}), \mu^*|_{\sigma(\mathcal{A})})$. In fact, since $\mu^*|_{\sigma(\mathcal{A})}$ is unique, we may as well simply denote it by μ. It turns out that $(X, \mathcal{A}, \mu)$ can be uniquely extended to an even bigger measure space $(X, \mathcal{M}, \mu^*|_{\mathcal{M}})$. Let us define what this bigger measure space is.

Definition 3.2. Let μ be an outer measure on a set X. A subset A of X is said to be μ-**measurable** if

$$\mu^*(S) \geq \mu^*(S \cap A) + \mu^*(S \backslash A) \tag{7}$$

holds for every subset S of X. Note that (7) is equivalent to the equation

$$\mu^*(S) = \mu^*(S \cap A) + \mu^*(S \backslash A)$$

since μ^* is finitely subadditive. Note also that (7) holds for every subset S of X is equivalent to the assertion that this inequality holds for every subset S of X for which $\mu^*(S) < \infty$. We denote by $\mathcal{M}_\mu$ (or more simply just $\mathcal{M}$ if μ is clear from context) to be the set of all μ-measurable sets.

3.4.1 $\mathcal{M}_{\mu^*}$ contains $\mathcal{A}$

Proposition 3.9. *Let $(X, \mathcal{A}, \mu)$ be a premeasure space. Then $\mathcal{A} \subseteq \mathcal{M}_{\mu^*}$.*

Proof. Let $A \in \mathcal{A}$, let $S \in \mathcal{P}(X)$, and let $\varepsilon > 0$. Choose a covering $\{A_n\} \in \mathcal{A}$ of S such that

$$\sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(S) + \varepsilon.$$

Then $\{A_n \cap A\}$ is a covering of $S \cap A$ and $\{A_n \backslash A\}$ is a covering of $S \backslash A$, and so

$$\begin{aligned}
\mu^*(S) &\geq \sum_{n=1}^{\infty} \mu(A_n) - \varepsilon \\
&= \sum_{n=1}^{\infty} \mu((A_n \cap A) \cup (A_n \backslash A)) - \varepsilon \\
&= \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \mu(A_n \backslash A) - \varepsilon \\
&\geq \mu^*(S \cap A) + \mu^*(S \backslash A) - \varepsilon.
\end{aligned}$$

Taking $\varepsilon \to 0$ gives us

$$\mu^*(S) \geq \mu^*(S \cap A) + \mu^*(S \backslash A).$$

Therefore A is μ^*-measurable since S was arbitrary. □

3.4.2 $(X, \mathcal{M}_{\mu^*}, \mu^*|_{\mathcal{M}_{\mu^*}})$ is a measure space

Proposition 3.10. *Let $(X, \mathcal{A}, \mu)$ be a premeasure space. Then $(X, \mathcal{M}_{\mu^*}, \mu^*|_{\mathcal{M}_{\mu^*}})$ is a measure space.*

Proof. To clean notation in what follows, we denote $\mathcal{M} = \mathcal{M}_{\mu^*}$ and $\mu^* = \mu^*|_{\mathcal{M}_{\mu^*}}$. We prove this proposition in several steps:

Step 1: We first show $\mathcal{M}$ is an algebra. First we show it is closed under finite unions. Let $A, B \in \mathcal{M}$ and let $S \in \mathcal{P}(X)$. Then

$$\begin{aligned}
\mu^*(S) &= \mu^*(S \cap A) + \mu^*(S \backslash A) \\
&= \mu^*(S \cap A) + \mu^*((S \backslash A) \cap B) + \mu^*((S \backslash A) \backslash B) \\
&\geq \mu^*((S \cap A) \cup ((S \backslash A) \cap B)) + \mu^*((S \backslash A) \backslash B) \\
&= \mu^*(S \cap (A \cup B)) + \mu^*(S \backslash (A \cup B))
\end{aligned}$$

Therefore $A \cap B \in \mathcal{M}$.

Next we shows it is closed under complements. Let $A \in \mathcal{M}$ and let $S \in \mathcal{P}(X)$. Then

$$\begin{aligned}
\mu^*(S) &\geq \mu^*(S \cap A) + \mu^*(S \backslash A) \\
&= \mu^*(S \backslash (X \backslash A)) + \mu^*(S \backslash A) \\
&= \mu^*(S \backslash (X \backslash A)) + \mu^*(S \cap (X \backslash A)).
\end{aligned}$$

Therefore $X \backslash A \in \mathcal{M}$.

Step 2: We show that μ^* restricts to a measure on $\mathcal{M}$. To do this, we just need to show that μ^* is finitely additive on $\mathcal{M}$. In fact, we claim that for any $S \in \mathcal{P}(X)$ and pairwise disjoint $A_1, \dots, A_n \in \mathcal{M}$, we have

$$\mu^*\left(S \cap \left(\bigcup_{m=1}^{n} A_m\right)\right) = \sum_{m=1}^{n} \mu^*(S \cap A_m). \tag{8}$$

We prove (8) by induction on n. The equality holds trivially for $n = 1$. For the induction step, assume that it holds for some $n \geq 1$. Let S be a subset of X and let$A_1, \dots, A_{n+1}$ be a finite sequence of members in $\mathcal{M}$. Then

$$\begin{aligned}
\mu^*\left(S \cap \left(\bigcup_{m=1}^{n+1} A_m\right)\right) &\geq \mu^*\left(S \cap \left(\bigcup_{m=1}^{n+1} A_m\right) \cap A_{n+1}\right) + \mu^*\left(S \cap \left(\bigcup_{m=1}^{n+1} A_m\right) \cap (X \backslash A_{n+1})\right) \\
&= \mu^*(S \cap A_{n+1}) + \mu^*\left(S \cap \left(\bigcup_{m=1}^{n} A_m\right)\right) \\
&= \mu^*(S \cap A_{n+1}) + \sum_{m=1}^{n} \mu^*(S \cap A_m) \\
&= \sum_{m=1}^{n+1} \mu^*(S \cap A_m).
\end{aligned}$$

This establishes (8). Setting $S = X$ in (8) gives us finite additivity of μ^* on $\mathcal{M}$.

Step 3: We prove that $\mathcal{M}$ is a σ-algebra. Since $\mathcal{M}$ was already shown to be an algebra, it suffices to show that $\mathcal{M}$ is closed under countable unions. Let (A_n) be a sequence in $\mathcal{M}$. Disjointify the sequence (A_n) to the sequence

(D_n): set $D_1 = A_1$ and $D_n = A_n \backslash \bigcup_{m=1}^{n-1} A_m$ for all $n > 1$. Note that (D_n) is a sequence in $\mathcal{M}$ since $\mathcal{M}$ is algebra. Let $S \in \mathcal{P}(X)$ and $n \in \mathbb{N}$. Observe that

$$\begin{aligned}
\mu^*(S) &\geq \mu^*\left(S \cap \left(\bigcup_{m=1}^{n} D_m\right)\right) + \mu^*\left(S \backslash \left(\bigcup_{m=1}^{n} D_m\right)\right) \\
&\geq \mu^*\left(S \cap \left(\bigcup_{m=1}^{n} D_m\right)\right) + \mu^*\left(S \backslash \left(\bigcup_{n \in \mathbb{N}} A_n\right)\right) \\
&= \sum_{m=1}^{n} \mu^*\left(S \cap D_m\right) + \mu^*\left(S \backslash \left(\bigcup_{n \in \mathbb{N}} A_n\right)\right),
\end{aligned}$$

where we applied finite-additivity of μ^* to the first term on the right-hand side and we applied monotonicity of μ^* to the second term on the right-hand side. Taking the limit as $n \to \infty$. We obtain

$$\begin{aligned}
\mu^*(S) &\geq \sum_{m=1}^{\infty} \mu^*\left(S \cap D_m\right) + \mu^*\left(S \backslash \left(\bigcup_{n \in \mathbb{N}} A_n\right)\right) \\
&\geq \mu^*\left(\bigcup_{n \in \mathbb{N}} \left(S \cap D_m\right)\right) + \mu^*\left(S \backslash \left(\bigcup_{n \in \mathbb{N}} A_n\right)\right) \\
&= \mu^*\left(S \cap \bigcup_{n \in \mathbb{N}} D_m\right) + \mu^*\left(S \backslash \left(\bigcup_{n \in \mathbb{N}} A_n\right)\right) \\
&= \mu^*\left(S \cap \left(\bigcup_{n \in \mathbb{N}} A_m\right)\right) + \mu^*\left(S \backslash \left(\bigcup_{n \in \mathbb{N}} A_n\right)\right),
\end{aligned}$$

where we applied countable subadditivity of μ^* to the first expression on the right-hand side. Thus $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$. □

Remark 4. One should compare the proof that $(X, \sigma(\mathcal{A}), \mu^*|_{\sigma(\mathcal{A})})$ is a measure space with the proof that $(X, \mathcal{M}, \mu^*|_{\mathcal{M}})$ is a measure space. They both have three similar steps. However in the former case, we needed to use continuity of μ^* to prove these steps, whereas in the latter case, we needed to use μ^*-measurability to prove these steps. Actually, we did prove a slightly stronger result in the former case, namely that $\overline{\mathcal{A}} = \sigma(\mathcal{A})$.

In general, $\mathcal{M}_{\mu^*}$ strictly contains $\sigma(\mathcal{A})$. The σ-algebra $\mathcal{M}$ has the following additional property: if E is a member of $\mathcal{M}$ and $\mu^*(E) = 0$, then every subset of E is also a member of $\mathcal{M}$. Indeed, if $S \subseteq E$, then for any $A \in \mathcal{A}$, we have

$$\begin{aligned}
\mu^*(S \cap A) + \mu^*(S \backslash A) &\leq \mu^*(E) + \mu^*(S \backslash A) \\
&= \mu^*(S \backslash A) \\
&\leq \mu^*(S),
\end{aligned}$$

which implies $S \in \mathcal{M}$. This property is not necessarily true for $\sigma(\mathcal{A})$. In the special case $\mathcal{A} = \mathcal{I}$ interval algebra, we call this σ-algebra $\mathcal{M}$ the σ-algebra of **Lebesgue measurable sets**. Recall that $\sigma(\mathcal{I})$ is called the σ-algebra of **Borel measurable sets**. It is known that not every Lebesgue measurable set is Borel measurable. The extension of the length measure m from $\mathcal{I}$ to $\sigma(\mathcal{I})$ or $\mathcal{M}$ is called *the* **Lebesgue measure**. By *a* **Borel measure**, we mean any measure defined on $\sigma(\mathcal{I})$. There are many Borel measures but there is only one Lebesgue measure.

3.5 The Borel Algebra

For an algebra $\mathcal{A}$, we define $\mathcal{A}_\sigma$ to be the collection of all countable unions of elements in $\mathcal{A}$. We define $\mathcal{A}_\delta$ to be the collection of all countable intersections of elements in $\mathcal{A}$. We define $\mathcal{A}_{\delta\sigma} = (\mathcal{A}_\delta)_\sigma$ and in the same way, we define $\mathcal{A}_{\sigma\delta} = (\mathcal{A}_\sigma)_\delta$. In general,

$$\mathcal{A}_{\sigma\delta\sigma\delta\cdots} \quad \text{and} \quad \mathcal{A}_{\delta\sigma\delta\sigma\cdots}$$

are not equal to $\sigma(\mathcal{A})$.

Proposition 3.11. *Let $E \in \sigma(\mathcal{A})$. Then*

1. *for all $\varepsilon > 0$, there exists $G \in \mathcal{A}_\delta$ and $F \in A_\supset$ such that*
$$G \subseteq E \subseteq F$$
and $\mu(F \backslash E) < \varepsilon$ and $\mu(E \backslash G) < \varepsilon$.

2. *there exists $B \in \mathcal{A}_{\supset\circ}$ such that $E \subseteq B$ and $\mu(E) = \mu(B)$.*

Proof. 1. Let $\varepsilon > 0$. Choose a cover $\{A_n\} \subseteq \mathcal{A}$ of E such that
$$\sum_{n=1}^{\infty} \mu(A_n) < \mu(E) + \varepsilon.$$
where we denote by $\mu\colon \sigma(\mathcal{A}) \to [0, \infty]$ to be the unique extension of $\mu\colon \mathcal{A} \to [0, \infty]$. Now set
$$F = \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_\sigma.$$
Then $E \subseteq F$ and
$$\begin{aligned} \mu(F \backslash E) &= \mu(F) - \mu(E) \\ &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) - \mu(E) \\ &\leq \sum_{n=1}^{\infty} \mu(A_n) - \mu(E) \\ &< \varepsilon. \end{aligned}$$
Now we apply what we just proved to the set $E^c \in \sigma(\mathcal{A})$. This means we can find a set $F \in \mathcal{A}_\sigma$ such that $E^c \subseteq F$ and
$$\mu(F \backslash E^c) < \varepsilon.$$
Now we set $G = F^c$. Then observe that $G \in \mathcal{A}_\delta$ and
$$\begin{aligned} \mu(E \backslash G) &= \mu(E \cap G^c) \\ &= \mu(E \cap (F^c)^c) \\ &= \mu(F \cap (E^c)^c) \\ &= \mu(F \backslash E^c) \\ &< \varepsilon. \end{aligned}$$

2. For each $n \in \mathbb{N}$, choose $F_n \in \mathcal{A}_\sigma$ such that $E \subseteq F_n$ and
$$\mu(F_n \backslash E) < \frac{1}{n}.$$
Let $B := \bigcap_{n=1}^{\infty} F_n$. Then $B \in \mathcal{A}_{\sigma\delta}$ and $E \subseteq B$ since $E \subseteq F_n$ for each $n \in \mathbb{N}$. Finally, observe that
$$\begin{aligned} \mu(B \backslash E) &\leq \mu(F_n \backslash E) \\ &< \frac{1}{n} \end{aligned}$$
for all $n \in \mathbb{N}$. This implies $\mu(B \backslash E) = 0$. □

3.5.1 Borel σ-Algebra on $\mathbb{R}$

Definition 3.3. The **Borel σ-algebra** on $\mathbb{R}$ is defined to be

$$\mathcal{B}(\mathbb{R}) = \{E \subseteq \mathbb{R} \mid E \cap [n, n+1) \text{ is a Borel set in } [n, n+1) \text{ for all } n \in \mathbb{Z}\}.$$

For $E \in \mathcal{B}(\mathbb{R})$, we define

$$\mathrm{m}(E) = \sum_{n \in \mathbb{Z}} \mathrm{m}(E \cap [n, n+1)).$$

Denote by $\mathcal{I}_n$ the interval algebra on $[n, n+1)$. Denote by $\mathcal{B}_n = \sigma(\mathcal{I}_n)$ the Borel σ-algebra on $[n, n+1)$. It is easy to show that

$$\mathcal{B}(\mathbb{R}) = \sigma\left(\bigcup_{n=-\infty}^{\infty} \mathcal{B}_n\right).$$

Indeed, we have

$$\mathcal{B}(\mathbb{R}) \supseteq \sigma\left(\bigcup_{n=-\infty}^{\infty} \mathcal{B}_n\right),$$

since $\bigcup_n \mathcal{B}_n \subseteq \mathcal{B}(\mathbb{R})$ for all $n \in \mathbb{Z}$ and $\sigma\left(\bigcup_n \mathcal{B}_n\right)$ is the smallest σ-algebra which contains all the $\bigcup_n \mathcal{B}_n$. For the reverse inclusion, let $E \in \mathcal{B}(\mathbb{R})$. Then $E \cap [n, n+1) \in \mathcal{B}_n$ for all $n \in \mathbb{Z}$. Therefore

$$\begin{aligned} E &= \bigcup_{n=-\infty}^{\infty} (E \cap [n, n+1)) \\ &\in \bigcup_{n=-\infty}^{\infty} \mathcal{B}_n. \end{aligned}$$

Thus

$$\mathcal{B}(\mathbb{R}) \subseteq \sigma\left(\bigcup_{n=-\infty}^{\infty} \mathcal{B}_n\right).$$

3.5.2 Translation Invariance

Proposition 3.12. *For all $E \in \mathcal{B}(\mathbb{R})$ and $t \in \mathbb{R}$, we have*

$$\mathrm{m}(E) = \mathrm{m}(E + t).$$

Proof. Define $\mu \colon \mathcal{B}(\mathbb{R}) \to [0, \infty]$ by

$$\mu(E) = \mathrm{m}(E + t)$$

for all $E \in \mathcal{B}(\mathbb{R})$. We need to show that $\mathrm{m}(E) = \mu(E)$ for all $E \in \mathcal{B}(\mathbb{R})$. Let $[c, d] \subseteq [n, n+1)$ (where $[c, d]$ can be open or half-open). Then $\mathrm{m}([c, d]) = d - c$. Note

$$\begin{aligned} \mu([c, d]) &= \mathrm{m}([c + t, d + t]) \\ &= \sum_{k=-\infty}^{\infty} \mathrm{m}([c + t, d + t] \cap [n, n+1)). \end{aligned}$$

Suppose that $[c + t, d + t] \subseteq [k, k+1)$ for some $k \in \mathbb{Z}$. Then

$$\mathrm{m}([c + t, d + t] \cap [n, n+1)) = \begin{cases} d - c & \text{if } n = k \\ 0 & \text{if } n \neq k. \end{cases}$$

So $\mu([c, d]) = d - c$. Now suppose that $[c + t, d + t] \subseteq [k, k+1) \cup [k+1, k+2)$ for some $k \in \mathbb{N}$. Then

$$\begin{aligned} \mathrm{m}([c + t, d + t]) &= \mathrm{m}([c + t, d + t] \cap [k, k+1]) + \mathrm{m}([c + t, d + t] \cap [k+1, k+2]) \\ &= \mathrm{m}([c + t, k+1]) + \mathrm{m}([k+1, d + t]) \\ &= k + 1 - (c + t) + d + t - (k+1) \\ &= d - c \\ &= \mathrm{m}([c, d]). \end{aligned}$$

So for each interval $[c,d] \subseteq [n,n+1)$, we have

$$\mu([c,d]) = \mathrm{m}([c,d]).$$

By finite additivity of m and μ, this implies $\mu(E) = \mathrm{m}(E)$ for all $E \in \mathcal{I}_n$. Since $\mu|_{\mathcal{B}_n}$ and $\mathrm{m}|_{\mathcal{B}_n}$ are both finite measures that coincide on $\mathcal{I}_n$, they must coincide on $\sigma(\mathcal{I}_n) = \mathcal{B}_n$ by the uniqueness part of the extension theorem. Now for every $E \in \mathcal{B}(\mathbb{R})$, we have

$$\begin{aligned}
\mathrm{m}(E) &= \sum_{n \in \mathbb{Z}} \mathrm{m}(E \cap [n,n+1)) \\
&= \sum_{n \in \mathbb{Z}} \mu(E \cap [n,n+1)) \\
&= \mu(E).
\end{aligned}$$

□

3.5.3 Existence of non-Borel sets

Lemma 3.7. *Let $E \subseteq [0,1)$ be a Borel set. For $t \in [0,1)$, define the* ***cyclic translation***

$$E \oplus t = \{x \oplus t \mid x \in E\}$$

where

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

The set $E \oplus t$ is a Borel set for all $t \in [0,1)$, and

$$\mathrm{m}(E \oplus t) = \mathrm{m}(E).$$

Proof. Let $E_1 = E \cap [1-t,1)$ and $E_2 = E \cap [0,1-t)$. Clearly $E = E_1 \cup E_2$ and E_1 and E_2 are disjoint. Since $E_2 + t \subseteq [0,1)$, we have

$$\begin{aligned}
\mathrm{m}(E_2 + t) &= \mathrm{m}(E_2 \oplus t) \\
&= \mathrm{m}(E_2)
\end{aligned}$$

Also $E_1 + t = [1, 1+t)$ implies $E_1 \oplus t = E_1 + t - 1$, which implies

$$\begin{aligned}
\mathrm{m}(E_1 \oplus t) &= \mathrm{m}(E_1 + (t-1)) \\
&= \mathrm{m}(E_1).
\end{aligned}$$

Thus since $(E_1 \oplus t) \cap (E_2 \oplus t) = \emptyset$, we have

$$\begin{aligned}
\mathrm{m}(E \oplus t) &= \mathrm{m}(E_1 \oplus t) + \mathrm{m}(E_2 \oplus t) \\
&= \mathrm{m}(E_1) + \mathrm{m}(E_2) \\
&= \mathrm{m}(E_1 \cup E_2) \\
&= \mathrm{m}(E).
\end{aligned}$$

□

Proposition 3.13. *There exists a set $N \subseteq [0,1)$ which is not a Borel set.*

Proof. Define a relation $\sim$ on $[0,1)$ by $x \sim y$ if $x - y \in \mathbb{Q}$. It's easy to see that $\sim$ is an equivalence relation. Define a set N by picking one element in each equivalence class of this relation. Denote by (q_n) the sequence of all rational numbers in $[0,1)$. Notice that

$$(N \oplus q_n) \cap (N \oplus q_m) = \emptyset$$

for all $q_n \neq q_m$. Furthermore, we have

$$[0,1) = \bigcup_{n=1}^{\infty} (N \oplus q_n).$$

Indeed, one inclusion is clear, so let's show the other inclusion. Assume for a contradiction that $[0,1)$ is not contained in $\bigcup_n N \oplus q_n$. Choose $x \in [0,1)$ such that $x \notin \bigcup_n (N \oplus q_n)$. Let $y \in N$ such that $y \sim x$. □

4 Integration

4.1 Simple Functions

Definition 4.1. Let $(X, \mathcal{M})$ be a measurable space. A function $\varphi\colon X \to \mathbb{R}$ is said to be a **simple function** (with respect to the measure space $(X, \mathcal{M})$) if it is a linear combination of indicators of measurable sets, that is, it has the form

$$\varphi = \sum_{i=1}^{n} a_i 1_{A_i} \tag{9}$$

where $a_i \in \mathbb{R}$ and $A_i \in \mathcal{M}$ for all $1 \leq i \leq n$. We call the expression (9) a **representation** of φ. If $i \neq i'$ implies $A_i \cap A_{i'} = \emptyset$, then we call (9) a **disjoint representation** of φ. If $i \neq i'$ implies $a_i \neq a_{i'}$ and $A_i \cap A_{i'} = \emptyset$, then we call (9) *the* **canonical representation** of φ. It is clear that this representation uniquely determines φ (which is why we say it is *the canonical representation*).

4.1.1 Integrating nonnegative simple functions

Definition 4.2. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $\varphi\colon X \to [0, \infty)$ be a nonnegative simple function with canonical representation

$$\varphi = \sum_{i=1}^{n} a_i 1_{A_i}.$$

We define the **integral of** φ by

$$\int_X \varphi \mathrm{d}\mu := \sum_{i=1}^{n} a_i \mu(A_i). \tag{10}$$

Note that there are no issues with well-definedness of (10) since the canonical representation is uniquely determined by φ. It will actually turn out that it doesn't matter how we represent φ; if $\varphi = \sum_{i=1}^{m} c_i 1_{E_i}$ is any representation of φ (canonical or not) then we will have $\int_X \varphi \mathrm{d}\mu = \sum_{i=1}^{m} c_i \mu(E_i)$. Let us show this for disjoint representations of φ.

Lemma 4.1. *Let φ be a simple function and suppose*

$$\varphi = \sum_{i=1}^{m} c_i 1_{E_i} \tag{11}$$

is a disjoint representation of φ. Then

$$\int_X \varphi \mathrm{d}\mu = \sum_{i=1}^{m} c_i \mu(E_i).$$

Proof. Note that the representation (11) is not necessarily canonical since we may have $c_i = c_{i'}$ for some $i \neq i'$. So express φ in terms of its canonical representation, say $\varphi = \sum_{j=1}^{n} a_j 1_{A_j}$. For each j we have

$$A_j = \bigcup_{i | c_i = a_j} E_i,$$

and so

$$\begin{aligned} a_j \mu(A_j) &= a_j \sum_{i | c_i = a_j} \mu(E_i) \\ &= \sum_{i | c_i = a_j} a_j \mu(E_i) \\ &= \sum_{i | c_i = a_j} c_i \mu(E_i). \end{aligned}$$

Now sum over all j to get

$$\begin{aligned}\int_X \varphi \mathrm{d}\mu(x) &= \sum_{j=1}^{n} a_j \mu(A_j) \\ &= \sum_{j=1}^{n} a_j \sum_{i \mid c_i = a_j} \mu(E_i) \\ &= \sum_{j=1}^{n} \sum_{i \mid c_i = a_j} c_i \mu(E_i) \\ &= \sum_{i=1}^{m} c_i \mu(E_i).\end{aligned}$$

□

4.1.2 $\mathbb{R}_{\geq 0}$-Scaling, Monotonicity, and Additivity of Integration for Nonnegative Simple Functions

Proposition 4.1. *Let $\varphi, \psi \colon X \to [0, \infty)$ be nonnegative simple functions and let $a \geq 0$. Then we have*

1. $\mathbb{R}_{\geq 0}$*-scaling of integration for nonnegative simple functions.:*

$$\int_X a\varphi \mathrm{d}\mu = a \int_X \varphi \mathrm{d}\mu. \tag{12}$$

2. *Monotonicity of integration for nonnegative simple functions: if $\varphi \leq \psi$, then*

$$\int_X \varphi \mathrm{d}\mu \leq \int_X \psi \mathrm{d}\mu.$$

3. *Additivity of integration for nonnegative simple functions:*

$$\int_X (\varphi + \psi) \mathrm{d}\mu = \int_X \varphi \mathrm{d}\mu + \int_X \psi \mathrm{d}\mu.$$

Proof. 1. Express φ in terms of its canonical representation, say $\varphi = \sum_{i=1}^{n} a_i 1_{A_i}$. Then $a\varphi$ is a nonnegative simple function with canonical representation $a\varphi = \sum_{i=1}^{n} a a_i 1_{A_i}$. Therefore

$$\begin{aligned}a \int_X \varphi \mathrm{d}\mu &= a \sum_{i=1}^{n} a_i \mu(A_i) \\ &= \sum_{i=1}^{n} a a_i \mu(A_i) \\ &= \int_X a\varphi \mathrm{d}\mu\end{aligned}$$

Thus we have $\mathbb{R}_{\geq 0}$-scaling of integration for nonnegative simple functions.

2. Assume $\varphi \leq \psi$. Express φ and ψ in terms of their canonical representations, say $\varphi = \sum_{i=1}^{m} a_i 1_{A_i}$ and $\psi = \sum_{j=1}^{n} b_j 1_{B_j}$. Since $\varphi \leq \psi$, we have

$$\begin{aligned}\varphi &= \sum_{i=1}^{m} a_i 1_{A_i} \\ &= \sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \min(a_i, b_j) 1_{A_i \cap B_j} \\ &\leq \sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \max(a_i, b_j) 1_{A_i \cap B_j} \\ &= \sum_{j=1}^{n} b_j 1_{B_j} \\ &= \psi.\end{aligned}$$

Therefore

$$\begin{aligned}\int_X \varphi \mathrm{d}\mu &= \sum_{\substack{1\le i\le m\\1\le j\le n}} \min(a_i, b_j)\mu(A_i\cap B_j)\\ &\le \sum_{\substack{1\le i\le m\\1\le j\le n}} \max(a_i, b_j)\mu(A_i\cap B_j)\\ &= \int_X \psi \mathrm{d}\mu.\end{aligned}$$

Thus we have monotonicity of integration for nonnegative simple functions.

3. Express φ and ψ in terms of their canonical representations, say $\varphi = \sum_{i=1}^m a_i 1_{A_i}$ and $\psi = \sum_{j=1}^n b_j 1_{B_j}$. Then $\varphi+\psi$ is a nonnegative simple function with a disjoint representation given by

$$\varphi+\psi = \sum_{\substack{1\le i\le m\\1\le j\le n}} (a_i+b_j)1_{A_i\cap B_j}$$

Therefore

$$\begin{aligned}\int_X (\varphi+\psi)\mathrm{d}\mu &= \sum_{\substack{1\le i\le m\\1\le j\le n}} (a_i+b_j)\mu(A_i\cap B_j)\\ &= \sum_{\substack{1\le i\le m\\1\le j\le n}} a_i\mu(A_i\cap B_j) + \sum_{\substack{1\le i\le m\\1\le j\le n}} b_j\mu(A_i\cap B_j)\\ &= \sum_{1\le i\le m} a_i\mu\left(\bigcup_{j=1}^n (A_i\cap B_j)\right) + \sum_{1\le j\le n} b_j\mu\left(\bigcup_{i=1}^m A_i\cap B_j\right)\\ &= \sum_{1\le i\le m} a_i\mu(A_i) + \sum_{1\le j\le n} b_j\mu(B_j)\\ &= \int_X \varphi\mathrm{d}\mu + \int_X \psi\mathrm{d}\mu.\end{aligned}$$

Thus we have additivity of integration for nonnegative simple functions. □

4.2 Measurable Functions

We want to be able to integrate more general nonnegative functions. The functions that we want to consider now are called measurable functions. Before we explain what these functions are how to integrate them, let's define them in a more general context first.

Definition 4.3. Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces and let $f\colon X\to Y$ be a function. We say f is **measurable with respect to** $\mathcal{M}$ **and** $\mathcal{N}$ if $f^{-1}(\mathcal{N})\subseteq\mathcal{M}$ where

$$f^{-1}(\mathcal{N}) = \{f^{-1}(B) \mid B\in\mathcal{N}\}.$$

In other words, f is measurable with respect to $\mathcal{M}$ and $\mathcal{N}$ if for all $B\in\mathcal{N}$ we have

$$f^{-1}(B) = \{x\in X \mid f(x)\in B\}\in\mathcal{M}.$$

Thus f is measurable with respect to $\mathcal{M}$ and $\mathcal{N}$ if the inverse image of a measurable set is a measurable set. If $\mathcal{M}=\mathcal{N}$, then we will just say f is measurable with respect to $\mathcal{M}$. If the σ-algebras $\mathcal{M}$ and $\mathcal{N}$ are clear from context, then we will just say f is measurable.

For our purposes, we will mostly be interested in the case where $Y=\mathbb{R}$ and $\mathcal{N}=\mathcal{B}(\mathbb{R})$ (or $Y=[0,\infty]$ and $\mathcal{N}=\mathcal{B}[0,\infty]$). We want to find another criterion for a function $f\colon X\to\mathbb{R}$ to be measurable (rather than the inverse image of a Borel-measurable set is a $\mathcal{M}$-measurable set). We state this criterion as a corollary following the next two propositions:

Proposition 4.2. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces and let $f\colon X\to Y$ be a function. Suppose that $\mathcal{N}$ is generated as a σ-algebra by the collection $\mathcal{C}$ of subsets of Y. Then $f^{-1}(\mathcal{N})\subseteq\mathcal{M}$ if and only if $f^{-1}(\mathcal{C})\subseteq\mathcal{M}$.*

Proof. One direction is clear, so we just prove the other direction. Suppose $f^{-1}(\mathcal{C}) \subseteq \mathcal{M}$. Observe that

$$\{B \in \mathcal{P}(Y) \mid f^{-1}(B) \in \mathcal{M}\}$$

is a σ-algebra which contains $\mathcal{C}$. Indeed, it is a σ-algebra since f^{-1} maps the emptyset set to the emptyset and maps the whole space Y to the whole space X, and since f^{-1} commutes with unions and complements. Furthemore, this σ-algebra contains $\mathcal{C}$ since $f^{-1}(\mathcal{C}) \subseteq \mathcal{M}$. Since $\mathcal{N}$ is the *smallest* σ-algebra which contains $\mathcal{C}$, it follows that

$$\mathcal{N} \subseteq \{B \in \mathcal{P}(Y) \mid f^{-1}(B) \in \mathcal{M}\}.$$

In particular, if $B \in \mathcal{N}$, then $f^{-1}(B) \in \mathcal{M}$. Thus f is measurable. □

Proposition 4.3. *Let* $\mathcal{C} = \{(-\infty, c) \mid c \in \mathbb{R}\}$. *Then* $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{C})$.

Proof. Let $\mathcal{I}_n$ be the collection of all subintervals of $[n, n+1)$ and let $\mathcal{B}_n = \sigma(\mathcal{I}_n)$. So

$$\mathcal{B}(\mathbb{R}) = \{E \subseteq \mathbb{R} \mid E \cap [n, n+1) \in \mathcal{B}_n \text{ for all } n \in \mathbb{Z}\}.$$

Let $c \in \mathbb{R}$. Then since $(-\infty, c) \cap [n, n+1]$ is a subinterval of $[n, n+1]$ for all $n \in \mathbb{Z}$, it follows that $(-\infty, c) \in \mathcal{B}$ for all $n \in \mathbb{Z}$. Thus $\mathcal{C} \subseteq \mathcal{B}$ which implies $\sigma(\mathcal{C}) \subseteq \mathcal{B}$ (as $\sigma(\mathcal{C})$ is the *smallest* σ-algebra which contains $\mathcal{C}$). Conversely, note that $\sigma(\mathcal{C})$ contains all subintervals of $[n, n+1)$ for all $n \in \mathbb{Z}$. Thus $\sigma(\mathcal{C}) \supseteq \mathcal{B}_n$ for all $n \in \mathbb{Z}$ (as $\mathcal{B}_n$ is the *smallest* σ-algebra which contains all subintervals of $[n, n+1)$. Since $\mathcal{B}(\mathbb{R}) = \sigma\left(\bigcup_{n\in\mathbb{Z}} \mathcal{B}_n\right)$, it follows that $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C})$. □

Corollary 1. *Let* $(X, \mathcal{M})$ *be a measurable space and let* $f\colon X \to \mathbb{R}$ *be a function. Then* f *is measurable with respect to* $\mathcal{M}$ *and* $\mathcal{B}(\mathbb{R})$ *if and only if* $\{f < c\} \in \mathcal{M}$ *for all* $c \in \mathbb{R}$.

Proof. Follows from Proposition (11.3) and Proposition (11.2). □

Remark 5. Here $\{f < c\}$ denotes the set $\{x \in X \mid f(x) < c\}$. In fact, we shall often use this shorthand notation for other sets like this. For instance, we write

$$\begin{aligned}
\{f < c\} &= \{x \in X \mid f(x) < c\} \\
\{f \leq c\} &= \{x \in X \mid f(x) \leq c\} \\
\{f > c\} &= \{x \in X \mid f(x) > c\} \\
\{f \geq c\} &= \{x \in X \mid f(x) \geq c\} \\
\{f = c\} &= \{x \in X \mid f(x) = c\} \\
\{f \neq c\} &= \{x \in X \mid f(x) \neq c\}.
\end{aligned}$$

This is notation is very common throughout the literature.

4.2.1 Combining Measurable Functions to get More Measurable Functions

Proposition 4.4. *Let* $f, g\colon X \to \mathbb{R}$ *be measurable functions and let* $a \in \mathbb{R}$. *Then* af, $f+g$, f^2, $|f|$, fg, $\max\{f,g\}$, *and* $\min\{f,g\}$ *are all measurable.*

Proof. We first show af is measurable. If $a = 0$, then af is the zero function, which is measurable, so assume $a \neq 0$. Then for any $c \in \mathbb{R}$ we have

$$\{af < c\} = \begin{cases} \{f < c/a\} \in \mathcal{M} & \text{if } a > 0 \\ \{f > c/a\} \in \mathcal{M} & \text{if } a < 0 \end{cases}$$

It follows that af is measurable.

Next we show that $f+g$ is measurable. Observe that for any $c \in \mathbb{R}$, we have

$$\begin{aligned}
x \in \{f + g < c\} &\iff f(x) + g(x) < c \\
&\iff \text{there exists an } r \in \mathbb{Q} \text{ such that } f(x) < r \text{ and } r < c - g(x) \\
&\iff \text{there exists an } r \in \mathbb{Q} \text{ such that } x \in \{f < r\} \cap \{g < c - r\}. \\
&\iff x \in \bigcup_{r\in\mathbb{Q}} \{f < r\} \cap \{g < c - r\}.
\end{aligned}$$

Therefore

$$\{f+g<c\} = \bigcup_{r\in\mathbb{Q}} \{f<r\}\cap\{g<c-r\} \in \mathcal{M}.$$

It follows that $f+g$ is measurable.

Next we show that f^2 is measurable. For any $c\in\mathbb{R}$, we have

$$\{f^2>c\} = \begin{cases} \{f>\sqrt{c}\}\cup\{f<-\sqrt{c}\}\in\mathcal{M} & c\geq 0 \\ X\in\mathcal{M} & c<0 \end{cases}$$

It follows that f^2 is measurable.

Next we show that $|f|$ is measurable. For any $c\in\mathbb{R}$, we have

$$\{|f|>c\} = \begin{cases} \{f>c\}\cup\{f<-c\}\in\mathcal{M} & c\geq 0 \\ X\in\mathcal{M} & c<0 \end{cases}$$

It follows that $|f|$ is measurable.

Finally, note that the remaining functions can be expressed as combinations of the previous ones. Indeed,

$$\begin{aligned} fg &= \frac{1}{4}\left((f+g)^2-(f-g)^2\right) \\ \max\{f,g\} &= \frac{1}{2}(|f+g|+|f-g|) \\ \min\{f,g\} &= \frac{1}{2}(|f+g|-|f-g|) \end{aligned}$$

It follows that they are all measurable. □

Proposition 4.5. *Let $(f_n\colon X\to\mathbb{R})$ be a sequence of measurable functions. Then $\sup f_n$, $\inf f_n$, $\limsup f_n$, and $\liminf f_n$ are all measurable. In particular, if*

$$\lim_{n\to\infty} f_n(x)$$

exists for all $x\in X$. The corresponding function is also measurable.

Proof. Let $c\in\mathbb{R}$. Then we have

$$\{\sup f_n>c\} = \bigcup_{n\geq 1}\{f_n>c\}.$$

This implies $\sup f_n$ is measurable. Similarly, we have

$$\{\inf f_n<c\} = \bigcup_{n\geq 1}\{f_n<c\}.$$

This implies $\inf f_n$ is measurable. Finally since

$$\limsup f_n = \inf_{N\geq 1}\sup_{n\geq N} f_n \qquad \text{and} \qquad \liminf f_n = \sup_{N\geq 1}\inf_{n\geq N} f_n,$$

we see that both $\limsup f_n$ and $\liminf f_n$ are measurable. □

4.2.2 Criterion for Nonnegative Function to be Measurable

If $f\colon X\to\mathbb{R}$ is a nonnegative function, then we have another criterion for f to be measurable.

Proposition 4.6. *Let $f\colon X\to[0,\infty]$ be a nonnegative function (which possibly takes infinite values). The following are equivalent:*

1. *f is measurable.*
2. *There exists an increasing sequence of nonnegative simple functions $(\varphi_n\colon X\to[0,\infty])$ such that (φ_n) converges pointwise to f.*

Proof. Proposition (4.5) gives us 2 implies 1 (since simple functions are measurable!), so we just need to show 1 implies 2. Suppose f is measurable. For each $n \in \mathbb{N}$ and for each $1 \leq i \leq 2^n$, we define

$$E_n = \{f \geq n\} \text{ and } E_{n,i} = \left\{ \frac{i-1}{2^n} \leq f < \frac{i}{2^n} \right\}.$$

Note that $E_n, E_{n,i} \in \mathcal{M}$. For each $n \in \mathbb{N}$ we define $\varphi_n \colon X \to [0, \infty]$ by

$$\varphi_n = n 1_{E_n} + \sum_{i=1}^{2^n} \left(\frac{i-1}{2^n} \right) 1_{E_{n,i}}$$

Therefore each φ_n is a simple function. It's easy to check that (φ_n) is an increasing sequence of nonnegative functions. Let us check that (φ_n) converges pointwise to f. Let $x \in X$. If $f(x) = \infty$, then $\varphi_n(x) = n$ for all $n \in \mathbb{N}$, and thus

$$\begin{aligned} \varphi_n(x) &= n \\ &\to \infty \\ &= f(x) \end{aligned}$$

as $n \to \infty$, so assume $f(x) < \infty$. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $2^{-N} < \varepsilon$. Then $n \geq N$ implies

$$\begin{aligned} f(x) - \varphi_n(x) &< 2^{-n} \\ &\leq 2^{-N} \\ &< \varepsilon. \end{aligned}$$

This implies $\varphi_n(x) \to f(x)$ as $n \to \infty$. Since x was arbitrary, we see that (φ_n) converges pointwise to f. □

4.3 The Integral of a Nonnegative Measurable Function

We are now ready to describe how to integrate nonnegative measurable functions.

Definition 4.4. Let $f \colon X \to [0, \infty]$ be a nonnegative measurable function. The **integral of** f is defined to be

$$\int_X f \mathrm{d}\mu := \sup \left\{ \int_X \varphi \mathrm{d}\mu \mid \varphi \text{ is a nonnegative simple function and } \varphi \leq f \right\}.$$

Recall that integrating nonnegative simple functions satisfies three very nice properties; namely $\mathbb{R}_{\geq 0}$-scaling, monotonicity, and additivity. In fact, these three properties continue to hold when integrating nonnegative measurable functions. The first two are easy to show, but additivity requires a little more effort (it will follow from the so-called Monotone Convergence Theorem).

4.3.1 Monotone Convergence Theorem

Theorem 4.2. *(MCT) Let $(f_n \colon X \to [0, \infty])$ be an increasing sequence of nonnegative measurable functions which converges pointwise to a nonnegative function $f \colon X \to [0, \infty]$. Then*

$$\lim_{n \to \infty} \int_X f_n \mathrm{d}\mu = \int_X f \mathrm{d}\mu.$$

Proof. The fact that $f \colon X \to [0, \infty]$ is measurable follows from (Proposition (4.5). Therefore $\int_X f \mathrm{d}\mu$ is defined. Since $f_n \leq f$ for all n, we have

$$\lim_{n \to \infty} \int_X f_n \mathrm{d}\mu \leq \int_X f \mathrm{d}\mu.$$

by monotonicity of integration for nonnegative functions. For the reverse inequality, we just need to show that

$$\lim_{n \to \infty} \int_X f_n \mathrm{d}\mu \geq \int_X \varphi \mathrm{d}\mu \tag{13}$$

for any nonnegative simple function φ such that $\varphi \leq f$, so let φ be such a nonnegative simple function and let $c \in (0,1)$. For each $n \in \mathbb{N}$, define

$$A_n := \{f_n - c\varphi > 0\}.$$

Then A_n is measurable since $f_n - c\varphi$ is measurable for all n. Also, (A_n) is an ascending sequence of sets such that

$$\bigcup_{n=1}^{\infty} A_n = X$$

since f_n converges pointwise to f. Therefore

$$\begin{aligned}
\lim_{n\to\infty} \int_X f_n \mathrm{d}\mu &\geq \lim_{n\to\infty} \int_X f_n 1_{A_n} \mathrm{d}\mu \\
&\geq \lim_{n\to\infty} \int_X c\varphi 1_{A_n} \mathrm{d}\mu \\
&= c \lim_{n\to\infty} \int_X \varphi 1_{A_n} \mathrm{d}\mu \\
&= c \int_X \varphi \mathrm{d}\mu
\end{aligned}$$

where we obtained the fourth line from the third line from the fact that the function $\nu\colon \mathcal{M} \to [0,\infty]$ defined by

$$\nu(E) = \int_X \varphi 1_E \mathrm{d}\mu$$

for all $E \in \mathcal{M}$ is a measure. In particular

$$\begin{aligned}
\lim_{n\to\infty} \int_X \varphi 1_{A_n} \mathrm{d}\mu &= \lim_{n\to\infty} \nu(A_n) \\
&= \nu\left(\bigcup_{n=1}^{\infty} A_n\right) \\
&= \nu(X) \\
&= \int_X \varphi 1_X \mathrm{d}\mu \\
&= \int_X \varphi \mathrm{d}\mu.
\end{aligned}$$

Now we take $c \to 1$ to get (13). □

4.3.2 $\mathbb{R}_{\geq 0}$-Scaling, Monotonicity, and Additivity of Integration for Nonnegative Measurable Functions

Proposition 4.7. *Let $f, g\colon X \to [0,\infty]$ be measurable and let $a \geq 0$. Then we have*

1. *$\mathbb{R}_{\geq 0}$-scaling of integration for nonnegative measurable functions.:*

$$\int_X af \mathrm{d}\mu = a \int_X f \mathrm{d}\mu. \tag{14}$$

2. *Monotonicity of integration for nonnegative measurable functions: if $f \leq g$, then*

$$\int_X f \mathrm{d}\mu \leq \int_X g \mathrm{d}\mu.$$

3. *Additivity of integration for nonnegative measurable functions.:*

$$\int_X (f+g) \mathrm{d}\mu = \int_X f \mathrm{d}\mu + \int_X g \mathrm{d}\mu.$$

Proof. 1) If $a = 0$, then (14) is obvious, so assume $a > 0$. Let $\varepsilon > 0$ and choose a nonnegative simple function $\varphi\colon X \to [0,\infty]$ such that $\varphi \leq f$ and

$$\int_X f \mathrm{d}\mu - \varepsilon < \int_X \varphi \mathrm{d}\mu.$$

Then $a\varphi$ is a nonnegative simple function such that $a\varphi \leq af$. Furthermore, we have

$$\begin{aligned}\int_X af\mathrm{d}\mu &\geq \int_X a\varphi\mathrm{d}\mu \\ &= a\int_X \varphi\mathrm{d}\mu \\ &> a\left(\int_X f\mathrm{d}\mu - \varepsilon\right) \\ &= a\int_X f\mathrm{d}\mu - a\varepsilon.\end{aligned}$$

Taking $\varepsilon \to 0$ gives us

$$\int_X af\mathrm{d}\mu \geq a\int_X f\mathrm{d}\mu.$$

This gives us one inequality.

For the reverse inequality, let $\varepsilon > 0$ and choose a nonnegative simple function $\varphi\colon X \to [0,\infty]$ such that $\varphi \leq af$ and

$$\int_X af\mathrm{d}\mu - \varepsilon < \int_X \varphi\mathrm{d}\mu.$$

Then $a^{-1}\varphi$ is a nonnegative simple function such that $a^{-1}\varphi \leq f$. Furthermore, we have

$$\begin{aligned}a\int_X f\mathrm{d}\mu &\geq a\int_X a^{-1}\varphi\mathrm{d}\mu \\ &= \int_X \varphi\mathrm{d}\mu \\ &> \int_X af\mathrm{d}\mu - \varepsilon\end{aligned}$$

Taking $\varepsilon \to 0$ gives us

$$\int_X af\mathrm{d}\mu \leq a\int_X f\mathrm{d}\mu.$$

Thus we have $\mathbb{R}_{\geq 0}$-scaling of integration for nonnegative measurable functions.

2. Assume $f \leq g$. Let $\varphi\colon X \to [0,1]$ be a nonnegative simple function such that $\varphi \leq f$. Then $\varphi \leq g$ since $f \leq g$. This implies

$$\begin{aligned}\int_X f\mathrm{d}\mu &:= \sup\left\{\int_X \varphi\mathrm{d}\mu \mid \varphi \text{ is a nonnegative simple function and } \varphi \leq f\right\} \\ &\leq \sup\left\{\int_X \varphi\mathrm{d}\mu \mid \varphi \text{ is a nonnegative simple function and } \varphi \leq g\right\} \\ &:= \int_X g\mathrm{d}\mu.\end{aligned}$$

Thus we have monotonicity of integration nonnegative measurable functions.

3. Choose an increasing sequence $(\varphi_n\colon X \to [0,\infty])$ of nonnegative simple functions which converges pointwise to f (we can do this since f is a nonnegative measurable function). Similarly, choose an increasing sequence $(\psi_n\colon X \to [0,\infty])$ of nonnegative simple functions which converges pointwise to g (we can do this since g is a nonnegative measurable function). Then $(\varphi_n + \psi_n)$ is an increasing sequence of nonnegative simple functions which converges pointwise to $f+g$. It follows from MCT that

$$\begin{aligned}\int_X (f+g)\mathrm{d}\mu &= \lim_{n\to\infty}\int_X (\varphi_n + \psi_n)\mathrm{d}\mu \\ &= \lim_{n\to\infty}\int_X \varphi_n\mathrm{d}\mu + \lim_{n\to\infty}\int_X \psi_n\mathrm{d}\mu \\ &= \int_X f\mathrm{d}\mu + \int_X g\mathrm{d}\mu.\end{aligned}$$

Thus we have additivity of integration for nonnegative measurable functions. □

Remark 6. Note that in each proof for the three statements in Proposition (4.7), we needed to use the fact that these three statements hold when integrating nonnegative *simple* functions.

4.3.3 Fatou's Lemma

Proposition 4.8. *(Fatou's Lemma) Let $(f_n \colon X \to [0,\infty])$ be a sequence of nonnegative measurable functions. Then*

$$\int_X \liminf f_n \mathrm{d}\mu \leq \liminf \int_X f_n \mathrm{d}\mu.$$

In particular, if (f_n) converges to a function $f \colon X \to [0,\infty]$ pointwise, then

$$\int_X f \mathrm{d}\mu \leq \liminf \int_X f_n \mathrm{d}\mu. \tag{15}$$

Proof. For each $N \in \mathbb{N}$, set $g_N = \inf_{n \geq N} f_n$. Then (g_N) is an increasing sequence of nonnegative measurable functions which converges pointwise to $\liminf f_n$. Indeed, for any $x \in X$, we have

$$\begin{aligned} \lim_{N\to\infty} g_N(x) &= \lim_{N\to\infty} \inf_{n\geq N} f_n(x) \\ &:= \liminf f_n(x) \end{aligned}$$

It follows from the MCT that

$$\begin{aligned} \int_X \liminf f_n \mathrm{d}\mu &= \lim_{N\to\infty} \int_X g_N \mathrm{d}\mu \\ &= \lim_{N\to\infty} \int_X \inf_{n\geq N} f_n \mathrm{d}\mu \\ &\leq \lim_{N\to\infty} \inf_{n\geq N} \int_X f_n \mathrm{d}\mu \\ &:= \liminf \int_X f_n \mathrm{d}\mu, \end{aligned}$$

where we obtained the third line from the second line from monotonicity of integration (for any $m \geq N$ we have $f_m \geq \inf_{n\geq N} f_n$ which implies $\int_X f_m \mathrm{d}\mu \geq \int_X \inf_{n\geq N} f_n \mathrm{d}\mu$ which implies $\inf_{n\geq N} \int_X f_n \mathrm{d}\mu \geq \int_X \inf_{n\geq N} f_n \mathrm{d}\mu$). The identity (15) follows from the fact that if (f_n) converges to f pointwise, then $f = \liminf f_n$. □

4.4 Integrable Functions

Definition 4.5. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f \colon X \to [-\infty, \infty]$ be a measurable function. The **positive part** of f, denoted f^+, and the **negative part** of f, denoted f^-, are defined by

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

for all $x \in X$. Note that both f^+ and f^- are both nonnegative measurable functions. Furthermore, we have

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

We say f is **integrable** if

$$\int_X f^+ \mathrm{d}\mu < \infty \quad \text{and} \quad \int_X f^- \mathrm{d}\mu < \infty.$$

Since $|f| = f^+ + f^-$, this is equivalent to saying

$$\int_X |f| \mathrm{d}\mu < \infty.$$

In this case, we define the **integral** of f to be

$$\int_X f \mathrm{d}\mu := \int_X f^+ \mathrm{d}\mu - \int_X f^- \mathrm{d}\mu,$$

4.4.1 $\mathbb{R}$-Scaling, Monotonicity, and Additivity of Integration for Integrable Functions

Proposition 4.9. *Let $f, g\colon X \to \mathbb{R}$ be integrable functions and let $a \in \mathbb{R}$. Then we have*

1. *$\mathbb{R}$-scaling of integration for integrable functions:*
$$\int_X af\mathrm{d}\mu = a\int_X f\mathrm{d}\mu. \tag{16}$$
2. *Additivity of integration for integrable functions:*
$$\int_X (f+g)\mathrm{d}\mu = \int_X f\mathrm{d}\mu + \int_X g\mathrm{d}\mu.$$
3. *Monotonicity of integration for integrable functions: if $f \leq g$, then*
$$\int_X f\mathrm{d}\mu \leq \int_X g\mathrm{d}\mu.$$

Proof. 1. If $a \geq 0$, then $(af)^+ = af^+$ and $(af)^- = af^-$. Therefore
$$\begin{aligned}\int_X af\mathrm{d}\mu &= \int_X af^+\mathrm{d}\mu - \int_X af^-\mathrm{d}\mu \\ &= a\int_X f^+\mathrm{d}\mu - a\int_X f^-\mathrm{d}\mu \\ &= a\left(\int_X f^+\mathrm{d}\mu - \int_X f^-\mathrm{d}\mu\right) \\ &= a\int_X f\mathrm{d}\mu.\end{aligned}$$
If $a < 0$, then $(af)^+ = -af^-$ and $(af)^- = -af^+$. Therefore
$$\begin{aligned}\int_X af\mathrm{d}\mu &= \int_X -af^-\mathrm{d}\mu - \int_X -af^+\mathrm{d}\mu \\ &= -a\int_X f^-\mathrm{d}\mu + a\int_X f^+\mathrm{d}\mu \\ &= a\left(\int_X f^+\mathrm{d}\mu - \int_X f^-\mathrm{d}\mu\right) \\ &= a\int_X (f^+ - f^-)\mathrm{d}\mu \\ &= a\int_X f\mathrm{d}\mu.\end{aligned}$$
Thus we have $\mathbb{R}$-scaling of integration for integrable functions.

2. Observe that
$$\begin{aligned}f+g &= f^+ - f^- + g^+ - g^- \\ &= f^+ + g^+ - (f^- + g^-).\end{aligned}$$
It follows from Proposition (**??**) that
$$\begin{aligned}\int_X (f+g)\mathrm{d}\mu &= \int_X (f^+ + g^+)\mathrm{d}\mu - \int_X (f^- + g^-)\mathrm{d}\mu \\ &= \int_X f^+\mathrm{d}\mu + \int_X g^+\mathrm{d}\mu - \int_X f^-\mathrm{d}\mu - \int_X g^-\mathrm{d}\mu \\ &= \int_X f^+\mathrm{d}\mu - \int_X f^-\mathrm{d}\mu + \int_X g^+\mathrm{d}\mu - \int_X g^-\mathrm{d}\mu \\ &= \int_X f\mathrm{d}\mu + \int_X g\mathrm{d}\mu.\end{aligned}$$
Thus we have additivity of integration for integrable functions.

3. Assume $f \leq g$. Observe that $f \leq g$ implies $g - f \geq 0$. It follows from 1 and 2 that.

$$\begin{aligned}\int_X g\mathrm{d}\mu - \int_X f\mathrm{d}\mu &= \int_X (g-f)\mathrm{d}\mu \\ &\geq 0.\end{aligned}$$

This implies $\int_X g\mathrm{d}\mu \geq \int_X f\mathrm{d}\mu$. □

Remark 7. Note that in each proof for the three statements in Proposition (4.9), we needed to use the fact that these three statements hold when integrating nonnegative *measurable* functions.

4.4.2 Lebesgue Dominated Convergence Theorem

Theorem 4.3. *(DCT) Let $g \colon X \to [0,\infty]$ be a nonnegative integrable function. Suppose $(f_n \colon X \to \mathbb{R})$ is a sequence of integrable functions such that*

1. *(f_n) converges pointwise to $f \colon X \to \mathbb{R}$.*
2. *$|f_n| \leq g$ pointwise for all $n \in \mathbb{N}$.*

Then

$$\lim_{n\to\infty} \int_X f_n \mathrm{d}\mu = \int_X f\mathrm{d}\mu.$$

Proof. Since $|f_n| \leq g$ for all $n \in \mathbb{N}$, we have (by taking limits) $|f| \leq g$. Thus f is integrable by monotonicity of integration for integrable functions. Observe that $(g - f_n)$ is a sequence of nonnegative measurable functions. Thus by Fatou's Lemma, we have

$$\begin{aligned}\int_X g\mathrm{d}\mu - \int_X f\mathrm{d}\mu &= \int_X (g-f)\mathrm{d}\mu \\ &\leq \liminf_{n\to\infty} \int_X (g-f_n)\mathrm{d}\mu \\ &= \int_X g\mathrm{d}\mu - \limsup_{n\to\infty} \int f_n\mathrm{d}\mu.\end{aligned}$$

In other words, we have

$$\limsup_{n\to\infty} \int_X f_n\mathrm{d}\mu \leq \int_X f\mathrm{d}\mu.$$

Now we apply the same argument with functions $g + f_n$ in place of $g - f_n$, and we obtain

$$\liminf_{n\to\infty} \int_X f_n\mathrm{d}\mu \geq \int_X f\mathrm{d}\mu.$$

□

4.4.3 Chebyshev-Markov Inequality

Proposition 4.10. *(C-M) Let $f \colon X \to \mathbb{R}$ be an integrable function and let $c > 0$. Then*

$$\mu\{|f| \geq c\} \leq \frac{1}{c}\int_X |f|\mathrm{d}\mu.$$

Proof. We have

$$\begin{aligned}\int_X |f|\mathrm{d}\mu &\geq \int_X |f| 1_{\{|f|\geq c\}}\mathrm{d}\mu \\ &\geq \int_X c1_{\{|f|\geq c\}}\mathrm{d}\mu \\ &= c\mu\{|f| \geq c\}.\end{aligned}$$

□

Proposition 4.11. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let f be a measurable function. Then $\mu\{f \neq 0\} = 0$ if and only if $\int_X |f| \mathrm{d}\mu = 0$.*

Proof. We first note that that $\{f \neq 0\} = \{|f| \neq 0\}$. Now suppose $\mu\{f \neq 0\} = 0$. Then we have

$$\begin{aligned}
\int_X |f| \mathrm{d}\mu &= \int_{\{f=0\}} |f| \mathrm{d}\mu + \int_{\{f\neq 0\}} |f| \mathrm{d}\mu \\
&= \int_{\{f\neq 0\}} |f| \mathrm{d}\mu \\
&= \sup\{\int_{\{f\neq 0\}} \varphi \mathrm{d}\mu \mid \varphi \leq |f| \text{ is nonnegative simple function}\} \\
&= \sup\{0 \mid \varphi \leq |f| \text{ is nonnegative simple function}\} \\
&= 0.
\end{aligned}$$

Conversely, suppose $\int_X |f| \mathrm{d}\mu = 0$. Then by Chebyshev-Markov's inequality, we have

$$\begin{aligned}
\mu\{f \neq 0\} &= \mu\{|f| \neq 0\} \\
&= \mu\left(\bigcup_{n=1}^{\infty} \{|f| \geq 1/n\}\right) \\
&= \lim_{n\to\infty} \mu(\{|f| \geq 1/n\}) \\
&\leq \lim_{n\to\infty} n \int_X |f| \mathrm{d}\mu \\
&= 0.
\end{aligned}$$

□

Definition 4.6. A property is said to hold μ**-almost everywhere** (or more simpwith respect to μ (denoted μ a.e.) if it is true everywhere except on a set of measure zero. If the measure μ is understood from context, then we will simply write "almost everywhere" instead of "μ-almost everywhere".

Corollary 2. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let f and g be integrable functions. Then $f = g$ be almost everywhere if and only if $\int_X |f| \mathrm{d}\mu = \int_X |g| \mathrm{d}\mu$.*

Proof. Suppose $f = g$ almost everywhere. Then also $|f| = |g|$ almost everywhere, and so we have

$$\begin{aligned}
\int_X |f| \mathrm{d}\mu &= \int_{\{f=g\}} |f| \mathrm{d}\mu + \int_{\{f\neq g\}} |f| \mathrm{d}\mu \\
&= \int_{\{f=g\}} |f| \mathrm{d}\mu \\
&= \int_{\{f=g\}} |g| \mathrm{d}\mu \\
&= \int_{\{f=g\}} |g| \mathrm{d}\mu + \int_{\{f\neq g\}} |g| \mathrm{d}\mu \\
&= \int_X |g| \mathrm{d}\mu.
\end{aligned}$$

Conversely, suppose $\int_X |f| \mathrm{d}\mu = \int_X |g| \mathrm{d}\mu$. Then we have

$$\begin{aligned}
0 &= \int_X |f| \mathrm{d}\mu - \int_X |g| \mathrm{d}\mu \\
&= \int_X (|f| - |g|) \mathrm{d}\mu.
\end{aligned}$$

It follows that $|f| - |g| = 0$ almost everywhere, which implies $f = g$ almost everywhere. □

4.5 L^1-Spaces

Throughout this subsection, let $(X, \mathcal{M}, \mu)$ be a measure space. The set of all integrable function $f\colon X \to \mathbb{R}$ is denoted $\mathrm{Int}^1(X, \mathcal{M}, \mu)$. By Proposotion (4.9) we see that $\mathrm{Int}^1(X, \mathcal{M}, \mu)$ is an $\mathbb{R}$-vector space. We define a relation $\sim$ on $\mathrm{Int}^1(X, \mathcal{M}, \mu)$ as follows, if $f, g \in \mathrm{Int}^1(X, \mathcal{M}, \mu)$, then we set $f \sim g$ if and only if $f = g$ almost everywhere. One checks that $\sim$ is in fact an equivalence relation. We will denote

$$[\mathrm{Int}^1(X, \mathcal{M}, \mu)] := \mathrm{Int}^1(X, \mathcal{M}, \mu)/\!\sim .$$

Thus $[\mathrm{Int}^1(X, \mathcal{M}, \mu)]$ is the set of all integrable functions from X to $\mathbb{R}$ where two such functions are identified if they agree almost everywhere. Technically speaking, elements in $[\mathrm{Int}^1(X, \mathcal{M}, \mu)]$ are *cosets* of functions. Thus an element in $[\mathrm{Int}^1(X, \mathcal{M}, \mu)]$ should be expressed like $[f]$, where f is an integrable function which represents the coset $[f]$. In practice however, we tend to abuse notation by dropping the square brackets around $[f]$ altogether.

We can give $\mathrm{Int}^1(X, \mathcal{M}, \mu)$ the structure of a pseudo-normed space as follows: We define $\|\cdot\|_1\colon \mathrm{Int}^1(X, \mathcal{M}, \mu) \to [0, \infty)$ by

$$\|f\|_1 := \int_X |f| \mathrm{d}\mu$$

for all $f \in \mathrm{Int}^1(X, \mathcal{M}, \mu)$. Observe that monotonicity of integration implies subadditivity of $\|\cdot\|_1$. Also linearity of integration combined with absolute homogeneity of $|\cdot|$ implies absolute homogeneity of $\|\cdot\|_1$. The reason why $\|\cdot\|_1$ is just a pseudo-norm and not a norm is because it lacks the positive-definiteness property: there exists $f \in \mathrm{Int}^1(X, \mathcal{M}, \mu)$ such that $f \neq 0$ but $\|f\|_1 = 0$. On the other hand, $\|\cdot\|_1$ induces a genuine norm when we pass to the quotient space $[\mathrm{Int}^1(X, \mathcal{M}, \mu)]$! Indeed, we define a norm on $[\mathrm{Int}^1(X, \mathcal{M}, \mu)]$, which again denote by $\|\cdot\|_1$, by

$$\|f\|_1 := \int_X |f| \mathrm{d}\mu \tag{17}$$

for all $f \in [\mathrm{Int}^1(X, \mathcal{M}, \mu)]$. Note that (17) is well-defined by Corollary (2). In fact, Corollary (2) tells us that $[\mathrm{Int}^1(X, \mathcal{M}, \mu)]$ is the normed linear space induced by $\mathrm{Int}^1(X, \mathcal{M}, \mu)$ with respect to the equivalence relation $\sim$ introduced above. We will denote this normed linear space by

$$L^1(X, \mathcal{M}, \mu) := ([\mathrm{Int}^1(X, \mathcal{M}, \mu)], \|\cdot\|_1).$$

We often refer to $\|\cdot\|_1$ as the **L^1-norm**.

4.5.1 L^1-Completeness

We now wish to show that $L^1(X, \mathcal{M}, \mu)$ is not just any normed linear space; it is a Banach space.

Theorem 4.4. *$L^1(X, \mathcal{M}, \mu)$ is a Banach space.*

To prove this we'll use the following criterion to test for completeness in a normed linear space.

Lemma 4.5. *Let $\mathcal{X}$ be a normed linear space. Then $\mathcal{X}$ is a Banach space if and only if every absolutely convergent series in $\mathcal{X}$ is convergent.*

Proof. Suppose first that every absolutely convergent series in $\mathcal{X}$ is convergent. Let (x_n) be a Cauchy sequence in $\mathcal{X}$. To show that (x_n) is convergent, it suffices to show that a subsequence of (x_n) is convergent, by Lemma (12.1). Choose a subsequence $(x_{\pi(n)})$ of (x_n) such that

$$\|x_{\pi(n)} - x_{\pi(n-1)}\| < \frac{1}{2^n}$$

and for all $n \in \mathbb{N}$ (we can do this since (x_n) is Cauchy). Then the series $\sum_{n=1}^{\infty}(x_{\pi(n)} - x_{\pi(n-1)})$ is absolutely convergent since

$$\begin{aligned} \sum_{n=1}^{\infty} \|x_{\pi(n)} - x_{\pi(n-1)}\| &< \sum_{n=1}^{\infty} \frac{1}{2^n} \\ &= 1. \end{aligned}$$

Therefore it must be convergent, say $\sum_{n=1}^{\infty}(x_{\pi(n)} - x_{\pi(n-1)}) \to x$. On the other hand, for each $n \in \mathbb{N}$, we have

$$x_{\pi(n)} - x_{\pi(1)} = \sum_{m=1}^{n}(x_{\pi(m)} - x_{\pi(1)}).$$

In particular, $x_{\pi(n)} \to x - x_{\pi(1)}$ as $n \to \infty$. Thus $(x_{\pi(n)})$ is a convergent subsequence of (x_n).

Conversely, suppose $\mathcal{X}$ is a Banach space and suppose $\sum_{n=1}^{\infty} x_n$ is absolutely convergent. Let $\varepsilon > 0$ and choose $K \in \mathbb{N}$ such that $N \geq M \geq K$ implies

$$\sum_{n=M}^{N} \|x_n\| < \varepsilon.$$

Then $N \geq M \geq K$ implies

$$\begin{aligned}\left\|\sum_{n=1}^{N} x_n - \sum_{n=1}^{M} x_n\right\| &= \left\|\sum_{n=M}^{N} x_n\right\| \\ &\leq \sum_{n=M}^{N} \|x_n\| \\ &< \varepsilon.\end{aligned}$$

It follows that the sequence of partial sums $(\sum_{n=1}^{N} x_n)_N$ is Cauchy. Since $\mathcal{X}$ is a Banach space, it follows that $\sum_{n=1}^{\infty} x_n$ is convergent. □

Now we prove Theorem (4.4).

Proof. By Lemma (12.2), it suffices to show that every absolutely convergent sequence is convergent. Suppose (f_n) is a sequence in $L^1(X, \mathcal{M}, \mu)$ such that $\sum_{n=1}^{\infty} \|f_n\|_1 < \infty$. Denote by

$$G_N = \sum_{n=1}^{N} |f_n| \quad \text{and} \quad G = \sum_{n=1}^{\infty} |f_n|.$$

Observe that (G_N) is an increasing sequence of nonnegative measurable functions which converges pointwise to G. Therefore by MCT we have

$$\begin{aligned}\|G\|_1 &= \lim_{N\to\infty} \|G_N\|_1 \\ &= \lim_{N\to\infty} \int_X G_N \mathrm{d}\mu \\ &= \lim_{N\to\infty} \int_X \sum_{n=1}^{N} |f_n| \mathrm{d}\mu \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \int_X |f_n| \mathrm{d}\mu \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \|f_n\|_1 \mathrm{d}\mu \\ &= \sum_{n=1}^{\infty} \|f_n\|_1 \\ &< \infty.\end{aligned}$$

It follows that $G \in L^1(X, \mathcal{M}, \mu)$. In particular, $\{G = \infty\}$ has measure zero, so if we denote $F_N = \sum_{n=1}^{N} f_n$ and if we define $F \colon X \to \mathbb{R}$ by

$$F(x) = \begin{cases} 0 & \text{if } G(x) = \infty \\ \sum_{n=1}^{\infty} f_n(x) & \text{if } G(x) < \infty \end{cases}$$

then we see that (F_N) converges pointwise to F almost everywhere, so the sequence $(F - F_N)$ converges pointwise to the zero function almost everywhere. Note that if $G(x) < \infty$, then $\sum_{n=1}^{\infty} f_n(x)$ converges since it converges absolutely, so the definition of F is well-posed. Also note that $F_N, F \in L^1(X, \mathcal{M}, \mu)$ since $|F_N|, |F| \leq G$ and

$G \in L^1(X, \mathcal{M}, \mu)$. Finally, we note that $(F - F_N)$ is dominated by the integrable function $2G$. It follows from the DCT that

$$\begin{aligned}\lim_{N\to\infty} \|F - F_N\|_1 &= \lim_{N\to\infty} \int_X |F_N - F| \mathrm{d}\mu \\ &= \int_X 0 \mathrm{d}\mu \\ &= 0.\end{aligned}$$

Therefore every absolutely convergent series is convergent, which implies $L^1(X, \mathcal{M}, \mu)$ is complete. □

What do we need to choose $(X, \mathcal{M}, \mu)$ to be such that $L^1(X, \mathcal{M}, \mu) = \ell^1(\mathbb{N})$. We need $X = \mathbb{N}$, $\mathcal{M} = \mathcal{P}(\mathbb{N})$, and μ to be counting measure. In this case, if $f \colon \mathbb{N} \to [0, \infty)$, then we have

$$\begin{aligned}\|f\|_1 &= \int_{\mathbb{N}} f \mathrm{d}\mu \\ &= \sum_{n=1}^{\infty} f(n)\mu(\{n\}) \\ &= \sum_{n=1}^{\infty} f(n) \\ &= \sum_{n=1}^{\infty} |f(n)).\end{aligned}$$

By using $f = f^+ - f^-$ we can easily see that $\|f\|_1 = \sum_{n=1}^{\infty} f(n)$.

4.5.2 Set of all Integrable Simple Functions is a Dense Subspace of $L^1(X, \mathcal{M}, \mu)$

We denote by $\mathrm{S}(X, \mathcal{M}, \mu)$ to be the set of all integrable simple functions where we identify two simple functions if they agree almost everywhere. It is easy to check that $\| \cdot \|_1$ restricts to a norm on $\mathrm{S}(X, \mathcal{M}, \mu)$ making it into a normed linear subspace of $L^1(X, \mathcal{M}, \mu)$. We now want to show that $\mathrm{S}(X, \mathcal{M}, \mu)$ is a dense subspace of $L^1(X, \mathcal{M}, \mu)$.

Proposition 4.12. $\mathrm{S}(X, \mathcal{M}, \mu)$ *is a dense subspace of* $L^1(X, \mathcal{M}, \mu)$.

Proof. Let $f \in L^1(X, \mathcal{M}, \mu)$. Decompose f into its positive and negative parts:

$$f = f^+ - f^-.$$

There exists an increasing sequence (φ_n) of nonnegative simple functions which converges to f^+ pointwise. Similarly, there exists an increasing sequence (ψ_n) of nonnegative simple functions which converges to f^- pointwise. Then $(\varphi_n + \psi_n)$ is an increasing sequence of nonnegative simple functions which converges pointwise to $|f|$. Also note that $(\varphi_n - \psi_n)$ is a sequence of simple functions which converges pointwise to f. Now set $s_n = \varphi_n - \psi_n$. We claim that $\|s_n - f\|_1 \to 0$ as $n \to \infty$. In other words, we claim that

$$\int_X |s_n - f| \mathrm{d}\mu \to 0$$

as $n \to \infty$. To this, we'll use DCT. Observe that for each $n \in \mathbb{N}$ we have

$$\begin{aligned}|s_n - f| &\leq |s_n| + |f| \\ &= |\varphi_n - \psi_n| + |f| \\ &\leq \varphi_n + \psi_n + |f| \\ &\leq |f| + |f| \\ &\leq 2|f|\end{aligned}$$

So $2|f|$ is a dominating function, which means we can apply DCT. Therefore

$$\begin{aligned}\lim_{n\to\infty} \|s_n - f\|_1 &= \lim_{n\to\infty} \int_X |s_n - f| \mathrm{d}\mu \\ &= \int_X 0 \mathrm{d}\mu \\ &= 0.\end{aligned}$$

□

4.5.3 $C[a,b]$ is a dense subspace of $L^1[a,b]$

Proposition 4.13. *The space of continuous functions* $C[a,b]$ *is a dense subspace of* $L^1[a,b]$.

Proof. By the previous proposition, it is enough to show that any simple function can be approximated arbitrarily well with a continuous function. Let φ be a simple function and write its canonical form as

$$\varphi = \sum_{k=1}^{n} c_k 1_{E_k}$$

where each E_k is a Borel subset of $[a,b]$. We proved earlier in the semester that for each $\varepsilon > 0$ and Borel set E_k there exists A_k in the interval algebra of $[a,b]$ such that $\mathrm{d}_m(E_k, A_k) < \varepsilon$ where m is the Lebesgue measure. In fact, we have

$$\mathrm{m}(E_k \Delta A_k) < \varepsilon \iff \int_a^b |1_{E_k} - 1_{A_k}| \mathrm{dm} < \varepsilon \iff \|1_{E_k} - 1_{A_k}\|_1 < \varepsilon.$$

Since each A_k in the interval algebra is a finite union of intervals, it is enough to show that 1_I, where I is a subinterval of $[a,b]$ can be approximated arbitrarily well by a continuous function. For this we can use trapezoid functions just like in HW1. Thus $L^1[a,b]$ is a completion of $C[a,b]$ with respect to $\|\cdot\|_1$ norm. □

4.6 L^p-Spaces

Throughout this subsection, let $(X, \mathcal{M}, \mu)$ be a measure space. Let $1 < p < \infty$. Just as we defined the notion of an integral function, we can also define the notion of a ***p*-integral** function. In particular, if f is a measurable function, then we say it is p-integral if

$$\int_X |f|^p \mathrm{d}\mu < \infty.$$

We denote by $\mathrm{Int}^p(X, \mathcal{M}, \mu)$ to be the set of all p-integral functions from X to $\mathbb{R}$. It is straightforward to check that $\mathrm{Int}^p(X, \mathcal{M}, \mu)$ is an $\mathbb{R}$-vector space just like in the case of $\mathrm{Int}^1(X, \mathcal{M}, \mu)$, and just like before, we denote $[\mathrm{Int}^p(X, \mathcal{M}, \mu)]$ to be the quotient space induced by $\sim$. We can equip $[\mathrm{Int}^p(X, \mathcal{M}, \mu)]$ with another norm, which we call the **L^p-norm**. It is defined by

$$\|f\|_p = \left(\int_X |f|^p \mathrm{d}\mu \right)^{\frac{1}{p}}$$

for all $f \in [\mathrm{Int}^p(X, \mathcal{M}, \mu)]$. We will denote this normed linear space by

$$L^p(X, \mathcal{M}, \mu) := ([\mathrm{Int}^p(X, \mathcal{M}, \mu)], \|\cdot\|_p).$$

Absolute homogeneity and positive-definiteness of $\|\cdot\|_p$ are straightforward to check. It turns out that subadditivity of $\|\cdot\|_p$ takes a little more work to show. The idea is to use the so called **Hölder's inequality**. However this inequality itself relies on another inequality called **Young's inequality**, so let's begin with that. We often refer to $\|\cdot\|_1$ as the **L^1-norm**.

4.6.1 Young's Inequality

Lemma 4.6. *Let x and y be nonnegative real numbers and let $0 < \gamma < 1$. Then we have*

$$x^\gamma y^{1-\gamma} \leq \gamma x + (1-\gamma) y. \tag{18}$$

Proof. We may assume that $x, y > 0$ since otherwise it is trivial. Set $t = x/y$ and rewrite (35) as

$$t^\gamma - \gamma t \leq 1 - \gamma. \tag{19}$$

Thus, to show (35) for all $x, y > 0$, we just need to show (36) for all $t > 0$. To see why (36) holds, define $f \colon \mathbb{R}_{>0} \to \mathbb{R}$ by

$$f(t) = t^\gamma - \gamma t$$

for all $t \in \mathbb{R}_{>0}$. Observe that f is a smooth function on $\mathbb{R}_{>0}$, with it's first derivative and second derivative given by

$$f'(t) = \gamma t^{\gamma-1} - \gamma \quad \text{and} \quad f''(t) = \gamma(\gamma-1) t^{\gamma-2}$$

for all $t \in \mathbb{R}_{>0}$ respectively. Observe that

$$\begin{aligned} f'(t) = 0 &\iff \gamma t^{\gamma-1} = \gamma \\ &\iff t^{\gamma-1} = 1 \\ &\iff t = 1, \end{aligned}$$

where the last if and only if follows from the fact that t is a positive real number. Also, we clearly have $f''(t) < 0$ for all $t \in \mathbb{R}_{>0}$. Thus, since f is concave down on all of $\mathbb{R}_{>0}$, and $f'(t) = 0$ if and only if $t = 1$, it follows that f has a global maximum at $t = 1$. In particular, we have

$$\begin{aligned} t^\gamma - \gamma t &= f(t) \\ &\leq f(1) \\ &\leq 1^\gamma - \gamma \cdot 1 \\ &= 1 - \gamma \end{aligned}$$

for all $t \in \mathbb{R}_{>0}$. □

Proposition 4.14. *(Young's inequality) Let a and b be nonnegative real numbers and let $1 \leq p, q < \infty$ such that $1/p + 1/q = 1$. Then we have*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof. Set $\gamma = 1/p$ (so $1 - \gamma = 1/q$), $a = x^\gamma$, and $b = y^{1-\gamma}$. Then Young's Inequality becomes (18), which was proved in Lemma (4.6). □

4.6.2 Hölder's Inequality

Proposition 4.15. *(Hölder's inequality) Let $1 < p, q < \infty$ such that $1/p + 1/q = 1$, let $f \in L^p(X, \mathcal{M}, \mu)$, and let $g \in L^q(X, \mathcal{M}, \mu)$. Then $fg \in L^1(X, \mathcal{M}, \mu)$ and*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \tag{20}$$

Proof. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then one of them is equal to 0 almost everywhere. In this case $fg = 0$ almost everywhere. Thus the inquality is trivial in this case, so we may assume that $\|f\|_p \neq 0$ and $\|g\|_q \neq 0$. We will first show the inequality in the special case where $\|f\|_p = 1 = \|g\|_q$. Then the righthand side of (20) is 1, so we need to show $\|fg\|_1 \leq 1$. This follows immediately from Young's inequality:

$$\begin{aligned} \|fg\|_1 &= \int_X |fg| \mathrm{d}\mu \\ &\leq \int_X \left(\frac{|f|^p}{p} + \frac{|g|^q}{q} \right) \mathrm{d}\mu \\ &= \frac{1}{p} \int_X |f|^p \mathrm{d}\mu + \frac{1}{q} \int_X |g|^q \mathrm{d}\mu \\ &= \frac{1}{p} + \frac{1}{q} \\ &= 1. \end{aligned}$$

Now we prove the general case. Let $F = f/\|f\|_p$ and $G = g/\|g\|_q$. Then $\|F\|_p = 1 = \|G\|_q$. Applying the special case that we just proved, we have

$$\begin{aligned} 1 &\geq \int_X |FG| \mathrm{d}\mu \\ &= \int_X \left| \frac{f}{\|f\|_p} \frac{g}{\|g\|_q} \right| \mathrm{d}\mu \\ &= \frac{1}{\|f\|_p \|g\|_q} \int_X |fg| \mathrm{d}\mu \\ &= \frac{1}{\|f\|_p \|g\|_q} \|fg\|_1. \end{aligned}$$

After multiplying both sides by $\|f\|_p \|g\|_q$, we obtain Hölder's inequality. □

4.6.3 Minkowski's Inequality

For historical reasons, subadditivity of $\|\cdot\|_p$ is refered to as **Minkowski's inequality**.

Proposition 4.16. *(Minkowski's inequality) Let* $1 \le p < \infty$ *and let* $f, g \in L^p(X, \mathcal{M}, \mu)$. *Then*

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

Proof. We proved this for $p = 1$, so let $p > 1$. If $\|f+g\|_p = 0$, then the inequality is obvious, so we can assume $\|f+g\|_p > 0$. Observe that for q such that $1/p + 1/q = 1$, we have $(p-1)q = p$. Thus we have

$$\begin{aligned}
\|f+g\|_p &= \||f+g|^p\|_1^{1/p} \\
&= \||f+g|^p\|_1^{1-1/q} \\
&= \||f+g|^p\|_1 \||f+g|^p\|_1^{-1/q} \\
&= \||f+g||f+g|^{p-1}\|_1 \||f+g|^p\|_1^{-1/q} \\
&\le \left(\||f||f+g|^{p-1}\|_1 + \||g||f+g|^{p-1}\|_1\right) \||f+g|^p\|_1^{-1/q} \\
&\le \left(\|f\|_p \||f+g|^{p-1}\|_q + \|g\|_p \||f+g|^{p-1}\|_q\right) \||f+g|^p\|_1^{-1/q} \\
&= (\|f\|_p + \|g\|_p) \||f+g|^{p-1}\|_q \||f+g|^p\|_1^{-1/q} \\
&= (\|f\|_p + \|g\|_p) \||f+g|^{q(p-1)}\|_1^{1/q} \||f+g|^p\|_1^{-1/q} \\
&= (\|f\|_p + \|g\|_p) \||f+g|^p\|_1^{1/q} \||f+g|^p\|_1^{-1/q} \\
&= \|f\|_p + \|g\|_p.
\end{aligned}$$

□

4.7 Types of Convergences

In this subsection, we wish to to discuss various types of convergences of sequences of functions.

Definition 4.7. Let $(X, \mathcal{M}, \mu)$ be a measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, and let $f \colon X \to \mathbb{R}$ be a function.

1. We say (f_n) converges **pointwise** to f, denoted $f_n \xrightarrow{\text{pw}} f$, if for all $x \in X$ and for all $\varepsilon > 0$ there exists $N_{x,\varepsilon} \in \mathbb{N}$ (which depends on x and ε) such that $n \geq N_{x,\varepsilon}$ implies
$$|f_n(x) - f(x)| < \varepsilon.$$
2. We say (f_n) converges **uniformly** to f, denoted $f_n \xrightarrow{\text{u}} f$, if for all $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ (which only depends on ε) such that $n \geq N_\varepsilon$ implies
$$|f_n(x) - f(x)| < \varepsilon$$
3. We say (f_n) converges to f **almost uniformly**, denoted $f_n \xrightarrow{\text{au}} f$, if for all $\delta > 0$ there exists a set $E_\delta \subseteq X$ with $\mu(E_\delta) < \delta$ such that $f_n \to f$ uniformly on E_δ^c. In other words, (f_n) converges to f almost uniformly if for all $\varepsilon, \delta > 0$ there exists $E_\delta \subseteq X$ with $\mu(E_\delta) < \delta$ and $N_{\varepsilon,\delta} \in \mathbb{N}$ such that $n \geq N_\varepsilon$ implies
$$|f_n(x) - f(x)| < \varepsilon$$
for all $x \in E_\delta^c$.
4. We say (f_n) converges to f in **measure zero**, denoted $f_n \xrightarrow{\text{m}} f$, if for all $\varepsilon > 0$ we have
$$\mu\{f_n - f \geq \varepsilon\} \to 0$$
as $n \to \infty$. In other words, for all $\varepsilon, \delta > 0$, there exists $N_{\varepsilon,\delta} \in \mathbb{N}$ such that $n \geq N_{\varepsilon,\delta}$ implies
$$\mu\{f_n - f \geq \varepsilon\} < \delta.$$
5. A convergence is said to hold **almost everywhere** if there exists a set $E \in \mathcal{M}$ such that $\mu(E) = 0$ and such that the convergence holds on E^c.

4.7.1 Almost Pointwise is Equivalent to Pointwise Almost Everywhere

It may seem reasonable to make an additonal defintion: we say (f_n) converges **almost pointwise** to f if for all $\delta > 0$ there exists $E_\delta \in \mathcal{M}$ such that $\mu(E_\delta) < \delta$ and (f_n) converges pointwise to f on E_δ^c. It's clear that if (f_n) converges to f pointwise almost everywhere, then (f_n) converges almost pointwise to f. In fact, the converse holds too! So "almost pointwise" and "pointwise almost everywhere" are equivalent notions. Let us show this.

Proposition 4.17. *Let $(X, \mathcal{M}, \mu)$ be a measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, and let $f \colon X \to \mathbb{R}$ be a function. Then (f_n) converges almost pointwise to f if and only if (f_n) converges to f pointwise almost everywhere.*

Proof. We just need to show that almost pointwise implies pointwise almost everywhere, since the converse direction is trivial. Suppose (f_n) converges almost pointwise to f. For each $k \in \mathbb{N}$, choose $E_k \in \mathcal{M}$ such that $\mu(E_k) < 1/k$ and (f_n) converges to f pointwise on E_k^c. Set
$$E = \bigcap_{k=1}^{\infty} E_k.$$
Clearly we have $\mu(E) = 0$. We claim that (f_n) converges to pointwise to f on E^c. Indeed, let $x \in E^c$ and let $\varepsilon > 0$. Since
$$E^c = \bigcup_{k=1}^{\infty} E_k^c,$$
there exists a $k \in \mathbb{N}$ such that $x \in E_k^c$, so we choose such a $k \in \mathbb{N}$. Then since (f_n) converges pointwise to f on E_k^c, we can choose $N \in \mathbb{N}$ such that $n \geq N$ implies
$$|f_n(x) - f(x)| < \varepsilon.$$
It follows that (f_n) converges pointwise to f almost everywhere. Note that N depends on k, ε, and x, but this isn't a problem! □

4.7.2 Uniform Convergence on Finite Measure Space Implies L^p Convergence

Proposition 4.18. *Let $(X, \mathcal{M}, \mu)$ be a finite measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, let $f \colon X \to \mathbb{R}$ be a function, and let $0 < p < \infty$. Suppose (f_n) converges to f uniformly on X. Then (f_n) converges to f in the L^p-norm.*

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$|f_n(x) - f(x)| < \frac{\varepsilon^{1/p}}{\mu(X)}$$

for all $x \in X$. Then $n \geq N$ implies

$$\begin{aligned} \|f_n - f\|_p &= \int_X |f_n - f|^p \mathrm{d}\mu \\ &\leq \sup |f_n - f|^p \mu(X) \\ &< \frac{\varepsilon^{1/p}}{\mu(X)} \mu(X) \\ &= \varepsilon. \end{aligned}$$

It follows that (f_n) converges to f in the L^p-norm. □

4.7.3 Convergence in L^p Implies Convergence in Measure Zero

Proposition 4.19. *Let $(X, \mathcal{M}, \mu)$ be a measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, let $f \colon X \to \mathbb{R}$ be a function, and let $0 < p < \infty$. Suppose (f_n) converges to f in the L^p-norm. Then (f_n) converges to f in measure zero.*

Proof. Let $\varepsilon, \delta > 0$. Choose $N_{\varepsilon,\delta} \in \mathbb{N}$ such that $n \geq N_{\varepsilon,\delta}$ implies

$$\|f_n - f\|_p < \varepsilon^{1/p} \delta^{1/p}.$$

Then it follows from Chebyshev's inequality that

$$\begin{aligned} \mu\{|f_n - f| \geq \varepsilon^{1/p}\} &\leq \mu\{|f_n - f|^p \geq \varepsilon\} \\ &\leq \frac{1}{\varepsilon} \|f_n - f\|_p^p \\ &< \frac{1}{\varepsilon} \varepsilon\delta \\ &= \delta. \end{aligned}$$

Therefore (f_n) converges to f in meaure zero. □

4.7.4 Convergence in L^p Does Not Imply Convergence Pointwise Almost Everywhere and Vice-Versa

Example 4.1. Let $(X, \mathcal{M}, \mu)$ be a measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, let $f \colon X \to \mathbb{R}$ be a function, and let $0 < p < \infty$. Clearly pointwise almost everywhere convergence does not imply L^p-convergence. It also turns out that if (f_n) converges to f in the L^p-norm, then it is not necessarily the case that (f_n) converges to f pointwise almost everywhere. Similarly, if (f_n) converges to f pointwise almost everywhere, then it is not necessarily the case that (f_n) converges to f in the L^p-norm. To see this, consider the case where $X = [0, 1]$, $\mu = \mathrm{m}$ is the Lebesgue measure, and (f_n) is the sequence of functions which starts out as

$$\begin{aligned} f_1 &= 1_{[0,1]} \\ f_2 &= 1_{[0,1/2]} \\ f_3 &= 1_{[1/2,1]} \\ f_4 &= 1_{[0,1/4]} \\ f_5 &= 1_{[1/4,1/2]} \\ f_6 &= 1_{[1/2,3/4]} \\ f_7 &= 1_{[3/4,1]} \\ f_8 &= 1_{[0,1/8]} \\ &\vdots \end{aligned}$$

and so on. It is easy to check that (f_n) converges to the 0 function in the L^1-norm. However it does not converge pointwise almost everywhere to any function! Indeed, for any $x \in [0,1]$, we have $f_n(x) = 0$ for infinitely many n and $f_n(x) = 1$ for infinitely many n.

4.7.5 Convergence in Measure Zero Implies a Subsequence Converges Pointwise Almost Everywhere

Lemma 4.7. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a sequence in $\mathcal{M}$ such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then*

$$\mu\left(\limsup E_n\right) = 0.$$

Proof. Note that the sequence

$$\left(\bigcup_{n \geq N} E_n\right)_{N \in \mathbb{N}}$$

is a descending sequence in N. This together with the fact that $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) < \infty$ implies

$$\begin{aligned} \mu\left(\limsup E_n\right) &= \mu\left(\bigcap_{N=1}^{\infty}\left(\bigcup_{n \geq N} E_n\right)\right) \\ &= \lim_{N \to \infty} \mu\left(\bigcup_{n \geq N} E_n\right) \\ &\leq \lim_{N \to \infty} \sum_{n=N}^{\infty} \mu(E_n) \\ &= 0, \end{aligned}$$

where the last equality follows since $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. □

Proposition 4.20. *Let $(X, \mathcal{M}, \mu)$ be a measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, and let $f \colon X \to \mathbb{R}$ be a function. Suppose (f_n) converges to f in measure zero. Then there exists a subsequence $(f_{\pi(n)})$ of (f_n) such that $(f_{\pi(n)})$ converges to f pointwise almost everywhere.*

Proof. For each $m \in \mathbb{N}$ choose $\pi(m) \geq m$ such that $n \geq \pi(m)$ implies

$$\mu\{|f_n - f| \geq 1/2^m\} < 1/2^m. \tag{21}$$

In fact, we don't need to know that (21) holds for *every* $n \geq \pi(m)$, we just need to know that it holds for $n = \pi(m)$! Indeed, for each $m \in \mathbb{N}$, denote

$$E_m = \{|f_{\pi(m)} - f| \geq 1/2^m\}$$

Then $\mu(E_m) < 1/2^m$ since this is a special case of (21) where $n = \pi(m)$. Next, denote

$$E = \limsup E_m = \bigcap_{M \geq 1} \bigcup_{m \geq M} E_m$$

Observe that $\sum_{m=1}^{\infty} \mu(E_m) < 1$. It follows from Lemma (4.7) that $\mu(E) = 0$. We claim that (f_n) converges pointwise to f on E^c. To see this, let $x \in E^c$. Then there exists some M such that $x \notin \bigcup_{m \geq M} E_m$. In other words, there exists some M such that $x \notin E_m$ for any $m \geq M$. In other words, there exists some M such that

$$|f_{\pi(m)}(x) - f(x)| < 1/2^m$$

for all $m \geq M$. Therefore $f_{\pi(m)}(x) \to f(x)$ as $m \to \infty$. This implies $(f_{\pi(m)})$ converges to f pointwise almost everywhere. □

Corollary 3. *Let $(X, \mathcal{M}, \mu)$ be a measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, and let $f \colon X \to \mathbb{R}$ be a function. Suppose (f_n) converges to f in the L^p-norm. Then there exists a subsequence $(f_{\pi(n)})$ of (f_n) such that $(f_{\pi(n)})$ converges to f pointwise almost everywhere.*

4.7.6 Convergence Pointwise Almost Everywhere on a Finite Measure Space Implies Almost Uniform Convergence

Proposition 4.21. *Let $(X, \mathcal{M}, \mu)$ be a finite measure space, let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions, let $f \colon X \to \mathbb{R}$ be a function, and let $0 < p < \infty$. Suppose (f_n) converges to f pointwise almost everywhere. Then (f_n) converges to f almost uniformly.*

Proof. Let $\varepsilon, \delta > 0$. We must find an $E_\delta \in \mathcal{M}$ and an $N_{\varepsilon,\delta} \in \mathbb{N}$ such that $\mu(E_\delta) < \delta$ and $n \geq N_{\varepsilon,\delta}$ implies

$$|f_n(x) - f(x)| < \frac{1}{k}$$

for all $x \in E_\delta^c$. Before we begin with the proof, let us state up front what E_δ and $N_{\varepsilon,\delta}$ will be. We will have

$$E_\delta = \bigcup_{k=1}^{\infty} \bigcup_{n \geq \pi(k)} \{|f_n - f| \geq 1/k\}.$$

Here, $\pi(k) \geq k$ will be chosen large enough such that $\mu(E_\delta) < \delta$. So $x \in E_\delta$ if and only if there exists $k \in \mathbb{N}$ and an $n \geq \pi(k)$ such that

$$|f_n(x) - f(x)| \geq \frac{1}{k}$$

and $x \in E_\delta^c$ if and only if for all $k \in \mathbb{N}$ we have

$$|f_n(x) - f(x)| < \frac{1}{k}$$

for all $n \geq \pi(k)$. Thus if we choose $k \in \mathbb{N}$ such that $1/k < \varepsilon$, then we set $N_{\varepsilon,\delta} = \pi(k)$. It will then follow that $n \geq N_{\varepsilon,\delta}$ implies

$$|f_n(x) - f(x)| < \varepsilon$$

for all $x \in E_\delta^c$. This will imply almost uniform convergence. So the main goal of the proof is to find $\pi(k)$ for each k such that $\mu(E_\delta) < \delta$.

For each $k, N \in \mathbb{N}$, let

$$E_N(k) = \bigcup_{n \geq N} \{|f_n - f| \geq 1/k\}.$$

Observe that $(E_N(k))_{N \in \mathbb{N}}$ is a descending sequence of sets in N. Furthermore, we observe that

$$\mu\left(\bigcap_{N=1}^{\infty} E_N(k)\right) = 0.$$

Indeed, $x \in \bigcap_{N=1}^{\infty} E_N(k)$ if and only if there exists a subsequence $(f_{\rho(n)})$ of (f_n) such that

$$|f_{\rho(n)}(x) - f(x)| \geq 1/k$$

for all $n \in \mathbb{N}$ if and only if $f_n(x) \not\to f(x)$. This set has measure zero by assumption. Now, since $\mu(X) < \infty$, we have $\mu(E_1(k)) < \infty$, and thus

$$\begin{aligned} \lim_{N \to \infty} \mu(E_N(k)) &= \mu\left(\bigcap_{N=1}^{\infty} E_N(k)\right) \\ &= 0. \end{aligned}$$

It follows that for each $k \in \mathbb{N}$, we can choose $\pi(k) \in \mathbb{N}$ such that

$$\mu(E_{\pi(k)}(k)) < \frac{\delta}{2^k}.$$

Now as we discuss earlier, we set $E_\delta = \bigcup_{k=1}^\infty E_{\pi(k)}(k)$. Then we have

$$\begin{aligned}
\mu(E_\delta) &= \mu\left(\bigcup_{k=1}^\infty E_{\pi(k)}(k)\right) \\
&\leq \sum_{k=1}^\infty \mu\left(E_{\pi(k)}(k)\right) \\
&< \sum_{k=1}^\infty \frac{\delta}{2^k} \\
&= \delta.
\end{aligned}$$

□

4.7.7 Convergences Diagram

We summarize our findings in the convergence implications diagram below.

$$\begin{tikzcd}
& f_n \xrightarrow{\mathrm{u}} f \arrow[dl] \arrow[dr, "\mu(X)<\infty"] & \\
f_n \xrightarrow{\mathrm{au}} f \arrow[d, shift left] \arrow[drr] & & f_n \xrightarrow{L^1} f \arrow[d] \\
f_n \xrightarrow{\mathrm{pwae}} f \arrow[u, shift left, "\mu(X)<\infty"] & & f_n \xrightarrow{\mathrm{m}} f \arrow[ll, dashed, "\text{subseq}"] \arrow[ull, dashed, "\text{subseq}"']
\end{tikzcd}$$

An unlabeled undashed arrow indicates implication. For instance, $f_n \xrightarrow{\mathrm{u}} f$ implies $f_n \xrightarrow{\mathrm{au}} f$. A labeled undashed arrows indicates implication if the labeled condition holds. For instance, if $\mu(X) < \infty$, then $f_n \xrightarrow{\mathrm{u}} f$ implies $f_n \xrightarrow{L^1} f$. Finally a labeled dashed arrow indicaties partial implication. For instance, $f_n \xrightarrow{\mathrm{m}} f$ implies $f_{\pi(n)} \xrightarrow{\mathrm{pwae}} f$ for some subsequence $(f_{\pi(n)})$ of (f_n).

5 Product Measures

Recall the following HW problem:

$$\sum_{n=1}^\infty \int_X f_n \mathrm{d}\mu = \int_X \left(\sum_{n=1}^\infty f_n\right) \mathrm{d}\mu \tag{22}$$

whenever f_n is a nonnegative measurable function. We can think of $\sum_{n=1}^\infty$ as an integral on $\mathbb{N}$ with respect to the counting measure. If we denote the counting measure by σ, we can write (22) as

$$\int_{\mathbb{N}} \int_X f(x,n) \mathrm{d}\mu(x) \mathrm{d}\sigma(n) = \int_X \int_{\mathbb{N}} f(x,n) \mathrm{d}\sigma(n) \mathrm{d}\mu(x).$$

Thus we can exchange the order of integration whenever $f(x,n) \geq 0$.

5.1 Definining the Product σ-Algebra

Throughout the rest of this section, let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two finite measure spaces. We want to equip $X \times Y$ with a σ-algebra from the σ-algebras $\mathcal{M}$ and $\mathcal{N}$ and we want to equip a measure to this σ-algebra from the measures μ and ν. Clearly any set of the form $A \times B$ where $A \in \mathcal{M}$ and $B \in \mathcal{N}$ shoud be included in this σ-algebra that we want to construct. These sets are called **measurable rectangles**. Let us denote by $\mathcal{R}(\mathcal{M}, \mathcal{N})$ to the collection of all measurable rectangles. If $\mathcal{M}$ and $\mathcal{N}$ are understood from context (as in our case at the moment), then we simply denote this by $\mathcal{R}$ rather than $\mathcal{R}(\mathcal{M}, \mathcal{N})$. Observe that $\mathcal{R}$ forms a semialgebra. Indeed,

1. It is closed under finite intersections: $(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2)$

2. Complements can be expressed as finite pairwise disjoint unions: $(A \times B)^c = (X \times B^c) \cup (A^c \times Y)$

If $\mathcal{M}$ and $\mathcal{N}$ are understood from context, then we simply denote this by $\mathcal{R}$ rather than $\mathcal{R}(\mathcal{M}, \mathcal{N})$. Unfortunately $\mathcal{R}$ does not form a σ-algebra. The **product σ-algebra** $\mathcal{M} \otimes \mathcal{N}$ is defined to be the smallest σ-algebra which contains all the measurable rectangles. Thus

$$\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{R}).$$

On the other hand, $\mathcal{R}$ forms a semi-algebra:

1. It is closed under finite intersections: $(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2)$

2. Complements can be expressed as finite pairwise disjoint unions: $(A \times B)^c = (X \times B^c) \cup (A^c \times B)$

Therefore the collection $\mathcal{A}$ formed by the collection of all finite disjoint unions of members of $\mathcal{R}$ forms an algebra. Whenever we start with a finite premeasure on a semialgebra $\mathcal{R}$, we can always extend it (uniquely!) to a finite premeasure to $\mathcal{A}$, and then extend it (uniquely!) to a finite measure on $\sigma(\mathcal{A})$. With this in mind, we define the **product σ-algebra**, denoted $\mathcal{M} \otimes \mathcal{N}$, to be the smallest σ-algebra which contains all the measurable rectangles. Thus

$$\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{A}).$$

We define a finite measure on $\mathcal{M} \otimes \mathcal{N}$, denoted $\mu \otimes \nu \colon \mathcal{M} \otimes \mathcal{N} \to [0, \infty]$, by first defining on $\mathcal{R}$ by

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$$

for all $A \times B \in \mathcal{R}$, and then extending it uniquely to all $\mathcal{A}$, and then finally extending it uniquely to all of $\mathcal{M} \otimes \mathcal{N}$. The extension from $\mathcal{A}$ to $\mathcal{M} \otimes \mathcal{N}$ is obtained from Theorem (3.1). How do we extend it from $\mathcal{R}$ to $\mathcal{A}$? We do this as follows: let $A_1 \times B_1, \ldots, A_n \times B_n$ be a pairwise disjoint sequence of measurable rectangles. Then we define

$$(\mu \otimes \nu)\left(\bigcup_{i=1}^{n} A_i \times B_i\right) = \sum_{i=1}^{n} \mu(A_i)\nu(B_i). \tag{23}$$

One must be careful, as we need to check that (23) is well-defined. In particular, suppose $A \times B = \bigcup_{i=1}^{n}(A_i \times B_i)$ where $A_1 \times B_1, \ldots, A_n \times B_n$ is a pairwise disjoint sequence of measurable rectangles. Then one must show that

$$\mu(A)\nu(B) = \sum_{i=1}^{n} \mu(A_i)\nu(B_i).$$

We leave this as an exercise.

5.2 Sections

Definition 5.1. Let $E \subseteq X \times Y$ and let $f\colon X \times Y \to \mathbb{R}$.

1. For each $x \in X$, we define the x**-section** of E, denoted E_x, to be the set

$$E_x = \{y \in Y \mid (x, y) \in E\}.$$

   We also define the x-**section** of f, denoted f_x, to be the function $f_x\colon Y \to \mathbb{R}$ given by

$$f_x(y) = f(x, y)$$

   for all $y \in Y$. Observe that if $g\colon X \times Y \to \mathbb{R}$ is another function, then we have $f_x + g_x = (f + g)_x$. Indeed, for any $y \in Y$, we have

$$\begin{aligned}(f_x + g_x)(y) &= f_x(y) + g_x(y) \\ &= f(x, y) + g(x, y) \\ &= (f + g)(x, y) \\ &= (f + g)_x(y).\end{aligned}$$

   Similarly, if $a \in \mathbb{R}$, then we have $(af)_x = af_x$. Indeed, for any $y \in Y$, we have

$$\begin{aligned}(af)_x(y) &= (af)(x, y) \\ &= af(x, y) \\ &= af_x(y).\end{aligned}$$

   Thus we can view $(-)_x$ as an $\mathbb{R}$-linear map from the set of functions from $X \times Y \to \mathbb{R}$ to itself. Another important property that we observe which is easy to prove is that $(1_E)_x = 1_{E_x}$.

2. For each $y \in Y$, we define the y**-section** of E, denoted E^y, to be the set

$$E^y = \{x \in X \mid (x, y) \in E\}.$$

   We also define the y-**section** of f, denoted f^y, to be the function $f^y\colon X \to \mathbb{R}$ given by

$$f^y(x) = f(x, y)$$

   for all $x \in X$. Similar to what mentioned above, we can view $(-)^y$ an $\mathbb{R}$-linear map from the set of functions from $X \times Y \to \mathbb{R}$ to itself, and we also have $(1_E)^y = E^y$.

Proposition 5.1.

Proposition 5.2. *The following statements hold:*

1. *For any $E \in \mathcal{M} \otimes \mathcal{N}$, we have $E_x \in \mathcal{N}$ for all $x \in X$. Similarly, for any $E \in \mathcal{M} \otimes \mathcal{N}$, we have $E^y \in \mathcal{M}$ for all $y \in Y$.*

2. *For any $E \in \mathcal{M} \otimes \mathcal{N}$, the function $\nu(E_{(-)})\colon X \to \mathbb{R}$ defined by*

$$\nu(E_{(-)})(x) = \nu(E_x)$$

for all $x \in X$ is $\mathcal{M}$-measurable. Similarly, for any $E \in \mathcal{M} \otimes \mathcal{N}$, the function $\mu(E^{-})\colon Y \to \mathbb{R}$ defined by

$$\mu(E^{(-)})(y) = \mu(E^y)$$

for all $y \in Y$ is $\mathcal{N}$-measurable.

3. *For any $E \in \mathcal{M} \otimes \mathcal{N}$, we have*

$$\int_X \nu(E_{(-)})\mathrm{d}\mu = (\mu \otimes \nu)(E).$$

Similarly, for any $E \in \mathcal{M} \otimes \mathcal{N}$, we have

$$(\mu \otimes \nu)(E) = \int_Y \mu(E^{(-)})\mathrm{d}\nu.$$

Proof. 1. It suffices to show the first part of this statement since the proof of the second part is completely symmetrical to the proof the first part. Let $x \in X$ and define

$$\mathcal{C} = \{E \in \mathcal{M} \otimes \mathcal{N} \mid E_x \in \mathcal{N}\}.$$

We will show that $\mathcal{C}$ is a σ-algebra which contains $\mathcal{R}$, and this will force $\mathcal{C} = \mathcal{M} \otimes \mathcal{N}$ (which implies $E_x \in \mathcal{N}$ for all $E \in \mathcal{M} \otimes \mathcal{N}$ and for all $x \in X$). First, let us show that $\mathcal{C}$ contains R. Let $A \times B \in \mathcal{R}$ (so $A \in \mathcal{M}$ and $B \in \mathcal{N}$). Then

$$(A \times B)_x = \begin{cases} B & \text{if } x \in A \\ \varnothing & \text{if } x \in A^c \end{cases}$$

In either case, we see that $(A \times B)_x \in \mathcal{N}$, and thus $A \times B \in \mathcal{C}$. Thus $\mathcal{C}$ contains $\mathcal{R}$.

Now we will show that $\mathcal{C}$ is a σ-algebra. It is nonempty since $\mathcal{R} \subseteq \mathcal{C}$. Let us show that $\mathcal{C}$ is closed under countable unions. Let (E_n) be a countable sequence of members of $\mathcal{C}$. Observe that

$$\begin{aligned} y \in \left(\bigcup_{n=1}^{\infty} E_n\right)_x &\iff (x,y) \in \bigcup_{n=1}^{\infty} E_n \\ &\iff (x,y) \in E_n \text{ for some } n \\ &\iff y \in (E_n)_x \text{ for some } n \\ &\iff y \in \bigcup_{n=1}^{\infty} (E_n)_x. \end{aligned}$$

Therefore

$$\left(\bigcup_{n=1}^{\infty} E_n\right)_x = \bigcup_{n=1}^{\infty} (E_n)_x \in \mathcal{N}.$$

This shows that $\mathcal{C}$ is closed under countable unions. Now let us show that $\mathcal{C}$ is closed under complements. Let $E \in \mathcal{C}$. Then observe that

$$\begin{aligned} y \in (E^c)_x &\iff (x,y) \in E^c \\ &\iff (x,y) \notin E \\ &\iff y \notin E_x \\ &\iff y \in (E_x)^c. \end{aligned}$$

Therefore

$$(E^c)_x = (E_x)^c \in \mathcal{N}.$$

This shows that $\mathcal{C}$ is closed under complements. Thus we have shown $\mathcal{C}$ is a σ-algebra which contains $\mathcal{R}$.

2. It suffices to show the first part of this statement since the proof of the second part is completely symmetrical to the proof the first part. We proceed as in the proof of part 1. Let

$$\mathcal{C} = \{E \in \mathcal{M} \otimes \mathcal{N} \mid \nu(E_{(-)}) \text{ is } \mathcal{M}\text{-measurable}\}.$$

We will show that $\mathcal{C}$ is a σ-algebra which contains $\mathcal{R}$ (just like in the proof of part 1). First we show it contains $\mathcal{R}$. Let $A \times B \in \mathcal{R}$ and let $c \in \mathbb{R}$. First note that for any $x \in X$, we have

$$\nu((A \times B)_x) = \begin{cases} \nu(B) & \text{if } x \in A \\ 0 & \text{if } x \in A^c \end{cases}$$

In particular, we see that $\nu(A \times B)_{(-)}$ is a simple function, namely the simple function

$$\nu(A \times B)_{(-)} = \nu(B)1_A.$$

It follows easily from this that $\nu(A \times B)_{(-)}$ is measurable. Thus, since $A \times B \in \mathcal{R}$ is arbitrary, we see that $\mathcal{C}$ contains $\mathcal{R}$.

Now we will show that $\mathcal{C}$ is a σ-algebra. Let us first show that it is closed under complements. Let $E \in \mathcal{C}$. Then observe that for all $x \in X$, we have

$$\begin{aligned}\nu((E^c)_x) &= \nu((E_x)^c) \\ &= \nu(Y) - \nu(E_x),\end{aligned}$$

where we used the fact that ν if finite. Thus, $\nu((E^c)_{(-)}) = \nu(Y) - \nu(E_{(-)})$, which implies $\nu((E^c)_{(-)})$ is a $\mathcal{M}$-measurable function. Thus $E^c \in \mathcal{C}$ and hence $\mathcal{C}$ is closed under complements. Now let us show that $\mathcal{C}$ is closed under ascending unions. Let (E_n) be an ascending sequence in $\mathcal{C}$. Then observe that for all $x \in X$, we have

$$\begin{aligned}\nu\left(\left(\bigcup_{n=1}^{\infty} E_n\right)_x\right) &= \nu\left(\bigcup_{n=1}^{\infty} (E_n)_x\right) \\ &= \lim_{n\to\infty} \nu((E_n)_x)\end{aligned}$$

In particular, the sequence of $\mathcal{M}$-measurable functions $(\nu(E_n)_{(-)})$ converges pointwise to the function $\nu\left((\bigcup_{n=1}^{\infty} E_n)_{(-)}\right)$. Thus $\nu\left((\bigcup_{n=1}^{\infty} E_n)_{(-)}\right)$ is a $\mathcal{M}$-measurable function, which implies $\bigcup_{n=1}^{\infty} E_n \in \mathcal{C}$. Thus we have shown $\mathcal{C}$ is a σ-algebra which contains $\mathcal{R}$.

3. It suffices to show the first part of this statement since the proof of the second part is completely symmetrical to the proof the first part. We proceed as in the proof of parts 1 and 2. Let

$$\mathcal{C} = \{E \in \mathcal{M} \otimes \mathcal{N} \mid \int_X \nu(E_{(-)}) = (\mu \otimes \nu)(E)\}.$$

We will show that $\mathcal{C}$ is a σ-algebra which contains $\mathcal{R}$ (just like in the proofs of part 1 and 2). First we show it contains $\mathcal{R}$. Let $A \times B \in \mathcal{R}$. Then we have

$$\begin{aligned}\int_X \nu((A \times B)_{(-)})\mathrm{d}\mu &= \int_X \nu(B)1_A \mathrm{d}\mu \\ &= \mu(A)\nu(B) \\ &= (\mu \otimes \nu)(A \times B).\end{aligned}$$

Thus, since $A \times B \in \mathcal{R}$ is arbitrary, we see that $\mathcal{C}$ contains $\mathcal{R}$.

Now we will show that $\mathcal{C}$ is a σ-algebra. Let us first show that it is closed under complements. Let $E \in \mathcal{C}$. Then we have

$$\begin{aligned}\int_X \nu((E^c)_{(-)})\mathrm{d}\mu &= \int_X (\nu(Y) - \nu(E_{(-)}))\mathrm{d}\mu \\ &= \int_X \nu(Y)\mathrm{d}\mu - \int_X \nu(E_{(-)})\mathrm{d}\mu \\ &= \mu(X)\nu(Y) - (\mu \otimes \nu)(E) \\ &= (\mu \otimes \nu)(X \times Y) - (\mu \otimes \nu)(E) \\ &= (\mu \otimes \nu)(E^c).\end{aligned}$$

Thus $E^c \in \mathcal{C}$, and hence $\mathcal{C}$ is closed under complements. Here, we need to remind the the reader that the manipulations we did above crucially depend on the fact that both $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are finite measure spaces. For instance, to get from the second line to the first line, we needed to use the fact that $\nu(E_-)$ is integrable (and not just $\mathcal{M}$-measurable). The fact $\nu(E_-)$ is integrable follows from the fact that both $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are finite measure spaces.

Now let us show that $\mathcal{C}$ is closed under ascending unions. Let (E_n) be an ascending sequence in $\mathcal{C}$. Then we have

$$\begin{aligned}\int_X \nu\left(\left(\bigcup_{n=1}^{\infty} E_n\right)_{(-)}\right) \mathrm{d}\mu &= \int_X \lim_{n\to\infty} \nu((E_n)_{(-)}) \mathrm{d}\mu \\ &= \lim_{n\to\infty} \int_X \nu((E_n)_{(-)}) \mathrm{d}\mu \\ &= \lim_{n\to\infty} (\mu \otimes \nu)(E_n) \\ &= (\mu \otimes \nu)\left(\bigcup_{n=1}^{\infty} E_n\right),\end{aligned}$$

where we used MCT to get from the first line to the second line (since $(\nu((E_n)_{(-)}))$ is an increasing sequence of nonnegative measurable functions which converges pointwise to $\nu\left((\bigcup_{n=1}^{\infty} E_n)_{(-)}\right)$). Thus we have shown $\mathcal{C}$ is a σ-algebra which contains $\mathcal{R}$. $\square$

5.3 Tonelli's Theorem and Fubini's Theorem

Theorem 5.1. *The following statements hold.*

1. *(Tonelli) Let $f\colon X \times Y \to [0, \infty]$ be a nonnegative $\mathcal{M} \otimes \mathcal{N}$-measurable function. Define the function $\int_Y f_{(-)} \mathrm{d}\nu \colon X \to [0, \infty]$ by*

$$\left(\int_Y f_{(-)} \mathrm{d}\nu\right)(x) = \int_Y f_x \mathrm{d}\nu$$

for all $x \in X$. Then $\int_Y f_{(-)} \mathrm{d}\nu$ is $\mathcal{M}$-measurable. Furthermore,

$$\int_X \left(\int_Y f_{(-)} \mathrm{d}\nu\right) \mathrm{d}\mu = \int_{X\times Y} f \mathrm{d}(\mu \otimes \nu)$$

Similarly, define the function $\int_X f^{(-)} \mathrm{d}\mu \colon Y \to [0, \infty]$ by

$$\left(\int_X f^{(-)} \mathrm{d}\mu\right)(y) = \int_X f^y \mathrm{d}\mu$$

for all $y \in Y$. Then $\int_X f^{(-)} \mathrm{d}\mu$ is $\mathcal{N}$-measurable. Furthermore,

$$\int_{X\times Y} f \mathrm{d}(\mu \otimes \nu) = \int_Y \left(\int_X f^{(-)} \mathrm{d}\mu\right) \mathrm{d}\nu.$$

2. *(Fubini) Let $f\colon X \times Y \to \mathbb{R}$ be a $\mathcal{M} \otimes \mathcal{N}$-integrable function. Then the function $f_{(-)}$ is $\mathcal{N}$-integrable. Furthemore,*

$$\int_X \left(\int_Y f_{(-)} \mathrm{d}\nu\right) \mathrm{d}\mu = \int_{X\times Y} f \mathrm{d}(\mu \otimes \nu).$$

Similarly, for almost every $y \in Y$, the function $f^{(-)}$ is $\mathcal{M}$-integrable. Furthermore,

$$\int_{X\times Y} f \mathrm{d}(\mu \otimes \nu) = \int_Y \left(\int_X f^{(-)} \mathrm{d}\mu\right) \mathrm{d}\nu$$

Proof. 1. (Tonelli) It suffices to show the first part of this statement since the proof of the second part is completely symmetrical to the proof the first part. First we prove this for nonnegative simple functions. Let $\varphi\colon X \times Y \to$

$[0,\infty]$ be a nonnegative simple function, and suppose its canonical form is given by $\varphi = \sum_{i=1}^n c_i 1_{E_i}$. Then note that

$$\begin{aligned}
\int_Y \varphi_{(-)} \mathrm{d}\nu &= \int_Y \left(\sum_{i=1}^n c_i 1_{E_i}\right)_{(-)} \mathrm{d}\nu \\
&= \int_Y \sum_{i=1}^n c_i (1_{E_i})_{(-)} \mathrm{d}\nu \\
&= \int_Y \sum_{i=1}^n c_i 1_{(E_i)_{(-)}} \mathrm{d}\nu \\
&= \sum_{i=1}^n c_i \left(\int_Y 1_{(E_i)_{(-)}} \mathrm{d}\nu\right) \\
&= \sum_{i=1}^n c_i \nu((E_i)_{(-)}).
\end{aligned}$$

Thus $\int_Y \varphi_{(-)} \mathrm{d}\nu$ is a sum of $\mathcal{M}$-measurable functions, and is hence $\mathcal{M}$-measurable. Furthermore, we have

$$\begin{aligned}
\int_X \left(\int_Y \varphi_{(-)} \mathrm{d}\nu\right) \mathrm{d}\mu &= \int_X \sum_{i=1}^n c_i \nu((E_i)_{(-)}) \mathrm{d}\mu \\
&= \sum_{i=1}^n c_i \left(\int_X \nu((E_i)_{(-)}) \mathrm{d}\mu\right) \\
&= \sum_{i=1}^n c_i (\mu \otimes \nu)(E_i) \\
&= \int_{X \times Y} \varphi \mathrm{d}(\mu \otimes \nu).
\end{aligned}$$

Now we prove it for more generally, let $f \colon X \times Y \to [0,\infty]$ be a nonnegative $\mathcal{M} \otimes \mathcal{N}$-measurable function. Choose an increasing sequence $(\varphi_n \colon X \times Y \to [0,\infty])$ of nonngetative simple functions such that (φ_n) converges to f pointwise. Then note that for each $x \in X$, we have

$$\begin{aligned}
\int_Y f_x \mathrm{d}\nu &= \int_Y \lim_{n\to\infty} ((\varphi_n)_x) \mathrm{d}\nu \\
&= \lim_{n\to\infty} \int_Y ((\varphi_n)_x) \mathrm{d}\nu
\end{aligned}$$

where we used MCT to get from the first line to the second line. Since $x \in X$ was arbitary, it follows that the sequence $(\int_Y \varphi_{(-)} \mathrm{d}\nu)$ of nonnegative $\mathcal{M}$-measurable converges pointwise to the function $\int_Y f_{(-)} \mathrm{d}\nu$. Hence $\int_Y f_{(-)} \mathrm{d}\nu$ is a nonnegative $\mathcal{M}$-measurable function. Furthermore, we have

$$\begin{aligned}
\int_X \left(\int_Y f_{(-)} \mathrm{d}\nu\right) \mathrm{d}\mu &= \int_X \left(\lim_{n\to\infty} \int_Y (\varphi_n)_{(-)} \mathrm{d}\nu\right) \mathrm{d}\mu \\
&= \lim_{n\to\infty} \left(\int_X \left(\int_Y (\varphi_n)_{(-)} \mathrm{d}\nu\right) \mathrm{d}\mu\right) \\
&= \lim_{n\to\infty} \int_{X\times Y} \varphi_n \mathrm{d}(\mu \otimes \nu) \\
&= \int_{X\times Y} \lim_{n\to\infty} \varphi_n \mathrm{d}(\mu \otimes \nu) \\
&= \int_{X\times Y} f \mathrm{d}(\mu \otimes \nu),
\end{aligned}$$

where used MCT twice.

2. (Fubini) It suffices to show the first part of this statement since the proof of the second part is completely symmetrical to the proof the first part. To prove Fubini, we write $f = f^+ - f^-$ and apply Tonelli to both f^+ and

f^-. We have

$$\begin{aligned}\int_{X\times Y} f\mathrm{d}(\mu\otimes\nu) &= \int_{X\times Y} f^+\mathrm{d}(\mu\otimes\nu) - \int_{X\times Y} f^-\mathrm{d}(\mu\otimes\nu). \\ &= \int_X\left(\int_Y (f^+)_{(-)}\mathrm{d}\nu\right)\mathrm{d}\mu - \int_X\left(\int_Y (f^-)_{(-)}\mathrm{d}\nu\right)\mathrm{d}\mu \\ &= \int_X\left(\int_Y (f^+)_{(-)}\mathrm{d}\nu - \int_Y (f^-)_{(-)}\mathrm{d}\nu\right)\mathrm{d}\mu \\ &= \int_X\left(\int_Y f_{(-)}\mathrm{d}\nu\right)\mathrm{d}\mu.\end{aligned}$$

We need to justify some of the steps that we used above. The first line is simply the definition of $\int_{X\times Y} f\mathrm{d}(\mu\otimes\nu)$. We obtained the second line from the first line from Tonelli. To get the third line from the second line, we needed to use the fact that $\int_Y f_{(-)}\mathrm{d}\nu$ is $\mathcal{M}$-integrable. To see why $\int_Y f_{(-)}\mathrm{d}\nu$ is $\mathcal{M}$-integrable, it suffices to show that $(\int_Y f_{(-)}\mathrm{d}\nu)^+$ (the positive part of $\int_Y f_{(-)}\mathrm{d}\nu$) and $(\int_Y f_{(-)}\mathrm{d}\nu)^-$ (the negative part of $\int_Y f_{(-)}\mathrm{d}\nu$) are $\mathcal{M}$-integrable. To see why $(\int_Y f_{(-)}\mathrm{d}\nu)^+$ is $\mathcal{M}$-integrable, observed that

$$\begin{aligned}\int_X\left(\int_Y f_{(-)}\mathrm{d}\nu\right)^+\mathrm{d}\mu &\leq \int_X\left(\int_Y (f_{(-)})^+\mathrm{d}\nu\right)\mathrm{d}\mu \\ &= \int_X\left(\int_Y (f^+)_{(-)}\mathrm{d}\nu\right)\mathrm{d}\mu \\ &= \int_{X\times Y} f^+\mathrm{d}(\mu\otimes\nu) \\ &< \infty.\end{aligned}$$

A similar proof shows that $(\int_Y f_{(-)}\mathrm{d}\nu)^-$ is $\mathcal{M}$-integrable. Thus $\int_Y f_{(-)}\mathrm{d}\nu$ is $\mathcal{M}$-integrable. The last step that we need to justify, is how we obtained fourth line from the third line. To see why this holds, observe that for each $x\in X$, we have

$$\begin{aligned}\int_Y (f^+)_x\mathrm{d}\nu - \int_Y (f^-)_x\mathrm{d}\nu &= \int_Y (f_x)^+\mathrm{d}\nu - \int_Y (f_x)^-\mathrm{d}\nu \\ &= \int_Y f_x\mathrm{d}\nu,\end{aligned}$$

where we applied the definition of $\int_Y f_x\mathrm{d}\nu$ to obtain the second from the first line. Note that this makes sense because f_x is ν-integrable almost everywhere. Indeed, since

$$\int_X\left(\int_Y (f_{(-)})^+\mathrm{d}\nu\right)\mathrm{d}\mu < \infty,$$

it follows that $\int_Y (f_x)^+\mathrm{d}\nu < \infty$ for almost all $x\in X$. Similarly, $\int_Y (f_x)^-\mathrm{d}\nu < \infty$ for almost all $y\in Y$. Thus $\int_Y f_{(-)}\mathrm{d}\nu$ is $\mathcal{N}$-integrable almost everywhere. □

Example 5.1. Assume you need to compute

$$\int_X\left(\int_Y f(x,y)\mathrm{d}\nu(y)\right)\mathrm{d}\mu(x).$$

Assume you know how to compute

$$\int_Y\left(\int_X f(x,y)\mathrm{d}\mu(x)\right)\mathrm{d}\nu(y).$$

Assume f is *not* nonnegative. To apply Fubini, we need $f\in L^1(X\times Y)$. Since $|f|$ is a nonnegative measurable function, we can apply Tonelli to $|f|$ to get

$$\int_{X\times Y}|f|\mathrm{d}(\mu\otimes\nu) = \int_Y\left(\int_X |f(x,y)|\mathrm{d}\mu(x)\right)\mathrm{d}\nu(y).$$

If we can compute the last itterated integral and show it is $<\infty$, then $f\in L^1(X\times Y)$. So we can use Fubini.

6 Signed Measures

We'll discuss only *finite* signed measures.

Definition 6.1. Let $(X, \mathcal{M})$ be a measurable space. A function $\nu\colon \mathcal{M} \to \mathbb{R}$ is said to be a **finite signed measure** if $\nu(\emptyset) = 0$ and if ν is countably additive: if (E_n) is sequence of pairwise disjoint sets in $\mathcal{M}$, then

$$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \nu(E_n)$$

where the series converges absolutely.

Example 6.1. If μ_1 and μ_2 are usual (positive) finite measures on $\mathcal{M}$, then $\nu = \mu_1 - \mu_2$ is a signed measure.

Example 6.2. If μ is a measure on $(X, \mathcal{M})$ and $f \in L^1(X, \mathcal{M}, \mu)$, then

$$\nu(E) = \int_X 1_E f \,\mathrm{d}\mu$$

is a signed measure.

Proposition 6.1. *Let ν be a finite signed measure on $(X, \mathcal{M})$ and let (E_n) be a sequence in $\mathcal{M}$, Then*

1. *if (E_n) is an ascending sequence, then $\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \nu(E_n)$.*
2. *if (E_n) is a descending sequence, then $\nu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \nu(E_n)$.*

Proof. Same proof as for usual measures. $\square$

Definition 6.2. Let $E \in \mathcal{M}$. We say

1. E is a ν**-positive set** if $\nu(F) \geq 0$ for all $F \subseteq E$ with $F \in \mathcal{M}$.
2. E is ν**-negative set** if $\nu(F) \leq 0$ for all $F \subseteq E$ with $F \in \mathcal{M}$.
3. E is a ν**-null set** if $\nu(F) = 0$ for all $F \subseteq E$ with $F \in \mathcal{M}$.

Example 6.3. Let $a, b \in X$. We define

$$\delta_a(E) = \begin{cases} 0 & \text{if } a \notin E \\ 1 & \text{if } a \in E \end{cases}$$

We also define $\nu = \delta_a - \delta_b$. Then ν is signed measure. If $a \in E$ and $b \notin E$, then $\nu(E) = 1$. If $a \notin E$ and $b \in E$, then $\nu(E) = -1$. If $a, b \in E$, then $\nu(E) = 0$.

Lemma 6.1. *If $(P_n) \subseteq \mathcal{M}$ is a sequence of ν-positive sets, then $\bigcup_{n=1}^{\infty} P_n$ is also a ν-positive set.*

Proof. By disjointifying if necessary, we may assume that the sequence (P_n) is pairwise disjoint. Let $E \in \mathcal{M}$ such that $E \subseteq \bigcup_{n=1}^{\infty} P_n$. Then

$$\begin{aligned} \nu(E) &= \nu\left(E \cap \left(\bigcup_{n=1}^{\infty} P_n\right)\right) \\ &= \nu\left(\bigcup_{n=1}^{\infty} (E \cap P_n)\right) \\ &= \sum_{n=1}^{\infty} \nu(E \cap P_n) \\ &\geq 0. \end{aligned}$$

$\square$

6.0.1 Hahn Decomposition Theorem

Theorem 6.2. *(Hahn decomposition theorem) If ν is a signed measure on $(X, \mathcal{M})$, then there exists $P \in \mathcal{M}$ ν-positive and $Q \in \mathcal{M}$ ν-negative such that $X = P \cup Q$ and $P \cap Q = \emptyset$. If P', Q' is another such pair, then $P \Delta P'$ and $Q \Delta Q'$ are both ν-null sets.*

Proof. Let $m = \sup\{\nu(E) \mid E \text{ is } \nu\text{-positive}\}$. Then for all $n \in \mathbb{N}$, there exists a ν-positive set P_n such that

$$m \geq \nu(P_n) > m - \frac{1}{n}.$$

Then $\lim_{n\to\infty} \nu(P_n) = m$. Take $P = \bigcup_{n=1}^{\infty} P_n$. We showed last time P is also ν-positive and

$$\begin{aligned} \nu(P) &= \nu\left(\bigcup_{n=1}^{\infty} P_n\right) \\ &= \lim_{n\to\infty} \nu(P_n) \\ &= m. \end{aligned}$$

Set $N = P^c$. We need to show N is ν-negative. First, we make two observations. First, if $E \subseteq N$ is a ν-positive set, then E must be a ν-null set. Indeed, assume $\nu(E) > 0$. Then

$$\begin{aligned} \nu(E \cup P) &= \nu(E) + \nu(P) \\ &= \nu(P) \\ &= m \end{aligned}$$

and $E \cup P$ is ν-positive, but this is a contradiction. This establishes our first claim. The next claim we make is that if $A \subseteq N$ and $\nu(A) > 0$, then there exists $B \subseteq A$ with $\nu(B) > \nu(A)$. Indeed, we just proved that A cannot be ν-positive. Therefore there exists $C \subseteq A$ such that $\nu(C) < 0$. Take $B = A \backslash C$. Then

$$\begin{aligned} \nu(B) &= \nu(A) - \nu(C) \\ &> \nu(A). \end{aligned}$$

This establishes our second claim.

Now we construct two sequence $(n_k) \subseteq \mathbb{N}$ and (A_k) of subsets of N in the following way: Set

$$n_1 = \min\{k \in \mathbb{N} \mid \exists B \subseteq N \text{ with } \nu(B) > 1/k\}.$$

Next we define A_1 to be one such subset B, so $A_1 \subseteq N$ such that $\nu(A_1) > 1/n_1$. Next, we define

$$n_2 = \min\{k \in \mathbb{N} \mid \exists B \subseteq A_1 \text{ with } \nu(B) > \nu(A_1) + 1/k\}.$$

Next we define A_2 to be one such subset B, so $A_2 \subseteq A_1$ such that $\nu(A_2) > \nu(A_1) + 1/n_2$. Continuing in this way, we obtain (n_k) and (A_k). Let $A = \bigcap_{n=1}^{\infty} A_n$. Then

$$\begin{aligned} \nu(A) &= \lim_{k\to\infty} \nu(A_k) \\ &\geq \sum_{k=1}^{\infty} \frac{1}{n_k}. \end{aligned}$$

Since $\nu(A)$ is finite, we see that this series must be convergent. Thus $1/n^k \to 0$ as $k \to \infty$. Then there exists $B \subseteq A$ with the property that $\nu(B) > \nu(A)$. Let $N \in \mathbb{N}$ such that $1/N < \nu(B) - \nu(A)$. There exists j such that

$$\frac{1}{n_j} < \frac{1}{N} < \nu(B) - \nu(A).$$

That is, $\nu(B) > \nu(A) + 1/n_j$. Also

$$\begin{aligned} B &\subseteq A \\ &= \bigcap_{k=1}^{\infty} A_k \\ &\subseteq A_j. \end{aligned}$$

So our assumption that N is not negative is wrong. Thus N is positive. The uniqueness part is simpler and left as an exercise. $\square$

6.0.2 Mutually Singular Signed Measures

Definition 6.3. The signed measures ν_1, ν_2 on $(X, \mathcal{M})$ are said to be **mutually singular**, denoted $\nu_1 \perp \nu_2$, if there exists $E, F \in \mathcal{M}$ such that $E \cap F = \emptyset$, $E \cup F = X$, E is a ν_1-null set and F is a ν_2-null set.

6.0.3 Jordan Decomposition Theorem

Theorem 6.3. *(Jordan decomposotion theorem) If ν is a signed measure on $(X, \mathcal{M})$, then there exists a unique pair of measures ν^+ and ν^- such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.*

Proof. Let $X = P \cup N$ be a Hahn decomposition of ν. Define

$$\nu^+(E) = \nu(E \cap P) \quad \text{and} \quad \nu^-(E) = -\nu(E \cap N)$$

for all $E \in \mathcal{M}$. It is straightforward to check that ν^+ and ν^- are both measures on $(X, \mathcal{M})$. Also,

$$\begin{aligned} \nu(E) &= \nu(E \cap P) + \nu(E \cap N) \\ &= \nu^+(E) - \nu^-(E). \end{aligned}$$

Thus $\nu = \nu^+ - \nu^-$. Also P is a ν^--null set. Indeed, if $F \subseteq P$, then

$$\begin{aligned} \nu^-(F) &= -\nu(F \cap N) \\ &= -\nu(\emptyset) \\ &= 0. \end{aligned}$$

Similarly, N is a ν^+-null set. It only remains to show uniqueness of such a pair. So suppose $\nu = \mu^+ - \mu^-$ for another such pair (μ^+, μ^-) with $\mu^+ \perp \mu^-$. Since $\mu^+ \perp \mu^-$, there exists $E, F \in X$ such that $X = E \cup F$, $E \cap F = \emptyset$, E is a μ^--null set, and F is a μ^+-null set. Then $X = E \cup F$ is another Hahn decomposition (since E is a μ-positive set and F is a ν-negative set). Therefore $P \Delta E$ and $Q \Delta F$ are both ν-null sets. Therefore if $A \in \mathcal{M}$, we have

$$\begin{aligned} \mu^+(A) &= \mu^+(A \cap E) + \mu^+(A \cap F) \\ &= \nu(A \cap E) + 0 \\ &= \nu(A \cap P) \\ &= \nu^+(A), \end{aligned}$$

so $\mu^+ = \nu^+$. Similarly, $\mu^- = \nu^-$. □

6.1 Banach Space of Signed Measures

Let $(X, \mathcal{M})$ be a measurable space. We denote by $M(X, \mathcal{M})$ to be the set of all finite signed measures on $\mathcal{M}$. First let us give $M(X, \mathcal{M})$ the structure of an $\mathbb{R}$-vector space as follows: Let $a \in \mathbb{R}$ and let $\mu, \nu \in M(X, \mathcal{M})$. We define addition $\mu + \nu \colon \mathcal{M} \to \mathbb{R}$ and scalar multiplication $a\mu \colon \mathcal{M} \to \mathbb{R}$ by

$$\begin{aligned} (\mu + \nu)(E) &= \mu(E) + \nu(E) \\ (a\mu)(E) &= a\mu(E) \end{aligned}$$

for all $E \in \mathcal{M}$. It is straightforward to check that addition and scalar multiplication defined in this way gives $M(X, \mathcal{M})$ the structure of an $\mathbb{R}$-vector space.

Next let us give $M(X, \mathcal{M})$ the structure of a normed linear space as follows: Let $\mu \in M(X, \mathcal{M})$ and let $\mu = \mu^+ - \mu^-$ be the Jordan decomposition of μ. We define the norm $\|\cdot\| \colon M(X, \mathcal{M}) \to \mathbb{R}_{\geq 0}$ by

$$\|\mu\| = \mu^+(X) + \mu^-(X).$$

Again it is staightforward to check that this gives $M(X, \mathcal{M})$ the structure of a normed linear space.

6.2 Absolute Continuity For Signed Measures

Definition 6.4. Let ν be a signed measure on $(X, \mathcal{M})$ and let μ be a usual measure on $(X, \mathcal{M})$. We say ν is **absolutely continuous** with respect to μ if for all $E \in \mathcal{M}$, we have $\mu(E) = 0$ implies $\nu(E) = 0$. We denote this by $\nu \ll \mu$.

It's esay to prove that

$$\begin{aligned} \nu \ll \mu &\iff \nu^+ \ll \mu \text{ and } \nu^- \ll \mu \\ &\iff |\nu| \ll \mu. \end{aligned}$$

Proposition 6.2. *$\nu \ll \mu$ if and only if for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $E \in \mathcal{M}$ with $\mu(E) < \delta$, we have $\nu(E) < \delta$.*

Proof. ($\Longleftarrow$) is easy. ($\Longrightarrow$) assume $\nu \ll \mu$ but doesn't hold. Then there exists $\varepsilon > 0$ fo rall $n \in \mathbb{N}$ ther exists $E_n \in \mathcal{M}$ with $\mu(E_n) < 1/2^n$ buth $|\nu|(E_n) \geq \varepsilon$. On the other hand, since $|\nu|(E_n) \geq \varepsilon$, we have $|\nu|(F) \geq 0$ which contradicts

$$|\nu| \ll \mu \iff \nu \ll \mu.$$

□

6.2.1 Radon-Nikodym

Theorem 6.4. *(Radon-Nikodym) $\nu \ll \mu$ implies there exists $h \in L^1(\mu)$ such that*

$$\nu(E) = \int_X 1_E h \mathrm{d}\mu.$$

Proof. Assume first that ν is the usual measure. Consider the Hilbert space $L^2(\mu + \nu)$. The map $\ell \colon L^2(\mu + \nu) \to \mathbb{R}$ defined by

$$\ell(f) = \int_X f \mathrm{d}\mu$$

for all $f \in L^2(\mu + \nu)$ is a linear functional and

$$\begin{aligned} |\ell(f)| &= \left| \int_X f \mathrm{d}\mu \right| \\ &\leq \int_X |f| \mathrm{d}\mu \\ &\leq \sqrt{\int_X |f|^2 \mathrm{d}(\mu + \nu)} \sqrt{\int_X 1^2 \mathrm{d}(\mu + \nu)} \\ &= \|f\|_2 (\mu(X) + \nu(X)). \end{aligned}$$

So by the Riesz representation theorem, there exists $g \in L^2(\mu + \nu)$ such that

$$\ell(f) = \langle f, g \rangle = \int_X fg \mathrm{d}(\mu + \nu).$$

So

$$\int_X f \mathrm{d}\mu = \int_X fg \mathrm{d}(\mu + \nu)$$

which is equivalent to

$$\int_X f(1 - g) \mathrm{d}\mu = \int_X fg \mathrm{d}\nu.$$

We claim that $0 < g \leq 1$ for μ almost everywhere. Define $F \in \mathcal{M}$ by

$$F = \{x \in X \mid g(x) \leq 0\}.$$

Take $f = 1_F$. Then

$$\begin{aligned} \int_X 1_F (1 - g) \mathrm{d}\mu &= \int 1_F g \mathrm{d}\nu \\ &\leq 0 \end{aligned}$$

which implies $\mu(0) \leq 0$ which implies $\mu(F) = 0$ which implies $0 < g$ for μ a.e. Similarly, by considering the set

$$G = \{x \in X \mid g(x) > 1\},$$

we can get $\mu(G) = 0$ which implies $g \leq 1$ for μ a.e.

Now we can define $h = (1-g)/g$. For $E \in \mathcal{M}$, pick $f = 1_E \cdot (1/g)$ plug in

$$\int_X 1_E \frac{1}{g}(1-g)\mathrm{d}\mu = \int 1_E \mathrm{d}\nu$$

this implies

$$\int 1_E h \mathrm{d}\mu = \int 1_E \mathrm{d}\nu = \nu(E)$$

Thus

$$\nu(E) = \int_X 1_E h \mathrm{d}\mu.$$

For general signed measures ν, we write $\nu = \nu^+ - \nu^-$ and we use $\nu \ll \mu$ to imply $\nu^+ \ll \mu$ and $\nu^- \ll \mu$ to imlpies

$$\nu^+(E) = \int 1_E h_1 \mathrm{d}\mu \quad \text{and } \nu^-(E) = \int 1_E h_2 \mathrm{d}\mu$$

which implies

$$\nu(E) = \nu^+(E) - \nu^-(E) = \int 1_E (h_1 - h_2)\mathrm{d}\mu.$$

So we are done. □

Part II
Homework

7 Homework 1

Throughout this homework, let X be a set and let $\mathcal{P}(X)$ denote the set of all subsets of X.

7.1 Charactertistic Function Identities

Proposition 7.1. *Let $A, B \in \mathcal{P}(X)$. Then*

1. $1_A = 1_B$ *if and only if* $A = B$*;*
2. $1_{A\cap B} = 1_A 1_B$*;*
3. $1_{A\cup B} = 1_A + 1_B - 1_A 1_B$*;*
4. $1_{A^c} = 1 - 1_A$*;*
5. $1_{A\setminus B} = 1_A - 1_B$ *if and only if* $B \subseteq A$*;*
6. $1_A + 1_B \equiv 1_{A\Delta B} \mod 2$.

Proof.
1. Suppose $1_A = 1_B$ and let $x \in A$. Then

$$\begin{aligned} 1 &= 1_A(x) \\ &= 1_B(x) \end{aligned}$$

implies $x \in B$. Thus $A \subseteq B$. Similarly, if $x \in B$, then

$$\begin{aligned} 1 &= 1_B(x) \\ &= 1_A(x) \end{aligned}$$

implies $x \in A$. Thus $B \subseteq A$.

Conversely, suppose $A = B$ and let $x \in X$. If $x \in A$, then $x \in B$, hence

$$\begin{aligned} 1_A(x) &= 1 \\ &= 1_B(x). \end{aligned}$$

If $x \notin A$, then $x \notin B$, hence

$$\begin{aligned} 1_A(x) &= 0 \\ &= 1_B(x). \end{aligned}$$

Therefore the indicator functions 1_A and 1_B agree on all of X, and hence must be equal to each other.

2. Let $x \in X$. If $x \in A \cap B$, then $x \in A$ and $x \in B$, and thus we have

$$\begin{aligned} 1_{A\cap B}(x) &= 1 \\ &= 1 \cdot 1 \\ &= 1_A(x)1_B(x). \end{aligned}$$

If $x \notin A \cap B$, then either $x \notin A$ or $x \notin B$. Without loss of generality, say $x \notin A$. Then we have

$$\begin{aligned} 1_{A\cap B}(x) &= 0 \\ &= 0 \cdot 1_B(x) \\ &= 1_A(x)1_B(x). \end{aligned}$$

Therefore the functions $1_{A\cap B}$ and $1_A 1_B$ agree on all of X, and hence must be equal to each other.

3. Let $x \in X$. If $x \in A \cup B$, then either $x \in A$ or $x \in B$. Without loss of generality, say $x \in A$. Then we have

$$\begin{aligned} 1_{A\cup B}(x) &= 1 \\ &= 1 + 1_B(x) - 1_B(x) \\ &= 1 + 1_B(x) - 1 \cdot 1_B(x) \\ &= 1_A(x) + 1_B(x) - 1_A(x)1_B(x). \end{aligned}$$

If $x \notin A \cup B$, then $x \notin A$ and $x \notin B$. Therefore we have

$$\begin{aligned} 1_{A\cup B}(x) &= 0 \\ &= 0 + 0 - 0 \cdot 0 \\ &= 1_A(x) + 1_B(x) - 1_A(x)1_B(x). \end{aligned}$$

Thus the functions $1_{A\cup B}$ and $1_A + 1_B - 1_A 1_B$ agree on all of X, and hence must be equal to each other.

4. Let $x \in X$. If $x \in A$, then $x \notin A^c$, hence

$$\begin{aligned} 1_{A^c}(x) &= 0 \\ &= 1 - 1 \\ &= 1 - 1_A(x). \end{aligned}$$

If $x \notin A$, then $x \in A^c$, hence

$$\begin{aligned} 1_{A^c}(x) &= 1 \\ &= 1 - 0 \\ &= 1 - 1_A(x). \end{aligned}$$

Therefore the functions 1_{A^c} and $1 - 1_A$ agree on all of X, and hence must be equal to each other.

5. Suppose $B \subseteq A$ and let $x \in X$. If $x \in A \backslash B$, then $x \in A$ and $x \notin B$, hence

$$\begin{aligned} 1_{A \backslash B}(x) &= 1 \\ &= 1 - 0 \\ &= 1_A(x) - 1_B(x). \end{aligned}$$

If $x \in B$, then $x \in A$ but $x \notin A \backslash B$, hence

$$\begin{aligned} 1_{A \backslash B}(x) &= 0 \\ &= 1 - 1 \\ &= 1_A(x) - 1_B(x). \end{aligned}$$

If $x \notin A$, then $x \notin B$ and $x \notin A \backslash B$, hence

$$\begin{aligned} 1_{A \backslash B}(x) &= 0 \\ &= 0 - 0 \\ &= 1_A(x) - 1_B(x). \end{aligned}$$

Therefore the functions $1_{A \backslash B}$ and $1_A - 1_B$ agree on all of X, and hence must be equal to each other.
For converse direction, we prove the contrapositive statement. Suppose $B \not\subseteq A$. Choose $b \in B$ such that $b \notin A$. Then

$$\begin{aligned} 1_{A \backslash B}(b) &= 0 \\ &\neq -1 \\ &= 0 - 1 \\ &= 1_A(b) - 1_B(b). \end{aligned}$$

Therefore $1_{A \backslash B} \neq 1_A - 1_B$.

5. Let $x \in X$. If $x \in A$, then $x \notin A^c$, hence

$$\begin{aligned} 1_{A^c}(x) &= 0 \\ &= 1 - 1 \\ &= 1 - 1_A(x). \end{aligned}$$

If $x \notin A$, then $x \in A^c$, hence

$$\begin{aligned} 1_{A^c}(x) &= 1 \\ &= 1 - 0 \\ &= 1 - 1_A(x). \end{aligned}$$

Therefore the functions 1_{A^c} and $1 - 1_A$ agree on all of X, and hence must be equal to each other.

6. We have

$$\begin{aligned} 1_{A\Delta B} &= 1_{(A \backslash B) \cup (B \backslash A)} \\ &= 1_{A \backslash B} + 1_{B \backslash A} - 1_{A \backslash B} 1_{B \backslash A} \\ &= 1_{A \backslash A \cap B} + 1_{B \backslash A \cap B} - 1_{(A \backslash B) \cap (B \backslash A)} \\ &= 1_A - 1_{A \cap B} + 1_B - 1_{A \cap B} - 1_{\emptyset} \\ &= 1_A + 1_B - 2 \cdot 1_{A \cap B} \\ &\equiv 1_A + 1_B \bmod 2. \end{aligned}$$

□

7.2 Cauchy Sequence in $(C[a,b], \|\cdot\|_1)$ Converging Pointwise and in L^1 to Indicator Function

Proposition 7.2. *Let I be a subinterval of $[a,b]$. Then there exists a Cauchy sequence (f_n) in $(C[a,b], \|\cdot\|_1)$ such that (f_n) converges pointwise to 1_I on $[a,b]$ and moreover*

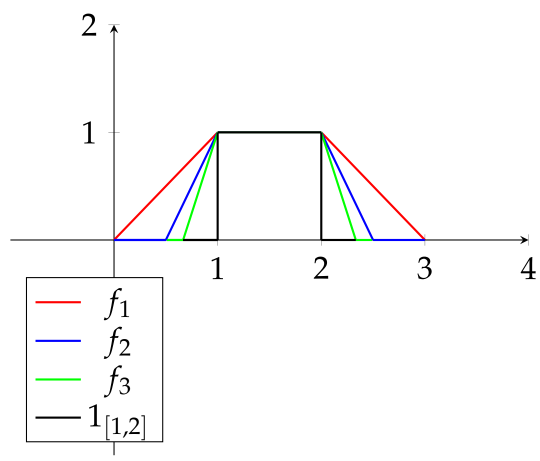
$$\lim_{n\to\infty} \|f_n\|_1 = \text{length}(I).$$

Proof. If $I = \emptyset$, then we take $f_n = 0$ for all $n \in \mathbb{N}$. Thus assume I is a nonempty subinterval of $[a,b]$. We consider two cases; namely $I = (c,d)$ and $I = [c,d]$. The other cases ($I = (c,d]$ and $I = [c,d)$) will easily be seen to be a mixture of these two cases.

Case 1: Suppose $I = [c,d]$. For each $n \in \mathbb{N}$, define

$$f_n(x) = \begin{cases} 0 & \text{if } a \le x < c - \left(\frac{c-a}{n}\right) \\ \frac{n}{c-a}(x-c)+1 & \text{if } c - \left(\frac{c-a}{n}\right) \le x < c \\ 1 & \text{if } c \le x \le d \\ \frac{n}{d-b}(x-d)+1 & \text{if } d < x \le d + \left(\frac{b-d}{n}\right) \\ 0 & \text{if } d + \left(\frac{b-d}{n}\right) < x \le b \end{cases}$$

The image below gives the graphs for f_1, f_2, and f_3 in the case where $[a,b] = [0,3]$ and $[c,d] = [1,2]$.



For each $n \in \mathbb{N}$, the function f_n is continuous since each of its segments is continuous and are equal on their boundaries.

Let us check that (f_n) converges pointwise to 1_I: If $x \in [a,c)$, then we choose $N \in \mathbb{N}$ such that

$$x \le c - \left(\frac{c-a}{N}\right).$$

Then $f_n(x) = 0$ for all $n \ge N$. Thus

$$\lim_{n\to\infty} f_n(x) = 0 = 1_I(x).$$

Similarly, if $x \in (d,b]$, then we choose $N \in \mathbb{N}$ such that

$$x \ge d + \left(\frac{b-d}{N}\right).$$

Then $f_n(x) = 0$ for all $n \ge N$. Thus

$$\lim_{n\to\infty} f_n(x) = 0 = 1_I(x).$$

Finally, if $x \in [c,d]$, then $f_n(x) = 1$ for all $n \in \mathbb{N}$ by definition and thus

$$\lim_{n\to\infty} f_n(x) = 0 = 1_I(x).$$

Let us check that (f_n) is Cauchy in $(C[a,b], \|\cdot\|_1)$. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that

$$\frac{c-a+b-d}{n} < \varepsilon$$

for all $n \geq N$. Then $n \geq m \geq N$ implies

$$\begin{aligned}
\|f_n - f_m\|_1 &= \int_a^b |f_n(x) - f_m(x)| \mathrm{d}x \\
&= \int_a^b (f_n(x) - f_m(x)) \mathrm{d}x \\
&= \int_{c - \left(\frac{c-a}{m}\right)}^{c} (f_n(x) - f_m(x)) \mathrm{d}x + \int_d^{d + \left(\frac{b-d}{m}\right)} (f_n(x) - f_m(x)) \mathrm{d}x \\
&\leq \int_{c - \left(\frac{c-a}{m}\right)}^{c} \mathrm{d}x + \int_d^{d + \left(\frac{b-d}{m}\right)} \mathrm{d}x \\
&= \frac{c-a}{m} + \frac{b-d}{m} \\
&= \frac{c-a+b-d}{m} \\
&< \varepsilon.
\end{aligned}$$

Thus the sequence (f_n) is Cauchy in $(C[a,b], \|\cdot\|_1)$.

Finally, we check that $\|f_n\|_1 \to \text{length}(I)$ as $n \to \infty$. We have

$$\begin{aligned}
d - c &\leq \|f_n\|_1 \\
&= \int_a^b |f_n(x)| \mathrm{d}x \\
&= \int_a^b f_n(x) \mathrm{d}x \\
&= \int_{c - \left(\frac{c-a}{n}\right)}^{c} f_n(x) \mathrm{d}x + \int_c^d \mathrm{d}x + \int_d^{d + \left(\frac{b-d}{n}\right)} f_n(x) \mathrm{d}x \\
&\leq \int_{c - \left(\frac{c-a}{n}\right)}^{c} \mathrm{d}x + \int_c^d \mathrm{d}x + \int_d^{d + \left(\frac{b-d}{n}\right)} \mathrm{d}x \\
&= \frac{c-a}{n} + d - c + \frac{b-d}{n} \\
&\to d - c.
\end{aligned}$$

Thus for each $n \in \mathbb{N}$, we have

$$d - c \leq \|f_n\|_1 \leq d - c + \frac{c-a+b-d}{n}. \tag{24}$$

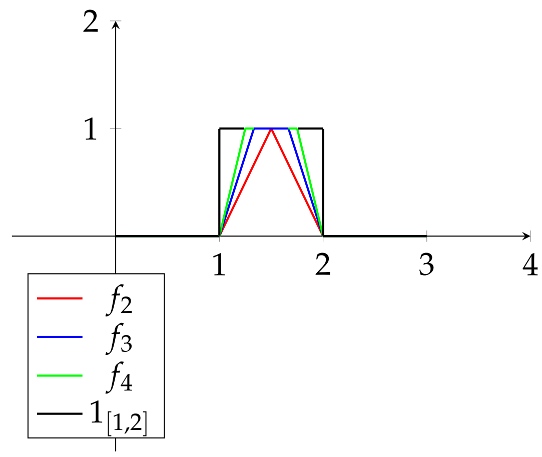
By taking $n \to \infty$ in (24), we see that

$$\begin{aligned}
\lim_{n \to \infty} \|f_n\|_1 &= d - c \\
&= \text{length}(I).
\end{aligned}$$

Case 2: Suppose $I = (c,d)$. For each $n \geq 2$, define

$$f_n(x) = \begin{cases} 0 & \text{if } a \leq x \leq c \\ \frac{n}{d-c}(x-c) & \text{if } c < x \leq c + \left(\frac{d-c}{n}\right) \\ 1 & \text{if } c + \left(\frac{d-c}{n}\right) \leq x \leq d - \left(\frac{d-c}{n}\right) \\ \frac{n}{c-d}(x-d) & \text{if } d - \left(\frac{d-c}{n}\right) \leq x \leq d \\ 0 & \text{if } d \leq x \leq b \end{cases}$$

The image below gives the graphs for f_2, f_3, and f_4 in the case where $[a,b] = [0,3]$ and $(c,d) = (1,2)$.



That (f_n) is a Cauchy sequence of continuous funtions in $(C[a,b], \|\cdot\|_1)$ which converges pointwise to 1_I and $\|f_n\|_1 \to \text{length}(I)$ as $n \to \infty$ follows from similar arguments used in case 1. □

7.3 Algebra of Subsets of X Closed under Symmetric Differences and Relative Compliments

Proposition 7.3. *Let $\mathcal{A}$ be an algebra of subsets of X. Then*

1. *$\mathcal{A}$ is closed under finite unions: if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.*
2. *$\mathcal{A}$ is closed under relative compliments: if $A, B \in \mathcal{A}$, then $A \backslash B \in \mathcal{A}$.*
3. *$\mathcal{A}$ is closed under symmetric differences: if $A, B \in \mathcal{A}$, then $A \Delta B \in \mathcal{A}$.*

Proof.
1. Let $A, B \in \mathcal{A}$. Then

$$\begin{aligned} A \cup B &= ((A \cup B)^c)^c \\ &= (A^c \cap B^c)^c \\ &\in \mathcal{A}. \end{aligned}$$

2. Let $A, B \in \mathcal{A}$. Then

$$\begin{aligned} A \backslash B &= A \cap B^c \\ &\in \mathcal{A}. \end{aligned}$$

3. Let $A, B \in \mathcal{A}$. Then it follows from 1 and 2 that

$$\begin{aligned} A \Delta B &= (A \backslash B) \cup (B \backslash A) \\ &\in \mathcal{A}. \end{aligned}$$

□

7.4 Collection of all Subintervals of $[a, b]$ Forms a Semialgebra

Definition 7.1. A nonempty collection $\mathcal{E}$ of subsets of X is said to be a **semialgebra** of sets if it satisfies the following properties:

1. $\emptyset \in \mathcal{E}$;
2. If $E, F \in \mathcal{E}$, then $E \cap F \in \mathcal{E}$;
3. If $E \in \mathcal{E}$, then E^c is a finite disjoint union of sets from $\mathcal{E}$.

Problem 4.a

Proposition 7.4. *The collection of all subintervals of $[a,b]$ forms a semialgebra of sets.*

Proof. Let $\mathcal{I}$ denote the collection of all subintervals of $[a,b]$. We have $\emptyset \in \mathcal{I}$ since $\emptyset = (c,c)$ for any $c \in [a,b]$.

Now we show $\mathcal{I}$ is closed under finite intersections. Let I_1 and I_2 be subintervals of $[a,b]$. Taking the closure of I_1 and I_2 gives us closed intervals, say

$$\overline{I}_1 = [c_1, d_1] \quad \text{and} \quad \overline{I}_2 = [c_2, d_2].$$

Assume without loss of generality that $c_1 \leq c_2$. If $d_1 < c_2$, then $I_1 \cap I_2 = \emptyset$, so assume that $d_1 \geq c_2$. If $d_1 \geq d_2$, then $I_1 \cap I_2 = I_2$, so assume that $d_1 < d_2$. If $c_1 = c_2$, then $I_1 \cap I_2 = I_1$, so assume that $c_2 > c_1$. So we have reduced the case to where

$$c_1 < c_2 \leq d_1 < d_2.$$

With these assumptions in mind, we now consider four cases:

Case 1: If $I_1 = [c_1, d_1]$ or $I_1 = (c_1, d_1]$ and $I_2 = [c_2, d_2]$ or $I_2 = [c_2, d_2)$, then $I_1 \cap I_2 = [c_2, d_1]$.

Case 2: If $I_1 = [c_1, d_1)$ or $I_1 = (c_1, d_1)$ and $I_2 = [c_2, d_2]$ or $I_2 = [c_2, d_2)$, then $I_1 \cap I_2 = [c_2, d_1)$.

Case 3: If $I_1 = [c_1, d_1]$ or $I_1 = (c_1, d_1)$ and $I_2 = (c_2, d_2]$ or $I_2 = (c_2, d_2)$, then $I_1 \cap I_2 = (c_2, d_1)$.

Case 4: If $I_1 = [c_1, d_1)$ or $I_1 = (c_1, d_1)$ and $I_2 = (c_2, d_2]$ or $I_2 = (c_2, d_2)$, then $I_1 \cap I_2 = (c_2, d_1)$.

In all cases, we see that $I_1 \cap I_2$ is a subinterval of $[a,b]$.

Now we show that compliments can be expressed as finite disjoint unions. Let I be a subinterval of $[a,b]$ and write $\overline{I} = [c,d]$. We consider four cases:

Case 1: If $I = [c,d]$, then $I^c = [a,c) \cup (d,b]$.

Case 2: If $I = (c,d]$, then $I^c = [a,c] \cup (d,b]$.

Case 3: If $I = [c,d)$, then $I^c = [a,c) \cup [d,b]$.

Case 4: If $I = (c,d)$, then $I^c = [a,c] \cup [d,b]$.

Thus in all cases, we can express I^c as a disjoint union of intervals since $a \leq c \leq d \leq b$. $\square$

Problem 4.b

Proposition 7.5. *Let $\mathcal{I}$ be the collection of all subintervals of $\mathbb{R} \cup \{\infty\}$ of the form $(a,b]$. Then $\mathcal{I}$ forms a semialgebra of sets.*

Proof. We have $\emptyset \in \mathcal{I}$ since $\emptyset = (c,c]$ for any $c \in \mathbb{R} \cup \{\infty\}$.

Now we show $\mathcal{I}$ is closed under finite intersections. Let $I_1 = (c_1, d_1]$ and $I_2 = (c_2, d_2]$. Assume without loss of generality that $c_1 \leq c_2$. Then

$$I_1 \cap I_2 = \begin{cases} (c_2, d_1] & \text{if } c_2 \leq d_1 \\ \emptyset & \text{else} \end{cases}$$

Now we show that compliments can be expressed as finite disjoint unions. Let $I = (c,d]$. Then

$$I^c = (-\infty, c] \cup (d, \infty],$$

where the union is disjoint since $c \leq d$. $\square$

Problem 4.c

Proposition 7.6. *Let $\mathcal{E}$ be a semialgebra of sets. Then the collection $\mathcal{A}$ consisting of all sets which are finite disjoint union of sets in $\mathcal{E}$ forms an algebra of sets.*

Proof. We have $\emptyset \in \mathcal{A}$ since $\emptyset \in \mathcal{E}$.

Next we show that $\mathcal{A}$ is closed under finite intersections. Let $A, A' \in \mathcal{A}$. Express A and A' as a disjoint union of members of $\mathcal{E}$, say

$$A = E_1 \cup \cdots \cup E_n \quad \text{and} \quad A' = E_1' \cup \cdots \cup E_{n'}'.$$

Then we have

$$\begin{aligned} A \cap A' &= \left(\bigcup_{i=1}^{n} E_i \right) \cap \left(\bigcup_{i'=1}^{n'} E_{i'}' \right) \\ &= \bigcup_{i'=1}^{n'} \left(\left(\bigcup_{i=1}^{n} E_i \right) \cap E_{i'}' \right) \\ &= \bigcup_{i'=1}^{n'} \left(\bigcup_{i=1}^{n} E_i \cap E_{i'}' \right) \\ &= \bigcup_{\substack{1 \leq i \leq n \\ 1 \leq i' \leq n'}} E_i \cap E_{i'}' \end{aligned}$$

where the union is disjoint since the E_i and $E_{i'}'$ are disjoint from one another.

Lastly we show that $\mathcal{A}$ is closed under compliments. Let $A \in \mathcal{A}$. Express A as a disjoint union of members of $\mathcal{E}$, say

$$A = E_1 \cup \cdots \cup E_n.$$

Then we have

$$\begin{aligned} A^c &= (E_1 \cup \cdots \cup E_n)^c \\ &= E_1^c \cap \cdots \cap E_n^c. \end{aligned}$$

Since the E_i^c belong to $\mathcal{A}$ and $\mathcal{A}$ is closed under finite intersections, it follows that $A^c \in \mathcal{A}$. □

7.5 Collection of Subsets of $\mathbb{Z}$ Forms Algebra Under Certain Conditions

Proposition 7.7. *Let $\mathcal{A}$ be a collection of subsets of $\mathbb{Z}$ such that*

1. *X is a member of $\mathcal{A}$;*
2. *$\mathcal{A}$ is closed under relative compliments: $A \backslash B \in \mathcal{A}$ for all $A, B \in \mathcal{A}$.*

Then $\mathcal{A}$ is an algebra.

Proof. We have $\emptyset \in \mathcal{A}$ since $\emptyset = X \backslash X \in \mathcal{A}$. Clearly $\mathcal{A}$ is closed under compliments since it is closed under relative compliments, so we just need to show that $\mathcal{A}$ is closed under finite intersections. Let $A, B \in \mathcal{A}$. Then

$$\begin{aligned} A \cap B &= A \cap (B^c)^c \\ &= A \backslash B^c \\ &\in \mathcal{A}. \end{aligned}$$

□

7.6 Finite Complement Algebra

Proposition 7.8. *Let $\mathcal{A}$ be the collection of subsets of X which satisfies the property that if $A \in \mathcal{A}$ then either A or A^c is finite. Then $\mathcal{A}$ forms an algebra.*

Proof. We have $\emptyset \in \mathcal{A}$ since $\emptyset$ is finite. Clearly $\mathcal{A}$ is closed under compliments since $A \in \mathcal{A}$ implies either A or A^c is finite which implies $A^c \in \mathcal{A}$. It remains to show that $\mathcal{A}$ is closed under finite intersections. Let $A, B \in \mathcal{A}$ and suppose that $A \cap B$ is infinite. We must show that $(A \cap B)^c = A^c \cup B^c$ is finite. In other words, we need to show that both A^c and B^c are finite. Assume for a contradiction that A^c is infinite. Then A must be finite since $A \in \mathcal{A}$. But this implies $A \cap B$ is finite, which is a contradiction. Thus A^c must be finite. Similarly, we can prove by contradiction that B^c is finite too. $\square$

7.7 Ascending Sequence of Algebras is an Algebra

Proposition 7.9. *Let $(\mathcal{A}_n)$ be an ascending sequence of algebras over X, that is, $\mathcal{A}_n$ is an algebra of subsets of X and $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ for all $n \in \mathbb{N}$. Then*

$$\mathcal{A} := \bigcup_{n \in \mathbb{N}} \mathcal{A}_n$$

is an algebra.

Proof. We have $\emptyset \in \mathcal{A}$ since $\emptyset \in A_1 \subseteq \mathcal{A}$. Next we show that $\mathcal{A}$ is closed under finite intersections. Let $A, B \in \mathcal{A}$. Then $A \in A_i$ and $B \in \mathcal{A}_j$ for some $i, j \in \mathbb{N}$. Without loss of generality, assume that $i \leq j$. Then $A \in \mathcal{A}_i \subseteq \mathcal{A}_j$. Thus $A \cap B \in \mathcal{A}_j \subseteq \mathcal{A}$. Lastly we show that $\mathcal{A}$ is closed under compliments. Let $A \in \mathcal{A}$. Then $A \in A_i$ for some $i \in \mathbb{N}$. Thus $A^c \in \mathcal{A}_i \subseteq \mathcal{A}$. $\square$

Remark 8. The ascending condition is not necessary. Indeed, consider $X = \{a, b, c\}$ and

$$\begin{aligned}
\mathcal{A} &= \{\emptyset, X, \{a\}, \{b, c\}\} \\
\mathcal{B} &= \{\emptyset, X, \{b\}, \{a, c\}\} \\
\mathcal{C} &= \{\emptyset, X, \{c\}, \{a, b\}\}
\end{aligned}$$

Then $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are algebras over X, and $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C} = \mathcal{P}(X)$ is an algebra over X, but none of the $\mathcal{A}$, $\mathcal{B}$, or $\mathcal{C}$ contain one another.

8 Homework 2

Throughout this homework, let $(X, \mathcal{M}, \mu)$ be a measure space. We say that $(X, \mathcal{M}, \mu)$ is a **finite** measure space if $\mu(X) < \infty$. Observe that in this case, we have $\mu(A) < \infty$ for all $A \in \mathcal{M}$, by monotonicity of μ.

8.1 Countable Subadditivity of Finite Measure

Proposition 8.1. *Let (E_n) be a sequence of sets in $\mathcal{M}$. Then*

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

Proof. Disjointify[3] (E_n) into the sequence (D_n); set $D_1 := E_1$ and

$$D_n := E_n \backslash \left(\bigcup_{i=1}^{n-1} E_n\right)$$

for all $n > 1$. Then we have

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{\infty} E_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} D_n\right) \\
&= \sum_{n=1}^{\infty} \mu(D_n) \\
&\leq \sum_{n=1}^{\infty} \mu(E_n),
\end{aligned}$$

where we use countable additivity of μ to get from the first line to the second line and where we used monotonicity of μ to get from the second line to the third line. $\square$

[3] See Appendix for details on disjointification.

8.2 Inverse Image of σ-Algebra is σ-Algebra

Proposition 8.2. *Let $(Y, \mathcal{N}, \nu)$ be a measure space and suppose $f\colon X \to Y$ is a function. Then $(X, f^{-1}(\mathcal{N}), f^{-1}\mu)$ is a measure space, where*

$$f^{-1}(\mathcal{N}) = \{f^{-1}(B) \subseteq X \mid B \in \mathcal{N}\}$$

and where $f^{-1}\mu\colon f^{-1}(\mathcal{N}) \to [0,\infty]$ is defined by

$$(f^{-1}\mu)(A) = \mu^*(f(A)).$$

for all $A \in f^{-1}(\mathcal{N})$.

Proof. We first show that $f^{-1}(\mathcal{N})$ is a σ-algebra. This follows from the fact that f^{-1} commutes with unions and compliments:

$$f^{-1}\left(\bigcup_{j\in J} B_j\right) = \bigcup_{j\in J} f^{-1}\left(B_j\right) \quad \text{and} \quad f^{-1}\left(Y \backslash B\right) = f^{-1}(Y)\backslash f^{-1}(B)$$

for all subsets B and B_j of Y for all $j \in J$. Indeed, we have

$$\begin{aligned} x \in \bigcup_{j\in J} f^{-1}\left(B_j\right) &\iff x \in f^{-1}(B_j) \text{ for some } j \in J \\ &\iff f(x) \in B_j \text{ for some } j \in J \\ &\iff f(x) \in \bigcup_{j\in J} B_j \\ &\iff x \in f^{-1}\left(\bigcup_{j\in J} B_j\right) \end{aligned}$$

and we have

$$\begin{aligned} x \in f^{-1}\left(Y\backslash B\right) &\iff f(x) \in Y\backslash B \\ &\iff f(x) \in Y \text{ and } f(a) \notin B \\ &\iff x \in f^{-1}(Y) \text{ and } a \notin f^{-1}(B) \\ &\iff x \in f^{-1}(Y)\backslash f^{-1}(B). \end{aligned}$$

Now we show that the function $f^*\mu$ is a measure. We have

$$\begin{aligned} (f^{-1}\mu)(\varnothing) &= \inf\{\mu(B) \mid f(\varnothing) \subset B\} \\ &= \inf\{\mu(B) \mid \varnothing \subset B\} \\ &\leq \mu(\varnothing) \\ &= 0. \end{aligned}$$

Next, let (A_n) be a sequence of members of $f^{-1}(\mathcal{N})$. Then

$$\begin{aligned} (f^{-1}\mu)\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu^*\left(f\left(\bigcup_{n=1}^{\infty} A_n\right)\right) \\ &= \mu^*\left(\bigcup_{n=1}^{\infty} f(A_n)\right) \\ &\leq \sum_{n=1}^{\infty} \mu^*(f(A_n)) \\ &= \sum_{n=1}^{\infty} (f^{-1}\mu)(A_n). \end{aligned}$$

Thus $f^{-1}\mu$ is countably subadditive.

Finally, let A and A' be two members in $f^{-1}(\mathcal{N})$ such that $A \cap A' = \emptyset$. Then

$$\begin{aligned}(f^{-1}\mu)(A \cup A') &= \mu^*(f(A) \cup f(A')) \\ &= \\ &= \end{aligned}$$

. ☐

We first show that

$$\sum_{n=1}^{\infty} \inf \{\mu(B_n) \mid f(A_n) \subset B_n\} \leq \inf \left\{ \mu(B) \mid \bigcup_{n=1}^{\infty} f(A_n) \subset B \right\}.$$

Let $\varepsilon > 0$. Choose

8.3 Locally Measurable Sets

Definition 8.1. A set $E \subseteq X$ is called **locally measurable** if $E \cap A \in \mathcal{M}$ for all $A \in \mathcal{M}$ with finite measure.

Proposition 8.3. *Suppose $(X, \mathcal{M}, \mu)$ is a finite measure space. Let $\widetilde{\mathcal{M}}$ be the collection of all locally measurable subsets of X. Then $\mathcal{M} = \widetilde{\mathcal{M}}$.*

Proof. Let first show that $\mathcal{M} \subseteq \widetilde{\mathcal{M}}$. Let $E \in \mathcal{M}$. Then $E \cap A \in \mathcal{M}$ for all $A \in \mathcal{M}$ since $\mathcal{M}$ is closed under finite intersections. In particular, this implies $E \in \widetilde{\mathcal{M}}$. Thus $\mathcal{M} \subseteq \widetilde{\mathcal{M}}$.

Now we show the reverse inclusion $\mathcal{M} \supseteq \widetilde{\mathcal{M}}$. Let $E \in \widetilde{\mathcal{M}}$. Since $\mu(X) < \infty$ and E is locally measurable, we have

$$\begin{aligned} E &= E \cap X \\ &\in \mathcal{M}. \end{aligned}$$

Thus $\mathcal{M} \supseteq \widetilde{\mathcal{M}}$. ☐

8.4 If $\mu(A) < \infty$, then $\mu(B \backslash A) = \mu(B) - \mu(A)$

Lemma 8.1. *Let $A, B \in \mathcal{M}$ such that $A \subseteq B$. If $\mu(A) < \infty$, then*

$$\mu(B \backslash A) = \mu(B) - \mu(A).$$

Proof. By finite additivity of μ, we have

$$\begin{aligned} \mu(B) &= \mu((B \backslash A) \cup A) \\ &= \mu(B \backslash A) + \mu(A). \end{aligned}$$

If moreover $\mu(A) < \infty$, then we may subtract $\mu(A)$ from both sides to obtain

$$\mu(B \backslash A) = \mu(B) - \mu(A).$$

☐

Problem 4.a

Proposition 8.4. *Suppose $(X, \mathcal{M}, \mu)$ is a finite measure space. Then*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

for all $A, B \in \mathcal{M}$.

Proof. Let $A, B \in \mathcal{M}$. Then by finite additivity of μ, we have

$$\begin{aligned} \mu(A \cup B) &= \mu(A \cup (B \backslash A)) \\ &= \mu(A) + \mu(B \backslash A) \\ &= \mu(A) + \mu(B \backslash (A \cap B)) \\ &= \mu(A) + \mu(B) - \mu(A \cap B), \end{aligned}$$

where the last equality follows Lemma (8.1) since $(X, \mathcal{M}, \mu)$ is a finite measure space. ☐

8.5 Countable Additivity of μ for "Almost Pairwise Disjoint" Sets

Proposition 8.5. *Let (A_n) be a sequence of "almost pairwise disjoint" members of $\mathcal{M}$, in the sense that $\mu(A_i \cap A_j) = 0$ whenever $i \neq j$. Then*

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

Proof. First note that countable subadditivity of μ implies

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

so it suffices to show the reverse inequality. Before doing so, we first prove by induction on $N \geq 1$, that

$$\mu\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mu(A_n). \tag{25}$$

The base case $N = 1$ holds trivially. Assume that we have shown (25) holds for some $N > 1$. Then

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{N+1} A_n\right) &= \mu\left(\left(\bigcup_{n=1}^{N} A_n\right) \cup A_{N+1}\right) \\
&= \mu\left(\bigcup_{n=1}^{N} A_n\right) + \mu(A_{N+1}) - \mu\left(\left(\bigcup_{n=1}^{N} A_n\right) \cap A_{N+1}\right) \\
&= \sum_{n=1}^{N} \mu\left(A_N\right) + \mu(A_{N+1}) - \mu\left(\bigcup_{n=1}^{N} (A_n \cap A_{N+1})\right) \\
&\geq \sum_{n=1}^{N+1} \mu(A_n) - \sum_{n=1}^{N} \mu(A_n \cap A_{N+1}) \\
&= \sum_{n=1}^{N+1} \mu(A_n) - \sum_{n=1}^{N} 0 \\
&= \sum_{n=1}^{N+1} \mu(A_n),
\end{aligned}$$

where we used the induction hypothesis to get from the second line to the third line, and where we used finite subadditivity of μ to get from the third line to the fourth line. We already have

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

by finite subadditivity of μ, and so it follows that

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

Therefore (25) holds for all $N \in \mathbb{N}$ induction.

Now we prove the reverse inequality: for each $N \in \mathbb{N}$, we have

$$\begin{aligned}
\sum_{n=1}^{N} \mu(A_n) &= \mu\left(\bigcup_{n=1}^{N} A_n\right) \\
&\subseteq \mu\left(\bigcup_{n=1}^{\infty} A_n\right)
\end{aligned}$$

by monotonicity of μ. By taking $N \to \infty$, we see that

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

$\square$

8.6 Nonuniqueness of Extension of Algebra to σ-Algebra

Proposition 8.6. *Let $\mathcal{A}$ be the collection of all finite unions of sets of the form $(a,b] \cap \mathbb{Q}$ where $-\infty \leq a \leq b \leq \infty$. Then*

1. *$\mathcal{A}$ is an algebra of subsets of $\mathbb{Q}$;*
2. *$\sigma(\mathcal{A}) = \mathcal{P}(\mathbb{Q})$ where $\mathcal{P}(\mathbb{Q})$ is the collection of all subsets of $\mathbb{Q}$;*
3. *the function $\mu \colon \mathcal{A} \to [0,\infty]$ defined by $\mu(\varnothing) = 0$ and $\mu(A) = \infty$ for all nonempty $A \in \mathcal{A}$ is a measure on $\mathcal{A}$;*
4. *there is more than one measure on $\sigma(\mathcal{A})$ whose restriction to $\mathcal{A}$ is μ;*

Proof.
1. In Homework 1, it was shown that $(\mathbb{R} \cup \{\infty\}, \mathcal{T})$ was a semialgebra, where $\mathcal{T}$ consisted of all subintervals of $\mathbb{R} \cup \{\infty\}$ of the form $(a,b]$ where $-\infty \leq a \leq b \leq \infty$. If we let $\iota \colon \mathbb{Q} \to \mathbb{R} \cup \{\infty\}$ denote the inclusion map, then we see that $\iota^{-1}(\mathcal{T}) = \mathcal{S}$, where $\mathcal{S}$ denotes the collection of all subintervals of $\mathbb{Q}$ of the form $(a,b] \cap \mathbb{Q}$. It follows easily from Proposition (8.2) that $\mathcal{S}$ is a semialgebra of subsets of $\mathbb{Q}$.[4]
Therefore the set of all finite disjoint unions of members of $\mathcal{S}$ forms an algebra, and as any finite union of members of $\mathcal{S}$ can be expressed as a finite disjoint union of members of $\mathcal{S}$ (since $\mathcal{S}$ is a semialgebra), we see that $\mathcal{A}$ is an algebra.

2. Clearly $\mathcal{P}(\mathbb{Q}) \supseteq \sigma(\mathcal{A})$. Let us prove the reverse inclusion. We first observe that $\{r\} \in \sigma(\mathcal{A})$ for all $r \in \mathbb{Q}$. Indeed, if $r \in \mathbb{Q}$, then we have

$$\{r\} = \bigcap_{n \in \mathbb{N}} (r - 1/n, r] \cap \mathbb{Q} \in \sigma(\mathcal{A})$$

Now let $S \in \mathcal{P}(\mathbb{Q})$. Then since S is countable, we have

$$S = \bigcup_{s \in S} \{s\} \in \sigma(\mathcal{A}).$$

3. We have $\mu(\varnothing) = 0$ by definition. Let (A_n) be a sequence of pairwise disjoint members of $\mathcal{A}$ whose union also belongs to $\mathcal{A}$. If $\bigcup_{n=1}^{\infty} A_n \neq \varnothing$, then $A_n \neq \varnothing$ for some $n \in \mathbb{N}$, and thus

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \infty \\ &= \mu(A_n) \\ &= \sum_{n=1}^{\infty} \mu(A_n). \end{aligned}$$

Similarly, if $\bigcup_{n=1}^{\infty} A_n = \varnothing$, then $A_n = \varnothing$ for all $n \in \mathbb{N}$, and thus

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= 0 \\ &= \sum_{n=1}^{\infty} 0 \\ &= \sum_{n=1}^{\infty} \mu(A_n). \end{aligned}$$

In both cases, we have

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

[4] Technically we showed that the inverse image of a σ-algebra is a σ-algebra. However the same reasoning used in that proof shows that the inverse image of a semialgebra is a semialgebra: namely f^{-1} commutes with complements and unions.

4. We define $\mu_1 \colon \mathcal{P}(\mathbb{Q}) \to [0,\infty]$ and $\mu_2 \colon \mathcal{P}(\mathbb{Q}) \to [0,\infty]$ by

$$\mu_1(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ \infty & \text{else} \end{cases} \quad \text{and} \quad \mu_2(A) = \begin{cases} 0 & \text{if } A = \varnothing \\ \infty & \text{if } A \neq \varnothing \end{cases}$$

for all $A \in \mathcal{P}(\mathbb{Q})$. Both μ_1 and μ_2 restrict to μ as functions since every member of $\mathcal{A}$ is infinite. They are also both distinct as functions since, for example, $\mu_1(\{x\}) = 1$ and $\mu_2(\{x\}) = \infty$ for any $x \in \mathbb{Q}$. Thus it suffices to show that they are measures. That μ_2 is a measure follows from a similar argument as in the case of μ, so we just show that μ_1 is a measure. We have $\mu_1(\varnothing) = 0$ since $|\varnothing| = 0$. Next we show it is finitely additive. Let A and B be members of $\mathcal{P}(\mathbb{Q})$ such that $A \cap B = \varnothing$. If $A = \varnothing$, then

$$\begin{aligned} \mu_1(A \cup B) &= \mu_1(\varnothing \cup B) \\ &= \mu_1(B) \\ &= 0 + \mu_1(B) \\ &= \mu_1(\varnothing) + \mu_1(B) \\ &= \mu_1(A) + \mu_1(B). \end{aligned}$$

Similarly, if $B = \varnothing$, then $\mu_1(A \cup B) = \mu_1(A) \cup \mu_1(B)$. So assume neither A nor B is the emptyset. Write them as

$$A = \{x_1, \dots x_m\} \quad \text{and} \quad B = \{y_1, \dots, y_n\}.$$

Then

$$A \cup B = \{x_1, \dots, x_m, y_1, \dots, y_n\},$$

and so

$$\begin{aligned} \mu_1(A \cup B) &= m + n \\ &= \mu_1(A) + \mu_1(B). \end{aligned}$$

It follows that μ_1 is finitely additive.

Now, let (A_n) be a sequence of pairwise disjoint members of $\mathcal{P}(\mathbb{Q})$. Suppose that $A_n \neq \varnothing$ for only finitely many n, say $n_1, \dots, n_k$. Then it follows from finite additivity of μ_1 and the fact that $\mu(\varnothing) = 0$ that

$$\begin{aligned} \mu_1\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu_1\left(\bigcup_{i=1}^{k} A_{n_i}\right) \\ &= \sum_{i=1}^{k} \mu(A_{n_i}) \\ &= \sum_{n=1}^{\infty} \mu(A_n). \end{aligned}$$

Now suppose that $A_n \neq \varnothing$ for infinitely many n. By taking a subsequence of (A_n) if necessary, we may assume that $A_n \neq \varnothing$ for all n. Then $\bigcup_{n=1}^{\infty} A_n$ is infinite, and so

$$\begin{aligned} \mu_1\left(\bigcup_{n=1}^{\infty} A_n\right) &= \infty \\ &\geq \sum_{n=1}^{\infty} \mu_1(A_n) \\ &\geq \sum_{n=1}^{\infty} 1 \\ &= \infty. \end{aligned}$$

It follows that

$$\mu_1\left(\bigcup_{n=1}^{\infty} A_n\right) = \infty = \sum_{n=1}^{\infty} \mu_1(A_n).$$

Therefore μ_1 and μ_2 are distinct measures which restrict to μ. □

Remark 9. Note that the extension theorem does not apply here as μ is not a finite measure.

8.7 Symmetric Difference Identities

Proposition 8.7. *Let $A, B \in \mathcal{P}(X)$. Then the following properties hold*

1. $A\Delta A = \emptyset$;
2. $(A\Delta B)\Delta C = A\Delta(B\Delta C)$;
3. $(A\Delta B)\Delta(B\Delta C) = A\Delta C$;
4. $(A\Delta B)\Delta(C\Delta D) = (A\Delta C)\Delta(B\Delta D)$;
5. $|1_A - 1_B| = 1_{A\Delta B}$.

Proof.
1. We have

$$\begin{aligned} A\Delta A &= (A\backslash A) \cup (A\backslash A) \\ &= \emptyset \cup \emptyset \\ &= \emptyset. \end{aligned}$$

2. We have

$$\begin{aligned} (A\Delta B)\Delta C &= ((A\Delta B) \cup C) \cap ((A\Delta B) \cap C)^c \\ &= ((A\Delta B) \cup C) \cap ((A\Delta B)^c \cup C^c) \\ &= (((A \cup B) \cap (A \cap B)^c) \cup C)) \cap ((A \cap B^c) \cup (A^c \cap B))^c \cup C^c) \\ &= (((A \cup B) \cap (A^c \cup B^c)) \cup C)) \cap (((A \cap B^c)^c \cap (A^c \cap B)^c) \cup C^c) \\ &= (A \cup B \cup C) \cap (A^c \cup B^c \cup C) \cap ((A^c \cup B) \cap (A \cup B^c)) \cup C^c) \\ &= (A \cup B \cup C) \cap (A^c \cup B^c \cup C) \cap (A^c \cup B \cup C^c) \cap (A \cup B^c \cup C^c) \\ &= (B \cup C \cup A) \cap (B^c \cup C^c \cup A) \cap (B^c \cup C \cup A^c) \cap (B \cup C^c \cup A^c) \\ &= ((B \cup C \cup A) \cap (B^c \cup C^c \cup A)) \cap ((B^c \cup C) \cap (B \cup C^c)) \cup A^c) \\ &= ((B \cup C \cup A) \cap (B^c \cup C^c \cup A)) \cap (((B \cap C^c)^c \cap (B^c \cap C)^c) \cup A^c) \\ &= (((B \cup C) \cap (B \cap C)^c) \cup A)) \cap ((B \cap C^c) \cup (B^c \cap C))^c \cup A^c) \\ &= ((B\Delta C) \cup A) \cap ((B\Delta C)^c \cup A^c) \\ &= ((B\Delta C) \cup A) \cap ((B\Delta C) \cap A)^c \\ &= (B\Delta C)\Delta A \\ &= A\Delta(B\Delta C) \end{aligned}$$

3. We have

$$\begin{aligned} (A\Delta B)\Delta(B\Delta C) &= A\Delta B\Delta B\Delta C \\ &= A\Delta\emptyset\Delta C \\ &= A\Delta C. \end{aligned}$$

4. We have

$$\begin{aligned} (A\Delta B)\Delta(C\Delta D) &= A\Delta B\Delta C\Delta D \\ &= A\Delta C\Delta B\Delta D \\ &= (A\Delta C)\Delta(B\Delta D) \end{aligned}$$

5. Let $x \in X$. If $x \notin A \cup B$, then

$$\begin{aligned}|1_A(x) - 1_B(x)| &= |0-0| \\ &= 0 \\ &= 0-0 \\ &= 1_{A\cup B}(x) - 1_{A\cap B}(x) \\ &= 1_{A\Delta B}(x).\end{aligned}$$

If $x \in A\backslash B$, then

$$\begin{aligned}|1_A(x) - 1_B(x)| &= |1-0| \\ &= 1 \\ &= 1-0 \\ &= 1_{A\cup B}(x) - 1_{A\cap B}(x) \\ &= 1_{A\Delta B}(x).\end{aligned}$$

If $x \in B\backslash A$, then

$$\begin{aligned}|1_A(x) - 1_B(x)| &= |0-1| \\ &= 1 \\ &= 1-0 \\ &= 1_{A\cup B}(x) - 1_{A\cap B}(x) \\ &= 1_{A\Delta B}(x).\end{aligned}$$

If $x \in A \cap B$, then

$$\begin{aligned}|1_A(x) - 1_B(x)| &= |1-1| \\ &= 0 \\ &= 1-1 \\ &= 1_{A\cup B}(x) - 1_{A\cap B}(x) \\ &= 1_{A\Delta B}(x).\end{aligned}$$

Thus $|1_A(x) - 1_B(x)| = 1_{A\Delta B}(x)$ for all $x \in X$ and hence $|1_A - 1_B| = 1_{A\Delta B}$. □

8.8 More Symmetric Difference Identities

Proposition 8.8. *Let (A_n) and (B_n) be two sequences of sets. Then*

$$\left(\bigcup_{m=1}^{\infty} A_m\right) \Delta \left(\bigcup_{n=1}^{\infty} B_n\right) \subseteq \bigcup_{n=1}^{\infty} (A_n \Delta B_n) \quad \textit{and} \quad \left(\bigcap_{m=1}^{\infty} A_m\right) \Delta \left(\bigcap_{n=1}^{\infty} B_n\right) \subseteq \bigcup_{n=1}^{\infty} (A_n \Delta B_n).$$

Proof. We have

$$\begin{aligned}\left(\bigcup_{m=1}^{\infty} A_m\right) \Delta \left(\bigcup_{n=1}^{\infty} B_n\right) &= \left(\left(\bigcup_{m=1}^{\infty} A_m\right) \cup \left(\bigcup_{n=1}^{\infty} B_n\right)\right) \Big\backslash \left(\left(\bigcup_{m=1}^{\infty} A_m\right) \cap \left(\bigcup_{n=1}^{\infty} B_n\right)\right) \\ &= \left(\bigcup_{n=1}^{\infty} (A_n \cup B_n)\right) \Big\backslash \left(\bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} (A_m \cap B_n)\right) \\ &\subseteq \left(\bigcup_{n=1}^{\infty} (A_n \cup B_n)\right) \Big\backslash \left(\bigcup_{n=1}^{\infty} (A_n \cap B_n)\right) \\ &\subseteq \bigcup_{n=1}^{\infty} (A_n \cup B_n) \backslash (A_n \cap B_n) \\ &= \bigcup_{n=1}^{\infty} (A_n \Delta B_n).\end{aligned}$$

Similarly, we have

$$
\begin{aligned}
\left(\bigcap_{m=1}^{\infty} A_m\right) \Delta \left(\bigcap_{n=1}^{\infty} B_n\right) &= \left(\bigcap_{m=1}^{\infty} (A_m^c)^c\right) \Delta \left(\bigcap_{n=1}^{\infty} (B_n^c)^c\right) \\
&= \left(\bigcup_{m=1}^{\infty} A_m^c\right)^c \Delta \left(\bigcup_{n=1}^{\infty} B_n^c\right)^c \\
&= \left(\bigcup_{m=1}^{\infty} A_m^c\right) \Delta \left(\bigcup_{n=1}^{\infty} B_n^c\right) \\
&\subseteq \bigcup_{n=1}^{\infty} (A_n^c \Delta B_n^c) \\
&= \bigcup_{n=1}^{\infty} (A_n \Delta B_n).
\end{aligned}
$$

□

Appendix

Disjointification

Proposition 8.9. *Let $\mathcal{A}$ be an algebra of subsets of X and let (A_n) be a sequence of sets in $\mathcal{A}$. Then there exists a sequence (D_n) of sets in $\mathcal{A}$ such that*

1. *$D_n \subseteq A_n$ for all $n \in \mathbb{N}$.*
2. *$D_m \cap D_n = \varnothing$ for all $m, n \in \mathbb{N}$ such that $m \neq n$.*
3. *$\bigcup_{m=1}^{n} D_m = \bigcup_{m=1}^{n} A_m$ for all $n \in \mathbb{N}$.*

*We say the sequence (D_n) is the **disjointification** of the sequence (A_n) or that we **disjointify** the sequence (A_n) to the sequence (D_n).*

Proof. Set $D_1 := A_1$ and

$$D_n := A_n \setminus \left(\bigcup_{m=1}^{n-1} A_m\right)$$

for all $n > 1$. It is clear that $D_n \in \mathcal{A}$ and that $D_n \subseteq A_n$ for all $n \in \mathbb{N}$. Let us show that $D_m \cap D_n = \varnothing$ whenever $m \neq n$. Without loss of generality, we may assume that $m < n$. Then since $D_m \subseteq A_m$ and $D_n \cap A_m = \varnothing$, we have $D_m \cap D_n = \varnothing$. It remains to show

$$\bigcup_{m=1}^{n} D_m = \bigcup_{m=1}^{n} A_m$$

for all $n \in \mathbb{N}$. Let $n \in \mathbb{N}$. Since $D_m \subseteq A_m$ for all $m \leq n$, we have

$$\bigcup_{m=1}^{n} D_m \subseteq \bigcup_{m=1}^{n} A_m.$$

To show the reverse inclusion, let $x \in \bigcup_{m=1}^{n} A_m$. Then $x \in A_m$ for some $m = 1, \ldots, n$. Choose m to be the smallest natural number such that $x \in A_m$. Then x belongs to A_m but does not belong to $A_1, \ldots, A_{m-1}$. In other words,

$$x \in D_m \subseteq \bigcup_{k=1}^{n} D_k.$$

This implies the reverse inclusion

$$\bigcup_{m=1}^{n} D_m \supseteq \bigcup_{m=1}^{n} A_m.$$

□

9 Homework 3

Throughout this homework, let X be a set and let $\mathcal{P}(X)$ denote the power set of X.

9.1 Limsup and Liminf (of sets) Identities

Proposition 9.1. *Let (A_n) be a sequence in $\mathcal{P}(X)$. Then*

1. $(\liminf A_n)^c = \limsup A_n^c$;
2. $\liminf A_n = \{x \in X \mid \sum_{n=1}^{\infty} 1_{A_n^c}(x) < \infty\} = \{x \in X \mid x \in A_n$ *for all but finitely many* $n\}$.
3. $\limsup A_n = \{x \in X \mid \sum_{n=1}^{\infty} 1_{A_n}(x) = \infty\} = \{x \in X \mid x \in A_{\pi(n)}$ *for all n some subsequence* $(A_{\pi(n)})$ *of* $(A_n)\}$.
4. $\liminf A_n \subseteq \limsup A_n$;
5. $1_{\liminf A_n} = \liminf 1_{A_n}$ *and* $1_{\limsup A_n} = \limsup A_n$.

Proof. 1. We have

$$\begin{aligned}
(\liminf A_n)^c &= \left(\bigcup_{N=1}^{\infty}\left(\bigcap_{n\geq N} A_n\right)\right)^c \\
&= \bigcap_{N=1}^{\infty}\left(\left(\bigcap_{n\geq N} A_n\right)^c\right) \\
&= \bigcap_{N=1}^{\infty}\left(\bigcup_{n\geq N} A_n^c\right) \\
&= \limsup A_n^c.
\end{aligned}$$

2. First note that

$$\begin{aligned}
x \in \liminf A_n &\iff x \in \bigcup_{N=1}^{\infty}\left(\bigcap_{n\geq N} A_n\right) \\
&\iff x \in \bigcap_{n\geq N} A_n \text{ for some } N \in \mathbb{N} \\
&\iff x \in A_n \text{ for all } n \geq N \text{ for some } N \in \mathbb{N}.
\end{aligned}$$

Now if $x \in A_n$ for all $n \geq N$ for some $N \in \mathbb{N}$, then clearly $x \in A_n$ for all but finitely many n. Conversely, let $x \in A_n$ for all but finitely many n. Set $N = \max\{n \mid x \notin A_n\}$. Then $x \in A_n$ for all $n \geq N$. Thus

$$\liminf A_n = \{x \in X \mid x \in A_n \text{ for all but finitely many } n\}.$$

Similarly, if $x \in X$ such that

$$\sum_{n=1}^{\infty} 1_{A_n^c}(x) < \infty,$$

then $1_{A_n^c}(x) = 1$ for only finitely many n. In other words, $x \in A_n$ for all but finitely many n. Convsersely, if $x \in A_n$ for all but finitely many n, then $x \in A_n^c$ for only finitely many n, and thus

$$\sum_{n=1}^{\infty} 1_{A_n^c}(x) < \infty.$$

Therefore

$$\left\{x \in X \mid \sum_{n=1}^{\infty} 1_{A_n^c}(x) < \infty\right\} = \{x \in X \mid x \in A_n \text{ for all but finitely many } n\}.$$

3. First note that

$$\begin{aligned}x \in \limsup A_n &\iff x \in \bigcap_{N=1}^{\infty}\left(\bigcup_{n\geq N} A_n\right)\\ &\iff x \in \bigcup_{n\geq N} A_n \text{ for all } N \in \mathbb{N}\\ &\iff x \in A_n \text{ for some } n \geq N \text{ for all } N \in \mathbb{N}.\end{aligned}$$

In other words, $x \in \limsup A_n$ if and only if for each $n \in \mathbb{N}$ we can find a $\pi(n) \geq n$ such that $x \in A_{\pi(n)}$, or equivalently, if and only if $x \in A_{\pi(n)}$ for all $n \in \mathbb{N}$ where $(A_{\pi(n)})$ is a subsequence of (A_n). Thus

$$\limsup A_n = \{x \in X \mid x \in A_{\pi(n)} \text{ for some subsequence } (A_{\pi(n)}) \text{ of } (A_n)\}.$$

Similarly, suppose $x \in A_{\pi(n)}$ for all $n \in \mathbb{N}$ where $(A_{\pi(n)})$ is a subsequence of (A_n). Then

$$\begin{aligned}\sum_{n=1}^{\infty} 1_{A_n}(x) &\geq \sum_{n=1}^{\infty} 1_{A_{\pi(n)}}(x)\\ &= \infty.\end{aligned}$$

Conversely, if

$$\sum_{n=1}^{\infty} 1_{A_n}(x) = \infty,$$

then $x \in A_n$ for infinitely many n. Thus there is a subsequence $(A_{\pi(n)})$ of (A_n) such that $x \in A_{\pi(n)}$ for all $n \in \mathbb{N}$. Therefore

$$\left\{x \in X \mid \sum_{n=1}^{\infty} 1_{A_n}(x) = \infty\right\} = \{x \in X \mid x \in A_{\pi(n)} \text{ for some subsequence } (A_{\pi(n)}) \text{ of } (A_n)\}.$$

4. We have

$$\begin{aligned}x \in \liminf A_n &\iff x \in A_n \text{ for all } n \geq N \text{ for some } N\\ &\implies x \in A_n \text{ for infinitely many } n\\ &\iff x \in \limsup A_n.\end{aligned}$$

Thus

$$\liminf A_n \subseteq \limsup A_n.$$

5. We first show $1_{\liminf A_n} = \liminf 1_{A_n}$. Let $x \in X$. First assume that $x \in \liminf A_n$. Then $x \in A_n$ for all $n \geq N$ for some $N \in \mathbb{N}$. Then

$$\begin{aligned}1 &\geq \liminf(1_{A_n}(x))\\ &= \lim_{M\to\infty} \inf\{1_{A_m}(x) \mid m \geq M\}\\ &\geq \inf\{1_{A_n}(x) \mid n \geq N\}\\ &= \inf\{1 \mid n \geq N\}\\ &= 1\end{aligned}$$

implies

$$\begin{aligned}1_{\liminf A_n}(x) &= 1\\ &= \liminf(1_{A_n}(x))\\ &= (\liminf 1_{A_n})(x).\end{aligned}$$

Now assume that $x \notin \liminf A_n$. Then $x \notin A_n$ for infinitely many n. In particular, for each $N \in \mathbb{N}$, there exists a $\pi(N) \geq N$ such that $x \notin A_{\pi(N)}$. Then

$$\begin{aligned} 0 &\leq \liminf(1_{A_n}(x)) \\ &= \lim_{N\to\infty} \inf\{1_{A_n}(x) \mid n \geq N\} \\ &= \lim_{N\to\infty} 0 \\ &= 0 \end{aligned}$$

implies

$$\begin{aligned} 1_{\liminf A_n}(x) &= 0 \\ &= \liminf(1_{A_n}(x)) \\ &= (\liminf 1_{A_n})(x). \end{aligned}$$

Thus all cases we have $1_{\liminf A_n}(x) = (\liminf 1_{A_n})(x)$, and therefore

$$1_{\liminf A_n} = \liminf 1_{A_n}.$$

Now we will show $1_{\limsup A_n} = \limsup 1_{A_n}$. Let $x \in X$. First assume that $x \notin \limsup A_n$. Then $x \notin A_n$ for all $n \geq N$ for some $N \in \mathbb{N}$. Then

$$\begin{aligned} 0 &\leq \limsup(1_{A_n}(x)) \\ &= \lim_{M\to\infty} \sup\{1_{A_m}(x) \mid m \geq M\} \\ &\leq \sup\{1_{A_n}(x) \mid n \geq N\} \\ &= \sup\{0 \mid n \geq N\} \\ &= 0 \end{aligned}$$

implies

$$\begin{aligned} 1_{\limsup A_n}(x) &= 0 \\ &= \limsup(1_{A_n}(x)) \\ &= (\limsup 1_{A_n})(x). \end{aligned}$$

Now assume that $x \in \limsup A_n$. Then $x \in A_n$ for infinitely many n. In particular, for each $N \in \mathbb{N}$, there exists a $\pi(N) \geq N$ such that $x \in A_{\pi(N)}$. Then

$$\begin{aligned} 1 &\geq \limsup(1_{A_n}(x)) \\ &= \lim_{N\to\infty} \sup\{1_{A_n}(x) \mid n \geq N\} \\ &\geq \lim_{N\to\infty} 1 \\ &= 1 \end{aligned}$$

implies

$$\begin{aligned} 1_{\limsup A_n}(x) &= 0 \\ &= \limsup(1_{A_n}(x)) \\ &= (\limsup 1_{A_n})(x). \end{aligned}$$

Thus all cases we have $1_{\limsup A_n}(x) = (\limsup 1_{A_n})(x)$, and therefore

$$1_{\limsup A_n} = \limsup 1_{A_n}.$$

$\square$

9.2 Limsup, Liminf, and Symmetric Difference Identities

Problem 2.a

Proposition 9.2. *Let (A_n) be a sequence in $\mathcal{P}(X)$. Then*

$$\bigcup_{n=1}^{\infty}(A_n \Delta A_{n+1}) = \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m.$$

Proof. Suppose $x \in \bigcup_{n=1}^{\infty}(A_n \Delta A_{n+1})$. Choose $n \in \mathbb{N}$ such that $x \in A_n \Delta A_{n+1}$. Thus either $x \in A_n \backslash A_{n+1}$ or $x \in A_{n+1} \backslash A_n$. Without loss of generality, say $x \in A_n \backslash A_{n+1}$. Then since $x \in A_n$, we see that $x \in \bigcup_{n=1}^{\infty} A_n$ and since $x \notin A_{n+1}$, we see that $x \notin \bigcap_{n=1}^{\infty} A_n$. Therefore $x \in \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m$. This implies

$$\bigcup_{n=1}^{\infty}(A_n \Delta A_{n+1}) \subseteq \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m.$$

Conversely, suppose $x \in \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m$. Since $x \in \bigcup_{n=1}^{\infty} A_n$, there exists some $n \in \mathbb{N}$ such that $x \in A_n$. Since $x \notin \bigcap_{m=1}^{\infty} A_m$, there exists some $k \in \mathbb{N}$ such that $x \notin A_k$. Assume without loss of generality that $k < n$. Choose m to be the least natural number number such that $x \in A_m$, $x \notin A_{m-1}$, and $k < m \leq n$. Clearly this number exists since $x \notin A_k$ and $x \in A_n$. Then $x \in A_m \Delta A_{m-1}$, which implies $x \in \bigcup_{n=1}^{\infty}(A_n \Delta A_{n+1})$. Thus

$$\bigcup_{n=1}^{\infty}(A_n \Delta A_{n+1}) \supseteq \bigcup_{n=1}^{\infty} A_n \backslash \bigcap_{m=1}^{\infty} A_m.$$

□

Problem 2.b

Proposition 9.3. *Let (A_n) be a sequence in $\mathcal{P}(X)$. Then*

$$\limsup A_n \backslash \liminf A_n = \limsup(A_n \Delta A_{n+1}).$$

Proof. Suppose $x \in \limsup A_n \backslash \liminf A_n$. Then the sets

$$\{n \in \mathbb{N} \mid x \in A_n\} \quad \text{and} \quad \{n \in \mathbb{N} \mid x \notin A_n\}$$

are both infinite. We claim this implies that the set

$$\begin{aligned}\{n \in \mathbb{N} \mid x \in A_n \Delta A_{n+1}\} &= \{n \in \mathbb{N} \mid x \in (A_n \backslash A_{n+1}) \cup (A_{n+1} \backslash A_n)\} \\ &= \{n \in \mathbb{N} \mid x \in A_n \backslash A_{n+1}\} \cup \{n \in \mathbb{N} \mid x \in A_{n+1} \backslash A_n\}\end{aligned}$$

is infinite. To see this, we first assume without loss of generality that $x \in A_1$. Choose the least $\pi(1) > 1$ such that $x \notin A_{\pi(1)}$ and $x \in A_{\pi(1)-1}$. Observe that $\pi(1)$ exists since otherwise $\{n \in \mathbb{N} \mid x \notin A_n\}$ would be finite. Next, choose $\pi(2) > \pi(1)$ such that $x \in A_{\pi(2)}$ and $x \notin A_{\pi(2)-1}$. We again observe that $\pi(2)$ exists since otherwise $\{n \in \mathbb{N} \mid x \in A_n\}$ would be finite. Continuing in this manner, we obtain a strictly increasing sequence $(\pi(n))$ of natural numbers with

$$x \in A_{\pi(2n)} \backslash A_{\pi(2n)-1} \quad \text{and} \quad x \in A_{\pi(2n-1)-1} \backslash A_{\pi(2n-1)}$$

for all $n \geq 1$. In particular, both sets

$$\{n \in \mathbb{N} \mid x \in A_n \backslash A_{n+1}\} \quad \text{and} \quad \{n \in \mathbb{N} \mid x \in A_{n+1} \backslash A_n\}$$

are infinite. Thus $\{n \in \mathbb{N} \mid x \in A_n \Delta A_{n+1}\}$ is infinite, which implies $x \in \limsup(A_n \Delta A_{n+1})$. Therefore

$$\limsup A_n \backslash \liminf A_n \subseteq \limsup(A_n \Delta A_{n+1}).$$

Conversely, suppose $x \in \limsup(A_n \Delta A_{n+1})$. Then the set

$$\begin{aligned}\{n \in \mathbb{N} \mid x \in A_n \Delta A_{n+1}\} &= \{n \in \mathbb{N} \mid x \in (A_n \backslash A_{n+1}) \cup (A_{n+1} \backslash A_n)\} \\ &= \{n \in \mathbb{N} \mid x \in A_n \backslash A_{n+1}\} \cup \{n \in \mathbb{N} \mid x \in A_{n+1} \backslash A_n\}\end{aligned}$$

is infinite. This implies one of

$$\{n \in \mathbb{N} \mid x \in A_n \backslash A_{n+1}\} \quad \text{or} \quad \{n \in \mathbb{N} \mid x \in A_{n+1} \backslash A_n\}$$

is infinite. Without loss of generality, suppose $\{n \in \mathbb{N} \mid x \in A_n \backslash A_{n+1}\}$ is infinite. Thus there exists a strictly increasing sequence $(\pi(n))$ of natural numbers with $x \in A_{\pi(n)}$ and $x \notin A_{\pi(n)+1}$. In particular, both sets

$$\{n \in \mathbb{N} \mid x \in A_n\} \quad \text{and} \quad \{n \in \mathbb{N} \mid x \notin A_n\}$$

are infinite. Equivalently, we have $x \in \limsup A_n \backslash \liminf A_n$. Therefore

$$\limsup A_n \backslash \liminf A_n \supseteq \limsup (A_n \Delta A_{n+1}).$$

□

9.3 Measure of Intersection of a Descending Sequence of Sets

Proposition 9.4. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a descending sequence in $\mathcal{M}$ such that $\mu(E_1) < \infty$. Then*

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n) \tag{26}$$

Proof. The sequence $(E_1 \backslash E_n)_{n\in\mathbb{N}}$ is an ascending sequence in $\mathcal{M}$, hence

$$\begin{aligned}
\mu(E_1) - \lim_{n\to\infty} \mu(E_n) &= \lim_{n\to\infty} \mu(E_1) - \lim_{n\to\infty} \mu(E_n) \\
&= \lim_{n\to\infty} (\mu(E_1) - \mu(E_n)) \\
&= \lim_{n\to\infty} \mu(E_1 \backslash E_n) \\
&= \mu\left(\bigcup_{n=1}^{\infty} (E_1 \backslash E_n)\right) \\
&= \mu\left(E_1 \backslash \left(\bigcap_{n=1}^{\infty} E_n\right)\right) \\
&= \mu(E_1) - \mu\left(\bigcap_{n=1}^{\infty} E_n\right),
\end{aligned}$$

where we used the fact that $\mu(E_1) < \infty$ to get from the second line to the third line and also from fifth line to the sixth line. Also since $\mu(E_1) < \infty$, we can subtract $\mu(E_1)$ from both sides to obtain (26). □

9.4 Measure of Liminf is Less Than or Equal to Liminf of Measure

Proposition 9.5. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a sequence in $\mathcal{M}$. Then*

$$\mu(\liminf E_n) \le \liminf \mu(E_n)$$

Proof. Note that the sequence

$$\left(\bigcap_{n\ge N} E_n\right)_{N\in\mathbb{N}}$$

is an ascending sequence in N. Therefore we have

$$\begin{aligned}
\mu(\liminf E_n) &= \mu\left(\bigcup_{N=1}^{\infty}\left(\bigcap_{n\ge N} E_n\right)\right) \\
&= \liminf \mu\left(\bigcap_{n\ge N} E_n\right) \\
&\le \lim_{N\to\infty} \inf \{\mu(E_n) \mid n \ge N\} \\
&= \liminf \mu(E_n),
\end{aligned}$$

where we obtained the third line from the second line since

$$\mu\left(\bigcap_{n\geq N} E_n\right) \leq \mu(E_n)$$

for all $n \geq N$ by monotonicity of μ. $\square$

9.5 Measure of Limsup is Greater Than or Equal to Limsup of Measure (Assuming Some Finiteness Condition)

Proposition 9.6. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a sequence in $\mathcal{M}$ such that $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) < \infty$. Then*

$$\mu(\limsup E_n) \geq \limsup \mu(E_n)$$

Proof. Note that the sequence

$$\left(\bigcup_{n\geq N} E_n\right)_{N\in\mathbb{N}}$$

is a descending sequence in N. This together with the fact that $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) < \infty$ implies

$$\begin{aligned}
\mu(\limsup E_n) &= \mu\left(\bigcap_{N=1}^{\infty}\left(\bigcup_{n\geq N} E_n\right)\right) \\
&= \lim_{N\to\infty} \mu\left(\bigcup_{n\geq N} E_n\right) \\
&\geq \lim_{N\to\infty} \sup\{\mu(E_n) \mid n \geq N\} \\
&= \limsup \mu(E_n),
\end{aligned}$$

where we obtained the third line from the second line since

$$\mu\left(\bigcup_{n\geq N} E_n\right) \geq \mu(E_n)$$

for all $n \geq N$ by monotonicity of μ. $\square$

9.6 Assuming Some Finiteness Condition, Measure of Limsup is Zero

Proposition 9.7. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a sequence in $\mathcal{M}$ such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then*

$$\mu(\limsup E_n) = 0.$$

Proof. Note that the sequence

$$\left(\bigcup_{n\geq N} E_n\right)_{N\in\mathbb{N}}$$

is a descending sequence in N. This together with the fact that $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) < \infty$ implies

$$\begin{aligned}
\mu(\limsup E_n) &= \mu\left(\bigcap_{N=1}^{\infty}\left(\bigcup_{n\geq N} E_n\right)\right) \\
&= \lim_{N\to\infty} \mu\left(\bigcup_{n\geq N} E_n\right) \\
&\leq \lim_{N\to\infty} \sum_{n=N}^{\infty} \mu(E_n) \\
&= 0,
\end{aligned}$$

where the last equality follows since $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. $\square$

9.7 Our Measure Equivalence Relation

Let $\mathcal{A}$ be an algebra of subsets of X and let μ be a finite measure on $\mathcal{A}$. Let μ^* be the outer measure on X induced by μ. Define a relation $\sim$ on $\mathcal{P}(X)$ as follows: if $A, B \in \mathcal{P}(X)$, then

$$A \sim B \text{ if and only if } \mu^*(A\Delta B) = 0.$$

We also define the pseudometric d_μ on $\mathcal{P}(X)$ by

$$\mathrm{d}_\mu(A, B) = \mu^*(A\Delta B)$$

for all $A, B \in \mathcal{P}(X)$.

Problem 4.a

Proposition 9.8. *The relation $\sim$ is an equivalence relation.*

Proof. We first check reflexivity. Let $A \in \mathcal{P}(X)$. Then

$$\begin{aligned} \mu^*(A\Delta A) &= \mu^*(\emptyset) \\ &= 0 \end{aligned}$$

implies $A \sim A$. Next we check symmetry. Let $A, B \in \mathcal{P}(X)$ and suppose $A \sim B$. Then

$$\begin{aligned} \mu^*(B\Delta A) &= \mu^*(A\Delta B) \\ &= 0 \end{aligned}$$

implies $B \sim A$. Finally we check transitivity. Let $A, B, C \in \mathcal{P}(X)$ and suppose $A \sim B$ and $B \sim C$. Then

$$\begin{aligned} \mu^*(A\Delta C) &= \mu^*(A\Delta B\Delta B\Delta C) \\ &\leq \mu^*((A\Delta B) \cup (B\Delta C)) \\ &\leq \mu^*(A\Delta B) + \mu^*(B\Delta C) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

implies $A \sim C$. □

Problem 4.b

Proposition 9.9. *Let $A, B \in \mathcal{P}(X)$. If $A \sim B$, then $\mu^*(A) = \mu^*(B)$. The converse need not be true.*

Proof. Suppose that $A \sim B$. Then $\mu^*(A\Delta B) = 0$ implies

$$\begin{aligned} \mu^*(A) &= \mu^*(A) + \mu^*(A\Delta B) \\ &\geq \mu^*(A \cup (A\Delta B)) \\ &\geq \mu^*(A\Delta A\Delta B) \\ &= \mu^*(B). \end{aligned}$$

Similarly,

$$\begin{aligned} \mu^*(B) &= \mu^*(B) + \mu^*(B\Delta A) \\ &\geq \mu^*(B \cup (B\Delta A)) \\ &\geq \mu^*(B\Delta B\Delta A) \\ &= \mu^*(A). \end{aligned}$$

Thus $\mu^*(A) = \mu^*(B)$.

To see that the converse does not hold, consider the case where $X = \{a, b\}$ and μ is counting measure on this set. Then on the one hand, we have

$$\mu(\{a\}) = 1 = \mu(\{b\}),$$

but on the other hand, we have

$$\begin{aligned}\mu(\{a\}\Delta\{b\}) &= \mu(\{a,b\}) \\ &= 2 \\ &\neq 0.\end{aligned}$$

□

9.8 μ^*-Measurable Forms σ-Algebra

Let $\mathcal{A}$ be an algebra of subsets of X and let μ be a finite measure on $\mathcal{A}$. Let μ^* be the outer measure on X induced by μ. A set E is said to be μ^*-measurable if

$$\mu^*(S) \geq \mu^*(S \cap E) + \mu^*(S \backslash E)$$

for all $S \in \mathcal{P}(X)$. Note that by countable subadditivity of μ^*, this implies

$$\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \backslash E).$$

Denote by $\mathcal{M}$ to be the collection of all μ^*-measurable sets.

Problem 6.a

Proposition 9.10. *Let $A \in \mathcal{A}$. Then A is μ^*-measurable.*

Proof. Let $S \in \mathcal{P}(X)$. Assume for a contradiction that

$$\mu^*(S) < \mu^*(S \cap A) + \mu^*(S \backslash A).$$

Choose $\varepsilon > 0$ such that

$$\mu^*(S) < \mu^*(S \cap A) + \mu^*(S \backslash A) - \varepsilon.$$

Choose $B \in \mathcal{A}$ such that $S \subseteq B$ and

$$\mu(B) \leq \mu^*(S) + \varepsilon.$$

Then

$$\begin{aligned}\mu^*(S) &\geq \mu(B) - \varepsilon \\ &= \mu\left((B \cap A) \cup (B \backslash A)\right) - \varepsilon \\ &= \mu(B \cap A) + \mu(B \backslash A) - \varepsilon \\ &\geq \mu^*(S \cap A) + \mu^*(S \backslash A) - \varepsilon.\end{aligned}$$

This is a contradiction. □

Problem 6.b

Proposition 9.11. *$\mathcal{M}$ is a σ-algebra.*

Proof. We prove this in several steps:

Step 1: We first show $\mathcal{M}$ is an algebra. First we show it is closed under finite unions. Let $A, B \in \mathcal{M}$ and let $S \in \mathcal{P}(X)$. Then

$$\begin{aligned}\mu^*(S) &= \mu^*\left(S \cap A\right) + \mu^*\left(S \backslash A\right) \\ &= \mu^*\left(S \cap A\right) + \mu^*\left((S \backslash A) \cap B\right) + \mu^*\left((S \backslash A) \backslash B\right) \\ &\geq \mu^*\left((S \cap A) \cup ((S \backslash A) \cap B)\right) + \mu^*\left((S \backslash A) \backslash B\right) \\ &= \mu^*\left(S \cap (A \cup B)\right) + \mu^*\left(S \backslash (A \cup B)\right)\end{aligned}$$

Therefore $A \cap B \in \mathcal{M}$.

Next we shows it is closed under complements. Let $A \in \mathcal{M}$ and let $S \in \mathcal{P}(X)$. Then

$$
\begin{aligned}
\mu^*(S) &\geq \mu^*(S\cap A)+\mu^*(S\backslash A)\\
&= \mu^*\left(S\backslash(X\backslash A)\right)+\mu^*(S\backslash A)\\
&= \mu^*\left(S\backslash(X\backslash A)\right)+\mu^*\left(S\cap(X\backslash A)\right).
\end{aligned}
$$

Therefore $X\backslash A\in\mathcal{M}$.

Step 2: We show μ^* is finitely additive on $\mathcal{M}$. In fact, we claim that for any $S\in\mathcal{P}(X)$ and pairwise disjoint $A_1,\dots,A_n\in\mathcal{M}$, we have

$$
\mu^*\left(S\cap\left(\bigcup_{m=1}^{n}A_m\right)\right)=\sum_{m=1}^{n}\mu^*\left(S\cap A_m\right). \tag{27}
$$

We prove (27) by induction on n. The equality holds trivially for $n=1$. For the induction step, assume that it holds for some $n\geq 1$. Let S be a subset of X and let$A_1,\dots,A_{n+1}$ be a finite sequence of members in $\mathcal{M}$. Then

$$
\begin{aligned}
\mu^*\left(S\cap\left(\bigcup_{m=1}^{n+1}A_m\right)\right) &\geq \mu^*\left(S\cap\left(\bigcup_{m=1}^{n+1}A_m\right)\cap A_{n+1}\right)+\mu^*\left(S\cap\left(\bigcup_{m=1}^{n+1}A_m\right)\cap(X\backslash A_{n+1})\right)\\
&= \mu^*\left(S\cap A_{n+1}\right)+\mu^*\left(S\cap\left(\bigcup_{m=1}^{n}A_m\right)\right)\\
&= \mu^*\left(S\cap A_{n+1}\right)+\sum_{m=1}^{n}\mu^*(S\cap A_m)\\
&= \sum_{m=1}^{n+1}\mu^*(S\cap A_m).
\end{aligned}
$$

This establishes (27). Setting $S=X$ in (27) gives us finite additivity of μ^* on $\mathcal{M}$.

Step 3: We prove that $\mathcal{M}$ is a σ-algebra. Since $\mathcal{M}$ was already shown to be an algebra, it suffices to show that $\mathcal{M}$ is closed under countable unions. Let (A_n) be a sequence in $\mathcal{M}$. Disjointify the sequence (A_n) to the sequence (D_n): set $D_1=A_1$ and $D_n=A_n\backslash\bigcup_{m=1}^{n-1}A_m$ for all $n>1$. Note that (D_n) is a sequence in $\mathcal{M}$ since $\mathcal{M}$ is algebra. Let $S\in\mathcal{P}(X)$ and $n\in\mathbb{N}$. Observe that

$$
\begin{aligned}
\mu^*(S) &\geq \mu^*\left(S\cap\left(\bigcup_{m=1}^{n}D_m\right)\right)+\mu^*\left(S\backslash\left(\bigcup_{m=1}^{n}D_m\right)\right)\\
&\geq \mu^*\left(S\cap\left(\bigcup_{m=1}^{n}D_m\right)\right)+\mu^*\left(S\backslash\left(\bigcup_{n\in\mathbb{N}}A_n\right)\right)\\
&= \sum_{m=1}^{n}\mu^*\left(S\cap D_m\right)+\mu^*\left(S\backslash\left(\bigcup_{n\in\mathbb{N}}A_n\right)\right),
\end{aligned}
$$

where we applied finite-additivity of μ^* to the first term on the right-hand side and we applied monotonicity of μ^* to the second term on the right-hand side. Taking the limit as $n\to\infty$. We obtain

$$
\begin{aligned}
\mu^*(S) &\geq \sum_{m=1}^{\infty}\mu^*\left(S\cap D_m\right)+\mu^*\left(S\backslash\left(\bigcup_{n\in\mathbb{N}}A_n\right)\right)\\
&\geq \mu^*\left(\bigcup_{n\in\mathbb{N}}\left(S\cap D_m\right)\right)+\mu^*\left(S\backslash\left(\bigcup_{n\in\mathbb{N}}A_n\right)\right)\\
&= \mu^*\left(S\cap\bigcup_{n\in\mathbb{N}}D_m\right)+\mu^*\left(S\backslash\left(\bigcup_{n\in\mathbb{N}}A_n\right)\right)\\
&= \mu^*\left(S\cap\left(\bigcup_{n\in\mathbb{N}}A_m\right)\right)+\mu^*\left(S\backslash\left(\bigcup_{n\in\mathbb{N}}A_n\right)\right),
\end{aligned}
$$

where we applied countable subadditivity of μ^* to the first expression on the right-hand side. Thus $\bigcup_{n\in\mathbb{N}}A_n\in\mathcal{M}$.

□

Problem 6.c

Proposition 9.12. *We have $\sigma(\mathcal{A}) \subseteq \mathcal{M}$.*

Proof. By problem 6.a and 6.b, we see that $\mathcal{M}$ is a σ-algebra which contains $\mathcal{A}$. Since $\sigma(\mathcal{A})$ is the *smallest* σ-algebra which contains $\mathcal{A}$, we must have $\sigma(\mathcal{A}) \subseteq \mathcal{M}$. □

Problem 6.d

Proposition 9.13. *The outer measure μ^* restricted to $\mathcal{M}$ is a measure.*

Proof. In Proposition (9.11), we showed that μ^* is finitely additive on $\mathcal{M}$. We already know that μ^* is already countably subadditive on $\mathcal{M}$. Therefore μ^* is countably additive on $\mathcal{M}$ since

$$\text{finte additivity + countable subadditivity} \; = \; \text{countable additivity}.$$

To see this, let (A_n) be a sequence of pairwise disjoint members of $\mathcal{M}$. By countable subadditivity of μ^*, we have

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

For the reverse inequality, notat that for each $N \in \mathbb{N}$, finite additivity of μ^* imlpies

$$\begin{aligned} \mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\geq \mu^*\left(\bigcup_{n=1}^{N} A_n\right) \\ &= \sum_{n=1}^{N} \mu^*(A_n). \end{aligned}$$

Taking $N \to \infty$ gives us

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

□

Problem 6.e

Proposition 9.14. *Let $E \in \mathcal{M}$ such that $\mu^*(E) = 0$, and let $F \in \mathcal{P}(X)$ such that $F \subseteq E$. Then $F \in \mathcal{M}$.*

Proof. Let $S \in \mathcal{P}(X)$. Then

$$\begin{aligned} \mu^*(S) &\geq \mu^*(S \backslash F) \\ &= \mu^*(S \cap F) + \mu^*(S \backslash F), \end{aligned}$$

where we used the fact that $\mu^*(S \cap F) = 0$ since $S \cap F \subseteq E$ and $\mu^*(E) = 0$. □

More generally:

Proposition 9.15. *Let $E \in \mathcal{P}(X)$ such that $\mu^*(E) = 0$. Then $E \in \mathcal{M}$.*

Proof. Let $S \in \mathcal{P}(X)$. First note that

$$\begin{aligned} 0 &= \mu^*(E) \\ &\geq \mu^*(S \cap E) \end{aligned}$$

implies $\mu^*(S \cap E) = 0$ by monotonicity of μ^*. Therefore

$$\begin{aligned} \mu^*(S) &\geq \mu^*(S \backslash E) \\ &= \mu^*(S \cap E) + \mu^*(S \backslash E). \end{aligned}$$

This implies $E \in \mathcal{M}$. □

Throughout this homework, let $(X, \mathcal{M}, \mu)$ be a measure space.

10 Homework 4

10.1 Characteristic Function Identities

Proposition 10.1. *Let $A, B \in \mathcal{P}(X)$. Then*

1. $1_{A\cap B} = 1_A 1_B$;
2. $1_{A\cup B} = 1_A + 1_B - 1_A 1_B$;
3. $1_{A^c} = 1 - 1_A$;

Proof. 1. Let $x \in X$. If $x \in A \cap B$, then $x \in A$ and $x \in B$, and thus we have

$$\begin{aligned} 1_{A\cap B}(x) &= 1 \\ &= 1 \cdot 1 \\ &= 1_A(x) 1_B(x). \end{aligned}$$

If $x \notin A \cap B$, then either $x \notin A$ or $x \notin B$. Without loss of generality, say $x \notin A$. Then we have

$$\begin{aligned} 1_{A\cap B}(x) &= 0 \\ &= 0 \cdot 1_B(x) \\ &= 1_A(x) 1_B(x). \end{aligned}$$

Therefore the functions $1_{A\cap B}$ and $1_A 1_B$ agree on all of X, and hence must be equal to each other.

2. Let $x \in X$. If $x \in A \cup B$, then either $x \in A$ or $x \in B$. Without loss of generality, say $x \in A$. Then we have

$$\begin{aligned} 1_{A\cup B}(x) &= 1 \\ &= 1 + 1_B(x) - 1_B(x) \\ &= 1 + 1_B(x) - 1 \cdot 1_B(x) \\ &= 1_A(x) + 1_B(x) - 1_A(x) 1_B(x). \end{aligned}$$

If $x \notin A \cup B$, then $x \notin A$ and $x \notin B$. Therefore we have

$$\begin{aligned} 1_{A\cup B}(x) &= 0 \\ &= 0 + 0 - 0 \cdot 0 \\ &= 1_A(x) + 1_B(x) - 1_A(x) 1_B(x). \end{aligned}$$

Thus the functions $1_{A\cup B}$ and $1_A + 1_B - 1_A 1_B$ agree on all of X, and hence must be equal to each other.

3. Let $x \in X$. If $x \in A$, then $x \notin A^c$, hence

$$\begin{aligned} 1_{A^c}(x) &= 0 \\ &= 1 - 1 \\ &= 1 - 1_A(x). \end{aligned}$$

If $x \notin A$, then $x \in A^c$, hence

$$\begin{aligned} 1_{A^c}(x) &= 1 \\ &= 1 - 0 \\ &= 1 - 1_A(x). \end{aligned}$$

Therefore the functions 1_{A^c} and $1 - 1_A$ agree on all of X, and hence must be equal to each other. $\square$

Proposition 10.2. *A product of simple functions is a simple function.*

Proof. Let $\varphi = \sum_{i=1}^{m} a_i 1_{A_i}$ and $\psi = \sum_{j=1}^{n} b_j 1_{B_j}$ be two simple functions. Then

$$\begin{aligned}
\varphi \cdot \psi &= \left(\sum_{i=1}^{m} a_i 1_{A_i}\right)\left(\sum_{j=1}^{n} b_j 1_{B_j}\right) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} (a_i 1_{A_i})(b_j 1_{B_j}) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j (1_{A_i} 1_{B_j}) \\
&= \sum_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} a_i b_j 1_{A_i\cap B_j}.
\end{aligned}$$

Since $A_i \cap B_j \in \mathcal{M}$ for all i, j, it follows that $\varphi \cdot \psi$ is simple. □

10.2 E is Measurable if and only if 1_E is Measurable

Proposition 10.3. *Let $E \in \mathcal{P}(X)$. Then $E \in \mathcal{M}$ if and only if 1_E is measurable.*

Proof. Suppose 1_E is measurable. Then

$$E^c = \{x \in X \mid 1_E(x) < 1\}$$

is measurable. This implies E is measurable.

Conversely, suppose E is measurable. Let $c \in \mathbb{R}$. We have three cases

$$\{x \in X \mid 1_E(x) < c\} = \begin{cases} X & \text{if } 1 < c \\ E^c & \text{if } 0 < c \leq 1 \\ \emptyset & \text{if } c \leq 0 \end{cases}$$

In all three cases, we see that $1_E^{-1}(-\infty, c) \in \mathcal{M}$. This implies 1_E is is measurable. □

10.3 Defining Measure from Integral Weighted by Nonnegative Simple Function

Proposition 10.4. *Let $\phi\colon X \to [0,\infty)$ be a nonnegative simple function. Define a function $\nu\colon \mathcal{M} \to [0,\infty]$ by*

$$\nu(E) = \int_X \phi 1_E \mathrm{d}\mu$$

for all $E \in \mathcal{M}$. Then ν is a measure on $(X, \mathcal{M})$.

Proof. First note that

$$\begin{aligned}
\nu(\emptyset) &= \int_X \phi 1_{\emptyset} \mathrm{d}\mu \\
&= \int_X \phi \cdot 0 \cdot \mathrm{d}\mu \\
&= \int_X 0 \cdot \mathrm{d}\mu \\
&= 0.
\end{aligned}$$

Now we show that ν is finitely additive. Let $(E_n)_{n=1}^N$ be a finite sequence of pairwise disjoint sets in $\mathcal{M}$. Then

$$\begin{aligned}\nu\left(\bigcup_{n=1}^N E_n\right) &= \int_X \phi 1_{\bigcup_{n=1}^N E_n} \mathrm{d}\mu \\ &= \int_X \phi \sum_{n=1}^N 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^N \int_X \phi 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^N \nu(E_n),\end{aligned}$$

where we used the fact that each $\phi 1_{E_n}$ is a nonnegative simple function in order to commute the finite sum with the integral. Thus it follows that ν is finitely additive. It remains to show that ν is countably subadditive. Let (E_n) be a sequence of sets in $\mathcal{M}$. We want to show that

$$\int_X \phi 1_{\bigcup_{n=1}^\infty E_n} \mathrm{d}\mu \leq \sum_{n=1}^\infty \int_X \phi 1_{E_n} \mathrm{d}\mu. \tag{28}$$

To do this, we will show that the sum on the righthand side in (28) is greater than or equal to all integrals of the form $\int \varphi \mathrm{d}\mu$ where $\varphi\colon X \to [0,\infty]$ is a simple function such that $\varphi \leq \phi 1_{\bigcup_{n=1}^\infty E_n}$. Then the inequality (28) will follow from the fact that the integral on the lefthand side in (28) is the supremum of this set. So let $\varphi\colon X \to [0,\infty]$ be a simple function such that $\varphi \leq \phi 1_{\bigcup_{n=1}^\infty E_n}$. Write φ and ϕ in terms of their canonical forms, say

$$\varphi = \sum_{i=1}^k a_i 1_{A_i} \quad \text{and} \quad \phi = \sum_{j=1}^m b_j 1_{B_j}.$$

So $a_i \neq a_{i'}$ and $A_i \cap A_{i'} = \emptyset$ whenever $i \neq i'$ and $b_j \neq b_{j'}$ and $B_j \cap B_{j'} = \emptyset$ whenever $j \neq j'$. Observe that the canonical representation of $\phi 1_{\bigcup_{n=1}^\infty E_n}$ is given by

$$\begin{aligned}\phi 1_{\bigcup_{n=1}^\infty E_n} &= \left(\sum_{j=1}^m b_j 1_{B_j}\right) 1_{\bigcup_{n=1}^\infty E_n} \\ &= \sum_{j=1}^m b_j 1_{B_j} 1_{\bigcup_{n=1}^\infty E_n} \\ &= \sum_{j=1}^m b_j 1_{\bigcup_{n=1}^\infty B_j \cap E_n},\end{aligned}$$

where this representation is the canonical representation since $b_j \neq b_{j'}$ and

$$\left(\bigcup_{n=1}^\infty B_j \cap E_n\right) \cap \left(\bigcup_{n=1}^\infty B_{j'} \cap E_n\right) = \emptyset$$

whenever $j \neq j'$ (since $B_j \cap B_{j'} = \emptyset$). Therefore we have

$$\begin{aligned}
\int_X \varphi \mathrm{d}\mu &\leq \int_X \phi 1_{\bigcup_{n=1}^\infty E_n} \mathrm{d}\mu \\
&= \sum_{j=1}^m b_j \mu \left(\bigcup_{n=1}^\infty B_j \cap E_n \right) \\
&\leq \sum_{j=1}^m b_j \sum_{n=1}^\infty \mu(B_j \cap E_n) \\
&= \sum_{n=1}^\infty \sum_{j=1}^m b_j \mu(B_j \cap E_n) \\
&= \sum_{n=1}^\infty \sum_{j=1}^m \int_X b_j 1_{B_j \cap E_n} \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \sum_{j=1}^m b_j 1_{B_j \cap E_n} \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \sum_{j=1}^m b_j \left(1_{B_j} 1_{E_n} \right) \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \left(\sum_{j=1}^m b_j 1_{B_j} \right) 1_{E_n} \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \phi 1_{E_n} \mathrm{d}\mu,
\end{aligned}$$

where we used monotonicity of integration in the first line and where we used countable subadditivity of μ to get from the second line to the third line.

□

10.4 Equivalent Criterion for Function to be Measurable

Proposition 10.5. *Let $f \colon X \to \mathbb{R}$ be a function. The following are equivalent;*

1. $f^{-1}(-\infty, c) \in \mathcal{M}$ *for all* $c \in \mathbb{R}$.
2. $f^{-1}[c, \infty) \in \mathcal{M}$ *for all* $c \in \mathbb{R}$.
3. $f^{-1}(c, \infty) \in \mathcal{M}$ *for all* $c \in \mathbb{R}$.
4. $f^{-1}(-\infty, c] \in \mathcal{M}$ *for all* $c \in \mathbb{R}$.

Proof.

($1 \implies 2$) Let $c \in \mathbb{R}$. Then

$$\begin{aligned}
\left(f^{-1}[c, \infty) \right)^c &= f^{-1}(-\infty, c) \\
&\in \mathcal{M}.
\end{aligned}$$

It follows that $f^{-1}[c, \infty) \in \mathcal{M}$ for all $c \in \mathbb{R}$.

($2 \implies 3$). Let $c \in \mathbb{R}$. Then

$$\begin{aligned}
f^{-1}(c, \infty) &= \bigcup_{n=1}^\infty f^{-1}[c + 1/n, \infty) \\
&\in \mathcal{M}.
\end{aligned}$$

It follows that $f^{-1}(c, \infty) \in \mathcal{M}$ for all $c \in \mathbb{R}$.

($3 \implies 4$). Let $c \in \mathbb{R}$. Then

$$\begin{aligned}\left(f^{-1}(-\infty, c]\right)^c &= f^{-1}(c, \infty) \\ &\in \mathcal{M}.\end{aligned}$$

It follows that $f^{-1}(-\infty, c] \in \mathcal{M}$ for all $c \in \mathbb{R}$.

($4 \implies 1$). Let $c \in \mathbb{R}$. Then

$$\begin{aligned}f^{-1}(-\infty, c) &= \bigcap_{n=1}^{\infty} f^{-1}(-\infty, c+1/n] \\ &\in \mathcal{M}.\end{aligned}$$

It follows that $f^{-1}(-\infty, c) \in \mathcal{M}$ for all $c \in \mathbb{R}$. □

10.5 Simple Functions are Measurable

Proposition 10.6. *Let $\phi \colon X \to \mathbb{R}$ be a simple function. Then $\phi^{-1}(-\infty, c) \in \mathcal{M}$ for all $c \in \mathbb{R}$.*

Proof. Let $c \in \mathbb{R}$ and express ϕ in terms of its canonical representation, say

$$\phi = \sum_{i=1}^{n} a_i 1_{A_i}.$$

Then

$$\begin{aligned}\phi^{-1}(-\infty, c) &= \bigcup_{i \mid a_i < c} A_i \\ &\in \mathcal{M}.\end{aligned}$$

It follows that $\phi^{-1}(-\infty, c) \in \mathcal{M}$ for all $c \in \mathbb{R}$. □

10.6 Pointwise Convergence of Measurable Functions is Measurable

Proposition 10.7. *Let $(f_n \colon X \to \mathbb{R})$ be a sequence of functions which converges pointwise to a function $f \colon X \to \mathbb{R}$. Then*

$$f^{-1}(-\infty, c] = \bigcap_{k=1}^{\infty} \liminf_{n \to \infty} f_n^{-1}(-\infty, c+1/k).$$

Proof. We have

$$\begin{aligned}x \in \bigcap_{k=1}^{\infty} \liminf_{n \to \infty} f_n^{-1}(-\infty, c+1/k) &\iff x \in \liminf_{n \to \infty} f_n^{-1}(-\infty, c+1/k) \text{ for all } k \\ &\iff x \in f_{\pi_k(n)}^{-1}(-\infty, c+1/k) \text{ for all } k \text{ for some subsequence } (\pi_k(n))_{n \in \mathbb{N}} \text{ of } (n)_{n \in \mathbb{N}} \\ &\iff f_{\pi_k(n)}(x) < c+1/k \text{ for all } k \text{ for some subsequence } (\pi_k(n))_{n \in \mathbb{N}} \text{ of } (n)_{n \in \mathbb{N}} \\ &\iff f(x) \leq c+1/k \text{ for all } k \\ &\iff f(x) \leq c \\ &\iff x \in f^{-1}(-\infty, c]\end{aligned}$$

where we obtained the fourth line from the third line since a subsequence of a convergent sequence must converge to the same limit (so $f_{\pi_k(n)}(x) \to f(x)$ as $n \to \infty$). □

Problem 5.c

Proposition 10.8. *Let $(\phi_n : X \to \mathbb{R})$ be a sequence of simple functions which converges pointwise to a function $f : X \to \mathbb{R}$. Then $f^{-1}(-\infty, c) \in \mathcal{M}$ for all $c \in \mathbb{R}$.*

Proof. Let $c \in \mathbb{R}$. Then by 5.a and 5.b, and the fact that $\mathcal{M}$ is closed under taking intersections and liminf, we have

$$\begin{aligned} f^{-1}(-\infty, c] &= \bigcap_{k=1}^{\infty} \liminf_{n\to\infty} \phi_n^{-1}(-\infty, c+1/k) \\ &\in \mathcal{M}. \end{aligned}$$

Thus $f^{-1}(-\infty, c] \in \mathcal{M}$ for all $c \in \mathbb{R}$. It follows from Proposition (10.5) that $f^{-1}(-\infty, c) \in \mathcal{M}$ for all $c \in \mathbb{R}$. □

10.7 Max, Min, Addition, and Scalar Multiplication of Measurable Functions is Measurable

Proposition 10.9. *Let $f, g : X \to [0, \infty]$ be two nonnegative measurable functions and let $a \geq 0$. Then af, $f+g$, $\max\{f, g\}$, and $\min\{f, g\}$ are all measurable functions.*

Proof. Choose an increasing sequence of nonnegative simple functions (φ_n) which converges pointwise to f and choose an increasing sequence of nonnegative simple functions (ψ_n) which converges pointwise to g. For each $x \in X$, we have

$$\begin{aligned} \lim_{n\to\infty} (c\varphi_n(x)) &= c \lim_{n\to\infty} \varphi_n(x) \\ &= cf(x), \end{aligned}$$

and

$$\begin{aligned} \lim_{n\to\infty} (\varphi_n(x) + \psi_n(x)) &= \lim_{n\to\infty} \varphi_n(x) + \lim_{n\to\infty} \psi_n(x) \\ &= f(x) + g(x), \end{aligned}$$

and

$$\begin{aligned} \lim_{n\to\infty} (\varphi_n(x)\psi_n(x)) &= \lim_{n\to\infty} \varphi_n(x) \lim_{n\to\infty} \psi_n(x) \\ &= f(x)g(x). \end{aligned}$$

It follows that $(c\varphi_n)$, $(\varphi_n + \psi_n)$, and $(\varphi_n\psi_n)$ are increasing functions of simple functions which converges pointwise to cf, $f+g$, and fg respectively. Therefore cf, $f+g$, and fg are measurable functions.

It remains to show that $\max\{f, g\}$ and $\min\{f, g\}$ are measurable. First note that $\max\{\varphi, \psi\}$ and $\min\{\varphi, \psi\}$ are simple functions for any two simple functions φ and ψ. Indeed, if

$$\varphi = \sum_{i=1}^{m} a_i 1_{A_i} \quad \text{and} \quad \psi = \sum_{j=1}^{n} b_j 1_{B_j},$$

then

$$\max\{\varphi, \psi\} = \sum_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} \max\{a_i, b_j\} 1_{A_i\cap B_j} \quad \text{and} \quad \min\{\varphi, \psi\} = \sum_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} \min\{a_i, b_j\} 1_{A_i\cap B_j}$$

Thus $(\max\{\varphi_n, \psi_n\})$ and $(\min\{\varphi_n, \psi_n\})$ are both sequences of simple functions. They are also increasing sequences since both (φ_n) and (ψ_n) are increasing. We will show that they converge to $\max\{f, g\}$ and $\min\{f, g\}$ respectively.

Let $x \in X$ be arbitrary and assume without loss of generality that $f(x) \geq g(x)$. If $f(x) = g(x)$ then both $\varphi_n(x) \to f(x)$ and $\psi_n(x) \to f(x)$, and so clearly $\max\{\varphi_n(x), \psi_n(x)\} \to f(x)$ and $\min\{\varphi_n(x), \psi_n(x)\} \to f(x)$, so assume $f(x) > g(x)$. Set $\varepsilon = f(x) - g(x)$ and choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$|f(x) - \varphi_n(x)| < \frac{\varepsilon}{2} \quad \text{and} \quad |g(x) - \psi_n(x)| < \frac{\varepsilon}{2}.$$

Then $n \geq N$ implies $\varphi_n(x) > \psi_n(x)$. Therefore $n \geq N$ implies

$$\begin{aligned} |\max\{f(x), g(x)\} - \max\{\varphi_n(x), \psi_n(x)\}| &= |f(x) - \varphi_n(x)| \\ &< \frac{\varepsilon}{2} \\ &< \varepsilon \end{aligned}$$

and $n \geq N$ implies

$$\begin{aligned} |\min\{f(x), g(x)\} - \min\{\varphi_n(x), \psi_n(x)\}| &= |g(x) - \psi_n(x)| \\ &< \frac{\varepsilon}{2} \\ &< \varepsilon. \end{aligned}$$

Therefore $\max\{\varphi_n(x), \psi_n(x)\} \to f(x)$ and $\min\{\varphi_n(x), \psi_n(x)\} \to g(x)$. Since x was arbitrary, it follows that the sequences $(\max\{\varphi_n, \psi_n\})$ and $(\min\{\varphi_n, \psi_n\})$ converges pointwise to $\max\{f, g\}$ and $\min\{f, g\}$ respectively. □

11 Homework 5

11.1 Criterion for Function to be Measurable Using Rational Numbers

Proposition 11.1. *Let $f\colon X \to \mathbb{R}$ be a function. Then f is measurable if and only if for every $q \in \mathbb{Q}$ the set $f^{-1}(-\infty, q)$ is measurable.*

Proof. If f is measurable, then certainly $f^{-1}(-\infty, c) \in \mathcal{M}$ for any $c \in \mathbb{R}$ (and hence for any $c \in \mathbb{Q}$). Conversely, suppose $f^{-1}(-\infty, q) \in \mathcal{M}$ for any $q \in \mathbb{Q}$. Let $c \in \mathbb{R}$. For each $n \in \mathbb{N}$ choose $q_n \in \mathbb{Q}$ such that

$$c < q_n < c + \frac{1}{n}.$$

Such a choice for each n can be made since $\mathbb{Q}$ is dense in $\mathbb{R}$. We claim that

$$f^{-1}(-\infty, c] = \bigcap_{n=1}^{\infty} f^{-1}(-\infty, q_n)$$

To see this, first note that the inclusion

$$f^{-1}(-\infty, c] \subseteq \bigcap_{n=1}^{\infty} f^{-1}(-\infty, q_n)$$

is clear since each $f^{-1}(-\infty, q_n)$ contains $f^{-1}(-\infty, c)$ (as $c < q_n$). For the reverse inclusion, suppose $x \in \bigcap_{n=1}^{\infty} f^{-1}(-\infty, q_n)$, so $f(x) < q_n$ for all n. Since $q_n \to c$, this implies $f(x) \leq c$. Thus $x \in f^{-1}(-\infty, c]$. It follows that f is measurable. □

Remark 10. Note that we needed to use the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$ in order to prove this.

11.2 Alternative Definition for Measurability of Function

Before we answer this problem, we give a more general definition of what it means for a function to be measurable with respect to σ-algebras $\mathcal{M}$ and $\mathcal{N}$. Then we show that this more general definition is equivalent to the definition we've been using when $\mathcal{N}$ is the Borel σ-algebra on $\mathbb{R}$.

Definition 11.1. Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces and let $f\colon X \to Y$ be a function. We say f is **measurable with respect to $\mathcal{M}$ and $\mathcal{N}$** if $f^{-1}(\mathcal{N}) \subseteq \mathcal{M}$ where

$$f^{-1}(\mathcal{N}) = \{f^{-1}(B) \mid B \in \mathcal{N}\}.$$

In other words, f is measurable with respect to $\mathcal{M}$ and $\mathcal{N}$ if for all $B \in \mathcal{N}$ we have

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\} \in \mathcal{M}.$$

If $\mathcal{M} = \mathcal{N}$, then we will just say f is measurable with respect to $\mathcal{M}$. If the σ-algebras $\mathcal{M}$ and $\mathcal{N}$ are clear from context, then we will just say f is measurable.

Let us now show that when $Y = \mathbb{R}$ and $\mathcal{N} = \mathcal{B}(\mathbb{R})$, that this definition is equivalent to the definition we gave in class. We first prove the following two propositions:

Proposition 11.2. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces and let $f\colon X \to Y$ be a function. Suppose that $\mathcal{N}$ is generated as a σ-algebra by the collection $\mathcal{C}$ of subsets of Y. Then $f^{-1}(\mathcal{N}) \subseteq \mathcal{M}$ if and only if $f^{-1}(\mathcal{C}) \subseteq \mathcal{M}$.*

Proof. One direction is clear, so we just prove the other direction. Suppose $f^{-1}(\mathcal{C}) \subseteq \mathcal{M}$. Observe that

$$\{B \in \mathcal{P}(Y) \mid f^{-1}(B) \in \mathcal{M}\}$$

is a σ-algebra which contains $\mathcal{C}$. Indeed, it is a σ-algebra since f^{-1} maps the emptyset set to the emptyset and maps the whole space Y to the whole space X, and since f^{-1} commutes with unions and complements. Furthemore, this σ-algebra contains $\mathcal{C}$ since $f^{-1}(\mathcal{C}) \subseteq \mathcal{M}$. Since $\mathcal{N}$ is the *smallest* σ-algebra which contains $\mathcal{C}$, it follows that

$$\mathcal{N} \subseteq \{B \in \mathcal{P}(Y) \mid f^{-1}(B) \in \mathcal{M}\}.$$

In particular, if $B \in \mathcal{N}$, then $f^{-1}(B) \in \mathcal{M}$. Thus f is measurable. □

Proposition 11.3. *Let $\mathcal{C} = \{(-\infty, c) \mid c \in \mathbb{R}\}$. Then $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{C})$.*

Proof. Let $\mathcal{I}_n$ be the collection of all subintervals of $[n, n+1)$ and let $\mathcal{B}_n = \sigma(\mathcal{I}_n)$. So

$$\mathcal{B}(\mathbb{R}) = \{E \subseteq \mathbb{R} \mid E \cap [n, n+1) \in \mathcal{B}_n \text{ for all } n \in \mathbb{Z}\}.$$

Let $c \in \mathbb{R}$. Then since $(-\infty, c) \cap [n, n+1]$ is a subinterval of $[n, n+1]$ for all $n \in \mathbb{Z}$, it follows that $(-\infty, c) \in \mathcal{B}$ for all $n \in \mathbb{Z}$. Thus $\mathcal{C} \subseteq \mathcal{B}$ which implies $\sigma(\mathcal{C}) \subseteq \mathcal{B}$ (as $\sigma(\mathcal{C})$ is the *smallest* σ-algebra which contains $\mathcal{C}$). Conversely, note that $\sigma(\mathcal{C})$ contains all subintervals of $[n, n+1)$ for all $n \in \mathbb{Z}$. Thus $\sigma(\mathcal{C}) \supseteq \mathcal{B}_n$ for all $n \in \mathbb{Z}$ (as $\mathcal{B}_n$ is the *smallest* σ-algebra which contains all subintervals of $[n, n+1)$. Since $\mathcal{B}(\mathbb{R}) = \sigma\left(\bigcup_{n \in \mathbb{Z}} \mathcal{B}_n\right)$, it follows that $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C})$. □

Corollary 4. *Let $(X, \mathcal{M})$ be a measurable space and let $f\colon X \to \mathbb{R}$ be a function. Then f is measurable with respect to $\mathcal{M}$ and $\mathcal{B}(\mathbb{R})$ if and only if $f^{-1}(-\infty, c) \in \mathcal{M}$ for all $c \in \mathbb{R}$.*

Proof. Follows from Proposition (11.3) and Proposition (11.2). □

Corollary 5. *Let $(X, \mathcal{M})$ be a measurable space and let $\mathcal{B}(\mathbb{R})$ be the Borel σ-algebra on $\mathbb{R}$. Suppose that $\mathcal{C} \subseteq \mathcal{B}(\mathbb{R})$ is a collection of sets in $\mathcal{B}(\mathbb{R})$ such that $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C})$. Then $f\colon X \to \mathbb{R}$ is measurable with respect to $\mathcal{M}$ and $\mathcal{B}(\mathbb{R})$ if and only if $f^{-1}(C) \in \mathcal{M}$ for all $C \in \mathcal{C}$.*

Proof. Follows from Proposition (11.2) and from Corollary (4). □

11.3 Monotone Increasing Function and Continuous Function are Borel Measurable

Problem 3.i

Proposition 11.4. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a continuous function. Then f is measurable with respect to $\mathcal{B}(\mathbb{R})$.*

Proof. For each $q, r \in \mathbb{Q}$ and $n \in \mathbb{N}$, let

$$\mathrm{B}_{1/n}(q) = \{x \in \mathbb{R} \mid |x - q| < 1/n\}$$

Then the collection

$$\mathscr{B} = \{\mathrm{B}_{1/n}(q) \mid n \in \mathbb{N} \text{ and } q \in \mathbb{Q}\}$$

forms a countable basis for the usual topology on $\mathbb{R}$. In particular, if U be an open subset of $\mathbb{R}$, then we can express U as a union of the form

$$U = \bigcup_{\lambda \in \Lambda} B_\lambda$$

where $B_\lambda \in \mathscr{B}$ and where the index set Λ is *countable*. In particular, it follows that $\tau(\mathscr{B}) \subseteq \mathcal{B}(\mathbb{R})$, where $\tau(\mathscr{B})$ is the usual Euclidean topology on $\mathbb{R}$. Thus since $(-\infty, c)$ is an open subset of $\mathbb{R}$ for any $c \in \mathbb{R}$, it follows that $f^{-1}(-\infty, c)$ can be expressed a countable union of open subsets of $\mathbb{R}$ (by definition of what it means to be continuous, the inverse image of an open set under f is open). Since every open subset of $\mathbb{R}$ is Borel measurable, it follows that $f^{-1}(-\infty, c) \in \mathcal{B}(\mathbb{R})$. Thus f is measurable with respect to $\mathcal{B}(\mathbb{R})$. □

Problem 3.ii

Proposition 11.5. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a monotone increasing function. Then f is $\mathcal{B}(\mathbb{R})$-measurable.*

Proof. Let $c \in \mathbb{R}$. We want to show that $f^{-1}(-\infty, c) \in \mathcal{B}(\mathbb{R})$. If $f^{-1}(-\infty, c) = \emptyset$ or $f^{-1}(-\infty, c) = \mathbb{R}$, then we are done, so assume $f^{-1}(-\infty, c) \neq \emptyset$ and $f^{-1}(-\infty, c) \neq \mathbb{R}$. Choose $y \in \mathbb{R}$ such that $c \leq f(y)$. Observe that if $x \in f^{-1}(-\infty, c)$, then

$$f(x) < c \leq f(y),$$

which implies $x \leq y$ since f is monotone increasing. Thus y is an upper bound of the set $f^{-1}(-\infty, c)$. Since $f^{-1}(-\infty, c)$ is nonempty and bounded above, it follows that its supremum exists. Denote its supremum by y_0. So

$$y_0 = \sup\{x \in \mathbb{R} \mid f(x) < c\}.$$

We claim that

$$f^{-1}(-\infty, c) = \begin{cases} (-\infty, y_0) & \text{if } f(y_0) \geq c \\ (-\infty, y_0] & \text{if } f(y_0) < c. \end{cases}$$

Indeed, since y_0 is an upper bound $f^{-1}(-\infty, c)$, it must be greater than or equal to all elements in $f^{-1}(-\infty, c)$. In other words, if $x \in f^{-1}(-\infty, c)$, then $x \leq y_0$. Thus

$$f^{-1}(-\infty, c) \subseteq \begin{cases} (-\infty, y_0) & \text{if } f(y_0) \geq c \\ (-\infty, y_0] & \text{if } f(y_0) < c. \end{cases}$$

Conversely, suppose $x \in (-\infty, y_0)$, so $x < y_0$. Then x is not an upper bound of the set $f^{-1}(-\infty, c)$ (since y_0 is the *least* upper bound), which means that there exists an $x' \in \mathbb{R}$ such that $x \leq x'$ and $f(x') < c$. But since $f(x) \leq f(x')$, this implies $f(x) < c$, and hence $x \in f^{-1}(-\infty, c)$. Thus

$$f^{-1}(-\infty, c) \supseteq \begin{cases} (-\infty, y_0) & \text{if } f(y_0) \geq c \\ (-\infty, y_0] & \text{if } f(y_0) < c. \end{cases}$$

In any case, we see that $f^{-1}(-\infty, c) \in \mathcal{B}(\mathbb{R})$. □

11.4 Sum of Nonnegative Measurable Functions Commutes with Integral

Proposition 11.6. *Let $(X, \mathcal{M}, \mu)$ be a measure space. Suppose $(f_n\colon X \to [0, \infty])$ is a sequence of nonnegative measurable functions. Define $f\colon X \to [0, \infty]$ by*

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

for all $x \in X$. Then f is a nonnegative measurable function and

$$\int_X f \mathrm{d}\mu = \sum_{n=1}^{\infty} \int_X f_n \mathrm{d}\mu.$$

Proof. For each $N \in \mathbb{N}$, let $s_N = \sum_{n=1}^{N} f_n$. Then s_N converges pointwise to f since

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty} f_n(x) \\ &= \lim_{N \to \infty} \sum_{n=1}^{N} f_n(x) \\ &= \lim_{N \to \infty} s_N(x) \end{aligned}$$

for all $x \in X$. Each s_N is a nonnegative measurable function since it is a finite sum of nonnegative measurable functions, and so (s_N) is a sequence of nonnegative functions which converges pointwise to f. This implies f is a nonnegative measurable function. Furthermore, s_N is an increasing sequence since if $M \leq N$, then

$$\begin{aligned} s_M(x) &= \sum_{n=1}^{M} f_n(x) \\ &\leq \sum_{n=1}^{N} f_n(x) \\ &= s_N(x) \end{aligned}$$

for all $x \in X$, where the inequality follows from the fact that each f_n is nonnegative. Therefore we may apply the Monotone Convergence Theorem to obtain

$$\begin{aligned} \int_X f \mathrm{d}\mu &= \lim_{N\to\infty} \int_X s_N \mathrm{d}\mu \\ &= \lim_{N\to\infty} \int_X \sum_{n=1}^{N} f_n \mathrm{d}\mu \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \int_X f_n \mathrm{d}\mu \\ &= \sum_{n=1}^{\infty} \int_X f_n \mathrm{d}\mu, \end{aligned}$$

where we obtained the third line from the fourth line from the fact that this is a finite sum. $\square$

11.5 Defining Measure From Integral Weighted by Nonnegative Measurable Function

Proposition 11.7. *Let $(X, \mathcal{M}, \mu)$ be measure space and let $g\colon X \to [0, \infty)$ be a nonnegative measurable function. Define $\nu_g\colon \mathcal{M} \to [0, \infty]$ by*

$$\nu_g(E) = \int_X g 1_E \mathrm{d}\mu$$

for all $E \in \mathcal{M}$. Then ν is a measure on $(X, \mathcal{M})$.

Proof. First note that

$$\begin{aligned} \nu_g(\emptyset) &= \int_X g 1_\emptyset \mathrm{d}\mu \\ &= \int_X g \cdot 0 \cdot \mathrm{d}\mu \\ &= \int_X 0 \cdot \mathrm{d}\mu \\ &= 0. \end{aligned}$$

Next we show that ν_g is finitely additive. Let $(E_n)_{n=1}^N$ be a finite sequence of pairwise disjoint sets in $\mathcal{M}$. Then

$$\begin{aligned} \nu_g\left(\bigcup_{n=1}^{N} E_n\right) &= \int_X g 1_{\bigcup_{n=1}^{N} E_n} \mathrm{d}\mu \\ &= \int_X g \sum_{n=1}^{N} 1_{E_n} \mathrm{d}\mu \\ &= \int_X \sum_{n=1}^{N} g 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^{N} \int_X g 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^{N} \nu_g(E_n), \end{aligned}$$

where we used the fact that each $g1_{E_n}$ is a nonnegative measurable function in order to commute the finite sum with the integral. Thus it follows that ν_g is finitely additive.

It remains to show that ν_g is countably subadditive. In problem 3 in HW4, we showed that for any nonnegative simple function $\varphi\colon X \to [0,\infty)$, the function $\nu_\varphi\colon \mathcal{M} \to [0,\infty]$ defined by

$$\nu_\varphi(E) = \int_X \varphi 1_E \mathrm{d}\mu$$

is a measure. We will use this fact in what follows:

Choose an increasing sequence $(\varphi_n\colon X \to [0,\infty])$ of nonnegative simple functions which converges pointwise to g (we can do this since g is a nonnegative measurable function). Then $(\varphi_n 1_E)$ is an increasing sequence of nonnegative simple functions which converges pointwise to $g1_E$ for all $E \in \mathcal{M}$. It follows from the Monotone Convergence Theorem that

$$\begin{aligned}\nu_{\varphi_n}(E) &= \int_X \varphi_n 1_E \mathrm{d}\mu \\ &\to \int_X g 1_E \mathrm{d}\mu \\ &= \nu_g(E)\end{aligned}$$

for any $E \in \mathcal{M}$. In particular, for any $E \in \mathcal{M}$ and $\varepsilon > 0$, we can find a $N_{E,\varepsilon} \in \mathbb{N}$ (which depends on E and ε) such that

$$\nu_g(E) < \nu_{\varphi_n}(E) + \varepsilon \tag{29}$$

for all $n \geq N_{E,\varepsilon}$. However we will don't need estimate $\nu_g(E)$ to this level of precision. We just need to know thta for any $\varepsilon > 0$, we can find an $n \in \mathbb{N}$ such that (29) holds.

Now let (E_k) be a sequence of measurable sets and let $\varepsilon > 0$. Choose $n \in \mathbb{N}$ such that

$$\nu_g\left(\bigcup_{k=1}^\infty E_k\right) < \nu_{\varphi_n}\left(\bigcup_{k=1}^\infty E_k\right) + \varepsilon$$

Then we have

$$\begin{aligned}\nu_g\left(\bigcup_{k=1}^\infty E_k\right) &\leq \nu_{\varphi_n}\left(\bigcup_{k=1}^\infty E_k\right) + \varepsilon \\ &\leq \sum_{k=1}^\infty \nu_{\varphi_n}(E_k) + \varepsilon \\ &\leq \sum_{k=1}^\infty \nu_g(E_k) + \varepsilon\end{aligned}$$

where we obtained the second line from the first line using countable subadditivity of ν_{φ_n}, and where we obtained the third line from the second line from the fact that $\varphi_n 1_{E_k} \leq g 1_{E_k}$ for all k and from monotonicity of integration. Taking $\varepsilon \to 0$ gives us countable subadditivity of ν_g. □

11.6 Decreasing Version of MCT

Proposition 11.8. *Let $(X, \mathcal{M}, \mu)$ be a measure space. Suppose $(f_n\colon X \to [0,\infty])$ is a decreasing sequence of nonnegative measurable functions which converges pointwise to $f\colon X \to [0,\infty]$. If $\int_X f_1 \mathrm{d}\mu < \infty$, then*

$$\lim_{n\to\infty} \int_X f_n \mathrm{d}\mu = \int_X f \mathrm{d}\mu. \tag{30}$$

Proof. For each $n \in \mathbb{N}$, set $g_n = f_n - f_{n+1}$. Since (f_n) is a decreasing sequence, each g_n is nonnegative. Furthermore each g_n is measurable since it is a difference of two measurable functions. Set $g = \sum_{n=1}^\infty g_n$ and observe that

$$\begin{aligned}
g &= \sum_{n=1}^{\infty} g_n \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} g_n \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} (f_n - f_{n+1}) \\
&= \lim_{N\to\infty} (f_1 - f_{N+1}) \\
&= f_1 - f.
\end{aligned}$$

It follows from the monotone convergence theorem that

$$\begin{aligned}
\int_X f_1 \mathrm{d}\mu - \int_X f \mathrm{d}\mu &= \int_X (f_1 - f) \mathrm{d}\mu \\
&= \int_X g \mathrm{d}\mu \\
&= \sum_{n=1}^{\infty} \int_X g_n \mathrm{d}\mu && \text{MCT} \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_X g_n \mathrm{d}\mu \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_X (f_n - f_{n+1}) \mathrm{d}\mu \\
&= \lim_{N\to\infty} \int_X \sum_{n=1}^{N} (f_n - f_{n+1}) \mathrm{d}\mu \\
&= \lim_{N\to\infty} \int_X (f_1 - f_{N+1}) \mathrm{d}\mu \\
&= \int_X f_1 \mathrm{d}\mu - \lim_{N\to\infty} \int_X f_{N+1} \mathrm{d}\mu.
\end{aligned}$$

Since $\int_X f_1 \mathrm{d}\mu < 0$, we can cancel it from both sides to get (30). □

11.7 Generalized Fatou's Lemma

Proposition 11.9. *Fatou's Lemma remains valid if the hypothesis that all $f_n \colon X \to [0, \infty]$ are nonnegative measurable functions is replaced by the hypothesis that $f_n \colon X \to \mathbb{R}$ are measurable and there exists a nonnegative integrable function $g \colon X \to [0, \infty]$ such that $-g \leq f_n$ pointwise for all $n \in \mathbb{N}$.*

Proof. Observe that $(g + f_n)$ is a sequence of nonnegative measurable functions which converges pointwise to the nonnegative measurable function $g+f$. Then it follows from Fatou's Lemma that

$$\begin{aligned}\int_X g\mathrm{d}\mu + \int_X f\mathrm{d}\mu &= \int_X g\mathrm{d}\mu + \int_X ((g+f)-g)\mathrm{d}\mu \\ &= \int_X g\mathrm{d}\mu + \int_X (g+f)\mathrm{d}\mu - \int_X g\mathrm{d}\mu \\ &= \int_X (g+f)\mathrm{d}\mu \\ &\leq \liminf_{n\to\infty} \int_X (g+f_n)\mathrm{d}\mu \\ &= \liminf_{n\to\infty} (\int_X g\mathrm{d}\mu + \int_X (g+f_n)\mathrm{d}\mu - \int_X g\mathrm{d}\mu) \\ &= \liminf_{n\to\infty} (\int_X g\mathrm{d}\mu + \int_X ((g+f_n)-g)\mathrm{d}\mu) \\ &= \liminf_{n\to\infty} (\int_X g\mathrm{d}\mu + \int_X f_n\mathrm{d}\mu) \\ &= \int_X g\mathrm{d}\mu + \liminf_{n\to\infty} \int_X f_n\mathrm{d}\mu.\end{aligned}$$

Since $\int_X g\mathrm{d}\mu < \infty$, we can cancel it from both sides to obtain

$$\int_X f\mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n\mathrm{d}\mu.$$

□

11.8 Integral Computations (Using DCT and Decreasing MCT)

Exercise 1. Compute the following integrals

$$\lim_{n\to\infty} \int_0^\infty \frac{\sin(x/n)}{(1+x/n)^n}\mathrm{d}x \tag{31}$$

$$\lim_{n\to\infty} \int_0^1 \frac{1+nx^2}{(1+x^2)^n}\mathrm{d}x \tag{32}$$

Solution 1. We first compute (31). For each $n \in \mathbb{N}$, let $f_n = \sin(x/n)(1+x/n)^{-n}$. Observe that each f_n is measurable since each f_n is continuous (the denominator is only zero when $x = -n$). Let us check that each f_n is integrable: we have

$$\begin{aligned}\int_0^\infty |f_n|\mathrm{d}x &= \int_0^\infty \left|\frac{\sin(x/n)}{(1+x/n)^n}\right| \\ &\leq \int_0^\infty \left|(1+x/n)^{-n}\right| \mathrm{d}x \\ &= \int_0^\infty (1+x/n)^{-n}\mathrm{d}x \\ &\leq \int_0^\infty e^{-x}\mathrm{d}x \\ &= 1.\end{aligned}$$

for all $n \in \mathbb{N}$. Thus each f_n is integrable.

Next we observe that f_n converges pointwise to 0 since $(1+x/n)^n \to e^x$ and $\sin(x/n) \to 0$ as $n \to \infty$ for all $x \in \mathbb{R}$. Finally, note that $e^{-x} \geq |f_n|$ pointwise and e^{-x} is integrable ($\int_0^\infty |e^{-x}|\mathrm{d}x = 1$). It follows from the Lebesgue Dominated Convergence Theorem that

$$\lim_{n\to\infty} \int_0^\infty \frac{\sin(x/n)}{(1+x/n)^n}\mathrm{d}x = \int_0^\infty 0\mathrm{d}x = 0.$$

Next we compute (32). For each $n \in \mathbb{N}$, let $f_n = (1+nx^2)(1+x^2)^{-n}$. Observe that each f_n is measurable since each f_n is continuous (the denominator is never zero for any $x \in \mathbb{R}$). Furthermore, each f_n is nonnegative (since taking squares makes everything nonnegative). We claim that f_n is a decreasing sequence. Indeed

$$\begin{aligned}
\frac{f_n}{f_{n+1}} &= \left(\frac{1+nx^2}{(1+x^2)^n}\right)\left(\frac{(1+x^2)^{n+1}}{1+(n+1)x^2}\right) \\
&= \frac{(1+nx^2)(1+x^2)}{1+(n+1)x^2} \\
&= \frac{nx^4+(n+1)x^2+1}{(n+1)x^2+1} \\
&\geq \frac{(n+1)x^2+1}{(n+1)x^2+1} \\
&= 1.
\end{aligned}$$

Thus (f_n) is a decreasing sequence which is bounded below, so it must converge pointwise to some function. For $x = 0$, it's easy to see that $f_n(0) \to 0$. For $x \neq 0$, we use L'Hopital's rule to get

$$\begin{aligned}
\lim_{n\to\infty} f_n &= \lim_{n\to\infty} \frac{1+nx^2}{(1+x^2)^n} \\
&= \lim_{n\to\infty} \frac{x^2}{\ln(1+x^2)(1+x^2)^n} \\
&= 0.
\end{aligned}$$

Thus (f_n) converges pointwise to 0. Since

$$\begin{aligned}
\int_0^1 f_1 \mathrm{d}x &= \int_0^1 \frac{1+x^2}{1+x^2} \mathrm{d}x \\
&= \int_0^1 \mathrm{d}x \\
&= 1 \\
&< \infty,
\end{aligned}$$

it follows from problem 6 that

$$\begin{aligned}
\lim_{n\to\infty} \int_0^1 \frac{1+nx^2}{(1+x^2)^n} \mathrm{d}x &= \lim_{n\to\infty} \int_0^1 f_n \mathrm{d}x \\
&= \int_0^1 0 \mathrm{d}x \\
&= 0.
\end{aligned}$$

12 Homework 6

12.1 Necessary and Sufficient Condition For Integrable Function Being Zero Almost Everywhere

Proposition 12.1. *Let $f \in L^1(X, \mathcal{M}, \mu)$ and suppose that $\int_X f 1_E \mathrm{d}\mu = 0$ for every $E \in \mathcal{M}$. Then $f = 0$ almost everywhere.*

Proof. Let $A^+ = \{f^+ \neq 0\}$ and $A^- = \{f^- \neq 0\}$. Then A^+ and A^- are measurable sets since f^+ and f^- are measurable functions. Since f agrees with f^+ on A^+, we have

$$\begin{aligned}
\int_X f^+ \mathrm{d}\mu &= \int_X f^+ 1_{A^+} \mathrm{d}\mu \\
&= \int_X f 1_{A^+} \mathrm{d}\mu \\
&= 0.
\end{aligned}$$

Similarly, since $-f$ agrees with f^- on A^-, we have

$$\begin{aligned}\int_X f^- \mathrm{d}\mu &= \int_X f^- 1_{A^-} \mathrm{d}\mu \\ &= \int_X -f 1_{A^-} \mathrm{d}\mu \\ &= -\int_X f 1_{A^-} \mathrm{d}\mu \\ &= 0.\end{aligned}$$

It follows that

$$\begin{aligned}\int_X |f| \mathrm{d}\mu &= \int_X (f^+ + f^-) \mathrm{d}\mu \\ &= \int_X f^+ \mathrm{d}\mu + \int_X f^- \mathrm{d}\mu \\ &= 0.\end{aligned}$$

Thus $f = 0$ almost everywhere (by a proposition proved in class). □

12.2 Integrable Function Takes Value Infinity on a Set of Measure Zero

Proposition 12.2. *Let $f : X \to [0, \infty]$ be a nonnegative measurable function such that $\int_X f \mathrm{d}\mu < \infty$. Then*

1. $\mu(\{f = \infty\}) = 0$;
2. *f does not need to be bounded almost everywhere.*

Proof. 1. Assume for a contradiction that $\mu(\{f = \infty\}) > 0$. Then for any $M \in \mathbb{N}$, we have

$$M 1_{\{f=\infty\}} \leq M f.$$

Therefore

$$\begin{aligned}\infty &> \int_X f \mathrm{d}\mu \\ &\geq \int_X M 1_{\{f=\infty\}} \mathrm{d}\mu \\ &= M\mu(\{f = \infty\}).\end{aligned}$$

Taking $M \to \infty$ gives us a contradiction.

2. To see that f does not need to be bounded, consider $X = [0, 1]$ and $f(x) = x^{-1/2}$. Then

$$\int_0^1 x^{-1/2} \mathrm{d}x = 2,$$

but f is not bounded almost everywhere. Indeed, for any $M \in \mathbb{N}$, the set $[0, 1/M^2]$ has nonzero measure and $f|_{[0,1/M^2]} \geq M$. □

12.3 $L^p(X, \mathcal{M}, \mu)$ is the Completion of Space of Simple Functions

Problem 3.a

Lemma 12.1. *Let (X, d) be a metric space and let (x_n) be a Cauchy sequence in X. Suppose there exists a subsequence $(x_{\pi(n)})$ of the sequence (x_n) such that $x_{\pi(n)} \to x$ for some $x \in X$. Then $x_n \to x$.*

Proof. Let $\varepsilon > 0$. Since $(x_{\pi(n)})$ is convergent, there exists an $N \in \mathbb{N}$ such that $\pi(n) \geq N$ implies

$$\mathrm{d}(x_{\pi(n)}, x) < \frac{\varepsilon}{2}.$$

Since (x_n) is Cauchy, there exists $M \in \mathbb{N}$ such that $m, n \geq M$ implies

$$d(x_m, x_n) < \frac{\varepsilon}{2}.$$

Choose such M and N and assume without loss of generality that $N \geq M$. Then $n \geq N$ implies

$$\begin{aligned} d(x_n, x) &\leq d(x_{\pi(n)}, x_n) + d(x_{\pi(n)}, x) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

It follows that $x_n \to x$. □

Lemma 12.2. *Let X be a normed linear space. Then $\mathcal{X}$ is a Banach space if and only if every absolutely convergent series in $\mathcal{X}$ is convergent.*

Proof. Suppose first that every absolutely convergent series in $\mathcal{X}$ is convergent. Let (x_n) be a Cauchy sequence in $\mathcal{X}$. To show that (x_n) is convergent, it suffices to show that a subsequence of (x_n) is convergent, by Lemma (12.1). Choose a subsequence $(x_{\pi(n)})$ of (x_n) such that

$$\|x_{\pi(n)} - x_{\pi(n-1)}\| < \frac{1}{2^n}$$

and for all $n \in \mathbb{N}$ (we can do this since (x_n) is Cauchy). Then the series $\sum_{n=1}^{\infty}(x_{\pi(n)} - x_{\pi(n-1)})$ is absolutely convergent since

$$\begin{aligned} \sum_{n=1}^{\infty} \|x_{\pi(n)} - x_{\pi(n-1)}\| &< \sum_{n=1}^{\infty} \frac{1}{2^n} \\ &= 1. \end{aligned}$$

Therefore it must be convergent, say $\sum_{n=1}^{\infty}(x_{\pi(n)} - x_{\pi(n-1)}) \to x$. On the other hand, for each $n \in \mathbb{N}$, we have

$$x_{\pi(n)} - x_{\pi(1)} = \sum_{m=1}^{n} (x_{\pi(m)} - x_{\pi(1)}).$$

In particular, $x_{\pi(n)} \to x - x_{\pi(1)}$ as $n \to \infty$. Thus $(x_{\pi(n)})$ is a convergent subsequence of (x_n).

Conversely, suppose $\mathcal{X}$ is a Banach space and suppose $\sum_{n=1}^{\infty} x_n$ is absolutely convergent. Let $\varepsilon > 0$ and choose $K \in \mathbb{N}$ such that $N \geq M \geq K$ implies

$$\sum_{n=M}^{N} \|x_n\| < \varepsilon.$$

Then $N \geq M \geq K$ implies

$$\begin{aligned} \left\| \sum_{n=1}^{N} x_n - \sum_{n=1}^{M} x_n \right\| &= \left\| \sum_{n=M}^{N} x_n \right\| \\ &\leq \sum_{n=M}^{N} \|x_n\| \\ &< \varepsilon. \end{aligned}$$

It follows that the sequence of partial sums $(\sum_{n=1}^{N} x_n)_N$ is Cauchy. Since $\mathcal{X}$ is a Banach space, it follows that $\sum_{n=1}^{\infty} x_n$ is convergent. □

Proposition 12.3. *Let $1 < p < \infty$. Then $L^p(X, \mathcal{M}, \mu)$ is a Banach space.*

Proof. By Lemma (12.2), it suffices to show that every absolutely convergent series in $L^p(X, \mathcal{M}, \mu)$ is convergent. Suppose (f_n) is a sequence in $L^p(X, \mathcal{M}, \mu)$ such that $\sum_{n=1}^{\infty} \|f_n\|_p < \infty$. For each $N \in \mathbb{N}$, set $s_N = (\sum_{n=1}^{N} f_n)$. We want to show that (s_N) is convergent in $L^p(X, \mathcal{M}, \mu)$. For each $N \in \mathbb{N}$, define

$$G_N = \sum_{n=1}^{N} |f_n| \quad \text{and} \quad G = \sum_{n=1}^{\infty} |f_n|.$$

Observe that (G_N^p) is increasing sequence of nonnegative measurable (in fact integrable) functions which converges pointwise to G^p. Therefore by MCT it follows that

$$\begin{aligned}\|G\|_p &= \|G^p\|_1^{1/p} \\ &= \lim_{N\to\infty} \|G_N^p\|_1^{1/p} \\ &= \lim_{N\to\infty} \|G_N\|_p.\end{aligned}$$

In particular, since

$$\begin{aligned}\|G_N\|_p &\leq \sum_{n=1}^{N} \|f_n\|_p \\ &\leq \sum_{n=1}^{\infty} \|f_n\|_p\end{aligned}$$

for all N, we have

$$\begin{aligned}\|G\|_p &\leq \sum_{n=1}^{\infty} \|f_n\| \\ &< \infty.\end{aligned}$$

This implies $G \in L^p(X, \mathcal{M}, \mu)$. Since $\|G^p\|_1 = \|G\|_p^p < \infty$, Proposition (12.2) implies $\{G^p = \infty\}$ has measure zero, which implies $\{G = \infty\}$ has measure zero. Define $F \colon X \to \mathbb{R}$ by

$$F(x) = \begin{cases} 0 & \text{if } G(x) = \infty. \\ \sum_{n=1}^{\infty} f_n(x) & \text{if } G(x) < \infty. \end{cases}$$

for all $x \in X$. Observe that $F(x)$ lands in $\mathbb{R}$ since if $G(x) < \infty$, then $\sum_{n=1}^{\infty} f_n(x)$ is absolutely convergent (and hence convergent since $\mathbb{R}$ is complete). Since $|F| \leq G$ and $G \in L^p(X, \mathcal{M}, \mu)$, we see that $F \in L^p(X, \mathcal{M}, \mu)$. Finally, observe that

$$\begin{aligned}\lim_{N\to\infty} \|s_N - F\|_p^p &= \lim_{N\to\infty} \int_X |s_N - F|^p \mathrm{d}\mu \\ &= \lim_{N\to\infty} \int_X \left| \sum_{n=N+1}^{\infty} f_n \right|^p \mathrm{d}\mu. \\ &= \int_X \lim_{N\to\infty} \left| \sum_{n=N+1}^{\infty} f_n \right|^p \mathrm{d}\mu \\ &= \int_X 0 \mathrm{d}\mu \\ &= 0.\end{aligned}$$

where we applied DCT to get from the second step to the third step with G^p being the dominating function. □

Problem 3.b

Proposition 12.4. *Let $1 < p < \infty$. Then the set of simple functions in $L^p(X, \mathcal{M}, \mu)$ is a dense subspace of $L^p(X, \mathcal{M}, \mu)$.*

Proof. Let $f \in L^p(X, \mathcal{M}, \mu)$. Decompose f into its positive and negative parts

$$f = f^+ - f^-.$$

There exists an increasing sequence (φ_n) of nonnegative simple functions which converges to f^+ pointwise. Similarly, there exists an increasing sequence (ψ_n) of nonnegative simple functions which converges to f^- pointwise. Then $(\varphi_n + \psi_n)$ is an increasing sequence of nonnegative simple functions which converges pointwise to $|f|$. Also note that $(\varphi_n - \psi_n)$ is a sequence of simple functions which converges pointwise to f. We claim that $\|s_n - f\|_p \to 0$ as $n \to \infty$. Indeed, it suffices to show that $\||s_n - f|^p\|_1 \to 0$ since $\|s_n - f\|_p = \||s_n - f|^p\|_1^{1/p}$ for all

n. To this, we'll use DCT. Clearly $(|s_n - f|^p)$ is a sequence of measurable functions which converges pointwise to 0. Also observe that

$$
\begin{aligned}
|s_n - f|^p &\leq (|s_n| + |f|)^p \\
&= (|\varphi_n + \psi_n| + |f|)^p \\
&= (\varphi_n + \psi_n + |f|)^p \\
&\leq (|f| + |f|)^p \\
&\leq 2^p |f|^p.
\end{aligned}
$$

Thus $2^p|f|^p$ is a dominating function, which means we can apply DCT. Therefore

$$
\begin{aligned}
\lim_{n\to\infty} \int_X |s_n - f|^p \mathrm{d}\mu &= \int_X \lim_{n\to\infty} |s_n - f|^p \mathrm{d}\mu \\
&= \int_X 0 \mathrm{d}\mu \\
&= 0.
\end{aligned}
$$

□

12.4 If $\mu(X) < \infty$, then Uniform Convergence Implies Integral Convergence

Problem 4.a

Proposition 12.5. *Assume that $\mu(X) < \infty$. Suppose $(f_n \colon X \to \mathbb{R})$ is a sequence of integrable functions such that $f_n \to f$ uniformly. Then f is integrable and*

$$
\lim_{n\to\infty} \int_X f_n \mathrm{d}\mu = \int_X f \mathrm{d}\mu. \tag{33}
$$

Proof. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$
|f(x) - f_n(x)| < \frac{\varepsilon}{\mu(X)}
$$

for all $x \in X$. Then

$$
\begin{aligned}
\int_X |f| \mathrm{d}\mu &= \int_X |f_N + f - f_N| \mathrm{d}\mu \\
&\leq \int_X |f_N| \mathrm{d}\mu + \int_X |f - f_n| \mathrm{d}\mu \\
&< \int_X |f_N| \mathrm{d}\mu + \frac{\varepsilon}{\mu(X)} \mu(X) \\
&< \int_X |f_N| \mathrm{d}\mu + \varepsilon \\
&< \infty.
\end{aligned}
$$

It follows that f is integrable. Now observe that $n \geq N$ implies

$$
\begin{aligned}
\left| \int_X f \mathrm{d}\mu - \int_X f_n \mathrm{d}\mu \right| &= \left| \int_X (f - f_n) \mathrm{d}\mu \right| \\
&\leq \int_X |f - f_n| \mathrm{d}\mu \\
&< \frac{\varepsilon}{\mu(X)} \mu(X). \\
&= \varepsilon.
\end{aligned}
$$

This implies (33). □

12.5 If $\mu(X) < \infty$ and $1 \leq p < q < \infty$, then $L^q(X, \mathcal{M}, \mu) \subseteq L^p(X, \mathcal{M}, \mu)$

Problem 4.b

Proposition 12.6. *Assume that $\mu(X) < \infty$. Let $1 \leq p < q < \infty$. Then $L^q(X, \mathcal{M}, \mu) \subseteq L^p(X, \mathcal{M}, \mu)$.*

Proposition 12.7.

Proof. Let $f \in L^q(X, \mathcal{M}, \mu)$. We want to show that $f \in L^p(X, \mathcal{M}, \mu)$. Let

$$A = \{x \in X \mid |f|(x) > 1\}.$$

Then $|f|^p 1_A < |f|^q 1_A$, thus

$$\begin{aligned}
\int_X |f|^p \mathrm{d}\mu &= \int_X (|f|^p 1_A + |f|^p 1_{A^c}) \mathrm{d}\mu \\
&= \int_X |f|^p 1_A \mathrm{d}\mu + \int_X |f|^p 1_{A^c} \mathrm{d}\mu \\
&\leq \int_X |f|^q 1_A \mathrm{d}\mu + \int_X 1_{A^c} \mathrm{d}\mu \\
&\leq \|f\|_q + \mu(A^c) \\
&< \infty.
\end{aligned}$$

It follows that $f \in L^p(X, \mathcal{M}, \mu)$. □

12.6 Generalized Dominated Convergence Theorem

Proposition 12.8. *Let $(f_n \colon X \to \mathbb{R})$ and $(g_n \colon X \to [0, \infty))$ be two sequences of integrable functions which converge almost everywhere to integrable functions $f \colon X \to \mathbb{R}$ and $g \colon X \to \mathbb{R}$ respectively. Suppose $|f_n| \leq g_n$ for all n and $\|g_n\|_1 \to \|g\|_1$. Then*

$$\lim_{n \to \infty} \int_X f_n \mathrm{d}\mu = \int_X f \mathrm{d}\mu.$$

Proof. Observe that $(g_n - f_n)$ is a sequence of nonnegative measurable functions. Thus by Fatou's Lemma, we have

$$\begin{aligned}
\int_X g \mathrm{d}\mu - \int_X f \mathrm{d}\mu &= \int_X (g - f) \mathrm{d}\mu \\
&\leq \liminf_{n \to \infty} \int_X (g_n - f_n) \mathrm{d}\mu \\
&= \int_X g \mathrm{d}\mu - \limsup_{n \to \infty} \int f_n \mathrm{d}\mu,
\end{aligned}$$

where we used the fact that $\|g_n\|_1 \to \|g\|_1$ to get from the second line to the third line. Subtracting $\int_X g \mathrm{d}\mu$ from both sides and canceling the sign gives us

$$\limsup_{n \to \infty} \int_X f_n \mathrm{d}\mu \leq \int_X f \mathrm{d}\mu.$$

Now we apply the same argument with functions $g_n + f_n$ in place of $g_n - f_n$, and we obtain

$$\liminf_{n \to \infty} \int f_n \mathrm{d}\mu \geq \int_X f \mathrm{d}\mu.$$

□

12.7 Almost Everywhere Convergence Plus Integral Convergence Implies L^1 Convergence

Proposition 12.9. *Let $(f_n \colon X \to \mathbb{R})$ be a sequence of integrable functions that converge almost everywhere to an integrable function $f \colon X \to \mathbb{R}$. Then $\|f_n - f\|_1 \to 0$ if and only if $\|f_n\|_1 \to \|f\|_1$.*

Proof. Suppose $\|f_n - f\|_1 \to 0$. Then

$$\begin{aligned}\lim_{n\to\infty} |\|f_n\|_1 - \|f\|_1| &\le \lim_{n\to\infty} \|f_n - f\|_1 \\ &= 0.\end{aligned}$$

Thus $\|f_n\|_1 \to \|f\|$.

Conversely, suppose $\|f_n\|_1 \to \|f\|_1$. For each $n \in \mathbb{N}$, set $g_n = |f_n| + |f|$, and set $g = 2|f|$. Then $|f_n - f| \le g_n$, also g_n converges pointwise almost everywhere to g, also

$$\begin{aligned}\|g_n\|_1 &= \int_X (|f_n| + |f|)\mathrm{d}\mu \\ &= \int_X |f_n|\mathrm{d}\mu + \int_X |f|\mathrm{d}\mu \\ &= \|f_n\|_1 + \|f\|_1 \\ &\to 2\|f\|_1 \\ &= \|g\|,\end{aligned}$$

and $f_n - f$ converges pointwise almost everywhere to 0. It follows from problem 5 that

$$\begin{aligned}\|f_n - f\|_1 &\to \|0\|_1 \\ &= 0.\end{aligned}$$

□

If $f\colon X \to \mathbb{R}$ is Integral, then $n\mu(\{|f| > n\}) \to 0$ as $n \to \infty$

Proposition 12.10. *Let $f\colon X \to \mathbb{R}$ be an integral function. Then*

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

Proof. First we consider the case for integrable simple functions, say

$$\varphi = \sum_{i=1}^{n} a_i 1_{A_i}, \tag{34}$$

where (34) is expressed in canonical form. Being integral here means $\mu(A_i) \neq \infty$ for all $1 \le i \le n$. In particular, $|\varphi|$ is bounded above by some N. Thus $n \ge N$ implies

$$\begin{aligned}n\mu(\{\varphi > n\}) &\ge n \cdot 0 \\ &= 0.\end{aligned}$$

Therefore

$$\lim_{n\to\infty} n\mu(\{\varphi > n\}) = 0.$$

Now we prove it for any integral function $f\colon X \to \mathbb{R}$. First note that since $\mu(\{|f| > n\}) \ge \mu(\{f > n\})$, we may assume that f is nonnegative. Using the fact that the set of all integrable simple functions is dense in $L^1(X, \mathcal{M}, \mu)$, choose a nonnegative integrable simple function φ such that $\varphi \le f$ and $\|f - \varphi\|_1 < \varepsilon$. Let M be an upper bound for φ. Then we have

$$\begin{aligned}\lim_{n\to\infty} n\mu(\{f > n\}) &= \lim_{n\to\infty} n\mu(\{\varphi > n\} \cup \{f - \varphi \ge n - \varphi\}) \\ &\le \lim_{n\to\infty} n\mu(\{\varphi > n\} \cup \{f - \varphi \ge n - M\}) \\ &\le \lim_{n\to\infty} n\mu(\{\varphi > n\}) + \lim_{n\to\infty} n\mu(\{f - \varphi \ge n - M\}) \\ &= \lim_{n\to\infty} n\mu(\{f - \varphi \ge n - M\}) \\ &\le \lim_{n\to\infty} \frac{n}{n-M}\|f - \varphi\|_1 \\ &< \lim_{n\to\infty} \frac{n\varepsilon}{n-M} \\ &= \varepsilon.\end{aligned}$$

Taking $\varepsilon \to 0$ gives us our desired result. □

12.8 Young's Inequality

Exercise 2. Let $x, y \geq 0$ and $0 < \gamma < 1$. Prove that

$$x^\gamma y^{1-\gamma} \leq \gamma x + (1-\gamma) y. \tag{35}$$

Deduce the Young's Inequality.

Solution 2. We may assume that $x, y > 0$ since otherwise it is trivial. Set $t = x/y$ and rewrite (35) as

$$t^\gamma - \gamma t \leq 1 - \gamma. \tag{36}$$

Thus, to show (35) for all $x, y > 0$, we just need to show (36) for all $t > 0$. To see why (36) holds, define $f \colon \mathbb{R}_{>0} \to \mathbb{R}$ by

$$f(t) = t^\gamma - \gamma t$$

for all $t \in \mathbb{R}_{>0}$. Observe that f is a smooth function on $\mathbb{R}_{>0}$, with it's first derivative and second derivative given by

$$f'(t) = \gamma t^{\gamma-1} - \gamma \quad \text{and} \quad f''(t) = \gamma(\gamma-1)t^{\gamma-2}$$

for all $t \in \mathbb{R}_{>0}$. Observe that

$$\begin{aligned} f'(t) = 0 &\iff \gamma t^{\gamma-1} = \gamma \\ &\iff t^{\gamma-1} = 1 \\ &\iff t = 1, \end{aligned}$$

where the last if and only if follows from the fact that t is a positive real number. Also, we clearly have $f''(t) < 0$ for all $t \in \mathbb{R}_{>0}$. Thus, since f is concave down on all of $\mathbb{R}_{>0}$, and $f'(t) = 0$ if and only if $t = 1$, it follows that f has a global maximum at $t = 1$. In particular, we have

$$\begin{aligned} t^\gamma - \gamma t &= f(t) \\ &\leq f(1) \\ &\leq 1^\gamma - \gamma \cdot 1 \\ &= 1 - \gamma \end{aligned}$$

for all $t \in \mathbb{R}_{>0}$.

With (35) established, we now prove Young's Inequality: Let $a, b \geq 0$ and let $1 \leq p, q < \infty$ such that $1/p + 1/q = 1$. We want to show that

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Set $\gamma = 1/p$ (so $1 - \gamma = 1/q$), $a = x^\gamma$, and $b = y^{1-\gamma}$. Then Young's Inequality becomes (35), which was proved above.

Part III
Exams

13 Exam 2

13.1 Prove function is measurable. Prove set is measurable

Exercise 3. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f \colon X \to \mathbb{R}$ be a measurable function. Prove that $f/(1+f^2)$ is also a measurable function.

Solution 3. By a proposition proved in class (see Appendix for details), the product and sum of two measurable functions is measurable. Therefore both f and $1 + f^2$ are measurable. It remains to show that $f/(1+f^2)$ is measurable. To see this, note that for any strictly positive measurable function $h: X \to (0, \infty)$, the function $1/h$ is measurable. Indeed, for any $c > 0$, we have

$$\begin{aligned} x \in \left\{ \frac{1}{h} < c \right\} &\iff \frac{1}{h(x)} < c \\ &\iff 1 < ch(x) \\ &\iff x \in \left\{ h > \frac{1}{c} \right\}. \end{aligned}$$

Thus $\{1/h < c\} = \{h > 1/c\} \in \mathcal{M}$. If $c \leq 0$, then we have $\{1/h < c\} = \emptyset \in \mathcal{M}$ since h is strictly positive. In either case, we see that $1/h$ is measurable. In particular, since $1 + f^2$ is a strictly positive measurable function, we see that $1/(1 + f^2)$ is a measurable function. Therefore the product $f/(1 + f^2)$ is a measurable function.

Exercise 4. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f: X \to \mathbb{R}$ be a measurable function. Prove that $\{-1 \leq f \leq 1\}$ is a measurable function.

Solution 4. Since f is measurable, we have

$$\begin{aligned} \{f \leq 1\} &= \bigcap_{n=1}^{\infty} \left\{ f < 1 + \frac{1}{n} \right\} \\ &\in \mathcal{M}. \end{aligned}$$

Similarly we have

$$\begin{aligned} \{f \geq -1\} &= \{f < -1\}^c \\ &\in \mathcal{M}. \end{aligned}$$

Therefore

$$\begin{aligned} \{-1 \leq f \leq 1\} &= \{f \geq -1\} \cap \{f \leq 1\} \\ &\in \mathcal{M}. \end{aligned}$$

13.2 $f_n \xrightarrow{\text{pwae}} f$ implies $f_n \xrightarrow{\text{m}} f$

Exercise 5. Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Suppose $(f_n: X \to \mathbb{R})$ is a sequence of measurable functions and suppose $f: X \to \mathbb{R}$ is another measurable function. Prove that

$$\bigcup_{k=1}^{\infty} \limsup_{n \to \infty} \{|f_n - f| \geq 1/k\} = \{\lim_{n \to \infty} f_n \neq f\} \tag{37}$$

Solution 5. Observe that

$$\begin{aligned} x \in \bigcup_{k=1}^{\infty} \limsup_{n \to \infty} \{|f_n - f| \geq 1/k\} &\iff x \in \limsup_{n \to \infty} \{|f_n - f| \geq 1/k\} \text{ for some } k \\ &\iff x \in \{|f_{\pi_k(n)} - f| \geq 1/k\} \text{ for some } k \text{ for some subsequence } (\pi_k(n))_{n \in \mathbb{N}} \text{ of } (n)_{n \in \mathbb{N}} \\ &\iff |f_{\pi_k(n)}(x) - f(x)| \geq 1/k \text{ for some } k \text{ for some subsequence } (\pi_k(n))_{n \in \mathbb{N}} \text{ of } (n)_{n \in \mathbb{N}} \\ &\iff x \in \{\lim_{n \to \infty} f_n \neq f\} \end{aligned}$$

where the last if and only if follows from the fact that the distance $|f_n(x) - f(x)|$ is frequently greater than $1/k$, which means $f_n(x) \not\to f(x)$. This gives us (37).

Exercise 6. Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Suppose $(f_n \colon X \to \mathbb{R})$ is a sequence of measurable functions and suppose $f \colon X \to \mathbb{R}$ is another measurable function. Prove that if $\lim_{n\to\infty} f_n(x) = f(x)$ for a.e. $x \in X$, then for all $k \in \mathbb{N}$, we have

$$\lim_{N\to\infty} \mu\{|f_N - f| \geq 1/k\} = 0 \tag{38}$$

Solution 6. Let $k \in \mathbb{N}$. Then observe that

$$\begin{aligned}
0 &= \mu\left\{\lim_{n\to\infty} f_n \neq f\right\} \\
&= \mu\left\{\bigcup_{m=1}^{\infty} \limsup_{n\to\infty}\{|f_n - f| \geq 1/m\}\right\} \\
&\geq \mu\left\{\limsup\{|f_n - f| \geq 1/k\}\right\} \\
&= \mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n\geq N} \{|f_n - f| \geq 1/k\}\right) \\
&= \lim_{N\to\infty} \mu\left(\bigcup_{n\geq N} \{|f_n - f| \geq 1/k\}\right) \\
&\geq \lim_{N\to\infty} \mu\{|f_N - f| \geq 1/k\}
\end{aligned}$$

where we used the fact that $\mu(X) < \infty$ to get from the fourth line to the fifth line. This gives us (38).

Exercise 7. Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Suppose $(f_n \colon X \to \mathbb{R})$ is a sequence of measurable functions and suppose $f \colon X \to \mathbb{R}$ is another measurable function. Prove that if $f_n \to f$ a.e, then $f_n \to f$ in measure.

Solution 7. Suppose $f_n \to f$ a.e and let $\varepsilon > 0$. Choose $k \in \mathbb{N}$ such that $1/k < \varepsilon$. Then by part (b), we have

$$\begin{aligned}
\lim_{n\to\infty} \mu\{|f_n - f| \geq \varepsilon\} &\leq \lim_{n\to\infty} \mu\{|f_n - f| \geq 1/k\} \\
&= 0.
\end{aligned}$$

This implies $f_n \to f$ in measure.

Exercise 8. Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Suppose $(f_n \colon X \to \mathbb{R})$ is a sequence of measurable functions and suppose $f \colon X \to \mathbb{R}$ is another measurable function. Prove that if $f_n \xrightarrow{L^2} f$, then $f_n \to f$ in measure.

Solution 8. Let $g \in L^2(X, \mathcal{M}, \mu)$. Since $\mu(X) < \infty$, we also have $1_X \in L^2(X, \mathcal{M}, \mu)$. By Hölder's inequality, we have

$$\begin{aligned}
\|g\|_1 &\leq \|g\|_2 \cdot \|1_X\|_2 \\
&= \sqrt{\mu(X)}\|g\|_2.
\end{aligned}$$

In particular, $f_n \xrightarrow{L^2} f$ implies $f_n \xrightarrow{L^1} f$ which implies $f_n \to f$ in measure (proved in class).

13.3 Integral Computation (Using Descending MCT)

Exercise 9. Compute the following limit

$$\lim_{n\to\infty} \int_{(0,1)} \frac{1 + nx}{(1 + x)^n} \mathrm{d}x$$

Solution 9. For each $n \in \mathbb{N}$, let $f_n = (1 + nx)(1 + x)^{-n}$. Observe that each f_n is measurable since each f_n is continuous (the denominator is never zero for any $x \in \mathbb{R}$). Furthermore, each f_n is nonnegative (since taking squares makes everything nonnegative). We claim that f_n is a decreasing sequence. Indeed

$$\begin{aligned}\frac{f_n}{f_{n+1}} &= \left(\frac{1+nx}{(1+x)^n}\right)\left(\frac{(1+x)^{n+1}}{1+(n+1)x}\right)\\ &= \frac{(1+nx)(1+x)}{1+(n+1)x}\\ &= \frac{nx^2+(n+1)x+1}{(n+1)x+1}\\ &> 1\end{aligned}$$

Thus (f_n) is a decreasing sequence which is bounded below, so it must converge pointwise to some function. To see what it converges to, we use L'Hopital's rule:

$$\begin{aligned}\lim_{n\to\infty} f_n &= \lim_{n\to\infty} \frac{1+nx}{(1+x)^n}\\ &= \lim_{n\to\infty} \frac{x}{\ln(1+x)(1+x)^n}\\ &= 0.\end{aligned}$$

Thus (f_n) converges pointwise to 0. Since

$$\begin{aligned}\int_0^1 f_1 \mathrm{d}x &= \int_0^1 \frac{1+x}{1+x}\mathrm{d}x\\ &= \int_0^1 \mathrm{d}x\\ &= 1\\ &< \infty,\end{aligned}$$

it follows from (decreasing version of MCT) that

$$\begin{aligned}\lim_{n\to\infty} \frac{1+nx}{(1+x)^n}\mathrm{d}x &= \lim_{n\to\infty} \int_0^1 f_n \mathrm{d}x\\ &= \int_0^1 0\mathrm{d}x\\ &= 0.\end{aligned}$$

13.4 Strictly Positive Measurable Function Gives Rise to Strictly Positive Measure

Exercise 10. Let $(X, \mathcal{M}, \mu)$ be a measure space. Suppose that $f\colon X \to \mathbb{R}$ is a measurable function such that $f(x) > 0$ for all $x \in X$. Prove that $\int_X 1_E f \mathrm{d}\mu > 0$ for every measurable $E \in \mathcal{M}$ such that $\mu(E) > 0$.

Solution 10. Let $E \in \mathcal{M}$ such that $\mu(E) > 0$. For each $n \in \mathbb{N}$, define

$$F_n := \{f \geq 1/n\}.$$

Since $f(x) > 0$ for all $x \in X$, we have

$$\begin{aligned}0 &< \mu(E)\\ &= \mu\left(\bigcup_{n=1}^{\infty} F_n \cap E\right)\\ &\leq \sum_{n=1}^{\infty} \mu(F_n \cap E).\end{aligned}$$

The strict inequality implies $\mu(F_n \cap E) > 0$ for some $n \in \mathbb{N}$. Choose such an $n \in \mathbb{N}$, then we have

$$\begin{aligned}\int_X f 1_E \mathrm{d}\mu &\geq \int_X f 1_{E\cap F_n} \mathrm{d}\mu\\ &\geq \int_X \frac{1}{n} \cdot 1_{E\cap F_n} \mathrm{d}\mu\\ &= \mu(E \cap F_n)/n\\ &> 0.\end{aligned}$$

13.5 Problem Involving MCT

Exercise 11. Let $f\colon [0,1] \to [0,\infty)$ be a nonnegative measurable function. Prove that if

$$\lim_{n\to\infty} \int_{[0,1]} f 1_{[0,\frac{n}{n+1}]} \mathrm{d}x \leq 1$$

for all $n \in \mathbb{N}$, then f is integrable and $\int_{[0,1]} f \mathrm{d}x \leq 1$.

Solution 11. Observe that since $(n/(n+1))$ is an increasing sequence which converges to 1, the sequence $(f 1_{[0,\frac{n}{n+1}]})$ is an increasing sequence of nonnegative measurable functions which converges pointwise to f. It follows from MCT that

$$\begin{aligned}\int_{[0,1]} f \mathrm{d}x &= \lim_{n\to\infty} \int_{[0,1]} f 1_{[0,\frac{n}{n+1}]} \mathrm{d}x \\ &\leq 1.\end{aligned}$$

In particular, f is integrable and $\int_{[0,1]} f \mathrm{d}x \leq 1$.

13.6 Nonnegative Measurable Function Induces Finite Measure

Exercise 12. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f\colon X \to \mathbb{R}$ be a nonnegative measurable function such that $\int_X f \mathrm{d}\mu < \infty$. Prove that $\nu\colon \mathcal{M} \to \mathbb{R}$ defined by

$$\nu(E) = \int_X f 1_E \mathrm{d}\mu$$

is a finite measure on $(X, \mathcal{M})$.

Solution 12. This was proved in the homework, but we include it for completeness.

First we prove it for nonnegative simple functions:

Proposition 13.1. *Let $\phi\colon X \to [0,\infty)$ be a nonnegative simple function. Define a function $\nu\colon \mathcal{M} \to [0,\infty]$ by*

$$\nu(E) = \int_X \phi 1_E \mathrm{d}\mu$$

for all $E \in \mathcal{M}$. Then ν is a measure on $(X, \mathcal{M})$.

Proof. First note that

$$\begin{aligned}\nu(\emptyset) &= \int_X \phi 1_\emptyset \mathrm{d}\mu \\ &= \int_X \phi \cdot 0 \cdot \mathrm{d}\mu \\ &= \int_X 0 \cdot \mathrm{d}\mu \\ &= 0.\end{aligned}$$

Now we show that ν is finitely additive. Let $(E_n)_{n=1}^N$ be a finite sequence of pairwise disjoint sets in $\mathcal{M}$. Then

$$\begin{aligned}\nu\left(\bigcup_{n=1}^N E_n\right) &= \int_X \phi 1_{\bigcup_{n=1}^N E_n} \mathrm{d}\mu \\ &= \int_X \phi \sum_{n=1}^N 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^N \int_X \phi 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^N \nu(E_n),\end{aligned}$$

where we used the fact that each $\phi 1_{E_n}$ is a nonnegative simple function in order to commute the finite sum with the integral. Thus it follows that ν is finitely additive. It remains to show that ν is countably subadditive. Let (E_n) be a sequence of sets in $\mathcal{M}$. We want to show that

$$\int_X \phi 1_{\bigcup_{n=1}^\infty E_n} \mathrm{d}\mu \leq \sum_{n=1}^\infty \int_X \phi 1_{E_n} \mathrm{d}\mu. \tag{39}$$

To do this, we will show that the sum on the righthand side in (39) is greater than or equal to all integrals of the form $\int \varphi \mathrm{d}\mu$ where $\varphi\colon X \to [0,\infty]$ is a simple function such that $\varphi \leq \phi 1_{\bigcup_{n=1}^\infty E_n}$. Then the inequality (39) will follow from the fact that the integral on the lefthand side in (39) is the supremum of this set. So let $\varphi\colon X \to [0,\infty]$ be a simple function such that $\varphi \leq \phi 1_{\bigcup_{n=1}^\infty E_n}$. Write φ and ϕ in terms of their canonical forms, say

$$\varphi = \sum_{i=1}^k a_i 1_{A_i} \quad \text{and} \quad \phi = \sum_{j=1}^m b_j 1_{B_j}.$$

So $a_i \neq a_{i'}$ and $A_i \cap A_{i'} = \emptyset$ whenever $i \neq i'$ and $b_j \neq b_{j'}$ and $B_j \cap B_{j'} = \emptyset$ whenever $j \neq j'$. Observe that the canonical representation of $\phi 1_{\bigcup_{n=1}^\infty E_n}$ is given by

$$\begin{aligned}
\phi 1_{\bigcup_{n=1}^\infty E_n} &= \left(\sum_{j=1}^m b_j 1_{B_j} \right) 1_{\bigcup_{n=1}^\infty E_n} \\
&= \sum_{j=1}^m b_j 1_{B_j} 1_{\bigcup_{n=1}^\infty E_n} \\
&= \sum_{j=1}^m b_j 1_{\bigcup_{n=1}^\infty B_j \cap E_n},
\end{aligned}$$

where this representation is the canonical representation since $b_j \neq b_{j'}$ and

$$\left(\bigcup_{n=1}^\infty B_j \cap E_n \right) \cap \left(\bigcup_{n=1}^\infty B_{j'} \cap E_n \right) = \emptyset$$

whenever $j \neq j'$ (since $B_j \cap B_{j'} = \emptyset$). Therefore we have

$$\begin{aligned}
\int_X \varphi \mathrm{d}\mu &\leq \int_X \phi 1_{\bigcup_{n=1}^\infty E_n} \mathrm{d}\mu \\
&= \sum_{j=1}^m b_j \mu \left(\bigcup_{n=1}^\infty B_j \cap E_n \right) \\
&\leq \sum_{j=1}^m b_j \sum_{n=1}^\infty \mu(B_j \cap E_n) \\
&= \sum_{n=1}^\infty \sum_{j=1}^m b_j \mu(B_j \cap E_n) \\
&= \sum_{n=1}^\infty \sum_{j=1}^m \int_X b_j 1_{B_j \cap E_n} \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \sum_{j=1}^m b_j 1_{B_j \cap E_n} \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \sum_{j=1}^m b_j \left(1_{B_j} 1_{E_n} \right) \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \left(\sum_{j=1}^m b_j 1_{B_j} \right) 1_{E_n} \mathrm{d}\mu \\
&= \sum_{n=1}^\infty \int_X \phi 1_{E_n} \mathrm{d}\mu,
\end{aligned}$$

where we used monotonicity of integration in the first line and where we used countable subadditivity of μ to get from the second line to the third line.

$\square$

Now we prove it for more general nonnegative measurable functions

Proposition 13.2. *Let $(X, \mathcal{M}, \mu)$ be measure space and let $g\colon X \to [0,\infty)$ be a nonnegative measurable function. Define $\nu_g\colon \mathcal{M} \to [0,\infty]$ by*

$$\nu_g(E) = \int_X g 1_E \mathrm{d}\mu$$

for all $E \in \mathcal{M}$. Then ν is a measure on $(X, \mathcal{M})$. Furthermore, if $\int_X g\mathrm{d}\mu < \infty$, then $(X, \mathcal{M}, \nu)$ is a finite measure space.

Proof. First note that

$$\begin{aligned} \nu_g(\emptyset) &= \int_X g 1_\emptyset \mathrm{d}\mu \\ &= \int_X g \cdot 0 \cdot \mathrm{d}\mu \\ &= \int_X 0 \cdot \mathrm{d}\mu \\ &= 0. \end{aligned}$$

Next we show that ν_g is finitely additive. Let $(E_n)_{n=1}^N$ be a finite sequence of pairwise disjoint sets in $\mathcal{M}$. Then

$$\begin{aligned} \nu_g\left(\bigcup_{n=1}^N E_n\right) &= \int_X g 1_{\bigcup_{n=1}^N E_n} \mathrm{d}\mu \\ &= \int_X g \sum_{n=1}^N 1_{E_n} \mathrm{d}\mu \\ &= \int_X \sum_{n=1}^N g 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^N \int_X g 1_{E_n} \mathrm{d}\mu \\ &= \sum_{n=1}^N \nu_g(E_n), \end{aligned}$$

where we used the fact that each $g1_{E_n}$ is a nonnegative measurable function in order to commute the finite sum with the integral. Thus it follows that ν_g is finitely additive.

It remains to show that ν_g is countably subadditive. In problem 3 in HW4, we showed that for any nonnegative simple function $\varphi\colon X \to [0,\infty)$, the function $\nu_\varphi\colon \mathcal{M} \to [0,\infty]$ defined by

$$\nu_\varphi(E) = \int_X \varphi 1_E \mathrm{d}\mu$$

is a measure. We will use this fact in what follows:

Choose an increasing sequence $(\varphi_n\colon X \to [0,\infty])$ of nonnegative simple functions which converges pointwise to g (we can do this since g is a nonnegative measurable function). Then $(\varphi_n 1_E)$ is an increasing sequence of nonnegative simple functions which converges pointwise to $g1_E$ for all $E \in \mathcal{M}$. It follows from the Monotone Convergence Theorem that

$$\begin{aligned} \nu_{\varphi_n}(E) &= \int_X \varphi_n 1_E \mathrm{d}\mu \\ &\to \int_X g 1_E \mathrm{d}\mu \\ &= \nu_g(E) \end{aligned}$$

for any $E \in \mathcal{M}$. In particular, for any $E \in \mathcal{M}$ and $\varepsilon > 0$, we can find a $N_{E,\varepsilon} \in \mathbb{N}$ (which depends on E and ε) such that

$$\nu_g(E) < \nu_{\varphi_n}(E) + \varepsilon \tag{40}$$

for all $n \geq N_{E,\varepsilon}$. However we will don't need estimate $\nu_g(E)$ to this level of precision. We just need to know thta for any $\varepsilon > 0$, we can find an $n \in \mathbb{N}$ such that (40) holds.

Now let (E_k) be a sequence of measurable sets and let $\varepsilon > 0$. Choose $n \in \mathbb{N}$ such that

$$\nu_g\left(\bigcup_{k=1}^{\infty} E_k\right) < \nu_{\varphi_n}\left(\bigcup_{k=1}^{\infty} E_k\right) + \varepsilon$$

Then we have

$$\begin{aligned}
\nu_g\left(\bigcup_{k=1}^{\infty} E_k\right) &\leq \nu_{\varphi_n}\left(\bigcup_{k=1}^{\infty} E_k\right) + \varepsilon \\
&\leq \sum_{k=1}^{\infty} \nu_{\varphi_n}(E_k) + \varepsilon \\
&\leq \sum_{k=1}^{\infty} \nu_g(E_k) + \varepsilon
\end{aligned}$$

where we obtained the second line from the first line using countable subadditivity of ν_{φ_n}, and where we obtained the third line from the second line from the fact that $\varphi_n 1_{E_k} \leq g 1_{E_k}$ for all k and from monotonicity of integration. Taking $\varepsilon \to 0$ gives us countable subadditivity of ν_g.

Finally, for the last part, we note that

$$\begin{aligned}
\nu(X) &= \int_X g 1_X \mathrm{d}\mu \\
&= \int_X g \mathrm{d}\mu \\
&< \infty.
\end{aligned}$$

□

Exercise 13. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f \colon X \to \mathbb{R}$ be a nonnegative measurable function such that $\int_X f \mathrm{d}\mu < \infty$. Prove that for any sequence (E_n) of measurable sets such that $\sum_{n=1}^{\infty} \nu(E_n) < \infty$, we have

$$\lim_{n\to\infty} 1_{E_n} f = 0$$

for μ a.e. x.

Solution 13. First recall three propositions we proved in the homework:

Proposition 13.3. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a sequence in $\mathcal{M}$. Then*

$$\mu\left(\liminf E_n\right) \leq \liminf \mu(E_n)$$

Proof. Note that the sequence

$$\left(\bigcap_{n\geq N} E_n\right)_{N\in\mathbb{N}}$$

is an ascending sequence in N. Therefore we have

$$\begin{aligned}
\mu\left(\liminf E_n\right) &= \mu\left(\bigcup_{N=1}^{\infty}\left(\bigcap_{n\geq N} E_n\right)\right) \\
&= \liminf \mu\left(\bigcap_{n\geq N} E_n\right) \\
&\leq \lim_{N\to\infty} \inf\left\{\mu(E_n) \mid n \geq N\right\} \\
&= \liminf \mu(E_n),
\end{aligned}$$

where we obtained the third line from the second line since

$$\mu\left(\bigcap_{n\geq N} E_n\right) \leq \mu(E_n)$$

for all $n \geq N$ by monotonicity of μ. $\square$

Proposition 13.4. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a sequence in $\mathcal{M}$ such that $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) < \infty$. Then*

$$\mu\left(\limsup E_n\right) \geq \limsup \mu(E_n)$$

Proof. Note that the sequence

$$\left(\bigcup_{n\geq N} E_n\right)_{N\in\mathbb{N}}$$

is a descending sequence in N. This together with the fact that $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) < \infty$ implies

$$\begin{aligned}
\mu\left(\limsup E_n\right) &= \mu\left(\bigcap_{N=1}^{\infty}\left(\bigcup_{n\geq N} E_n\right)\right) \\
&= \lim_{N\to\infty} \mu\left(\bigcup_{n\geq N} E_n\right) \\
&\geq \lim_{N\to\infty} \sup\left\{\mu(E_n) \mid n \geq N\right\} \\
&= \limsup \mu(E_n),
\end{aligned}$$

where we obtained the third line from the second line since

$$\mu\left(\bigcup_{n\geq N} E_n\right) \geq \mu(E_n)$$

for all $n \geq N$ by monotonicity of μ. $\square$

Proposition 13.5. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let (E_n) be a sequence in $\mathcal{M}$ such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then*

$$\mu\left(\limsup E_n\right) = 0.$$

Proof. Note that the sequence

$$\left(\bigcup_{n\geq N} E_n\right)_{N\in\mathbb{N}}$$

is a descending sequence in N. This together with the fact that $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) < \infty$ implies

$$\begin{aligned}
\mu\left(\limsup E_n\right) &= \mu\left(\bigcap_{N=1}^{\infty}\left(\bigcup_{n\geq N} E_n\right)\right) \\
&= \lim_{N\to\infty} \mu\left(\bigcup_{n\geq N} E_n\right) \\
&\leq \lim_{N\to\infty} \sum_{n=N}^{\infty} \mu(E_n) \\
&= 0,
\end{aligned}$$

where the last equality follows since $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. $\square$

Now we prove part (b). Let (E_n) be a sequence of measurable sets such that

$$\sum_{n=1}^{\infty} \nu(E_n) < \infty.$$

Then observe that

$$\begin{aligned}\int_X \lim_{n\to\infty} f1_{E_n}\mathrm{d}\mu &= \int_X \liminf f1_{E_n}\mathrm{d}\mu \\ &\leq \liminf \int_X f1_{E_n}\mathrm{d}\mu \\ &= \liminf \nu(E_n) \\ &\leq \limsup \nu(E_n) \\ &\leq \nu(\limsup E_n) \\ &= 0.\end{aligned}$$

where we applied Fatou's Lemma to get the second line from the first line. It follows that $\lim_{n\to\infty} f1_{E_n} = 0$ almost everywhere (by a proposition proved in class).

Appendix

Problem 1

Proposition 13.6. *Let $f, g\colon X \to \mathbb{R}$ be measurable functions and let $a \in \mathbb{R}$. Then af, $|f|$, f^2, $f+g$, fg, $\max\{f,g\}$, and $\min\{f,g\}$ are all measurable.*

Proof. We first show af is measurable. If $a = 0$, then af is the zero function, which is measurable. So assume $a \neq 0$. Then we have

$$(af)^{-1}(-\infty, c) = \begin{cases} f^{-1}(-\infty, c/a) \in \mathcal{M} & \text{if } a > 0 \\ f^{-1}(c/a, \infty) \in \mathcal{M} & \text{if } a < 0 \end{cases}$$

αf is measurable αf is measurable.
Observe that

$$\begin{aligned} x \in (f+g)^{-1}(-\infty, c) &\iff f(x) + g(x) < c \\ &\iff \text{there exists an } r \in \mathbb{Q} \text{ such that } f(x) < r \text{ and } r < c - g(x) \\ &\iff \text{there exists an } r \in \mathbb{Q} \text{ such that } x \in f^{-1}(-\infty, r) \cap g^{-1}(-\infty, c-r). \\ &\iff x \in \bigcup_{r\in\mathbb{Q}} f^{-1}(-\infty, r) \cap g^{-1}(-\infty, c-r). \end{aligned}$$

Therefore

$$(f+g)^{-1}(-\infty, c) = \bigcup_{r\in\mathbb{Q}} f^{-1}(-\infty, r) \cap g^{-1}(-\infty, c-r) \in \mathcal{M}.$$

We first prove f^2 is measurable:

$$(f^2)^{-1}(c, \infty) = \begin{cases} f^{-1}(\sqrt{c}, \infty) \cup f^{-1}(-\infty, -\sqrt{c}) \in \mathcal{M} & c \geq 0 \\ E \in \mathcal{M} & c < 0 \end{cases}$$

Next, note that

$$fg = \frac{1}{4}\left((f+g)^2 - (f-g)^2\right) \in \mathcal{M}.$$

Finally note that

$$\max\{f, g\} = \frac{1}{2}(|f+g| + |f-g|)$$

and

$$\min\{f, g\} = \frac{1}{2}(|f+g| - |f-g|).$$

□

Proposition 13.7. *Let $(f_n\colon X \to \mathbb{R})$ be a sequence of measurable functions. Then $\sup f_n$, $\inf f_n$, $\limsup f_n$, and $\liminf f_n$ are all measurable. In particular, of*

$$\lim_{n\to\infty} f_n(x)$$

exists for all $x \in X$. The corresponding function is also measurable.

Proof. Let $c \in \mathbb{R}$. Then we have

$$(\sup f_n)^{-1}(c,\infty) = \bigcup_n f_n^{-1}(c,\infty) \in \mathcal{M}.$$

Similarly, we have

$$(\inf f_n)^{-1}(-\infty,c) = \bigcup_n f_n^{-1}(-\infty,c) \in \mathcal{M}.$$

Also we have

$$\limsup f_n = \inf_k \sup_{n \geq k} \in \mathcal{M}.$$

Similarly, we have

$$\liminf f_n = \sup_k \inf_{n \geq k} f_n \in \mathcal{M}$$

□

Problem 2

Proposition 13.8. *If $f_n \xrightarrow{L^1} f$, then $f_n \xrightarrow{\mathrm{m}} f$.*

Proof. Suppose $f_n \xrightarrow{L^1} f$ and let $\varepsilon, \delta > 0$. Choose $N_{\varepsilon,\delta} \in \mathbb{N}$ such that $n \geq N_{\varepsilon,\delta}$ implies

$$\|f_n - f\|_1 < \varepsilon\delta.$$

Then it follows from Chebyshev's inequality that

$$\begin{aligned}
\mu\left(\{\{x \in X \mid |f_n(x) - f(x)| \geq \varepsilon\}\}\right) &\leq \frac{1}{\varepsilon}\|f_n - f\|_1 \\
&< \frac{1}{\varepsilon}\varepsilon\delta \\
&= \delta.
\end{aligned}$$

Thus $f_n \xrightarrow{\mathrm{m}} f$. □

Problem 3

Proposition 13.9. *Let $(X, \mathcal{M}, \mu)$ be a measure space. Suppose $(f_n \colon X \to [0,\infty])$ is a decreasing sequence of nonnegative measurable functions which converges pointwise to $f \colon X \to [0,\infty]$. If $\int_X f_1 \mathrm{d}\mu < \infty$, then*

$$\lim_{n\to\infty} \int_X f_n \mathrm{d}\mu = \int_X f \mathrm{d}\mu. \tag{41}$$

Proof. For each $n \in \mathbb{N}$, set $g_n = f_{n+1} - f_n$. Since (f_n) is a decreasing sequence, each g_n is nonnegative. Furthermore each g_n is measurable since it is a difference of two measurable functions. Set $g = \sum_{n=1}^\infty g_n$ and observe that

$$\begin{aligned}
g &= \sum_{n=1}^{\infty} g_n \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} g_n \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} (f_{n+1} - f_n) \\
&= \lim_{N\to\infty} (f_N - f_1) \\
&= f - f_1.
\end{aligned}$$

It follows from problem 4 that

$$\begin{aligned}
\int_X f \mathrm{d}\mu - \int_X f_1 \mathrm{d}\mu &= \int_X (f - f_1)\mathrm{d}\mu \\
&= \int_X g \mathrm{d}\mu \\
&= \sum_{n=1}^{\infty} \int_X g_n \mathrm{d}\mu \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_X g_n \mathrm{d}\mu \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \int_X (f_{n+1} - f_n)\mathrm{d}\mu \\
&= \lim_{N\to\infty} \int_X \sum_{n=1}^{N} (f_{n+1} - f_n)\mathrm{d}\mu \\
&= \lim_{N\to\infty} \int_X (f_N - f_1)\mathrm{d}\mu \\
&= \lim_{N\to\infty} \int_X f_N \mathrm{d}\mu - \int_X f_1 \mathrm{d}\mu.
\end{aligned}$$

Since $\int_X f_1 \mathrm{d}\mu < 0$, we can cancel it from both sides to get (30). □

Problem 6

Proposition 13.10. *If* $\int_X |f| \mathrm{d}\mu = 0$, *then* $\mu(\{f \neq 0\}) = 0$.

Proof. Note that $\{f \neq 0\} = \{|f| \neq 0\}$. Also $\{|f| \neq 0\} = \bigcup_{n=1}^{\infty} \{|f| \geq 1/n\}$. Thus

$$\begin{aligned}
\mu(\{|f| \neq 0\}) &= \mu\left(\bigcup_{n=1}^{\infty} \{|f| \geq 1/n\}\right) \\
&= \lim_{n\to\infty} \mu(\{|f| \geq 1/n\}) \\
\text{(C-M)} &\leq \lim_{n\to\infty} n \int_X |f| \mathrm{d}\mu \\
&= 0.
\end{aligned}$$

□

Part IV
Appendix

14 Pseudometric Spaces

Recall that a metric space is a pair (X, d) where X is a set and where $\mathrm{d}\colon X \times X \to \mathbb{R}$ is a metric on X which satisfies

1. (Identity of Indiscernibles) $\mathrm{d}(x, y) = 0$ if and only if $x = y$;
2. (Symmetry) $\mathrm{d}(x, y) = \mathrm{d}(y, x)$ for all $x, y \in X$,
3. (Triangle Inequality) $\mathrm{d}(x, z) \leq \mathrm{d}(x, y) + \mathrm{d}(y, z)$ for all $x, y, z \in X$.

If we weaken the "Identity of Indiscernibles" axiom, then we get a pseudometric space:

Definition 14.1. A **pseudometric** on a set X is a function $\mathrm{d}\colon X \times X \to \mathbb{R}$ which satisfies the following three properties:

1. (Reflexivity) $\mathrm{d}(x,x) = 0$ for all $x \in X$;
2. (Symmetry) $\mathrm{d}(x,y) = \mathrm{d}(y,x)$ for all $x, y \in X$,
3. (Triangle Inequality) $\mathrm{d}(x,z) \le \mathrm{d}(x,y) + \mathrm{d}(y,z)$ for all $x, y, z \in X$.

If d is a pseudometric on a set X, then we call the pair (X, d) a **pseudometric space**. If the pseudometric is understood from context, then we often denote a pseudometric space by X instead of (X, d).

Remark 11. Given the three axioms above, we also have $\mathrm{d}(x,y) \ge 0$ for all $x, y \in X$. Indeed,

$$\begin{aligned} 0 &= \mathrm{d}(x,x) \\ &\le \mathrm{d}(x,y) + \mathrm{d}(y,x) \\ &= \mathrm{d}(x,y) + \mathrm{d}(x,y) \\ &= 2\mathrm{d}(x,y). \end{aligned}$$

This implies $\mathrm{d}(x,y) \ge 0$.

14.1 Topology Induced by Pseudometric Space

Proposition 14.1. *Let (X, d) be a pseudometric space. For each $x \in X$ and $r > 0$, define*

$$\mathrm{B}^{\mathrm{d}}_r(x) := \{y \in X \mid \mathrm{d}(x,y) < r\},$$

and let

$$\mathcal{B}^{\mathrm{d}} = \{\mathrm{B}_r(x) \mid x \in X \text{ and } r > 0\}.$$

Finally, let $\tau(\mathcal{B}^{\mathrm{d}})$ be the smallest topology on X which contains $\mathcal{B}^{\mathrm{d}}$. Then $\mathcal{B}^{\mathrm{d}}$ is a basis for $\tau(\mathcal{B}^{\mathrm{d}})$.

Remark 12. We often remove the d in the superscript in $\mathrm{B}^{\mathrm{d}}_r(x)$ and $\mathcal{B}^{\mathrm{d}}$ whenever context is clear.

Proof. First note that $\mathcal{B}$ covers X. Indeed, for any $r > 0$, we have

$$X \subseteq \bigcup_{x \in X} \mathrm{B}_r(x).$$

Next, let $\mathrm{B}_r(x)$ and $\mathrm{B}_{r'}(x')$ be two members of $\mathcal{B}$ which have nontrivial intersection and let $x'' \in \mathrm{B}_r(x) \cap \mathrm{B}_{r'}(x')$. Set

$$r'' = \min\{r' - \mathrm{d}(x', x''), r - \mathrm{d}(x, x'')\}.$$

We claim that $\mathrm{B}_{r''}(x'') \subseteq \mathrm{B}_r(x) \cap \mathrm{B}_{r'}(x')$. Indeed, assume without loss of generality that $r'' = r - \mathrm{d}(x, x'')$. Let $y \in \mathrm{B}_{r''}(x'')$. Then

$$\begin{aligned} \mathrm{d}(y,x) &\le \mathrm{d}(y,x'') + \mathrm{d}(x'',x) \\ &< r - \mathrm{d}(x,x'') + \mathrm{d}(x'',x) \\ &= r - \mathrm{d}(x'',x) + \mathrm{d}(x'',x) \\ &= r \end{aligned}$$

implies $y \in \mathrm{B}_r(x)$. Similarly,

$$\begin{aligned} \mathrm{d}(y,x') &\le \mathrm{d}(y,x'') + \mathrm{d}(x'',x') \\ &< r' - \mathrm{d}(x',x'') + \mathrm{d}(x'',x') \\ &= r' - \mathrm{d}(x'',x') + \mathrm{d}(x'',x') \\ &= r' \end{aligned}$$

implies $y \in \mathrm{B}_{r'}(x')$. Thus $\mathrm{B}_{r''}(x'') \subseteq \mathrm{B}_r(x) \cap \mathrm{B}_{r'}(x')$, and so $\mathcal{B}$ is a basis for $\tau(\mathcal{B})$. □

Definition 14.2. The topology $\tau(\mathcal{B})$ in Proposition (14.1) is called the **topology induced by the pseudometric** d. We also denote this topology by τ_{d}.

14.1.1 Subspace topology agrees with topology induced by pseudometric

Let (X, d) be a pseudometric space and let $A \subseteq X$. Then the pseudometric on X restricts to a pseudometric on A. We denote this restriction by $\mathrm{d}|_A$. Thus there are two natural topologies on A. One is the subspace is the subspace topology given by

$$\tau \cap A := \{U \cap A \mid U \in \tau\}.$$

The other is the topology induced by the pseudometric $\mathrm{d}|_A$ given by

$$\tau_{\mathrm{d}|_A} := \tau(\mathcal{B}^{\mathrm{d}|_A}).$$

The next proposition tells us that these are actually the same.

Proposition 14.2. *Let (X, d) be a pseudometric space and let $A \subseteq X$. Then*

$$\tau_{\mathrm{d}} \cap A = \tau_{\mathrm{d}|_A}.$$

Proof. Let $a \in A$ and $r > 0$. Then

$$\begin{aligned}
\mathrm{B}_r^{\mathrm{d}|_A}(a) &= \{b \in A \mid \mathrm{d}|_A(a, b) < r\} \\
&= \{b \in A \mid \mathrm{d}(a, b) < r\} \\
&= A \cap \{x \in X \mid \mathrm{d}(a, x) < r\} \\
&= A \cap \mathrm{B}_r^{\mathrm{d}}(a).
\end{aligned}$$

It follows that $\tau_{\mathrm{d}|_A}$ and $\tau_{\mathrm{d}} \cap A$ have the same basis, and hence $\tau_{\mathrm{d}} \cap A = \tau_{\mathrm{d}|_A}$. □

14.1.2 Convergence in (X, d)

Concepts like convergence and completion still make sense in pseudometric spaces. Let us state these definitions in the context of pseudometric spaces now.

Definition 14.3. Let (X, d) be a pseudometric space and let (x_n) be a sequence in X.

1. We say the sequence (x_n) converges to $x \in X$ if for all $\varepsilon > 0$ there exists an $N_\varepsilon \in \mathbb{N}$ (we write ε in the subscript because N_ε depends on ε, however we usually omit ε and just write N) such that

   $$n \geq N_\varepsilon \text{ implies } \mathrm{d}(x_n, x) < \varepsilon.$$

   In this case, we say (x_n) is a **convergent sequence** and that it **converges** to x. We denote this by $x_n \to x$ as $n \to \infty$, or $\lim\limits_{n\to\infty} x_n = x$, or even just $x_n \to x$.

2. We say the sequence (x_n) is **Cauchy** if for all $\varepsilon > 0$ there exists an $N_\varepsilon \in \mathbb{N}$ such that

   $$n, m \geq N_\varepsilon \text{ implies } \mathrm{d}(x_n, x_m) < \varepsilon.$$

14.1.3 Completeness in (X, d)

In a metric space, every Cauchy sequence is convergence but the converse may not hold. The same thing is true for pseudometric spaces.

Proposition 14.3. *Let (x_n) be a sequence in X, let $x \in X$, and suppose $x_n \to x$. Then (x_n) is Cauchy.*

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\mathrm{d}(x_n, x) < \varepsilon/2.$$

Then $n, m \geq N$ implies

$$\begin{aligned}
\mathrm{d}(x_n, x_m) &\leq \mathrm{d}(x_n, x) + \mathrm{d}(x, x_m) \\
&< \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon.
\end{aligned}$$

This implies (x_n) is Cauchy. □

Thus, the concept of completeness makes sense in a pseudometric space.

Definition 14.4. Let (X, d) be a pseudometric space. We say (X, d) is **complete** if every Cauchy sequence in (X, d) is a convergent.

14.2 Metric Obtained by Pseudometric

Unless otherwise specified, we let (X, d) be a pseudometric space throughout the remainder of this section. There is a natural way to obtain a metric space from (X, d) which we now describe as follows: define a relation $\sim$ on X by

$$x \sim y \text{ if and only if } \mathrm{d}(x, y) = 0.$$

Then $\sim$ is an equivalence relation. Indeed, we have reflexivity of $\sim$ since $\mathrm{d}(x, x) = 0$ for all $x \in X$, we have symmetry of $\sim$ since $\mathrm{d}(x, y) = \mathrm{d}(y, x)$ for all $x, y \in X$, and we have transitivity of $\sim$ since d satisfies the triangle inequality: if $x \sim y$ and $y \sim z$, then

$$\begin{aligned} \mathrm{d}(x, z) &\leq \mathrm{d}(x, y) + \mathrm{d}(y, z) \\ &= 0 + 0 \\ &= 0. \end{aligned}$$

Thus $\mathrm{d}(x, z) = 0$ which implies $x \sim z$.

Therefore we may consider the quotient space of X with respect to the equivalence relation above. We shall denote this quotient space by $[X] := X/\sim$. A coset in $[X]$ which is represented by $x \in X$ will be written as $[x]$. There is a natural **projection map** $\pi \colon X \to [X]$ that sends $x \in X$ to its equivalence class $[x]$. Since π is surjective, any subset of $[X]$ has the form

$$[A] = \{[a] \in [X] \mid a \in A\}$$

for some $A \subseteq X$. We are ready now to define the metric on $[X]$.

Theorem 14.1. *Define* $[\mathrm{d}] \colon [X] \times [X] \to \mathbb{R}$ *by*

$$[\mathrm{d}]([x], [y]) = \mathrm{d}(x, y) \tag{42}$$

for all $[x], [y] \in [X]$. *Then* $[\mathrm{d}]$ *is a metric on* $[X]$. *It is called the metric* ***induced*** *by the pseudometric* $[\mathrm{d}]$.

Proof. We first show that (42) is well-defined. Indeed, choose different coset representatives of $[x]$ and $[y]$, say x' and y' respectively (so $\mathrm{d}(x, x') = 0$ and $\mathrm{d}(y, y') = 0$). Then

$$\begin{aligned} [\mathrm{d}]([x'], [y']) &= \mathrm{d}(x', y') \\ &\leq \mathrm{d}(x', x) + \mathrm{d}(x, y) + \mathrm{d}(y, y') \\ &= \mathrm{d}(x, y) \\ &= [\mathrm{d}]([x], [y]). \end{aligned}$$

Thus $[\mathrm{d}]$ is well-defined.

Next we show that $[\mathrm{d}]$ is in fact a metric on $[X]$. First we check $[\mathrm{d}]$ is symmetric. Let $[x], [y] \in [X]$. Then

$$\begin{aligned} [\mathrm{d}]([x], [y]) &= \mathrm{d}(x, y) \\ &= \mathrm{d}(y, x) \\ &= [\mathrm{d}]([y], [x]). \end{aligned}$$

Thus $[\mathrm{d}]$ is symmetric. Next we check $[\mathrm{d}]$ satisfies triangle inequality. Let $[x], [y], [z] \in [X]$. Then

$$\begin{aligned} [\mathrm{d}]([x], [z]) &= \mathrm{d}(x, z) \\ &\leq \mathrm{d}(x, y) + \mathrm{d}(y, z) \\ &= [\mathrm{d}]([x], [y]) + [\mathrm{d}]([y], [z]). \end{aligned}$$

Thus $[\mathrm{d}]$ satisfies triangle inequality. Finally we check $[\mathrm{d}]$ satisfies identify of indiscernables. Let $[x], [y] \in [X]$ and suppose $[\mathrm{d}]([x], [y]) = 0$. Then

$$\begin{aligned} 0 &= [\mathrm{d}]([x], [y]) \\ &= \mathrm{d}(x, y) \end{aligned}$$

implies $x \sim y$ by definition. Therefore $[x] = [y]$. Thus $[\mathrm{d}]$ satisfies identify of indiscernables. $\square$

14.2.1 Completeness in (X, d) is equivalent to completeness in $([X], [\mathrm{d}])$

As in the case of the pseudometric d, the metric [d] induces a topology on $[X]$. We denote this topology by $\tau_{[\mathrm{d}]}$.

Proposition 14.4. (X, d) *is complete if and only if* $([X], [\mathrm{d}])$ *is complete.*

Proof. Suppose that (X, d) is complete. Let $([x_n])$ be a Cauchy sequence in $([X], [\mathrm{d}])$. We claim (x_n) is a Cauchy sequence in (X, d). Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies

$$[\mathrm{d}]([x_n], [x_m]) < \varepsilon.$$

Then $m, n \geq N$ implies

$$\begin{aligned} \mathrm{d}(x_n, x_m) &= [\mathrm{d}]([x_n], [x_m]) \\ &< \varepsilon. \end{aligned}$$

This implies (x_n) is a Cauchy sequence in (X, d). Since (X, d) is complete, the sequence converges to a (not necessarily unique) $x \in X$. Then we claim that $[x_n] \to [x]$. Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\mathrm{d}(x_n, x) < \varepsilon.$$

Then $n \geq N$ implies

$$\begin{aligned} [\mathrm{d}]([x_n], [x]) &= \mathrm{d}(x_n, x) \\ &< \varepsilon. \end{aligned}$$

This implies $[x_n] \to [x]$. Thus $([X], [\mathrm{d}])$ is complete.

Conversely, suppose $([X], [\mathrm{d}])$ is complete. Let (x_n) be a Cauchy sequence in (X, d). We claim $([x_n])$ is a Cauchy sequence in $([X], [\mathrm{d}])$. Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies

$$\mathrm{d}(x_n, x_m) < \varepsilon.$$

Then $m, n \geq N$ implies

$$\begin{aligned} [\mathrm{d}]([x_n], [x_m]) &= \mathrm{d}(x_n, x_m) \\ &< \varepsilon. \end{aligned}$$

This implies (x_n) is a Cauchy sequence in $([X], [\mathrm{d}])$. Since $([X], [\mathrm{d}])$ is complete, the sequence converges to a unique $[x] \in [X]$. We claim that $x_n \to x$ (in fact it converges to any $y \in X$ such that $y \sim x$). Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$[\mathrm{d}]([x_n], [x]) < \varepsilon.$$

Then $n \geq N$ implies

$$\begin{aligned} \mathrm{d}(x_n, x) &= [\mathrm{d}]([x_n], [x]) \\ &< \varepsilon. \end{aligned}$$

This implies $x_n \to x$. Thus (X, d) is complete. □

14.3 Quotient Topology

Recall that we view X as a topological space with toplogy τ_{d}; the topology induced by the pseudometric d. It turns out that there are two natural topologies on $[X]$. One such topology is $\tau_{[\mathrm{d}]}$; the topology induced by the metric [d]. The other topology is called the **quotient topology with respect to** $\sim$, and is denoted by $[\tau_{\mathrm{d}}]$, where $[\tau_{\mathrm{d}}]$ is defined by

$$[\tau_{\mathrm{d}}] = \{[A] \subseteq [X] \mid \pi^{-1}([A]) \in \tau_{\mathrm{d}}\}.$$

In other words, we declare a subset $[A]$ of $[X]$ to be $[\tau_{\mathrm{d}}]$-open in $[X]$ if and only if

$$\begin{aligned} \pi^{-1}([A]) &= \{x \in X \mid x \sim a \text{ for some } a \in A\} \\ &= \{x \in X \mid \mathrm{d}(x, a) = 0 \text{ for some } a \in A\} \end{aligned}$$

is open in X. Since $\pi^{-1}(\emptyset) = \emptyset$ and $\pi^{-1}([X]) = X$, we see that both $\emptyset$ and $[X]$ are open in $[X]$. Furthermore, since

$$\pi^{-1}\left(\bigcup_{i \in I}[A_i]\right) = \bigcup_{i \in I} \pi^{-1}([A_i]) \text{ and } \pi^{-1}\left(\bigcap_{i \in I}[A_i]\right) = \bigcap_{i \in I} \pi^{-1}([A_i]),$$

we see that the collection of open sets in $[X]$ is closed under arbitrary unions and finite intersections. Therefore $[\tau_{\mathrm{d}}]$ is indeed a topology on $[X]$. Note that $[\tau_{\mathrm{d}}]$ was defined in such a way that it makes the projection map $\pi\colon X \to [X]$ continous.

14.3.1 Universal Mapping Property For Quotient Space

Quotient spaces satisfy the following universal mapping property.

Proposition 14.5. *Let $f\colon X \to Y$ be any continuous function which is constant on each equivalence class. Then there exists a unique continuous function $[f]\colon [X] \to Y$ such that $f = [f] \circ \pi$.*

Proof. We define $[f]\colon [X] \to Y$ by

$$[f]([x]) = f(x) \tag{43}$$

for all $x \in X$. We first show that (43) is well-defined. Suppose x and x' are two different representatives of the same coset (so $x \sim x'$). Then $f(x) = f(x')$ as f was assumed to be constant on equivalence classes, and so

$$\begin{aligned} [f]([x']) &= f(x') \\ &= f(x) \\ &= [f]([x]). \end{aligned}$$

Thus (43) is well-defined.

Next we want to show that $[f]$ is continuous. Let V be an open set in Y. Then

$$f^{-1}(V) = \pi^{-1}\left([f]^{-1}(V)\right)$$

is open in X. By the definition of quotient topology, this implies $[f]^{-1}(V)$ is open in $[X]$. This implies $[f]$ is continuous.

Finally, we want to show that $f = [f] \circ \pi$ holds. Let $x \in X$. Then we have

$$\begin{aligned} ([f] \circ \pi)(x) &= [f](\pi(x)) \\ &= [f]([x]) \\ &= f(x). \end{aligned}$$

It follows that $[f] \circ \pi = f$. This establishes existence of f.

For uniqueness, assume for a contradiction that $\overline{f}\colon [X] \to Y$ is a continuous function such that $f = \overline{f} \circ \pi$ and such that $\overline{f} \neq [f]$. Choose $[x] \in [X]$ such that $\overline{f}[x] \neq [f][x]$. Then

$$\begin{aligned} f(x) &= (\overline{f} \circ \pi)(x) \\ &= \overline{f}(\pi(x)) \\ &= \overline{f}([x]) \\ &\neq [f]([x]) \\ &= f(x), \end{aligned}$$

which gives us a contradiction. □

It follows from Proposition (14.5) that we have the following bijection of sets

$$\{f\colon X \to Y \mid f \text{ is continuous and constant on equivalence classes}\} \cong \{\text{continuous functions from } [X] \text{ to } Y\}.$$

In particular, if we want to study continuous functions out of $[X]$, then we just need to study the continuous functions out of X which are constant on equivalence classes.

Proposition 14.6. *Suppose (Y, d_Y) is a metric space and $f\colon (X, \mathrm{d}) \to (Y, \mathrm{d}_Y)$ is continuous. The f is constant on equivalence classes.*

Proof. Let $x, x' \in X$ such that $x \sim x'$. Thus $\mathrm{d}(x, x') = 0$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that

$$\mathrm{d}(x, y) < \delta \text{ implies } \mathrm{d}_Y(f(x), f(y)) < \varepsilon.$$

We want to show that $f(x) = f(x')$. □

14.3.2 Open Equivalence Relation

An equivalence relation $\sim$ on a topological space X is said to be **open** if the projection map $\pi\colon X \to [X]$ is open. In other words, the equivalence relation $\sim$ on X is open if and only if for every open set U in X, the set

$$\pi^{-1}(\pi(U)) = \bigcup_{x \in U} [x]$$

of all points equivalent to some point of U is open. The imporance of open equivalence relations is that if $\mathscr{B}$ is a basis for X, then $[\mathscr{B}]$ is a basis for $[X]$.

Lemma 14.2. *Let $x \in X$ and $r > 0$. Then*

$$\mathrm{B}_r(x) = \pi^{-1}([\mathrm{B}_r(x)]).$$

In particular, π is an open mapping.

Proof. We have

$$\begin{aligned} \mathrm{B}_r(x) &\subseteq \pi^{-1}\left(\pi\left(\mathrm{B}_r(x)\right)\right) \\ &= \pi^{-1}([\mathrm{B}_r(x)]). \end{aligned}$$

For the reverse inclusion, let $y \in \pi^{-1}([\mathrm{B}_r(x)])$. Then $\mathrm{d}(y,z) = 0$ for some $z \in \mathrm{B}_r(x)$. Choose such a $z \in \mathrm{B}_r(x)$. Then

$$\begin{aligned} \mathrm{d}(y,x) &\leq \mathrm{d}(y,z) + \mathrm{d}(z,x) \\ &= \mathrm{d}(z,x) \\ &< r \end{aligned}$$

implies implies $y \in \mathrm{B}_r(x)$. Therefore

$$\pi^{-1}([\mathrm{B}_r(x)]) \subseteq \mathrm{B}_r(x).$$

Thus each subset in $[X]$ of the form $[\mathrm{B}_r(x)]$ is open in $[X]$.

To see that π is an open mapping, let U be an open set in X. Since the set of all open balls is a basis for τ_{d}, we can cover U by open balls, say

$$U = \bigcup_{i \in I} \mathrm{B}_{r_i}(x_i).$$

Then

$$\begin{aligned} \pi(U) &= \pi\left(\bigcup_{i \in I} \mathrm{B}_{r_i}(x_i)\right) \\ &= \bigcup_{i \in I} \pi\left(\mathrm{B}_{r_i}(x_i)\right) \\ &= \bigcup_{i \in I} [\mathrm{B}_{r_i}(x_i)] \\ &\in [\tau_{\mathrm{d}}]. \end{aligned}$$

Thus π is an open mapping. □

14.3.3 Quotient Topology Agrees With Metric Topology

Theorem 14.3. *With the notation as above, we have*

$$[\tau_{\mathrm{d}}] = \tau_{[\mathrm{d}]}.$$

Proof. We first note that for each $x \in X$ and $r > 0$, we have

$$\begin{aligned} [\mathrm{B}_r(x)] &= \{[y] \in [X] \mid y \in \mathrm{B}_r(x)\} \\ &= \{[y] \in [X] \mid \mathrm{d}(y,x) < r\} \\ &= \{[y] \in [X] \mid [\mathrm{d}]([y],[x]) < r\} \\ &= \mathrm{B}_r([x]). \end{aligned}$$

In particular, $\tau_{[\mathrm{d}]}$ and $[\tau_{\mathrm{d}}]$ share a common basis. Therefore $\tau_{[\mathrm{d}]} = [\tau_{\mathrm{d}}]$. □

15 Completing a Normed Linear Space

Recall from linear analysis that if $(\mathcal{X}, \|\cdot\|)$ is a normed linear space, then $\mathcal{X}$[5] may or may not be complete with respect to the topology induced by $\|\cdot\|$. In other words, there may exist a Cauchy sequence in $\mathcal{X}$ which does not converge in $\mathcal{X}$. For instance, $(\mathbb{Q}, |\cdot|)$ is not complete. Indeed, the sequence

$$3, 3.1, 3.14, 3.145, \dots$$

is Cauchy but does not converge in $\mathbb{Q}$; it converges to π which is irrational. All is not lost however because we can always **complete** a normed linear space:

Theorem 15.1. *Every normed linear space can be completed. More precisely, if $(X, \|\cdot\|)$ is a normed linear space, then there exists a normed linear space $(\widetilde{\mathcal{X}}, \widetilde{\|\cdot\|})$ such that*

1. *$(\widetilde{\mathcal{X}}, \widetilde{\|\cdot\|})$ is a complete normed linear space (that is, it is a Banach space).*
2. *There exists a isometric embedding $\iota \colon (\mathcal{X}, \|\cdot\|) \to (\widetilde{\mathcal{X}}, \widetilde{\|\cdot\|})$ with dense image. This means that*
   (a) ι is a linear map which preserves norms: $\|\iota(x)\| = \widetilde{\|x\|}$ for all $x \in \mathcal{X}$. In particular, this implies ι is injective;
   (b) the closure of $\iota(\mathcal{X})$ in $\widetilde{\mathcal{X}}$ is equal to all of $\widetilde{\mathcal{X}}$.

*We call $(\widetilde{\mathcal{X}}, \widetilde{\|\cdot\|})$ the **completion** of $(\mathcal{X}, \|\cdot\|)$.*

15.1 Constructing Completions

Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. We denote by $\mathscr{C}_{\mathcal{X}}$ to be the set of all Cauchy sequences in $\mathcal{X}$. We can give $\mathscr{C}_{\mathcal{X}}$ the structure of a $\mathbb{C}$-vector space as follows: let $(x_n), (y_n) \in \mathscr{C}_{\mathcal{X}}$ and let $\lambda, \mu \in \mathbb{C}$. Then we define

$$\lambda(x_n) + \mu(y_n) := (\lambda x_n + \mu y_n). \tag{44}$$

One easily checks that scalar multiplication and addition defined as in (44) gives $\mathscr{C}_{\mathcal{X}}$ the structure of a $\mathbb{C}$-vector space.

15.1.1 Seminorm

A natural contender for a norm on $\mathscr{C}_{\mathcal{X}}$ is the map $\|\cdot\| \colon \mathscr{C}_{\mathcal{X}} \to \mathbb{C}$ defined by

$$\|(x_n)\| := \lim_{n\to\infty} \|x_n\| \tag{45}$$

for all $(x_n) \in \mathscr{C}_{\mathcal{X}}$. In fact, (45) will not be a norm, but rather a seminorm[6]. Before we explain this however, let us first show that the righthand side of (45) converges in $\mathbb{C}$.

Proposition 15.1. *Let (x_n) be a Cauchy sequence of vectors in $\mathcal{X}$. Then $(\|x_n\|)$ is a Cauchy sequence of complex numbers in $\mathbb{C}$. In particular, (45) converges in $\mathbb{C}$.*

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies

$$\|x_n - x_m\| < \varepsilon.$$

Then $m, n \geq N$ implies

$$\begin{aligned} |\|x_n\| - \|x_m\|| &\leq \|x_n - x_m\| \\ &< \varepsilon. \end{aligned}$$

It follows that $(\|x_n\|)$ is a Cauchy sequence of complex numbers in $\mathbb{C}$. The latter statement in the proposition follows from the fact that $\mathbb{C}$ is complete. $\square$

[5]To simplify notation, we often write $\mathcal{X}$ rather than $(\mathcal{X}, \|\cdot\|)$. Context will make it clear which norm we are equipping $\mathcal{X}$ with.
[6]See Appendix for notes on pseudonorms

15.1.2 Quotienting Out To get an Inner-Product

As mentioned above, (45) is not a norm. It is what's called a seminorm:

Definition 15.1. Let V be a vector space over $\mathbb{C}$. A map $\|\cdot\|\colon V \to \mathbb{C}$ is called a **seminorm** if it satisfies the following properties:

1. Absolute homogeneity: $\|\lambda x\| = |\lambda|\|x\|$ for all $x \in V$ and $\lambda \in \mathbb{C}$;
2. Subadditivity: $\|x+y\| \leq \|x\| + \|y\|$ for all $x, y \in V$;
3. Semipositive definiteness: $\|x\| \geq 0$ for all $x \in V$.

To see why (45) is a seminorm, note that absolute homogeneity and subadditivity are clear. What makes (45) a seminorm and not a norm is that we have only have semipositive definiteness:

$$\begin{aligned}\|(x_n)\| &= \lim_{n\to\infty} \|x_n\| \\ &\geq 0.\end{aligned}$$

In particular, we may have $\|x_n\| \to 0$ with $(x_n) \neq 0$. To remedy this situation, we define

$$\mathscr{C}_{\mathcal{X}}^0 := \{(x_n) \in \mathscr{C}_{\mathcal{X}} \mid \|x_n\| \to 0\}.$$

Then $\mathscr{C}_{\mathcal{X}}^0$ is a subspace of $\mathscr{C}_{\mathcal{X}}$. Indeed, if $\lambda, \mu \in \mathbb{C}$ and $(x_n), (y_n) \in \mathscr{C}_{\mathcal{X}}^0$, then

$$\begin{aligned}\|\lambda x_n + \mu y_n\| &\leq \|\lambda x_n\| + \|\mu y_n\| \\ &= |\lambda|\|x_n\| + |\mu|\|y_n\| \\ &\to 0\end{aligned}$$

and hence $(\lambda x_n + \mu y_n) \in \mathscr{C}_{\mathcal{X}}^0$. Therefore we obtain a quotient space $\mathscr{C}_{\mathcal{X}}/\mathscr{C}_{\mathcal{X}}^0$. Now we claim that the pseudo-norm (45) induces a genuine norm, which we denote again by $\|\cdot\|$, on $\mathscr{C}_{\mathcal{X}}/\mathscr{C}_{\mathcal{X}}^0$, defined by

$$\|\overline{(x_n)}\| := \lim_{n\to\infty} \|x_n\|. \tag{46}$$

for all $\overline{(x_n)}$[7] in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}_{\mathcal{X}}^0$. We need to be sure that (46) is well-defined. Let (x_n') be different representative of the cosets $\overline{(x_n)}$ (so $\|x_n' - x_n\| \to 0$). Then

$$\begin{aligned}\|(x_n')\| &= \lim_{n\to\infty} \|x_n'\| \\ &= \lim_{n\to\infty} \|x_n' + x_n - x_n\| \\ &\leq \lim_{n\to\infty} \|x_n\| + \lim_{n\to\infty} \|x_n' - x_n\| \\ &= \lim_{n\to\infty} \|x_n\| \\ &= \|(x_n)\|.\end{aligned}$$

Thus (46) is well-defined (meaning it is independent of the choice of representatives of cosets).

Now absolute homogeneity of (46) and subadditivity of (46) are clear. This time however, we also have positive definiteness: if $\overline{(x_n)} \in \mathscr{C}_{\mathcal{X}}/\mathscr{C}_{\mathcal{X}}^0$ such that $\|\overline{(x_n)}\| = 0$, then

$$\begin{aligned}\lim_{n\to\infty} \|x_n\| &= \|\overline{(x_n)}\| \\ &= 0,\end{aligned}$$

which implies $\overline{(x_n)} = 0$ in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}_{\mathcal{X}}^0$.

[7] When we write $\overline{(x_n)}$ for a coset in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}_{\mathcal{X}}^0$, then it is implicitly understood that (x_n) is an element $\mathscr{C}_{\mathcal{X}}$ which represents the coset $\overline{(x_n)}$ in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}_{\mathcal{X}}^0$.

15.1.3 The map $\iota\colon \mathcal{X} \to \mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$

Let $\iota\colon \mathcal{X} \to \mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$ be defined by

$$\iota(x) = \overline{(x)}$$

for all $x \in \mathcal{X}$, where (x) is a constant sequence in $\mathscr{C}_{\mathcal{X}}$.

Definition 15.2. An **isometry** between normed linear space $(\mathcal{X}_1, \|\cdot\|_1)$ and $(\mathcal{X}_2, \|\cdot\|_2)$ is an operator $T\colon \mathcal{X}_1 \to \mathcal{X}_2$ such that

$$\|Tx\|_2 = \|x\|_1$$

for all $x \in \mathcal{X}_1$.

Remark 13. Note that an isometry is automatically injective. Indeed, let $x \in \ker T$. Then

$$\begin{aligned}\|x\| &= \|Tx\| \\ &= \|0\| \\ &= 0\end{aligned}$$

implies $x = 0$. Thus we sometimes call an isometry an **isometric embedding**.

Proposition 15.2. *The map $\iota\colon \mathcal{X} \to \mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$ is an isometry.*

Proof. Linearity of ι is clear. Let $x \in \mathcal{X}$. Then

$$\begin{aligned}\|\iota(x)\| &= \|\overline{(x)}\| \\ &= \lim_{n\to\infty} \|x\| \\ &= \|x\|.\end{aligned}$$

Thus ι is an isometry. □

Proposition 15.3. *The image of $\mathcal{X}$ under ι is dense in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$. In other words, the closure of $\iota(\mathcal{X})$ in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$ is all of $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$.*

Proof. Let $\overline{(x_n)}$ be a coset in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$. To show that the closure of $\iota(\mathcal{X})$ is is all of $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$, we construct a sequence of cosets in $\iota(\mathcal{X})$ which converges to $\overline{(x_n)}$. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies

$$\|x_n - x_m\| < \varepsilon/2.$$

Then $n, m \geq N$ implies

$$\begin{aligned}\|\iota(x_m) - \overline{(x_n)}\| &= \lim_{n\to\infty} \|x_m - x_n\| \\ &< \lim_{n\to\infty} \varepsilon \\ &= \varepsilon.\end{aligned}$$

Thus, $(\iota(x_m))$ is a sequence of cosets in $\iota(\mathcal{X})$ which converges to $\overline{(x_n)}$. □

15.1.4 $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$ is Complete

Proposition 15.4. *$\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$ is a Banach space.*

Proof. Let $(\overline{x}^n)$ be a Cauchy sequence of cosets in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$ where

$$\overline{x}^n = \overline{(x^n_k)}_{k\in\mathbb{N}}$$

for each $n \in \mathbb{N}$. Throughout the remainder of this proof, let $\varepsilon > 0$.

Since each $x^n = (x^n_k)_{k\in\mathbb{N}}$ is a Cauchy sequence of elements in $\mathcal{X}$, there exists a $\pi(n) \in \mathbb{N}$ such that $k, l \geq \pi(n)$ implies

$$\|x^n_k - x^n_l\| < \frac{\varepsilon}{3}.$$

For each $n \in \mathbb{N}$, choose such $\pi(n) \in \mathbb{N}$ in such a way so $\pi(n) \geq \pi(m)$ whenever $n \geq m$.

Step 1: We show that the sequence $(x^n_{\pi(n)})$ of elements in $\mathcal{X}$ is a Cauchy sequence. Since $(\overline{x}^n)$ is a Cauchy sequence of cosets in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$, there exists an $N \in \mathbb{N}$ such that $n \geq m \geq N$ implies

$$\|(\overline{x}^n) - (\overline{x}^m)\| = \lim_{k\to\infty} \|x^n_k - x^m_k\| < \frac{\varepsilon}{4}. \tag{47}$$

Choose such an $N \in \mathbb{N}$. It follows from (47) that for each $n \geq m \geq N$, there exists $\pi(n,m) \geq \pi(n)$ such that

$$\|x^n_k - x^m_k\| < \frac{\varepsilon}{3}$$

for all $k \geq \pi(n,m)$. Choose such $\pi(n,m)$ for each $n \geq m \geq N$. Then if $n \geq m \geq N$, we have

$$\begin{aligned}
\|x^n_{\pi(n)} - x^m_{\pi(m)}\| &= \|x^n_{\pi(n)} - x^n_{\pi(n,m)} + x^n_{\pi(n,m)} - x^m_{\pi(n,m)} + x^m_{\pi(n,m)} - x^m_{\pi(m)}\| \\
&\leq \|x^n_{\pi(n)} - x^n_{\pi(n,m)}\| + \|x^n_{\pi(n,m)} - x^m_{\pi(n,m)}\| + \|x^m_{\pi(n,m)} - x^m_{\pi(m)}\| \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\
&= \varepsilon.
\end{aligned}$$

Therefore $(x^n_{\pi(n)})$ is a Cauchy sequence of elements in $\mathcal{X}$ and hence represents a coset $\overline{(x^n_{\pi(n)})}$ in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$.

Step 2: Let $x = (x^k_{\pi(k)})$[8]. We want to show that the sequence $(\overline{x}^n)$ of cosets in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$ converges to the coset $\overline{x}$ in $\mathscr{C}_{\mathcal{X}}/\mathscr{C}^0_{\mathcal{X}}$. In particular, we need to find an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\|\overline{x}^n - \overline{x}\| = \lim_{k\to\infty} \|x^n_k - x^k_{\pi(k)}\| < \varepsilon$$

or in other words, $n \geq N$ implies

$$\|x^n_k - x^k_{\pi(k)}\| < \varepsilon$$

for all k sufficiently large. Since x is a Cauchy sequence of elements in $\mathcal{X}$, there exists an $M \in \mathbb{N}$ such that $n, k \geq M$ implies

$$\|x^n_{\pi(n)} - x^k_{\pi(k)}\| < 2\varepsilon/3.$$

Choose such an $M \in \mathbb{N}$. Then $n \geq M$ implies

$$\begin{aligned}
\|x^n_k - x^k_{\pi(k)}\| &\leq \|x^n_k - x^n_{\pi(n)}\| + \|x^n_{\pi(n)} - x^k_{\pi(k)}\| \\
&< \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}$$

for all $k \geq \max\{M, \pi(n)\}$. □

[8]Note the change in index from n to k.