An Introduction to Measure Theory

Hsin-Miao Chao

E-mail: grammcyao@icloud.com

Update: Monday 7th February, 2022

Contents

CHAPTER I

Measure theory

§1.1 Prologue: The problem of measure

1.1.1 Elementary measure

Before we discuss Jordan measure, we discuss the even simpler notion of *elementary measure*, which allows one to measure a very simple class of sets, namely the *elementary sets* (finite unions of boxes).

Definition 1.1.1 (Intervals, boxes, elementary sets)**.** An *interval* is a subset of $\mathbf{R}$ of the form $[a,b] := \{x \in \mathbf{R} : a \leq x \leq b\}$, $[a,b) := \{x \in \mathbf{R} : a \leq x < b\}$, $(a,b] := \{x \in \mathbf{R} : a < x \leq b\}$, or $(a,b) := \{x \in \mathbf{R} : a < x < b\}$, where $a \leq b$ are real numbers. We define the *length* $|I|$ of an interval $I = [a,b], [a,b), (a,b], (a,b)$ to be $|I| := b - a$. A *box* in $\mathbf{R}^d$ is a Cartesian product $B := \prod_{i=1}^d I_i$ of d intervals $I_1, \cdots, I_d$, thus for instance, an interval is a one-dimensional box. The *volume* $|B|$ of such a box B is defined as $|B| := \prod_{i=1}^d |I_i|$. An *elementary set* is any subset of $\mathbf{R}^d$ which is the union of a finite number of boxes.

Proposition 1.1.2 (Boolean closure). *If $E, F \subset \mathbf{R}^d$ are elementary sets, then the union $E \cup F$, the intersection $E \cap F$, and the set theoretic difference $E \setminus F := \{x \in E : x \notin F\}$, and the symmetric difference $E \triangle F := (E \setminus F) \cup (F \setminus E)$ are also elementary. If $x \in \mathbf{R}^d$, then the translate $E + x := \{y + x : y \in E\}$ is also an elementary set.*

Proof. Suppose that E and F can be represented as the finite union of $A_1, \cdots, A_m$ and the finite union of $B_1, \cdots, B_n$, respectively, where $m, n \leq d$.

Then we have

$$E \cup F = \Big(\bigcup_{i=1}^{m} A_i\Big) \cup \Big(\bigcup_{j=1}^{n} B_j\Big).$$

Thus $E \cup F$ is an elementary set for that it equals to the union of $m + n$ boxes.

For the intersection $E \cap F$, we have

$$E \cap F = \Big(\bigcup_{i=1}^{m} A_i\Big) \cap \Big(\bigcup_{j=1}^{n} B_j\Big) = \bigcup_{(i,j)\in\{1,\cdots,m\}\times\{1,\cdots,n\}} (A_i \cap B_j).$$

Let $A_i := I_{i,1} \times \cdots \times I_{i,d}$ and $B_j := J_{j,1} \times \cdots \times J_{j,d}$. We have

$$\begin{aligned} A_i \cap B_j &= (I_{i,1} \times \cdots \times I_{i,d}) \cap (J_{j,1} \times \cdots \times J_{j,d}) \\ &= (I_{i,1} \cap J_{j,1}) \times \cdots \times (I_{i,d} \cap J_{j,d}). \end{aligned}$$

Thus $A_i \cap B_j$ are boxes for all $(i, j) \in \{1, \cdots, m\} \times \{1, \cdots, n\}$. Therefore, $E \cap F$ is an elementary set for that it equals to the union of $m \times n$ boxes.

For the difference $E \setminus F$, for any $1 \leq j \leq n$ we have

$$E \setminus B_j = \Big(\bigcup_{i=1}^{m} A_i\Big) \setminus B_j = \bigcup_{i=1}^{m} (A_i \setminus B_j).$$

We want to show that if A and B are boxes, then $A \setminus B$ is elementary. For this, let $A := I_1 \times \cdots \times I_d$ and $B := J_1 \times \cdots \times J_d$. Then

$$A \setminus B = A \cap B^c$$

$$\begin{aligned} &= A \cap (K_1 \times \cdots \times K_d) \\ &= (I_1 \cap K_1) \times \cdots \times (I_d \cap K_d), \end{aligned}$$

where K_n equals to the union of bounded or unbounded intervals, then $I_n \cap K_n$ equals to the union of bounded intervals. Hence $A \setminus B$ equals to the union of some Cartesian sets, so that elementary.

Thus $E \setminus B_j$ is an elementary. Since

$$E \setminus F = E \setminus \Big(\bigcup_{j=1}^{n} B_j \Big) = \bigcap_{j=1}^{n} (E \setminus B_j),$$

by conclusion above, $E \setminus F$ is an elementary.

The symmetric difference $E \triangle F$ is an elementary set is immediately comes from above conclusions.

For the translation $E + x$, this is easy to see that

$$E + x = \bigcup_{i=1}^{m} (A_i + x),$$

where $A_i + x$ are boxes for all $1 \le i \le m$. □

We now give each elementary set a measure.

Lemma 1.1.3 (Measure of an elementary set)**.** *Let $E \subset \mathbf{R}^d$ be an elementary set.*

(i) *E can be expressed as the finite union of disjoint boxes.*

(ii) *If E is partitioned as the finite union $B_1 \cup \cdots \cup B_k$ of disjoint boxes, then the quantity $m(E) := |B_1| + \cdots + |B_k|$ is independent of the partition. In other words, given any partition $B_1' \cup \cdots \cup B_{k'}'$ of E, one has $|B_1| + \cdots + |B_k| = |B_1'| + \cdots + |B_{k'}'|$.*

We refer to $m(E)$ as the elementary measure of E.

Proof. We first prove (i) in the one-dimensional case $d = 1$. Given any finite collection of intervals $I_1, \cdots, I_k$, one can place the $2k$ endpoints of these intervals in increasing order (discarding repetitions). Looking at the open intervals between these endpoints, together with the endpoints themselves (viewed as intervals of length zero), we see that there exists a finite collection of disjoint intervals $J_1, \cdots, J_{k'}$ such that each of the $I_1, \cdots, I_k$ are a union of some subcollection of the $J_1, \cdots, J_{k'}$. This already gives (i) when $d = 1$. To prove the higher dimensional case, we express E as the union $B_1, \cdots, B_k$ of boxes $B_i = I_{i,1} \times \cdots \times I_{i,d}$. For each $j = 1, \cdots, d$, we use the one-dimensional argument to express $I_{1,j}, \cdots, I_{k,j}$ as the union of subcollections of a collection $J_{1,j}, \cdots, J_{k'_j,j}$ of disjoint intervals. Taking Cartesian products, we can express the $B_1, \cdots, B_k$ as finite unions of boxes $J_{i_1,1} \times \cdots \times J_{i_d,d}$, where $1 \le i_j \le k'_j$ for all $1 \le j \le d$. Such boxes are all disjoint, and the claim follows.

To prove (ii) we let $\mathcal{B}$ be the collection of $B_1, \cdots, B_k$ and $\mathcal{B}'$ be the collection of $B'_1, \cdots, B'_{k'}$. Then we can define the collection

$$\mathcal{B}\#\mathcal{B}' := \{B_i \cap B'_j : B_i \in \mathcal{B} \text{ and } B'_j \in \mathcal{B}'\}.$$

Define $|\mathcal{B}| := \sum_{B_i \in \mathcal{B}} |B_i|$, and define $|\mathcal{B}'|$ and $|\mathcal{B}\#\mathcal{B}'|$ by similar process. Now we want to show that $|\mathcal{B}| = |\mathcal{B}\#\mathcal{B}'|$ and $|\mathcal{B}'| = |\mathcal{B}\#\mathcal{B}'|$, this implies that $|\mathcal{B}| = |\mathcal{B}'|$.

Since

$$B_i = B_i \cap E = B_i \cap \left(\bigcup_{j=1}^{n} B'_j \right) = \bigcup_{j=1}^{n} (B_i \cap B'_j),$$

so that

$$E = \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} (B_i \cap B'_j),$$

where $B_i \cap B'_j$ are disjoint boxes for all $1 \le i \le m, 1 \le j \le n$. Then B_i can be expressed as the finite union of disjoint boxes $B_i \cap B'_j$, and we have

$$|B_j| = \sum_{j=1}^{n} |B_i \cap B'_j|.$$

Thus

$$m(E) = |B_1| + \cdots + |B_k| = \sum_{i=1}^{m} \sum_{j=1}^{n} |B_i \cap B'_j|,$$

i.e., $|\mathcal{B}| = |\mathcal{B}\#\mathcal{B}'|$. A similar argument shows that $|\mathcal{B}'| = |\mathcal{B}\#\mathcal{B}'|$. Thus $|\mathcal{B}| = |\mathcal{B}'|$, as desired. □

From definitions, the elementary measure obeys following properties:

Proposition 1.1.4 (The properties of elementary measure)**.**

(i) $m(\emptyset) = 0$.

(ii) *For all boxes* B*, we have* $m(B) = |B|$.

(iii) *(Non-negativity) For every elementary set* E*, we have* $m(E) \geq 0$.

(iv) *For every* E *and* F *are disjoint elementary sets, we have* $m(E \cup F) = m(E) + m(F)$.

(v) *(Finite additivity) Let* $E_1, \cdots, E_k$ *be a finite sequence of disjoint elementary sets, then* $m(\bigcup_{i=1}^k E_i) = \sum_{i=1}^k m(E_i)$.

(vi) *(Monotonicity) For every elementary sets* $E \subset F$*, we have* $m(E) \leq m(F)$.

(vii) *For arbitrary elementary sets* E *and* F*, we have* $m(E \cup F) \leq m(E) + m(F)$.

(viii) *(Finite subadditivity) Let* $E_1, \cdots, E_k$ *be a finite sequence of arbitrary elementary sets, then* $m(\bigcup_{i=1}^k E_i) \leq \sum_{i=1}^k m(E_i)$.

(ix) *(Translation invariance) For all elementary set* E *and* $x \in \mathbf{R}^d$*, we have* $m(E + x) = m(E)$.

Proof. Proof omitted. □

These properties in fact define elementary measure up to normalisation:

Theorem 1.1.5 (Uniqueness of elementary measure)**.** *Let* $d \geq 1$*. Let* $m' : \mathcal{E}(\mathbf{R}^d) \to \mathbf{R}^+$ *be a map from the collection* $\mathcal{E}(\mathbf{R}^d)$ *of elementary subsets of* $\mathbf{R}^d$ *to the non-negative reals that obeys the non-negativity, finite additivity, and translation invariance properties. Then there exists* *a constant $c \in \mathbf{R}^+$ such that $m'(E) = cm(E)$ for all elementary sets E. In particular, if we impose the additional normalisation $m'([0,1)^d) = 1$, then $m' \equiv m$.*

Proof. We first prove the statement in the one-dimensional case $d = 1$. This will give an intuition about the proof.

Let m' be a map from $\mathcal{E}(\mathbf{R})$ to $\mathbf{R}^+$ which obeys non-negativity, finite additivity, and translation invariance. Set $c := m'([0,1))$. Since $[0,1) = \bigcup_{k=1}^{n}[\frac{k-1}{n}, \frac{k}{n})$ where $[\frac{k-1}{n}, \frac{k}{n})$ are disjoint intervals for all $k \in [0,n]$, and we also have $[0, \frac{1}{n}) = [\frac{k-1}{n}, \frac{k}{n}) - \frac{k-1}{n}$, hence by finite additivity and translation invariance, we have

$$\begin{aligned} m'([0,1)) &= m'\Big(\bigcup_{k=1}^{n}\Big[\frac{k-1}{n}, \frac{k}{n}\Big)\Big) \\ &= \sum_{k=1}^{n} m'\Big(\Big[\frac{k-1}{n}, \frac{k}{n}\Big)\Big) \\ &= n \times m'([0, \frac{1}{n})). \end{aligned}$$

Thus for every $E := [0, \frac{1}{n}) \in \mathcal{E}(\mathbf{R})$, there is a $c = m'([0,1))$ such that $m'(E) = m'([0,1)) \times \frac{1}{n} = cm(E)$.

Now we extend $[0, \frac{1}{n})$ to rational case $[0, \frac{p}{q})$, where $p, q \in \mathbf{Z}^+$. Because $[0,p) = \bigcup_{k=1}^{q}[\frac{(k-1)p}{q}, \frac{kp}{q})$ and $[0, \frac{p}{q}) = [\frac{(k-1)p}{q}, \frac{kp}{q}) - \frac{(k-1)p}{q}$, by the finite additivity and translation invariance, we have

$$\begin{aligned} m'([0,p)) &= m'\Big(\bigcup_{k=1}^{q}\Big[\frac{(k-1)p}{q}, \frac{kp}{q}\Big)\Big) \\ &= \sum_{k=1}^{q} m'\Big(\Big[\frac{(k-1)p}{q}, \frac{kp}{q}\Big)\Big) \\ &= q \times m'([0, \frac{p}{q})). \end{aligned}$$

Thus $m'([0, \frac{p}{q})) = \frac{1}{q} m'([0,p))$. By finite additivity and translation invariance,

$$m'([0,p)) = m'\Big(\bigcup_{k=1}^{p}[k-1,k)\Big) = \sum_{k=1}^{p} m'([k-1,k)) = p \times m'([0,1)).$$

Thus for $E = [0, \frac{p}{q}) \in \mathcal{E}(\mathbf{R})$ we have $m'(E) = \frac{p}{q} m'([0,1)) = cm(E)$.

Furthermore, we extend to real interval $[0, x]$ where $x \in \mathbf{R}^+$. From non-negativity and finite additivity, we conclude the monotonicity property $m'(E) \leq m'(F)$ whenever $E \subset F$. Then for every $x \in \mathbf{R}^+$, there is a $r, s \in \mathbf{Q}^+$ where $0 < r < s$ such that $r \leq x \leq s$. ($\mathbf{Q}$ is dense in $\mathbf{R}$.) By monotonicity and conclusion above, we have

$$m'([0, r)) \leq m'([0, x)) \leq m'([0, s)),$$

this implies

$$cr \leq m'([0, x)) \leq cs.$$

Recall that we define reals as the limit of Cauchy sequence of rationals, then for $\varepsilon > 0$, we have

$$c(x - \varepsilon) \leq cr \leq m'([0, x)) \leq cs \leq c(x + \varepsilon).$$

Since ε is arbitrary, we have $m'([0, x)) = cx = cm([0, x))$. This is easy to see that every $E \in \mathcal{E}(\mathbf{R})$ can be expressed as the disjoint union of real intervals, using finite additivity and translation invariance, we have $m'(E) = cm(E)$ for all elementary set E, as desired.

Now we briefly show the higher dimensional case. Set $c := m'([0, 1)^d)$. Consider d-dimensional box $[0, 1)^d \in \mathcal{E}(\mathbf{R}^d)$, by Definition 1.1.1, it can be expressed as the Cartesian product of intervals $[0, 1)$. We dividing $[0, 1)$ into n disjoint intervals, then $[0, 1)^d$ is divided into n^d disjoint boxes (see Lemma 1.1.3). Hence by finite additivity and translation invariance,

$$m'([0, 1)^d) = n^d \times m'([0, \frac{1}{n})^d),$$

and we have $m'([0, \frac{1}{n})^d) = \frac{1}{n^d} m'([0, 1)^d) = cm([0, \frac{1}{n})^d)$.

Similarly, we have

$$m'([0, p)^d) = q^d \times m'([0, \frac{p}{q})^d)$$

and

$$m'([0, p)^d) = p^d \times m'([0, 1)^d).$$

This implies that $m'([0,\frac{p}{q})^d) = (\frac{p}{q})^d m'([0,1)^d) = cm([0,\frac{p}{q})^d)$.

And for real box case, we have

$$c(x-\varepsilon)^d \le cr^d \le m'([0,x)^d) \le cs^d \le c(x+\varepsilon)^d,$$

this implies $m'([0,x)^d) = cx^d = cm([0,x)^d)$.

Because every elementary set $E \in \mathcal{E}(\mathbf{R}^d)$ can be expressed as the disjoint union of $[0,x)^d$, we have $m'(E) = cm(E)$. In particular, if we normalize $m'([0,1)^d) = 1$, we have $m' \equiv m$. □

Lemma 1.1.6. *Let $d_1, d_2 \ge 1$, and let $E_1 \subset \mathbf{R}^{d_1}, E_2 \subset \mathbf{R}^{d_2}$ be elementary sets. Then $E_1 \times E_2 \subset \mathbf{R}^{d_1+d_2}$ is elementary, and $m^{d_1+d_2}(E_1 \times E_2) = m^{d_1}(E_1) \times m^{d_2}(E_2)$.*

Proof. Let E_1 equals to the finite disjoint union of $A_1, \cdots, A_k$ and E_2 equals to the finite disjoint union of $B_1, \cdots, B_{k'}$. Then for $1 \le i \le k$, we have

$$A_i \times E_2 = A_i \times \bigcup_{j=1}^{k'} B_j = \bigcup_{j=1}^{k'} (A_i \times B_j).$$

By Definition 1.1.1, we have $|A_i \times B_j| = |A_i| \times |B_j|$. Then by Proposition 1.1.4(ii), we have

$$\begin{aligned} m^{d_1+d_2}(A_i \times E_2) &= \sum_{j=1}^{k'} m^{d_1+d_2}(A_i \times B_j) \\ &= \sum_{j=1}^{k'} |A_i \times B_j| \\ &= \sum_{j=1}^{k'} |A_i| \times |B_j| \\ &= |A_i| \sum_{j=1}^{k'} |B_j| \\ &= m^{d_1}(A_i) \sum_{j=1}^{k'} m^{d_2}(B_j) \end{aligned}$$

$$= m^{d_1}(A_i)m^{d_2}(E_2).$$

Thus

$$\begin{aligned}
m^{d_1+d_2}(E_1 \times E_2) &= m^{d_1+d_2}\Big(\bigcup_{i=1}^{k}(A_i \times E_2)\Big) \\
&= \sum_{i=1}^{k} m^{d_1}(A_i)m^{d_2}(E_2) \\
&= m^{d_2}(E_2)\sum_{i=1}^{k} m^{d_1}(A_i) \\
&= m^{d_1}(E_1)m^{d_2}(E_2),
\end{aligned}$$

as desired. □

1.1.2 Jordan measure

The elementary sets are a very restrictive class of sets, far too small for most applications. For instance, a solid triangle or disk in the plane will not be elementary, or even a rotated box. On the other hand, as essentially observed long ago by Archimedes, such sets E can be approximated from within and without by elementary sets $A \subset E \subset B$, and the inscribing elementary set A and the circumscribing elementary set B can be used to give lower and upper bounds on the putative measure of E. As one makes the approximating sets A, B increasingly fine, one can hope that these two bounds eventually match. This gives rise to the following definitions.

Definition 1.1.7 (Jordan measure)**.** Let $E \subset \mathbf{R}^d$ be a bounded set.

- The *Jordan inner measure* $m_{*,(J)}(E)$ of E is defined as

$$m_{*,(J)}(E) := \sup_{A \subset E, A \text{ elementary}} m(A).$$

- The *Jordan outer measure* $m^{*,(J)}(E)$ of E is defined as

$$m^{*,(J)}(E) := \inf_{B \supset E, B \text{ elementary}} m(B).$$

- If $m_{*,(J)}(E) = m^{*,(J)}(E)$, then we say that E is *Jordan measurable*, and call $m(E) := m_{*,(J)}(E) = m^{*,(J)}(E)$ then *Jordan measure* of E. As before, we write $m(E)$ as $m^d(E)$ when we wish to emphasise the dimension d.

Jordan measurable sets are those sets which are "almost elementary" with respect to Jordan outer measure. More precisely, we have

Proposition 1.1.8 (Characterisation of Jordan measurability)**.** *Let $E \subset \mathbf{R}^d$ be bounded. Then following are equivalent:*

(i) *E is Jordan measurable.*

(ii) *For every $\varepsilon > 0$, there exist elementary sets $A \subset E \subset B$ such that $m(B \setminus A) \leq \varepsilon$.*

(iii) *For every $\varepsilon > 0$, there exists an elementary set A such that $m^{*,(J)}(A \triangle E) \leq \varepsilon$.*

Proof. (i) $\Rightarrow$ (ii). Since E is Jordan measurable, by Definition 1.1.7, there is elementary sets $A \subset E \subset B$ such that $m(E) \leq m(A) + \varepsilon/2$ and $m(E) \geq m(B) - \varepsilon/2$. Then $m(B \setminus A) = m(B) - m(A) \leq \varepsilon$, as desired.

(ii) $\Rightarrow$ (iii). There exist elementary sets $A \subset E \subset B$ such that $E \setminus A \subset B \setminus A$. Then by Definition 1.1.7, we have

$$m^{*,(J)}(A \triangle E) = m^{*,(J)}(E \setminus A) = \inf_{C \supset E \setminus A} m(B \setminus A) \leq m(B \setminus A) \leq \varepsilon.$$

(iii) $\Rightarrow$ (i). From definition, we obviously have $m_{*,(J)}(E) \leq m^{*,(J)}(E)$. This is sufficient to show that $m^{*,(J)}(E) \leq m_{*,(J)}(E)$ so that E is Jordan measurable with $m(E) = m_{*,(J)}(E) = m^{*,(J)}(E)$.

By hypothesis, for every $\varepsilon > 0$, there exists an elementary set A such that $m^{*,(J)}(A\triangle E)\leq\varepsilon$. Since

$$\begin{aligned}m^{*,(J)}(A\triangle E)&=m^{*,(J)}((A\setminus E)\cup(E\setminus A))\\&=\inf_{U\cup V\supset A\triangle E:U\supset A\setminus E,V\supset E\setminus A}m(U\cup V)\\&\geq\inf_{V\supset E\setminus A}m(V)\\&=m^{*,(J)}(E\setminus A),\end{aligned}$$

we have $m^{*,(J)}(E\setminus A)\leq\varepsilon$ where $E\setminus A\subset E$.

Since $A\cap E\subset E$, by definition, we have

$$m_{*,(J)}(E)+\varepsilon\geq m_{*,(J)}(A\cap E)+m^{*,(J)}(E\setminus A).$$

Now we want to show that $m_{*,(J)}(A\cap E)=m^{*,(J)}(A\cap E)$, i.e., $A\cap E$ is Jordan measurable. From $m(A\triangle E)\leq\varepsilon$, there exists an elementary set $S\supset A\triangle E$ such that $m(S)\leq\varepsilon$ (Why this is true? Use the Jordan outer measure of $A\triangle E$). Since we have $A\setminus S\subset A\cap E\subset A\cup S$, then

$$\begin{aligned}m^{*,(J)}(A\cap E)-m_{*,(J)}(A\cap E)&\leq m(A\cup S)-m(A\setminus S)\\&\leq m(A)+m(S)-m(A)+m(S)\\&\leq 2\varepsilon,\end{aligned}$$

for arbitrary ε. Thus we have $m_{*,(J)}(A\cap E)=m^{*,(J)}(A\cap E)$. Together our conclusions, we have

$$m_{*,(J)}(E)+\varepsilon\geq m^{*,(J)}(A\cap E)+m^{*,(J)}(E\setminus A).$$

Then we show that $m^{*,(J)}(A\cap E)+m^{*,(J)}(E\setminus A)\geq m^{*,(J)}(E)$. We can see that $A\cap E$ and $E\setminus A$ are disjoint, then for every elementary set D containing E can be divided into the union of elementary sets D_1 and D_2 where $D_1\supset A\cap E$ and $D_2\supset E\setminus A$. Notice that D_1 and D_2 are not necessary to be disjoint. Thus

$$\begin{aligned}m^{*,(J)}(E)&=\inf_{D\supset E}m(D)\\&\leq\inf_{D_1\cup D_2\supset E}(m(D_1)+m(D_2))\end{aligned}$$

$$= \inf_{D_1 \supset A \cap E} m(D_1) + \inf_{D_2 \supset E \setminus A} m(D_2)$$
$$= m^{*,(J)}(A \cap E) + m^{*,(J)}(E \setminus A).$$

Therefore, we have $m_{*,(J)}(E) + \varepsilon \geq m^{*,(J)}(E)$. Because ε is arbitrary, we have $m_{*,(J)}(E) = m^{*,(J)}(E)$, as desired. This shows that E is Jordan measurable and we complete the proof. □

Remark. As one corollary of this proposition, we see that every elementary set E is Jordan measurable, and that Jordan measure and elementary measure coincide for such sets; this justifies the use of $m(E)$ to denote both. In particular, we still have $m(\emptyset) = 0$.

Jordan measurability inherits many of properties of elementary measure:

Proposition 1.1.9 (The properties of Jordan measure)**.** *Let $E, F \subset \mathbf{R}^d$ be Jordan measurable sets.*

(i) *(Boolean closure) $E \cup F$, $E \cap F$, $E \setminus F$, and $E \triangle F$ are Jordan measurable.*

(ii) *(Non-negativity) $m(E) \geq 0$.*

(iii) *(Finite additivity) If E, F are disjoint, then $m(E \cup F) = m(E) + m(F)$.*

(iv) *(Monotonicity) If $E \subset F$, then $m(E) \leq m(F)$.*

(v) *(Finite subadditivity) $m(E \cup F) \leq m(E) + m(F)$.*

(vi) *(Translation invariance) For any $x \in \mathbf{R}^d$, $E + x$ is Jordan measurable, and $m(E + x) = m(E)$.*

Proof. (i) To show the union $E \cup F$ is Jordan measurable. Since E and F are Jordan measurable, by Proposition 1.1.8(ii), for every $\varepsilon > 0$ there are elementary sets $A \subset E \subset B$ and $C \subset F \subset D$ such that $m(B \setminus A) \leq \varepsilon/4$ and $m(D \setminus C) \leq \varepsilon/4$. Since $A \cup C \subset E \cup F \subset B \cup D$, and

$$(B \cup D) \setminus (A \cup C) = ((B \cup D) \setminus A) \cap ((B \cup D) \setminus C)$$
$$= ((B \setminus A) \cup (D \setminus A)) \cap ((B \setminus C) \cup (D \setminus C))$$

$$\begin{aligned}&= ((B \setminus A) \cap (B \setminus C)) \cup ((D \setminus A) \cap (D \setminus C)) \\ &\quad \cup ((B \setminus A) \cap (D \setminus C)) \cup ((D \setminus A) \cap (B \setminus C)) \\ &\subset (B \setminus A) \cup (D \setminus C) \cup (B \setminus A) \cup (D \setminus C),\end{aligned}$$

by monotonicity and finite subadditivity, there exist $A \cup C \subset E \cup F \subset B \cup D$ such that

$$m((B \cup D) \setminus (A \cup C)) \leq 2m(B \setminus A) + 2m(D \setminus C) \leq \varepsilon.$$

Thus by Proposition 1.1.8(ii), $E \cup F$ is Jordan measurable.

To show the intersection $E \cap F$ is Jordan measurable. From Proposition 1.1.8(ii), for every $\varepsilon > 0$ there is elementary sets $A \subset E \subset B$ and $C \subset F \subset D$ such that $m(B \setminus A) \leq \varepsilon/2$ and $m(D \setminus C) \leq \varepsilon/2$. Since $A \cap C \subset E \cap F \subset B \cap D$, and

$$\begin{aligned}(B \cap D) \setminus (A \cap C) &= ((B \cap D) \setminus A) \cup ((B \cap D) \setminus C) \\ &= ((B \setminus A) \cap (D \setminus A)) \cup ((B \setminus C) \cap (D \setminus C)) \\ &\subset (B \setminus A) \cup (D \setminus C),\end{aligned}$$

by monotonicity and finite subadditivity, there exist $A \cap C \subset E \cap F \subset B \cap D$ such that

$$m((B \cup D) \setminus (A \cup C)) \leq m(B \setminus A) + m(D \setminus C) \leq \varepsilon.$$

Thus by Proposition 1.1.8(ii), $E \cap F$ is Jordan measurable.

To show the difference $E \setminus F$. From Proposition 1.1.8(ii), for every $\varepsilon > 0$ there is elementary sets $A \subset E \subset B$ and $C \subset F \subset D$ such that $m(B \setminus A) \leq \varepsilon$ and $m(D \setminus C) \leq \varepsilon$. Then $A \setminus C \subset E \setminus F \subset B \setminus D$. Since

$$(B \setminus D) \setminus (A \setminus C) = (B \setminus A) \setminus (D \setminus C) \subset B \setminus A,$$

by monotonicity, there exist $A \setminus C \subset E \setminus F \subset B \setminus D$ such that

$$m((B \setminus D) \setminus (A \setminus C)) \leq m(B \setminus A) \leq \varepsilon.$$

Thus by Proposition 1.1.8(ii), $E \setminus F$ is Jordan measurable.

The symmetric difference $E \triangle F$ is Jordan measurable is immediately comes from above.

(ii) Since for every elementary set A we have $m(A) \geq 0$, by Definition 1.1.7, $m_{*,(J)}(E) = \sup_{A \subset E} m(A) \geq 0$. Thus for every Jordan measurable set E, we have $m(E) \geq 0$.

Since E and F are Jordan measurable, for every $\varepsilon > 0$, there exist elementary sets $A \subset E$ and $B \subset F$ such that $m(E) \leq m(A) + \varepsilon/2$ and $m(F) \leq m(B) + \varepsilon/2$. By finite additivity of elementary measure, we have

$$\begin{aligned} m(E \cup F) &= m_{*,(J)}(E \cup F) \\ &= \sup_{A \cup B \subset E \cup F} m(A \cup B) \\ &= \sup_{A \cup B \subset E \cup F} m(A) + m(B) \\ &\geq m(E) + m(F) - \varepsilon. \end{aligned}$$

For the other hand, for every $\varepsilon > 0$, there exist elementary sets $E \subset C$ and $F \subset D$ such that $m(E) \geq m(C) - \varepsilon/2$ and $m(F) \geq m(D) - \varepsilon/2$. Then

$$\begin{aligned} m(E \cup F) &= m^{*,(J)}(E \cup F) \\ &= \inf_{E \cup F \subset C \cup D} m(C \cup D) \\ &\leq \inf_{E \cup F \subset C \cup D} m(C) + m(D) \\ &\leq m(E) + m(F) + \varepsilon. \end{aligned}$$

Thus we conclude that for arbitrary $\varepsilon > 0$ we have

$$m(E) + m(F) - \varepsilon \leq m(E \cup F) \leq m(E) + m(F) + \varepsilon.$$

This means that $m(E \cup F) = m(E) + m(F)$, as desired.

(iv) This immediately comes from non-negativity and finite additivity of Jordan measure.

(v) Finite subadditivity immediately comes from monotonicity and finite additivity.

(vi) By Proposition 1.1.8(ii), for every $\varepsilon > 0$, there exist elementary sets $A \subset E \subset B$ such that $m(B \setminus A) \leq \varepsilon$. Then by translation invariance of elementary measure, for $x \in \mathbf{R}^d$, there exists elementary sets $A + x \subset E + x \subset B + x$ such that

$$m((B+x) \setminus (A+x)) = m(B \setminus A + x) = m(B \setminus A) \leq \varepsilon.$$

Thus $E + x$ is Jordan measurable. $\square$

Now we give some examples of Jordan measurable sets:

Example 1.1.10 (Regions under graphs are Jordan measurable)**.** Let B be a closed box in $\mathbf{R}^d$, and let $f : B \to \mathbf{R}$ be a continuous function.

(1) The graph $\{(x, f(x)) : x \in B\} \subset \mathbf{R}^{d+1}$ is Jordan measurable in $\mathbf{R}^{d+1}$ with Jordan measure zero.

(2) The set $\{(x,t) : x \in B; 0 \leq t \leq f(x)\} \subset \mathbf{R}^{d+1}$ is Jordan measurable.

Proof. (1) Denote the graph of f as graph(f). Since B is closed and bounded (see Definition 1.1.1), by Heine-Borel theorem, B is compact.

For the first assertion, since f is defined on a compact metric space, continuity of f is equivalent to uniform continuity. Then by the definition of uniform continuity, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x')| < \varepsilon$ whenever $x, x' \in B$ are such that $|x - x'| < \delta$. This means that every element $x \in B$ lies in some open ball $B_n := \{x \in B : |x_n - x| < \delta\}$, and we have $B = \bigcup_{n=1}^{\infty} B_n$. Since B is compact, there exists an N such that $B = \bigcup_{n=1}^{N} B_n$.

Then for every $\varepsilon > 0$, there exist a $\delta > 0$ and $N > 0$ such that $\emptyset \subset \text{graph}(f) \subset G_n$, where

$$G_n := \bigcup_{n=1}^{N} B_n \times (f(x_n) - \varepsilon, f(x_n) + \varepsilon),$$

such that

$$m(G_n \setminus \emptyset) = m(G_n) \leq 2\varepsilon m(B).$$

Thus by Proposition 1.1.8(ii), graph(f) is Jordan measurable. Since ε is arbitrary, by non-negativity and monotonicity, we have $m(\text{graph}(f)) \leq m(G_n) = 0$. Thus $m(\text{graph}(f)) = 0$, as desired.

(2) Denote the set as S. Since f is uniformly continuous, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\bigcup_{n=1}^{N} B_n \times [0, f(x_n) - \varepsilon] \subset S \subset \bigcup_{n=1}^{N} B_n \times [0, f(x_n) + \varepsilon]$$

such that

$$\begin{aligned}
& m\Big(\bigcup_{n=1}^{N} B_n \times [0, f(x_n) + \varepsilon] \setminus \bigcup_{n=1}^{N} B_n \times [0, f(x_n) - \varepsilon]\Big) \\
\leq {} & m\Big(\bigcup_{n=1}^{N} B_n \times \bigcup_{n=1}^{N} [0, f(x_n) + \varepsilon]\Big) - m\Big(\bigcup_{n=1}^{N} B_n \times \bigcup_{n=1}^{N} [0, f(x_n) - \varepsilon]\Big) \\
\leq {} & m(B) \sum_{n=1}^{N} (f(x_n) + \varepsilon) - m(B) \sum_{n=1}^{N} (f(x_n) - \varepsilon) \\
= {} & 2\varepsilon m(B).
\end{aligned}$$

Because ε is arbitrary, S is Jordan measurable. □

Example 1.1.11. Let A, B, C be three points in $\mathbf{R}^2$.

(1) The solid triangle with vertices A, B, C is Jordan measurable.

(2) The Jordan measure of the solid triangle is equal to $\frac{1}{2}|(B - A) \wedge (C - A)|$, where $|(a, b) \wedge (c, d)| := |ad - bc|$.

Proof. (1) Let $A = (x_1, y_1), B = (x_2, y_2), C = (x_3, y_3)$. Suppose that $x_1 < x_2 < x_3$ and $y_1 < y_3$. We first consider that the edge AB is horizontal to x-axis, i.e., $A = (x_1, y_1), B = (x_2, y_1)$. Define the edges AB, AC, and BC as the lines pass through the points and restricting the domain on the corresponding intervals. The solid triangle is the area under the graph as following:

$$\triangle ABC := \{(x, t) : x \in [x_1, x_3]; AB \leq t \leq AC, BC \leq t \leq AC\}.$$

Since

$$\begin{aligned}
U_1 &:= \{(x,t) : x \in [x_1, x_2]; 0 \le t \le AC\},\\
U_2 &:= \{(x,t) : x \in [x_1, x_2]; 0 \le t < AB\},\\
U_3 &:= \{(x,t) : x \in [x_2, x_3]; 0 \le t \le AC\},\\
U_4 &:= \{(x,t) : x \in [x_2, x_3]; 0 \le t < BC\},
\end{aligned}$$

are Jordan measurable from Exercise 1.1.10(2), we have

$$\triangle ABC = (U_1 \setminus U_2) \cup (U_3 \setminus U_4)$$

is Jordan measurable by Proposition 1.1.9.

Without loss of generality, let $AB : [x_1, y_1] \to \mathbf{R}$, $AC : [x_2, y_2] \to \mathbf{R}$, $BC : [x_3, y_3] \to \mathbf{R}$ be the lines segments pass through the corresponding vertices. Define

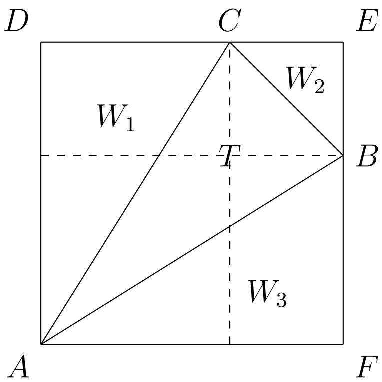
$$\begin{aligned}
V_1 &:= \{(x,t) : x \in [x_1, y_1] \cap [x_2, y_2]; t \in [AC, AB] \cup [AB, AC]\},\\
V_2 &:= \{(x,t) : x \in [x_2, y_2] \cap [x_3, y_3]; t \in [AC, BC] \cup [BC, AC]\},\\
V_3 &:= \{(x,t) : x \in [x_1, y_1] \cap [x_3, y_3]; t \in [AB, BC] \cup [BC, AC]\},
\end{aligned}$$

where V_1, V_2, V_3 are Jordan measurable (use Example 1.1.10). Thus

$$\triangle ABC = V_1 \cup V_2 \cup V_3$$

is Jordan measurable from Proposition 1.1.9.

(2) Let $A = (x_1, y_1)$, $B = (x_2, y_2)$, $C = (x_3, y_3)$. Let $a = x_2 - x_1$, $b = y_2 - y_1$, $c = x_3 - x_1$, $d = y_3 - y_1$. Then by translation invariance, we have $A = (0, 0)$, $B = (a, b)$, $C = (c, d)$, it would not change the areas. We have shown that the triangle is Jordan measurable, denote the square $ADEF$ as S, then we can see that $S = W_1 \cup W_2 \cup W_3 \cup T$ which is Jordan measurable. Since the edges have zero measure, we can just ignore them and $T = S \setminus (W_1 \cup W_2 \cup W_3)$.



This is easy to see that $m(S) = ad$, $m(W_1) = \frac{1}{2}cd$, $m(W_2) = \frac{1}{2}(a - c)(d - b)$, $m(W_3) = \frac{1}{2}ab$. Since $m(W_1 \cap W_2 \cap W_3) = 0$, we have

$$\begin{aligned} m(W_1 \cup W_2 \cup W_3) &= m(W_1) + m(W_2) + m(W_3) - m(W_1 \cap W_2 \cap W_3) \\ &= \frac{1}{2}(cd + (a - c)(d - b) + ab) \\ &= \frac{1}{2}(ad + cb). \end{aligned}$$

Thus

$$\begin{aligned} m(T) &= m(S \setminus (W_1 \cup W_2 \cup W_3)) \\ &= ad - \frac{1}{2}(ad + cb) \\ &= \frac{1}{2}(ad - cb) \\ &= \frac{1}{2}|(B - A) \wedge (C - A)|. \end{aligned}$$

Here, notice that $m(T \cap (W_1 \cup W_2 \cup W_3)) = 0$. □

Example 1.1.12. Every compact convex polytope[a] in $\mathbf{R}^d$ is Jordan measurable.

[a] A *closed convex polytope* is a subset of $\mathbf{R}^d$ formed by intersecting together finitely many closed half-spaces of the form $\{x \in \mathbf{R}^d : x \cdot v \leq c\}$, where $v \in \mathbf{R}^d, c \in$

$\mathbf{R}$, and $\cdot$ denotes the usual dot product on $\mathbf{R}^d$. A *compact convex polytope* is a closed convex polytope which is also bounded.

Proof. For arbitrary half-space $H := \{x \in \mathbf{R}^d : x \cdot v \leq c\}$ we have

$$\begin{aligned} H &= \{x \in \mathbf{R}^d : x \cdot v \leq c\} \\ &= \left\{ x \in \mathbf{R}^d : \sum_{i=1}^{d} x_i v_i \leq c \right\} \\ &= \left\{ x \in \mathbf{R}^d : x_d \leq \frac{1}{v_d} \left(c - \sum_{i=1}^{d-1} x_i v_i \right) \right\} \\ &= \{(x', x_d) \in \mathbf{R}^d : x_d \leq f(x')\}. \end{aligned}$$

Define the function $f : \mathbf{R}^{d-1} \to \mathbf{R}$ as following

$$f(x') := \frac{1}{v_d} \left(c - \sum_{i=1}^{d-1} x_i v_i \right),$$

where $x' = (x_1, \cdots, x_{d-1})$. This is easy to see that f is continuous and x' lies in some closed boxes. Then there is an $m \in \mathbf{R}$ such that $m \leq x_d$, and we have

$$\begin{aligned} H &= \{(x', x_d) \in \mathbf{R}^d : m \leq x_d \leq f(x')\} \\ &= \{(x', x_d) \in \mathbf{R}^d : 0 \leq x_d \leq f(x')\} \setminus \{(x', x_d) \in \mathbf{R}^d : 0 < x_d < m\} \end{aligned}$$

which is Jordan measurable from Exercise 1.1.10 and Proposition 1.1.9.

Thus by Proposition 1.1.9(i), the compact convex polytope in $\mathbf{R}^d$ is Jordan measurable. $\square$

Example 1.1.13.

(1) All open and closed Euclidean balls $B(x, r) := \{y \in \mathbf{R}^d : |y - x| < r\}$, $\overline{B(x, r)} := \{y \in \mathbf{R}^d : |y - x| \leq r\}$ in $\mathbf{R}^d$ are Jordan measurable, with Jordan measure $c_d r^d$ for some constant $c_d > 0$ depending only on d

(2) One has the crude bounds

$$\left(\frac{2}{\sqrt{d}}\right)^d \leq c_d \leq 2^d.$$

(An exact formula for c_d is $c_d = \frac{1}{d}\omega_d$, where $\omega_d := \frac{2\pi^{d/2}}{\Gamma(d/2)}$ is volume of the unit sphere $S^{d-1} \subset \mathbf{R}^d$ and Γ is the *Gamma function*, but we will not derive this formula here.)

Proof. (1) Consider the closed Euclidean ball $\overline{B(0,r)}$ we have

$$\begin{aligned}\overline{B(0,r)} &= \{y \in \mathbf{R}^d : |y| \leq r\} \\ &= \left\{y \in \mathbf{R}^d : \sum_{i=1}^{d} y_i^2 \leq r^2\right\} \\ &= \left\{(y', y_d) \in \mathbf{R}^d : y_d^2 \leq r^2 - \sum_{i=1}^{d-1} y_i^2\right\}.\end{aligned}$$

Then there is a continuous function $f : \mathbf{R}^{d-1} \to \mathbf{R}$ defined as

$$f(y') := \left(r^2 - \sum_{i=1}^{d-1} y_i^2\right)^{1/2}$$

where $y' = (y_1, \cdots, y_{d-1})$ lies in some closed boxes, such that

$$\overline{B(0,r)} = \{(y', y_d) \in \mathbf{R}^d : -f(y') \leq y_d \leq f(y')\}.$$

By Exercise 1.1.10(2), $\overline{B(0,r)}$ is Jordan measurable. Use invariance translation, $\overline{B(x,r)} = \overline{B(0,r)} - x$ is also Jordan measurable. Since $\partial(\overline{B(x,r)}) := \{(y', y_d) \in \mathbf{R}^d : y_d = f(y')\}$ is a graph under $\mathbf{R}^d$, so that is Jordan measurable with zero measure. We have $B(x,r) = \overline{B(x,r)} \setminus \partial(\overline{B(x,r)})$ is also Jordan measurable.

(2) By invariance translation, we only need to consider $B := B(0,r)$. Since $|y| \leq r$ implies that $|y_i| \leq r$, B is contained in some boxes, i.e., $B \subset \{y \in \mathbf{R}^d : |y_i| \leq r \text{ for all } i\}$. Thus we have $m(B) \leq 2^d r^d$. Other hand, for $|y| \leq r$, there is a $y' = \min\{y_i : 1 \leq i \leq d\}$. Then we have $|(y', \cdots, y')| =$

$\frac{1}{\sqrt{d}}r \leq |y| \leq r$. This implies that $B \supset \{y \in \mathbf{R}^d : |y_i| \leq \frac{1}{\sqrt{d}}r \text{ for all } i\}$, which is a box with length $\frac{2}{\sqrt{d}}r$. Thus we have $m(B) \geq (\frac{2}{\sqrt{d}})^d r^d$, as desired. □

Example 1.1.14. Define a *Jordan null set* to be a Jordan measurable set of Jordan measure zero. Then any subset of a Jordan null set is a Jordan null set.

Proof. Let A be a Jordan null set. We show that every subset E of a Jordan null set is Jordan measurable. Since we have $\emptyset \subset E \subset A$, by Proposition 1.1.8(ii), we have

$$m(A \setminus \emptyset) = m(A) = 0 \leq \varepsilon$$

for every $\varepsilon > 0$. Thus E is Jordan measurable. Then by monotonicity and non-negativity, E is Jordan null set. □

Example 1.1.15. One has

$$m(E) := \lim_{N\to\infty} \frac{1}{N^d} \#(E \cap \frac{1}{N}\mathbf{Z}^d)$$

where $\frac{1}{N}\mathbf{Z} := \{\frac{n}{N} : n \in \mathbf{Z}\}$ and $\#A$ denotes the cardinality of a finite set A, holds for all Jordan measurable $E \subset \mathbf{R}^d$.

Proof. Since E is Jordan measurable, by Proposition 1.1.8(ii) and definition of m, there is $A \subset E \subset B$ such that

$$\begin{aligned} m(B \setminus A) &= \lim_{N\to\infty} \frac{1}{N^d} \#((B \setminus A) \cap \frac{1}{N}\mathbf{Z}^d) \\ &= \lim_{N\to\infty} \frac{1}{N^d} \#((B \cap \frac{1}{N}\mathbf{Z}^d) \setminus (A \cap \frac{1}{N}\mathbf{Z}^d)) \\ &= \lim_{N\to\infty} \frac{1}{N^d} \#(B \cap \frac{1}{N}\mathbf{Z}^d) - \lim_{N\to\infty} \frac{1}{N^d} \#(A \cap \frac{1}{N}\mathbf{Z}^d) \\ &\leq \varepsilon. \end{aligned}$$

This implies $m(B) \leq m(A) + \varepsilon$. Then

$$m^{*,(J)}(E) = \inf_{B \supset E} m(B) \leq \sup_{A \subset E} m(A) = m_{*,(J)}(E).$$

Since we obviously have $m_{*,(J)}(E) \leq m^{*,(J)}(E)$ and

$$m(A) \leq \lim_{N\to\infty} \frac{1}{N^d} \#(E \cap \frac{1}{N}\mathbf{Z}^d) \leq m(B),$$

thus

$$m(E) = \lim_{N\to\infty} \frac{1}{N^d} \#(E \cap \frac{1}{N}\mathbf{Z}^d) = m_{*,(J)}(E) = m^{*,(J)}(E).$$

□

Example 1.1.16 (Metric entropy formulation of Jordan measurability)**.** Define a *dyadic cube* to be a half-open box of the form

$$\left[\frac{i_1}{2^n}, \frac{i_1+1}{2^n}\right) \times \left[\frac{i_d}{2^n}, \frac{i_d+1}{2^n}\right)$$

for some integers $n, i_1, \cdots, i_d$. Let $E \subset \mathbf{R}^d$ be a bounded set. For each integer n, let $\mathcal{E}_*(E, 2^{-n})$ denote the number of dyadic cubes of sidelength 2^{-n} that are contained in E, and let $\mathcal{E}^*(E, 2^{-n})$ be the number of dyadic cubes[a] of sidelength 2^{-n} that intersect E. Then E is Jordan measurable if and only if

$$\lim_{n\to\infty} 2^{-dn}(\mathcal{E}^*(E, 2^{-n}) - \mathcal{E}_*(E, 2^{-n})) = 0,$$

in which case one has

$$m(E) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}_*(E, 2^{-n}) = \lim_{n\to\infty} 2^{-dn}\mathcal{E}^*(E, 2^{-n}).$$

[a] This quantity could be called the (dyadic) *metric entropy* of E at scale 2^{-n}.

Proof.

Theorem 1.1.17 (Uniqueness of Jordan measure)**.** *Let $d \geq 1$. Let $m' : \mathcal{J}(\mathbf{R}^d) \to \mathbf{R}^+$ be a map from the collection $\mathcal{J}(\mathbf{R}^d)$ of Jordan-measurable subsets of $\mathbf{R}^d$ to the non-negative reals that obeys the non-negativity, finite additivity, and translation invariance properties. Then there exists a constant $c \in \mathbf{R}^+$ such that $m'(E) = cm(E)$ for all Jordan* *measurable sets E. In particular, if we impose the additional normalisation $m'([0,1]^d) = 1$, then $m' \equiv m$.*

Proof. Set $c := m'([0,1))$. Since every elementary sets are Jordan measurable, if E is an elementary set, then from Theorem 1.1.5 we have $m'(E) = cm(E)$ (recall that elementary measure and Jordan measure are coincide for each other for elementary sets). For arbitrary elementary sets A, B that $A \subset E \subset B$, we have $m(E) = \sup_{A \subset E} m(A) = \inf_{B \supset E} m(B)$. This implies that

$$cm(E) = \sup_{A \subset E} m'(A) = \inf_{B \supset E} m'(B).$$

This is trivial to show that m' obeys monotonicity, then for every $\varepsilon > 0$ there exist elementary sets A, B such that

$$\sup_{A \subset E} m'(A) - \varepsilon \leq m'(A) \leq m'(E) \leq m'(B) \leq \inf_{B \supset E} m'(B) + \varepsilon.$$

Because ε is arbitrary, thus we have $m'(E) = cm(E)$ for all Jordan measurable sets. $\square$

Lemma 1.1.18. *Let $d_1, d_2 \geq 1$, and let $E_1 \subset \mathbf{R}^{d_1}, E_2 \subset \mathbf{R}^{d_2}$ be Jordan measurable sets. Then $E_1 \times E_2 \subset \mathbf{R}^{d_1+d_2}$ is Jordan measurable, and $m^{d_1+d_2}(E_1 \times E_2) = m^{d_1}(E_1) \times m^{d_2}(E_2)$.*

Proof. For every $\varepsilon > 0$, there exist elementary sets $A_1 \subset E_1 \subset B_1$, $A_2 \subset E_2 \subset B_2$ such that $m(B_1 \setminus A_1) \leq \varepsilon m(B_2)/2$ and $m(B_2 \setminus A_2) \leq \varepsilon m(B_1)/2$. Since we have $A_1 \times A_2 \subset E_1 \times E_2 \subset B_1 \times B_2$, we have

$$\begin{aligned} m((B_1 \times B_2) \setminus (A_1 \times A_2)) &= m([(B_1 \setminus A_1) \times B_2] \cup [B_1 \times (B_2 \setminus A_2)]) \\ &\leq m((B_1 \setminus A_1) \times B_2) + m(B_1 \times (B_2 \setminus A_2)) \\ &= m(B_1 \setminus A_1) m(B_2) + m(B_1) m(B_2 \setminus A_2) \\ &\leq \varepsilon \end{aligned}$$

Thus $E_1 \times E_2$ is Jordan measurable.

By Lemma 1.1.6 there exist $A_1 \subset E_1$ and $A_2 \subset E_2$ such that

$$\begin{aligned} m^{d_1+d_2}(E_1 \times E_2) &\geq m^{d_1+d_2}(A_1 \times A_2) \\ &= m^{d_1}(A_1)m^{d_2}(A_2) \\ &\geq (m^{d_1}(E_1) - \varepsilon)(m^{d_2}(E_2) - \varepsilon). \end{aligned}$$

For the other hand, there is $B_1 \supset E_1$ and $B_2 \supset E_2$ such that

$$\begin{aligned} m^{d_1+d_2}(E_1 \times E_2) &\leq m^{d_1+d_2}(B_1 \times B_2) \\ &= m^{d_1}(B_1)m^{d_2}(B_2) \\ &\leq (m^{d_1}(E_1) + \varepsilon)(m^{d_2}(E_2) + \varepsilon). \end{aligned}$$

Because ε is arbitrary, we have $m^{d_1+d_2}(E_1 \times E_2) = m^{d_1}(E_1)m^{d_2}(E_2)$. □

Example 1.1.19. Let P, Q be two polytopes in $\mathbf{R}^d$. Suppose that P can be partitioned into finitely many sub-polytopes $P_1, \cdots, P_n$ which, after being rotated and translated, form new polytopes $Q_1, \cdots, Q_n$ which are an *almost disjoint cover* of Q, which means that $Q = Q_1 \cup \cdots \cup Q_n$, and for any $1 \leq i < j \leq n$, Q_i and Q_j only intersect at the boundary (i.e. the interior of Q_i is disjoint from the interior of Q_j). Then P and Q have the same Jordan measure. The converse statement is true in one and two dimensions $d = 1, 2$ (this is the *Bolyai-Gerwien theorem*), but false in higher dimensions (this was Dehn's negative answer to *Hilbert's third problem*).

Proof.

The above examples give a fairly large class of Jordan measurable sets. However, not every subset of $\mathbf{R}^d$ is Jordan measurable. First of all, the unbounded sets are not Jordan measurable, by construction. But there are also bounded sets that are not Jordan measurable:

Proposition 1.1.20. *Let $E \subset \mathbf{R}^d$ be a bounded set.*

(1) E and the closure[a] $\overline{E}$ of E have the same Jordan outer measure.

(2) E and the interior[b] E° of E have the same Jordan inner measure.

(3) E is Jordan measurable if and only if the topological boundary ∂E of E has Jordan outer measure zero.

(4) The bullet-riddled square $[0,1]^2 \backslash \mathbf{Q}^2$*, and set of bullets* $[0,1]^2 \cap \mathbf{Q}^2$*, both have Jordan inner measure zero and Jordan outer measure one. In particular, both sets are not Jordan measurable.*

[a]The *closure* of E is defined as the intersection of all closed sets containing E.
[b]The *interior* of E is defined as the union of all open sets contained in E.

Proof. (1) Let B_n be elementary sets that $B_n \supset E$, then $\bigcap_n B_n \supset \overline{E}$. By definition, we have

$$m^{*,(J)}(E) = \inf_{B \supset E} m(B) \leq \inf_{B \supset \overline{E}} m(B) = m^{*,(J)}(\overline{E}).$$

Other hand, for every $\varepsilon > 0$ there exist elementary set $B \supset E$ such that $m^{*,(J)}(E) \geq m(B) - \varepsilon$. See that we have $B \supset \bigcap_n B_n$, thus $m^{*,(J)}(E) \geq m(\bigcap_n B_n) - \varepsilon$. Since $m^{*,(J)}(\overline{E}) = \inf_{\bigcap_n B_n \supset \overline{E}} m(\bigcap_n B_n)$, we have

$$m^{*,(J)}(E) \geq m^{*,(J)}(\overline{E}) - \varepsilon.$$

Because ε is arbitrary, we have $m^{*,(J)}(E) \geq m^{*,(J)}(\overline{E})$. Thus $m^{*,(J)}(E) = m^{*,(J)}(\overline{E})$.

(2) Let A_n be elementary set that $A_n \subset E$, then $\bigcup_n A_n \subset E$. By definition, we have

$$m_{*,(J)}(E^\circ) = \sup_{A \subset E^\circ} m(A) \leq \sup_{A \subset E} m(A) = m_{*,(J)}(E).$$

Other hand, for every $\varepsilon > 0$ there exist elementary set $A \subset E$ such that $m_{*,(J)}(E) \leq m(A) + \varepsilon$. See that we have $A \subset \bigcup_n A_n$, thus $m_{*,(J)}(E) \leq$

$m(\bigcup_n A_n) + \varepsilon$. Since $m_{*,(J)}(E^\circ) = \sup_{\bigcap_n A_n \subset E^\circ} m(\bigcup_n A_n)$, we have

$$m_{*,(J)}(E) \leq m_{*,(J)}(E^\circ) + \varepsilon.$$

Because ε is arbitrary, we have $m_{*,(J)}(E) \leq m_{*,(J)}(E^\circ)$. Thus $m_{*,(J)}(E) = m_{*,(J)}(E^\circ)$.

(3) Suppose that E is Jordan measurable, for every $\varepsilon > 0$ there exist $A \subset E \subset B$ such that $m(B \setminus A) \leq \varepsilon$. From definition, we have $\overline{E} \subset B$ and $A \subset E^\circ$ for arbitrary $A \subset E \subset B$. Then $\partial E = \overline{E} \setminus E^\circ \subset B \setminus A$, so that

$$m^{*,(J)}(\partial E) = m^{*,(J)}(\overline{E} \setminus E^\circ) \leq m(B \setminus A) \leq \varepsilon$$

Because ε is arbitrary, we have $m^{*,(J)}(\partial E) = 0$.

Conversely, suppose that $m^{*,(J)}(\partial E) = 0$. Then we have

$$\begin{aligned} m^{*,(J)}(\partial E) &= m^{*,(J)}(\overline{E} \setminus E^\circ) \\ &= m^{*,(J)}(\overline{E}) - m^{*,(J)}(E^\circ) \\ &\leq m^{*,(J)}(E) - m_{*,(J)}(E) \end{aligned}$$

This implies that $m^{*,(J)}(E) \geq m_{*,(J)}(E)$.

Proposition 1.1.21 (Carathéodory type property)**.** *Let $E \subset \mathbf{R}^d$ be a bounded set, and let $F \subset \mathbf{R}^d$ be an elementary set. Then $m^{*,(J)}(E) = m^{*,(J)}(E \cap F) + m^{*,(J)}(E \setminus F)$. This result still hold when F is Jordan measurable instead of elementary.*

Proof.

1.1.3 Connection with the Riemann integral

Definition 1.1.22 (Riemann integrability)**.** Let $[a, b]$ be an interval of positive length, and let $f : [a, b] \to \mathbf{R}$ be a function. A *tagged partition* $\mathcal{P} = ((x_0, x_1, \cdots, x_n), (x_1^*, \cdots, x_n^*))$ of $[a, b]$ is a finite sequence of real numbers $a = x_0 < x_1 < \cdots < x_n = b$, together with additional numbers $x_{i-1} \leq x_i^* \leq x_i$ for each $i = 1, \cdots, n$. We abbreviate $x_i - x_{i-1}$ as δx_i. The quantity $\Delta(\mathcal{P}) := \sup_{1 \leq i \leq n} \delta x_i$ will be called the *norm* of the tagged partition. The *Riemann sum* $\mathcal{R}(f, \mathcal{P})$ of f with respect to the tagged partition $\mathcal{P}$ is defined as

$$\mathcal{R}(f, \mathcal{P}) := \sum_{i=1}^{n} f(x_i^*)\delta x_i.$$

We say that f is *Riemann integrable* on $[a, b]$ if there exists a real number, denote $\int_a^b f(x)dx$ and referred to as the *Riemann integral* of f on $[a, b]$, for which we have

$$\int_a^b f(x)dx = \lim_{\Delta(\mathcal{P})\to 0} \mathcal{R}(f, \mathcal{P})$$

by which we mean that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|\mathcal{R}(f, \mathcal{P}) - \int_a^b f(x)dx| \leq \varepsilon$ for every tagged partition $\mathcal{P}$ with $\Delta(\mathcal{P}) \leq \delta$.

If $[a, b]$ is an interval of zero length, we adopt the convention that every function $f : [a, b] \to \mathbf{R}$ is Riemann integrable, with a Riemann integral of zero.

Proposition 1.1.23 (Piecewise constant functions)**.** *Let $[a, b]$ be an interval. A* piecewise constant function *$f : [a, b] \to \mathbf{R}$ is a function for which there exists a partition of $[a, b]$ into finitely many intervals $I_1, \cdots, I_n$ such that f is equal to a constant c_i on each of the intervals I_i. If f is piecewise constant, then the expression*

$$\sum_{i=1}^{n} c_i|I_i|$$

is independent of the choice of partition used to demonstrate the piecewise constant nature of f. We denote this quantity by p.c. $\int_a^b f(x)dx$, and refer to it as the piecewise constant integral *of f on $[a, b]$.*

Proof. Since $[a, b]$ is elementary, by Proposition 1.1.3(ii), it is independent of the partition. Let $J_1, \cdots, J_m$ be a different partition of $[a, b]$ from $I_1, \cdots, I_n$ such that f be a piecewise constant function. Then $\{I_i \cap J_j : 1 \leq i \leq n, 1 \leq$

$j \leq m\}$ is also a partition of $[a, b]$, and we have

$$\sum_{i=1}^{n}\sum_{j=1}^{m}|I_i \cap J_j| = \sum_{i=1}^{n}|I_i| = \sum_{j=1}^{m}|J_j|.$$

We have $c_i = c_j$ for all $I_i \cap I_j$. Thus

$$\begin{aligned}\sum_{i=1}^{n} c_i|I_i| &= \sum_{i=1}^{n} c_i\Big(\sum_{j=1}^{m}|I_i \cap J_j|\Big)\\ &= \sum_{j=1}^{m}\Big(\sum_{i=1}^{n} c_i|I_i \cap J_j|\Big)\\ &= \sum_{i=1}^{n}\Big(\sum_{j=1}^{m} c_j|I_i \cap J_j|\Big)\\ &= \sum_{j=1}^{m} c_j\Big(\sum_{i=1}^{n}|I_i \cap J_j|\Big)\\ &= \sum_{j=1}^{m} c_j|J_j|.\end{aligned}$$

Thus $\sum_{i=1}^{n} c_i|I_i|$ is independent of the choice of partition and we denote this quantity as $\mathrm{p.c.}\int_a^b f(x)dx := \sum_{i=1}^{n} c_i|I_i|$. □

Theorem 1.1.24 (Basic properties of the piecewise constant integral)**.** *Let $[a, b]$ be an interval, and let $f, g : [a, b] \to \mathbf{R}$ be piecewise constant functions.*

(i) *(Linearity) For any real number c, cf and $f + g$ are piecewise constant, with $\mathrm{p.c.}\int_a^b cf(x)dx = c\,\mathrm{p.c.}\int_a^b f(x)dx$ and $\mathrm{p.c.}\int_a^b f(x) + g(x)dx = \mathrm{p.c.}\int_a^b f(x) + \mathrm{p.c.}\int_a^b g(x)dx$.*

(ii) *(Monotonicity) If $f \leq g$ pointwise (i.e., $f(x) \leq g(x)$ for all $x \in [a, b]$), then $\mathrm{p.c.}\int_a^b f(x)dx \leq \mathrm{p.c.}\int_a^b g(x)dx$.*

(iii) *(Indicator) If E is an elementary subset of $[a, b]$, then the indicator function $1_E : [a, b] \to \mathbf{R}$ (defined by setting $1_E(x) := 1$ when $x \in E$ and $1_E(x) := 0$ otherwise) is piecewise constant,* *and* $\text{p. c.} \int_a^b 1_E(x)dx = m(E)$.

Proof. (i) Let $I_1, \cdots, I_n$ and $J_1, \cdots, J_m$ be partitions of f and g, respectively. Then f and g are piecewise constant functions with respect to the partition $\{I_i \cap J_j : 1 \leq i \leq n, 1 \leq j \leq m\}$. In each $I_i \cap J_j$, f equals to a_i and g equals to b_j.

We can see that cf is a piecewise constant and in each $I_i \cap J_j$ equals to $c \cdot a_i$, and

$$\text{p. c.} \int_a^b cf(x)dx = \sum_{i=1}^n c \cdot a_i |I_i| = c \sum_{i=1}^n a_i |I_i| = c \cdot \text{p. c.} \int_a^b f(x)dx.$$

We also have $f + g$ is a piecewise constant and in each $I_i \cap J_j$ equals to $a_i + b_j$, and

$$\begin{aligned} \text{p. c.} \int_a^b f(x) + g(x)dx &= \sum_{i=1}^n \sum_{j=1}^m (a_i + b_j)|I_i \cap I_j| \\ &= \sum_{i=1}^n \sum_{j=1}^m a_i |I_i \cap I_j| + \sum_{j=1}^m \sum_{i=1}^n b_j |I_i \cap I_j| \\ &= \sum_{i=1}^n a_i |I_i| + \sum_{j=1}^m b_j |I_j| \\ &= \text{p. c.} \int_a^b f(x)dx + \text{p. c.} \int_a^b g(x)dx. \end{aligned}$$

(ii) Since $f \leq g$, we have $a_i \leq b_j$ for all $I_i \cap J_j$. Then

$$\begin{aligned} \text{p. c.} \int_a^b f(x)dx &= \sum_{i=1}^n \sum_{j=1}^m a_i |I_i \cap I_j| \\ &\leq \sum_{j=1}^m \sum_{i=1}^n b_j |I_i \cap I_j| \\ &= \text{p. c.} \int_a^b g(x)dx. \end{aligned}$$

(iii) Because E is elementary, then $[a, b] \setminus E$ is also elementary. Then there is disjoint sequences $I_1, \cdots, I_n$ and $J_1, \cdots, J_m$ such that $E = \bigcup_{i=1}^n I_i$

and $[a,b]\setminus E=\bigcup_{j=1}^m J_j$ (see Lemma 1.1.3). Then $I_1,\cdots,I_n,J_1,\cdots,J_m$ is a partition of $[a,b]$, and 1_E is a piecewise constant function. Thus

$$\mathrm{p.\,c.}\int_a^b 1_E(x)dx=\sum_{i=1}^n 1\cdot|I_i|+\sum_{j=1}^m 0\cdot|J_j|=\sum_{i=1}^n |I_i|=m(E).$$

□

Definition 1.1.25 (Darboux integral)**.** Let $[a,b]$ be an interval, and $f:[a,b]\to\mathbf{R}$ be a bounded function. The *lower Darboux integral* $\underline{\int_a^b} f(x)dx$ of f on $[a,b]$ is defined as

$$\underline{\int_a^b} f(x)dx := \sup_{g\le f,\text{piecewise constant}} \mathrm{p.\,c.}\int_a^b g(x)dx,$$

where g ranges over all piecewise constant functions that are pointwise bounded above by f. (The hypothesis that f is bounded ensures that the supremum is over a non-empty set.) Similarly, we define the *upper Darboux integral* $\overline{\int_a^b} f(x)dx$ of f on $[a,b]$ by the formula

$$\overline{\int_a^b} f(x)dx := \inf_{h\ge f,\text{piecewise constant}} \mathrm{p.\,c.}\int_a^b h(x)dx.$$

Clearly $\underline{\int_a^b} f(x)dx\le\overline{\int_a^b} f(x)dx$. If these two quantities are equal, we say that f is *Darboux integrable*, and refer to this quantity as the *Darboux integral* of f on $[a,b]$.

Remark. Note that the upper and lower Darboux integrals are related by the reflection identity

$$\overline{\int_a^b} -f(x)dx=-\underline{\int_a^b} f(x)dx.$$

Proposition 1.1.26. *Let $[a,b]$ be an interval, and $f:[a,b]\to\mathbf{R}$ be a bounded function. We say that f is Riemann integrable if and only if it is Darboux integrable, in which case the Riemann integral and Darboux integrals are equal.*

Proof. First we suppose that f is Riemann integrable. Since f is bounded, define $\underline{f}:[a,b]\to\mathbf{R}$ as $\underline{f}(x)=\inf_{x\in[x_{i-1},x_i]}f(x)$ and $\overline{f}:[a,b]\to\mathbf{R}$ as $\overline{f}(x)=\sup_{x\in[x_{i-1},x_i]}f(x)$. We can see that $\underline{f}$ and $\overline{f}$ are piecewise constant functions with respect to $\mathcal{P}$ such that $\underline{f}\le f\le\overline{f}$. Then for arbitrary piecewise constant function g that minorize f, we have $\mathrm{p.c.}\int_a^b g(x)dx\le\mathcal{R}(\underline{f},\mathcal{P})$, and piecewise constant function h that majorize f, we have $\mathrm{p.c.}\int_a^b h(x)dx\ge\mathcal{R}(\overline{f},\mathcal{P})$.

Since f is Riemann integrable, by Definition 1.1.22, for every $\varepsilon>0$ there exists $\delta>0$ such that

$$\left|\mathcal{R}(f,\mathcal{P})-\int_a^b f(x)dx\right|\le\varepsilon$$

for every $\mathcal{P}$ with $\Delta(\mathcal{P})\le\delta$. Then

$$\mathcal{R}(\underline{f},\mathcal{P})-\varepsilon\le\mathcal{R}(f,\mathcal{P})-\varepsilon\le\int_a^b f(x)dx\le\mathcal{R}(f,\mathcal{P})+\varepsilon\le\mathcal{R}(\overline{f},\mathcal{P})+\varepsilon,$$

so that

$$\mathrm{p.c.}\int_a^b g(x)dx-\varepsilon\le\int_a^b f(x)dx\le\mathrm{p.c.}\int_a^b h(x)dx+\varepsilon$$

Because ε is arbitrary, taking the supremum and infimum, we have

$$\underline{\int_a^b} f(x)dx\le\int_a^b f(x)dx\le\overline{\int_a^b} f(x)dx.$$

For the other hand, we have $\underline{\int_a^b}f(x)dx\le\mathcal{R}(\underline{f},\mathcal{P})$ and $\overline{\int_a^b}f(x)dx\ge\mathcal{R}(\overline{f},\mathcal{P})$ from $\mathrm{p.c.}\int_a^b g(x)dx\le\mathcal{R}(\underline{f},\mathcal{P})$ and $\mathrm{p.c.}\int_a^b h(x)dx\ge\mathcal{R}(\overline{f},\mathcal{P})$. Then

$$\underline{\int_a^b} f(x)dx-\overline{\int_a^b} f(x)dx\le\mathcal{R}(\underline{f},\mathcal{P})-\mathcal{R}(\overline{f},\mathcal{P})$$

$$\begin{aligned} &\leq \int_a^b f(x)dx + \varepsilon - \int_a^b f(x)dx + \varepsilon \\ &= 2\varepsilon \end{aligned}$$

for every $\mathcal{P}$ with $\Delta(\mathcal{P}) \leq \delta$. Since ε is arbitrary, we have $\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx$. Thus

$$\int_a^b f(x)dx = \underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx$$

and f is Darboux integrable.

Conversely, suppose that f is Darboux integrable. By Definition 1.1.25, we have

$$\int_a^b f(x)dx = \underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx.$$

We want to show that

$$\underline{\int_a^b} f(x)dx = \sup(\mathcal{R}(\underline{f}, \mathcal{P}))$$

and

$$\overline{\int_a^b} f(x)dx = \inf(\mathcal{R}(\overline{f}, \mathcal{P})).$$

Then Darboux integrability implies that

$$\int_a^b f(x)dx = \sup(\mathcal{R}(\underline{f}, \mathcal{P})) = \inf(\mathcal{R}(\overline{f}, \mathcal{P})).$$

This is sufficient to show that

$$\int_a^b f(x)dx = \lim_{\Delta(\mathcal{P}) \to 0} \mathcal{R}(f, \mathcal{P})$$

and f is Riemann integrable.

We already show that $\underline{\int_a^b} f(x)dx \leq \mathcal{R}(\underline{f}, \mathcal{P})$ and $\overline{\int_a^b} f(x)dx \geq \mathcal{R}(\overline{f}, \mathcal{P})$, taking the supremum and infimum, we have

$$\underline{\int_a^b} f(x)dx \leq \sup(\mathcal{R}(\underline{f}, \mathcal{P}))$$

and

$$\overline{\int_a^b} f(x)dx \geq \inf(\mathcal{R}(\overline{f}, \mathcal{P})).$$

To prove in the other direction, suppose for sake of contradiction that $\underline{\int_a^b} f(x)dx < \sup(\mathcal{R}(\underline{f}, \mathcal{P}))$. Then there exists some $\mathcal{P}$ such that

$$\underline{\int_a^b} f(x)dx < \mathcal{R}(\underline{f}, \mathcal{P}).$$

Since $\underline{f}$ minorizes f (i.e., $\underline{f} \leq f$ for all x), by Definition 1.1.25, we have

$$\mathcal{R}(\underline{f}, \mathcal{P}) \leq \underline{\int_a^b} f(x)dx,$$

a contradiction. Thus $\underline{\int_a^b} f(x)dx \geq \sup(\mathcal{R}(\underline{f}, \mathcal{P}))$. This implies that

$$\underline{\int_a^b} f(x)dx = \sup(\mathcal{R}(\underline{f}, \mathcal{P})).$$

A similar argument shows that

$$\overline{\int_a^b} f(x)dx = \inf(\mathcal{R}(\overline{f}, \mathcal{P})),$$

as desired. □

Proposition 1.1.27. *Any continuous function $f : [a, b] \to \mathbf{R}$ is Riemann integrable. More generally, any bounded, piecewise continuous*[a] *function $f : [a, b] \to \mathbf{R}$ is Riemann integrable.*

[a] A function $f : [a, b] \to \mathbf{R}$ is *piecewise continuous* if one can partition $[a, b]$ into finitely many intervals, such that f is continuous on each interval.

Proof. Since f is continuous and defined on $[a, b]$, it is bounded. By Proposition 1.1.26, we only need to show that f is Darboux integrable. From continuity, for every ε there exists a $\delta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ for all $x, y \in [a, b]$ such that $|x - y| \leq \delta$. By Archimedes principle, there exists an N such that $(b-a)/N < \delta$. Let $I_k = [a + \frac{(b-a)(k-1)}{N}, a + \frac{(b-a)k}{N}]$. Then $I_1, \cdots, I_N$ is a partition of $[a,b]$ with $|I_k| = (b-a)/N$.

From $\underline{\int_a^b} f(x)dx = \sup(\mathcal{R}(\underline{f}, \mathcal{P}))$ and $\overline{\int_a^b} f(x)dx = \inf(\mathcal{R}(\overline{f}, \mathcal{P}))$ we have

$$\underline{\int_a^b} f(x)dx \geq \sum_{k=1}^{N} (\inf_{x \in I_k} f(x))|I_k|$$

and

$$\overline{\int_a^b} f(x)dx \leq \sum_{k=1}^{N} (\sup_{x \in I_k} f(x))|I_k|,$$

so

$$\overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx \leq \sum_{k=1}^{N} (\sup_{x \in I_k} f(x) - \inf_{x \in I_k} f(x))|I_k|.$$

Since $|f(x) - f(y)| \leq \varepsilon$ is hold for all $x, y \in I_k$ for that $|x-y| \leq |I_k| < \delta$, we have

$$\overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx \leq \sum_{k=1}^{N} \varepsilon |I_k| = \varepsilon(b-a).$$

Since ε is arbitrary, we have

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx.$$

Thus by Definition 1.1.25 and Proposition 1.1.26, f is Riemann integrable.

$\square$

Now we connect the Riemann integral to Jordan measure in two ways. First, we connect the Riemann integral to one-dimensional Jordan measure:

Theorem 1.1.28 (Basic properties of the Riemann integral)**.** *Let $[a,b]$ be an interval, and let $f, g : [a,b] \to \mathbf{R}$ be Riemann integrable.*

(i) *(Linearity) For any real number c, cf and $f+g$ are Riemann integrable, with $\int_a^b cf(x)dx = c\int_a^b f(x)dx$ and $\int_a^b f(x)+$*

$g(x)dx = \int_a^b f(x) + \int_a^b g(x)dx$.

(ii) *(Monotonicity) If $f \leq g$ pointwise (i.e., $f(x) \leq g(x)$ for all $x \in [a,b]$) then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.*

(iii) *(Indicator) If E is a Jordan measurable of $[a,b]$, then the indicator function $1_E : [a,b] \to \mathbf{R}$ (defined by setting $1_E(x) := 1$ when $x \in E$ and $1_E(x) := 0$ otherwise) is Riemann integrable, and $\int_a^b 1_E(x)dx = m(E)$.*

These properties uniquely define the Riemann integral, in the sense that the function $f \mapsto \int_a^b f(x)dx$ is the only map from the space of Riemann integrable functions on $[a,b]$ to $\mathbf{R}$ which obeys all three of the above properties.

Proof. (i) Since f is Riemann integrable, we have

$$\int_a^b f(x)dx = \lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n f(x_i^*)\delta x_i.$$

Then for cf

$$\int_a^b cf(x)dx = \lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n cf(x_i^*) = c \lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n f(x_i^*)\delta x_i$$

which is convergent. Thus cf is Riemann integrable with $\int_a^b cf(x)dx = c\int_a^b f(x)dx$.

Similarly, we have

$$\begin{aligned}\int_a^b f(x)dx + \int_a^b g(x)dx &= \lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n f(x_i^*)\delta x_i + \lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n g(x_i^*)\delta x_i \\ &= \lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n (f(x_i^*) + g(x_i^*))\delta x_i\end{aligned}$$

which is convergent. Thus $f+g$ is Riemann integrable with $\int_a^b f(x)dx + \int_a^b g(x)dx = \int_a^b f(x) + g(x)dx$.

(ii) Since $f \leq g$, we have

$$\lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n f(x_i^*)\delta x_i \leq \lim_{\Delta(\mathcal{P})\to 0} \sum_{i=1}^n g(x_i^*)\delta x_i,$$

thus $\int_a^b f(x)dx \le \int_a^b g(x)dx$.

(iii) Since 1_E is bounded, by Proposition 1.1.26, we show that 1_E is Darboux integrable. By Theorem 1.1.24(iii) and Definition 1.1.25, for arbitrary elementary set $A \subset E$,

$$m(A) = \text{p. c.} \int_a^b 1_A(x)dx \le \underline{\int_a^b} 1_E(x)dx.$$

Taking the supremum, we have $m_{*,(J)}(E) \le \underline{\int_a^b} 1_E(x)dx$ by Definition 1.1.7. Similarly, for every elementary set $B \supset E$ where $B \subset [a,b]$, we have

$$m(B) = \text{p. c.} \int_a^b 1_B(x)dx \ge \overline{\int_a^b} 1_E(x)dx.$$

Taking the infimum, we have $m^{*,(J)}(E) \ge \overline{\int_a^b} 1_E(x)dx$.

Because E is Jordan measurable, from

$$m_{*,(J)}(E) \le \underline{\int_a^b} 1_E(x)dx \le \overline{\int_a^b} 1_E(x)dx \le m^{*,(J)}(E)$$

and $m_{*,(J)}(E) = m^{*,(J)}(E)$ we have $\underline{\int_a^b} 1_E(x)dx = \overline{\int_a^b} 1_E(x)dx$. Hence 1_E is Riemann integrable.

Finally, we prove that these three properties uniquely define the Riemann integral. Let $\mathcal{R}([a,b] \to \mathbf{R})$ be the space of Riemann integrable functions on $[a,b]$ to $\mathbf{R}$. The Riemann integral is the map $\mathcal{M} : \mathcal{R}([a,b] \to \mathbf{R}) \to \mathbf{R}$. Suppose that there is another map $\mathcal{M}' : \mathcal{R}([a,b] \to \mathbf{R}) \to \mathbf{R}$ satisfying above three properties where $\mathcal{M} \ne \mathcal{M}'$. Then the property (i) means that for every Riemann integrable functions $f, g \in \mathcal{R}([a,b] \to \mathbf{R})$ and constants $a, b \in \mathbf{R}$ we have $\mathcal{M}(af+bg) = a\mathcal{M}(f) + b\mathcal{M}(g)$. This is also holds for $\mathcal{M}'$.

We already know that every piecewise constant function f is Riemann integrable, i.e., $f \in \mathcal{R}([a,b] \to \mathbf{R})$. Let f be p.c. with respect to $I_1, \cdots, I_n$ with $c_1, \cdots, c_n$. Then we have

$$f(x) = \sum_{i=1}^n c_i 1_{I_i}(x),$$

so that

$$\begin{aligned}
\mathcal{M}'(f) &= \mathcal{M}'\Big(\sum_{i=1}^{n} c_i 1_{I_i}(x)\Big) \\
&= \sum_{i=1}^{n} c_i \mathcal{M}'(1_{I_i}(x)) \\
&= \sum_{i=1}^{n} c_i m(I_i) \\
&= \sum_{i=1}^{n} c_i \mathcal{M}(1_{I_i}(x)) \\
&= \sum_{i=1}^{n} \mathcal{M}(c_i 1_{I_i}(x)) \\
&= \mathcal{M}\Big(\sum_{i=1}^{n} c_i 1_{I_i}(x)\Big) \\
&= \mathcal{M}(f).
\end{aligned}$$

Since the Riemann integral depends on the piecewise constant integral, we have $\mathcal{M}(f) = \mathcal{M}'(f)$ for arbitrary Riemann integrable function, a contradiction. □

Next, we connect the integral to two-dimensional Jordan measure:

Proposition 1.1.29 (Area interpretation of the Riemann integral)**.** *Let $[a, b]$ be an interval, and let $f : [a, b] \to \mathbf{R}$ be a bounded function. Then f is Riemann integrable if and only if the sets $E_+ := \{(x, t) : x \in [a, b]; 0 \le t \le f(x)\}$ and $E_- := \{(x, t) : x \in [a, b]; f(x) \le t \le 0\}$ are both Jordan measurable in $\mathbf{R}^2$, in which case one has*

$$\int_a^b f(x)dx = m^2(E_+) - m^2(E_-),$$

where m^2 denotes two-dimensional Jordan measure.

Proof. We first suppose that f is non-negative, then $m^2(E_-) = 0$.

Suppose that f is Riemann integrable. Let $I_1, \cdots, I_n$ be a partition of $[a, b]$, let $\overline{f}(x) = \sup_{x \in I_k} f(x)$ and $\underline{f}(x) = \inf_{x \in I_k} f(x)$. Then we have

$$E_+ \subset \overline{E} := \bigcup_{k=1}^{n} (I_k \times [0, \overline{f}(x)])$$

and

$$E_+ \supset \underline{E} := \bigcup_{k=1}^{n} (I_k \times [0, \underline{f}(x)]).$$

By Proposition 1.1.18, we have

$$m^2(\overline{E}) = \sum_{k=1}^{n} \overline{f}(x)|I_k|$$

and

$$m^2(\underline{E}) = \sum_{k=1}^{n} \underline{f}(x)|I_k|$$

Since f is Riemann integrable, we have

$$m^2(\underline{E}) = m^2(\overline{E}) = \lim_{\Delta(\mathcal{P}) \to 0} \sum_{k=1}^{n} f(x_i^*)|I_k| = \int_a^b f(x)dx.$$

Thus from monotonicity of Jordan measure, we have $m^2(E_+) = \int_a^b f(x)dx$.

Conversely, suppose that

$$\int_a^b f(x)dx = m^2(E_+).$$

Let $\underline{E}, \overline{E}$ be arbitrary elementary sets such that $\underline{E} \subset E_+ \subset \overline{E}$.

Since the elementary sets' expression is independent of the choice of partition (see Proposition 1.1.3), we express elementary sets $\underline{E}$ and $\overline{E}$ as disjoint unions $\underline{E} = \bigcup_{i=1}^{n} A_i$ and $\overline{E} = \bigcup_{j=1}^{m} B_j$, where $A_i = I_{i,1} \times I_{i,2}$ and $B_j = J_{j,1} \times J_{j,2}$, such that $I_{1,1}, \cdots, I_{n,1}$ and $J_{1,1}, \cdots, J_{m,1}$ be two partition of $[a, b]$. Then we have $|I_{i,2}| \leq f(x)$ and $|J_{j,2}| \geq f(y)$ for all $x \in I_{i,1}$ and $y \in J_{j,1}$.

Now we define piecewise constant functions $g, h : [a, b] \to \mathbf{R}$ as $g(x) := |I_{i,2}|$ for $x \in I_{i,1}$ and $h(x) := |J_{j,2}|$ for $x \in J_{j,1}$. Then

$$\underline{\int_a^b} f(x)dx = \sup_{g \le f} \text{p. c.} \int_a^b g(x)dx = \sup_{g \le f} \sum_{i=1}^n g(x)|I_{i,1}| = \sup_{\underline{E} \subset E_+} m(\underline{E}).$$

and

$$\overline{\int_a^b} f(x)dx = \inf_{h \ge f} \text{p. c.} \int_a^b h(x)dx = \inf_{h \ge f} \sum_{i=1}^n h(x)|J_{i,1}| = \inf_{\overline{E} \supset E_+} m(\overline{E}).$$

Therefore,

$$\inf_{\overline{E} \supset E_+} m(\overline{E}) = \overline{\int_a^b} f(x)dx \le \underline{\int_a^b} f(x)dx = \sup_{\underline{E} \subset E_+} m(\underline{E}).$$

Because E_+ is Jordan measurable, by Definition 1.1.7, there exist elementary sets $\underline{E} \subset E_+ \subset \overline{E}$ such that

$$m(E_+) = \sup_{\underline{E} \subset E_+} m(\underline{E}) = \inf_{\overline{E} \supset E_+} m(\overline{E}).$$

Thus f is Riemann integrable.

A similar argument shows that when f is negative, we have

$$\int_a^b f(x)dx = -m^2(E_-)$$

if and only if f is Riemann integrable.

Define functions f^+ and f^- from $[a, b]$ to $\mathbf{R}$ by

$$f^+(x) := \begin{cases} f(x), & \text{if } f(x) \ge 0, \\ 0, & \text{if } f(x) < 0, \end{cases}$$

and

$$f^-(x) := \begin{cases} 0, & \text{if } f(x) \ge 0, \\ -f(x), & \text{if } f(x) < 0. \end{cases}$$

Then $f = f^+ - f^-$ and

$$\int_a^b f(x)dx = \int_a^b f^+(x)dx - \int_a^b f^-(x)dx = m^2(E_+) - m^2(E_-),$$

as desired. □

§1.2 Lebesgue measure

Following example shows that not all sets are Jordan measurable, even if one restricts attention to bounded sets.

Example 1.2.1. The countable union $\bigcup_{n=1}^{\infty} E_n$ or countable intersection $\bigcap_{n=1}^{\infty} E_n$ of Jordan measurable sets $E_1, E_2, \dots \subset \mathbf{R}$ need not be Jordan measurable, even when bounded.

Proof. We know that $[0,1] \setminus \mathbf{Q}$ is not Jordan measurable. Since $\mathbf{Q}$ is countable, then $[0,1] \setminus \mathbf{Q}$ can be represented as the countable union of open interval where each interval is Jordan measurable.

For the second assertion, suppose that $[0,1] \setminus \mathbf{Q} = \bigcup_{n=1}^{\infty} E_n$ for Jordan measurable sets $E_1, E_2, \cdots$. Then $[0,1] \setminus \bigcup_{n=1}^{\infty} E_n = \bigcap_{n=1}^{\infty}([0,1] \setminus E_n)$. Since $[0,1] \setminus E_n$ are Jordan measurable for all $n \geq 1$ and $[0,1] \setminus \bigcup_{n=1}^{\infty} E_n = [0,1] \cap \mathbf{Q}$ which is not Jordan measurable. $\square$

Definition 1.2.2 (Lebesgue measurability)**.** The *Lebesgue outer measure $m^*(E)$* of E is defined as

$$m^*(E) := \inf_{\bigcup_{n=1}^{\infty} B_n \supset E; B_n \text{ boxes}} \sum_{n=1}^{\infty} |B_n|.$$

A set $E \subset \mathbf{R}^d$ is said to be *Lebesgue measurable* if, for every $\varepsilon > 0$, there exists an open set $U \subset \mathbf{R}^d$ containing E such that $m^*(U \setminus E) \leq \varepsilon$. If E is Lebesgue measurable, we refer to $m(E) := m^*(E)$ as the Lebesgue measure of E. We also write $m(E)$ as $m^d(E)$ when we wish to emphasise the dimension d.

1.2.1 Properties of Lebesgue outer measure

We begin by studying the Lebesgue outer measure m^*, which was defined earlier, and takes values in the extended non-negative real axis $[0, +\infty]$. We first record three easy properties of Lebesgue outer measure, which we will use repeatedly in the sequel without further comment:

Lemma 1.2.3 (The outer measure axioms)**.**

(i) *(Empty set)* $m^*(\emptyset) = 0$.

(ii) *(Monotonicity) If* $E \subset F \subset \mathbf{R}^d$, *then* $m^*(E) \leq m^*(F)$.

(iii) *(Countable subadditivity) If* $E_1, E_2, \dots \subset \mathbf{R}^d$ *is a countable sequence of sets, then* $m^*(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} m^*(E_n)$.

Proof. (i) Since empty set is subset of any set, in particular, we have $B_n = \emptyset$ for all $n \geq 1$. By Definition 1.2.2, $m^*(\emptyset) = \sum_{n=1}^{\infty} |B_n| = 0$.

(ii) If $E \subset F$, every collection of boxes cover F must cover E, thus $m^*(E) \leq m^*(F)$.

(iii) If $m^*(E_n) = +\infty$ for some $n \geq 1$, the conclusion is trivial. We suppose that $m^*(E_n)$ is finite for all $n \geq 1$. Let $\varepsilon > 0$. For every $n \geq 1$ there exists a cover such that $E_n \subset \bigcup_{k=1}^{\infty} A_{n,k}$ where $A_{n,k}$ are boxes. By Definition 1.2.2, for each E_n, there is an k such that

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

Then we have $\bigcup_{n=1}^{\infty} E_n \subset \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{n,k}$ and

$$\begin{aligned} m^*\Big(\bigcup_{n=1}^{\infty} E_n\Big) &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} A_{n,k} \\ &\leq \sum_{n=1}^{\infty} \Big(m^*(E_n) + \frac{\varepsilon}{2^k}\Big) \\ &= \sum_{n=1}^{\infty} m^*(E_n) + \varepsilon. \end{aligned}$$

Since ε is arbitrary, we have

$$m^*\Big(\bigcup_{n=1}^{\infty} E_n\Big) \leq \sum_{n=1}^{\infty} m^*(E_n),$$

as desired. $\square$

Remark. Note that countable subadditivity, when combined with the empty set axiom, gives as a corollary the finite subadditivity property

$$m^*(E_1 \cup \cdots \cup E_k) \leq m^*(E_1) + \cdots + m^*(E_k)$$

for any $k \geq 0$.

It is natural to ask whether Lebesgue outer measure has the *finite additivity property*, that is to say that $m^*(E \cup F) = m^*(E) + m^*(F)$ whenever $E, F \subset \mathbf{R}^d$ are disjoint. The answer to this question is somewhat subtle: as we shall see later, we have finite additivity (and even countable additivity) when all sets involved are Lebesgue measurable, but that finite additivity (and hence also countable additivity) can break down in the non-measurable case.

Following lemma says that if disjoint sets E and F have a positive separation from each other, then the Lebesgue outer measure is finitely additive:

Lemma 1.2.4 (Finite additivity for separated sets)**.** *Let* $E, F \subset \mathbf{R}^d$ *be such that* $\operatorname{dist}(E, F) > 0$*, where*

$$\operatorname{dist}(E, F) := \inf\{|x - y| : x \in E, y \in F\}$$

is the distance[a] *between* E *and* F*. Then* $m^*(E \cup F) = m^*(E) + m^*(F)$*.*

[a] Recall from the preface that we use the usual Euclidean metric $|(x_1, \cdots, x_d)| := \sqrt{x_1^2 + \cdots + x_d^2}$ on $\mathbf{R}^d$.

Proof. Proof omitted. □

In general, disjoint sets E, F need not have a positive separation from each other (e.g. $E = [0, 1)$ and $F = [1, 2]$). But the situation improves when E, F are closed, and at least one of E, F is compact:

Lemma 1.2.5. *Let $E, F \subset \mathbf{R}^d$ be disjoint closed sets, with at least one of E, F being compact. Then* $\operatorname{dist}(E, F) > 0$.

Proof. Let E be compact and F closed. Suppose for sake of contradiction that $\operatorname{dist}(E, F) = 0$. Then there exist sequences $(x_n)_{n=1}^\infty$ in E and $(y_n)_{n=1}^\infty$ in F such that $\lim_{n\to\infty} |x_n - y_n| = 0$. Since E is compact, there is a subsequence $(x_{n_k})_{k=1}^\infty$ of $(x_n)_{n=1}^\infty$ which converges to x. Then

$$\lim_{k\to\infty} |x - y_{n_k}| \le \lim_{k\to\infty} (|x - x_{n_k}| + |x_{n_k} - y_{n_k}|) = 0.$$

This means that x is an adherent point of F. Since F is closed, it contains all of its adherent points, thus we have $x \in F$ so that $E \cap F \neq \emptyset$, a contradiction.

For the counterexample, since every singleton is closed in $\mathbf{R}^d$. We can see that $\mathbf{Z}^+$ and $E := \{n + \frac{1}{n} : n \in \mathbf{Z}^+\}$ are closed and disjoint. But $\operatorname{dist}(\mathbf{Z}^+, E) = \inf_{n\in\mathbf{N}} \frac{1}{n} = 0$. □

From definition we know that countable sets have Lebesgue outer measure zero. Now we start computing the outer measure of some other sets. We begin with *elementary sets*:

Lemma 1.2.6 (Outer measure of elementary sets)**.** *Let E be an elementary set. Then the Lebesgue outer measure $m^*(E)$ of E is equal to the elementary measure $m(E)$ of E, i.e., $m^*(E) = m(E)$.*

Proof. Proof omitted. □

Remark. The above lemma allows us to compute the Lebesgue outer measure of a finite union of boxes. From this and monotonicity we conclude that the Lebesgue outer measure of any set is bounded below by its Jordan inner measure. As it is also bounded above by the Jordan outer measure, we have

$$m_{*,(J)}(E) \le m^*(E) \le m^{*,(J)}(E) \tag{1.1}$$

for every $E \subset \mathbf{R}^d$.

Now we turn to *countable unions of boxes*. First we define almost disjoint:

Definition 1.2.7 (Almost disjoint)**.** We say that two boxes are *almost disjoint* if their interiors are disjoint.

With Lemma 1.2.6 has the following consequence:

Lemma 1.2.8 (Outer measure of countable unions of almost disjoint boxes)**.** *Let $E = \bigcup_{n=1}^{\infty} B_n$ be a countable union of almost disjoint boxes $B_1, B_2, \cdots$. Then*

$$m^*(E) = \sum_{n=1}^{\infty} |B_n|.$$

Thus, for instance, $\mathbf{R}^d$ itself has an infinite outer measure.

Proof. Proof omitted. □

Lemma 1.2.9. *If a set $E \subset \mathbf{R}^d$ is expressible as the countable union of almost disjoint boxes, then the Lebesgue outer measure of E is equal to the Jordan inner measure: $m^*(E) = m_{*,(J)}(E)$, where*

$$m_{*,(J)}(E) := \sup_{\bigcup_{n=1}^{k} B_n \subset E; B_n \, boxes} \sum_{n=1}^{k} |B_n|,$$

and $m_{*,(J)}(E) = +\infty$ when E is unbounded.

Proof. Let $E = \bigcup_{n=1}^{\infty} B_n$ be a countable union of almost disjoint boxes $B_1, B_2, \cdots$. Then for all $k \in \mathbf{N}$ we have $E \supset \bigcup_{n=1}^{k} B_n$. From Lemma 1.2.8 we have

$$m^*(E) = \sum_{n=1}^{\infty} |B_n| \geq \sum_{n=1}^{k} |B_n|. \tag{1.2}$$

Since the inequality is hold for all $k \in \mathbf{N}$, taking the supremum, we have $m^*(E) \geq m_{*,(J)}(E)$.

For the other hand, we want to show that $m^*(E) \leq m_{*,(J)}(E)$. Since

$$\sum_{n=1}^{k} |B_n| \leq m_{*,(J)}(E)$$

for all k. Letting $k \to \infty$, we have $m^*(E) = m_{*,(J)}(E)$, as desired. □

Lemma 1.2.10. *Let $E \subset \mathbf{R}^d$ be an open set. Then E can be expressed as the countable union of almost disjoint boxes (and, in fact, as the countable union of almost disjoint closed cubes).*

Proof. Proof omitted. □

Lemma 1.2.11 (Outer regularity)**.** *Let $E \subset \mathbf{R}^d$ be an arbitrary set. Then one has*

$$m^*(E) = \inf_{E \subset U; U \text{ open}} m^*(U). \tag{1.3}$$

Proof. Proof omitted. □

1.2.2 Lebesgue measurability

We now define the notion of a Lebesgue measurable set as one which can be efficiently contained in open sets in the sense of Definition 1.2.2, and set out their basic properties.

First, we show that there are plenty of Lebesgue measurable sets.

Lemma 1.2.12 (Existence of Lebesgue measurable sets)**.**

(i) *Every open set is Lebesgue measurable.*

(ii) *Every closed set is Lebesgue measurable.*

(iii) *Every set of Lebesgue outer measure zero is measurable. (Such sets are called* null sets*.)*

(iv) *The empty set* $\emptyset$ *is Lebesgue measurable.*

(v) *If* $E \subset \mathbf{R}^d$ *is Lebesgue measurable, then so is its complement* $\mathbf{R}^d \setminus E$.

(vi) *If* $E_1, E_2, \cdots \subset \mathbf{R}^d$ *are a sequence of Lebesgue measurable sets, then the union* $\bigcup_{n=1}^{\infty} E_n$ *is Lebesgue measurable.*

(vii) *If* $E_1, E_2, \cdots \subset \mathbf{R}^d$ *are a sequence of Lebesgue measurable sets, then the intersection* $\bigcap_{n=1}^{\infty} E_n$ *is Lebesgue measurable.*

Remark. The properties (iv), (v) and (vi) of Lemma 1.2.12 assert that the collection of Lebesgue measurable subsets of $\mathbf{R}^d$ form a σ*-algebra*.For the property (ii), notice that every closed set F can be represented as the union of compact sets, i.e., $F = \bigcup_{n=1}^{\infty} F \cap \overline{B(0,n)}$, where $\overline{B(0,n)}$ are closed balls with radius n and centered at origin. Lemma 1.2.5 shows that compact set satisfying the countable additivity property.

Theorem 1.2.13 (Criteria for measruability)**.** *Let* $E \subset \mathbf{R}^d$. *The following are equivalent:*

(i) E *is Lebesgue measurable.*

(ii) *(Outer approximation by open) For every* $\varepsilon > 0$, *there exists an open set* $U \supset E$ *with* $m^*(U \setminus E) \leq \varepsilon$

(iii) *(Almost open) For every* $\varepsilon > 0$, *there exists an open* U *such that* $m^*(U \triangle E) \leq \varepsilon$. *(In other words,* E *differs from an open set by a set of outer measure at most* ε*.)*

(iv) *(Inner approximation by closed) For every* $\varepsilon > 0$, *there exists a closed set* $F \subset E$ *with* $m^*(E \setminus F) \leq \varepsilon$.

(v) *(Almost closed) For every* $\varepsilon > 0$, *there exists a closed set* F *such that* $m^*(F \triangle E) \leq \varepsilon$. *(In other words,* E *differs from a closed set by a set of outer measure at most* ε*.)*

(vi) *(Almost measurable) For every* $\varepsilon > 0$, *there exists a Lebesgue measurable set* E_ε *such that* $m^*(E_\varepsilon \triangle E) \leq \varepsilon$. *(In other words,* E *differs from a measurable set by a set of outer measure at* *most ε.)*

Proof. By Definition 1.2.2, (i) and (ii) are equivalent.

(ii) $\Rightarrow$ (iii). For every $\varepsilon > 0$, there exists an open set $U \supset E$ with $m^*(U \setminus E) \leq \varepsilon$. Then

$$m^*(U \triangle E) = m^*((U \setminus E) \cup (E \setminus U)) = m^*(U \setminus E) \leq \varepsilon.$$

(iii) $\Rightarrow$ (vi). By Lemma 1.2.12(i), U is Lebesgue measurable set, then let $E_\varepsilon := U$ we have $m^*(E_\varepsilon \triangle E) \leq \varepsilon$.

(vi) $\Rightarrow$ (i). For every $n \in \mathbf{Z}^+$ and every $\varepsilon > 0$, there exists a Lebesgue measurable set $E_{\varepsilon/2^n}$ such that $m^*(E_{\varepsilon/2^n} \triangle E) \leq \varepsilon/2^n$. Let $F := \bigcup_{n=1}^\infty E_{\varepsilon/2^n}$. We have

$$m^*(E \setminus F) = m^*\Big(E \setminus \bigcup_{n=1}^\infty E_{\varepsilon/2^n}\Big) \leq m^*(E \setminus E_{\varepsilon/2^n}) \leq m^*(E_{\varepsilon/2^n} \triangle E) \leq \varepsilon/2^n,$$

since ε is arbitrary and the inequality is hold for all $n \geq 1$, we have $m^*(E \setminus F) = 0$. For the other hand, we have

$$\begin{aligned} m^*(F \setminus E) &= m^*\Big(\bigcup_{n=1}^\infty E_{\varepsilon/2^n} \setminus E\Big) \\ &= m^*\Big(\bigcup_{n=1}^\infty (E_{\varepsilon/2^n} \setminus E)\Big) \\ &\leq m^*\Big(\bigcup_{n=1}^\infty (E_{\varepsilon/2^n} \triangle E)\Big) \\ &\leq \sum_{n=1}^\infty m^*(E_{\varepsilon/2^n} \triangle E) \\ &= \varepsilon \end{aligned}$$

for arbitrary ε, thus $m^*(F \setminus E) = 0$. Hence, by Lemma 1.2.12(iii), $E \setminus F$ and $F \setminus E$ are Lebesgue measurable.

Now we want to show that there exists an open set $U \supset E$ with $m^*(U \setminus E) \leq \varepsilon$. Because $F \cup E = F \cup (E \setminus F)$, by Lemma 1.2.12(vi), $F \cup E$

is Lebesgue measurable. Thus for every $\varepsilon > 0$ there exists an open set $U \supset F \cup E$ with $m^*(U \setminus (F \cup E)) \leq \varepsilon$. Obviously, we have $U \supset E$. Then

$$\begin{aligned} m^*(U \setminus E) &= m^*((U \setminus (F \cup E)) \cup (F \setminus E)) \\ &\leq m^*(U \setminus (F \cup E)) + m^*(F \setminus E) \\ &= 2\varepsilon. \end{aligned}$$

Thus E is Lebesgue measurable.

(iv) $\Rightarrow$ (v). For every $\varepsilon > 0$, there exists a closed set $F \subset E$ with $m^*(E \setminus F) \leq \varepsilon$. Then

$$m^*(F \triangle E) \leq m^*(E \setminus F) \leq \varepsilon.$$

(v) $\Rightarrow$ (vi). By Lemma 1.2.12(ii), F is Lebesgue measurable. Then we have $m^*(E_\varepsilon \triangle E) \leq \varepsilon$ for $E_\varepsilon := F$.

(i) $\Rightarrow$ (iv). By Lemma 1.2.12(v), E is Lebesgue measurable implies that $\mathbf{R}^d \setminus E$ is also Lebesgue measurable. Then there exists an open set $U \supset \mathbf{R}^d \setminus E$ with

$$m^*(U \setminus (\mathbf{R}^d \setminus E)) = m^*(E \setminus (\mathbf{R}^d \setminus U)),$$

where $\mathbf{R}^d \setminus U$ is closed and $\mathbf{R}^d \setminus U \subset E$. This complete the proof. $\square$

Proposition 1.2.14. *Every Jordan measurable set is Lebesgue measurable.*

Proof. By Definition 1.2.2, elementary set is Lebesgue measurable. By Proposition 1.1.8(iii), for every Jordan measurable set E, there is an elementary set A such that $m^{*,(J)}(A \triangle E) \leq \varepsilon$. Then by Definition 1.2.2, there exists an elementary set (which is Lebesgue measurable) such that

$$m^*(A \triangle E) \leq m^{*,(J)}(A \triangle E) \leq \varepsilon.$$

Here, we use the inequality that $m^*(F) \leq m^{*,(J)}(F)$ for every $F \subset \mathbf{R}^d$. Thus by Theorem 1.2.13(vi), E is Lebesgue measurable. $\square$

Every countable set has outer measure 0. A reasonable question arises about whether the converse holds. In other words, is every set with outer measure 0 countable? unfortunately, this is not always true, the *Cantor set* gives a counterexample:

Example 1.2.15 (Middle thirds Cantor set)**.** Let $I_0 := [0,1]$ be the unit interval, let $I_1 := [0,1/3] \cup [2/3,1]$ be I_0 with the interior of the middle third interval removed, let $I_2 := [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]$ be I_1 with the interior of the middle third of each of the two intervals of I_1 removed, and so forth. More formally, write

$$I_n := \bigcup_{a_1,\cdots,a_n \in \{0,2\}} \left[\sum_{i=1}^{n} \frac{a_i}{3^i}, \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n} \right].$$

Let $C := \bigcap_{n=1}^{\infty} I_n$ be the intersection of all the elementary sets I_n. Then C is compact, uncountable, and a null set.

Proof. We first show that C is compact. Since every I_n is the union of finitely many closed intervals, I_n is closed for all $n \in \mathbf{N}$. Thus C is closed. For every $a_1, \cdots, a_n \in \{0,2\}$, we have

$$0 \leq \sum_{i=1}^{n} \frac{a_i}{3^i} \leq \sum_{i=1}^{n} \frac{a_i}{3^i} + \frac{1}{3^n} \leq 1.$$

So that $I_n \subset [0,1]$, and C is also bounded by $[0,1]$. Because C is closed and bounded, it is compact.

Then we show that C is uncountable. Consider the function $\Lambda : \{0,2\}^{\mathbf{N}} \to [0,1]$ where $(a_i)_{i=1}^{\infty} \in \{0,2\}^{\mathbf{N}}$ (recall that $a_n : \mathbf{N} \to \{0,2\}$) defined as

$$\Lambda((a_i)_{i=1}^{\infty}) := \sum_{i=1}^{\infty} \frac{a_i}{3^i}.$$

Because the series is convergent, function Λ is well-defined. We can see that for every $n \in \mathbf{N}$ we have $\sum_{i=1}^{n} \frac{a_i}{3^i} \in I_n$. Then

$$\Lambda((a_i)_{i=1}^{\infty}) = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{a_i}{3^i} \in \bigcap_{n=1}^{\infty} I_n = C.$$

We want to show that Λ is injective. Suppose for sake of contradiction that there exists two different sequences $(a_i)_{i=1}^{\infty}, (b_i)_{i=1}^{\infty} \in \{0,2\}^{\mathbf{N}}$ such that $\Lambda((a_i)_{i=1}^{\infty}) = \Lambda((b_i)_{i=1}^{\infty})$. Suppose that these two sequences are different from the kth term, i.e., $a_k \neq b_k$. Let $a_k > b_k$. Then from

$$\sum_{i=1}^{\infty} \frac{a_i}{3^i} - \sum_{i=1}^{\infty} \frac{b_i}{3^i} = \frac{2}{3^k} + \sum_{i=k+1}^{\infty} \frac{a_i - b_i}{3^i} = 0,$$

we have

$$\frac{2}{3^k} = \sum_{i=k+1}^{\infty} \frac{b_i - a_i}{3^i} \leq \sum_{i=k+1}^{\infty} \frac{2}{3^i}$$

This implies that

$$1 \leq \sum_{i=1}^{\infty} \frac{1}{3^i} = \frac{1}{2},$$

a contradiction, Λ is injective. Since $\{0,2\}^{\mathbf{N}}$ is uncountable, $\Lambda(\{0,2\}^{\mathbf{N}}) \subset C$ is also uncountable. Thus C is uncountable.

Now we show that C is a null set. Since $C \subset I_n$ for all $n \in \mathbf{N}$, we have $m(C) \leq m(I_n)$. We can see that I_n is the union of disjoint closed intervals, we have $m(I_n) = \frac{n}{3^n}$. As $n \to \infty$, we have $m(C) \leq m(I_n) = 0$. Thus C is a null set. $\square$

Example 1.2.16. The half-open interval $[0,1)$ cannot be expressed as the countable union of disjoint closed *intervals*. In general, $[0,1)$ cannot be expressed as the countable union of disjoint closed *sets*.

Now we look at the Lebesgue measure $m(E)$ of a Lebesgue measurable set E, which is defined to equal its Lebesgue outer measure $m^*(E)$. If E is

Jordan measurable, we see from (1.2) that the Lebesgue measure and the Jordan measure of E coincide, thus Lebesgue measure extends Jordan measure. This justifies the use of the notation $m(E)$ to denote both Lebesgue measure of a Lebesgue measurable set, and Jordan measure of a Jordan measurable set (as well as elementary measure of an elementary set).

Lebesgue measure obeys significantly better properties than Lebesgue outer measure, when restricted to lebesgue measurable sets:

Lemma 1.2.17 (The measure axioms)**.**

(i) *(Empty set)* $m(\emptyset) = 0$.

(ii) *(Countable additivity) If $E_1, E_2, \dots \subset \mathbf{R}^d$ is a countable sequence of disjoint Lebesgue measurable sets, then*

$$m\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} m(E_n).$$

Proof. Proof omitted. $\square$

Theorem 1.2.18 (Monotone convergence thm for measurable sets)**.**

(i) *(Upward monotone convergence) Let $E_1 \subset E_2 \subset \dots \subset \mathbf{R}^n$ be a countable non-decreasing sequence of Lebesgue measurable sets. Then we have $m(\bigcup_{n=1}^{\infty} E_n) = \lim_{n\to\infty} m(E_n)$.*

(ii) *(Downward monotone convergence) Let $\mathbf{R}^d \supset E_1 \supset E_2 \supset \cdots$ be a countable non-increasing sequence of Lebesgue measurable sets. If at least one of the $m(E_n)$ is finite, then we have $m(\bigcap_{n=1}^{\infty} E_n) = \lim_{n\to\infty} m(E_n)$.*

Proof. (i) Let $A_1 := E_1$ and let $A_n := E_n \setminus E_{n-1}$. Then for every $n \geq 1$, A_n are disjoint. Since $E_n = \bigcup_{i=1}^{n} A_i$ and $\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} A_n$, by Lemma 1.2.17(ii),

$$m\Big(\bigcup_{n=1}^{\infty} E_n\Big) = m\Big(\bigcup_{n=1}^{\infty} A_n\Big)$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} m(A_n) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} m(A_i) \\
&= \lim_{n\to\infty} m\Big(\bigcup_{i=1}^{n} A_i\Big) \\
&= \lim_{n\to\infty} m(E_n).
\end{aligned}
$$

(ii) Let $A_1 := \emptyset$ and $A_n = E_1 \setminus A_n$. Then we have $A_n \subset A_{n+1}$ for all $n \geq 1$. From above conclusion, we have

$$m\Big(\bigcup_{n=1}^{\infty} A_n\Big) = m\Big(\bigcup_{n=1}^{\infty} (E_1 \setminus E_n)\Big) = \lim_{n\to\infty} m(A_n) = \lim_{n\to\infty} m(E_1 \setminus E_n).$$

Since $\bigcup_{n=1}^{\infty}(E_1 \setminus E_n) = E_1 \setminus \bigcap_{n=1}^{\infty} E_n$, we have

$$m\Big(\bigcup_{n=1}^{\infty} (E_1 \setminus E_n)\Big) = m\Big(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} m(E_1 \setminus E_n).$$

Use additivity of Lebesgue measure, we see that RHS of the second equality above can be written as

$$\lim_{n\to\infty} m(E_1 \setminus E_n) = \lim_{n\to\infty} (m(E_1) - m(E_n)) = m(E_1) - \lim_{n\to\infty} m(E_n),$$

and LHS of the second equality can be written as

$$m\Big(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\Big) = m(E_1) - m\Big(\bigcap_{n=1}^{\infty} E_n\Big).$$

Thus we have

$$m(E_1) - \lim_{n\to\infty} m(E_n) = m(E_1) - m\Big(\bigcap_{n=1}^{\infty} E_n\Big).$$

This means that $m(\bigcap_{n=1}^{\infty} E_n) = \lim_{n\to\infty} m(E_n)$, as desired. $\square$

Corollary 1.2.19. *Any map $E \mapsto \mu(E)$ from Lebesgue measurable sets to elements of $[0, +\infty]$ that obeys the above empty set and countable additivity axioms will also obey the monotonicity and countable subadditivity axioms from Lemma 1.2.3, when restricted to Lebesgue measurable sets of course.*

Proof. Let μ be arbitrary map from Lebesgue measurable sets to $[0, +\infty]$ obeys the Lemma 1.2.17. Monotonicity and countable subadditivity are trivial when there is unbounded set. We assume that all of sets are bounded, so that with finite measure.

For Lebesgue measurable sets E, F where $E \subset F$, we have E and $F \setminus E$ are two disjoint Lebesgue measurable sets. From finite additivity (which is the implication of countable additivity), we have

$$\mu(F) = \mu(E \cup (F \setminus E)) = \mu(E) + \mu(F \setminus E) \geq \mu(E).$$

Thus μ is monotone.

For the countable subadditivity, let $E_1, E_2, \cdots$ be a sequence of countable Lebesgue measurable sets. Let $E := \bigcup_{n=1}^{\infty} E_n$. For every E_n we can find a subset $F_n \subset E_n$ which is defined as $F_n := E_n \setminus \bigcup_{i=1}^{n-1} E_i$. Then $(F_n)_{n=1}^{\infty}$ consisting a disjoint sequence and we have

$$\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} \left(E_n \setminus \bigcup_{i=1}^{n-1} E_i \right) = \bigcup_{n=1}^{\infty} E_n,$$

this holds for that $\bigcup_{i=1}^{0} E_i = \emptyset$.

Now consider the sequence $(E \cap E_n)_{n=1}^{\infty}$, we have

$$\begin{aligned} \mu(E) &= \mu\Big(\bigcup_{n=1}^{\infty} E \cap E_n \Big) \\ &\leq \mu\Big(\bigcup_{n=1}^{\infty} E_n \Big) \\ &= \mu\Big(\bigcup_{n=1}^{\infty} F_n \Big) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} \mu(F_n) \\
&= \sum_{n=1}^{\infty} \mu\Big(E_n \setminus \bigcup_{i=1}^{n-1} E_i\Big) \\
&\le \sum_{n=1}^{\infty} \mu(E_n),
\end{aligned}
$$

where the second line holds from the monotonicity; the fourth and fifth lines hold from the countable additivity. Thus μ is countably subadditive. $\square$

Theorem 1.2.20. *We say that a sequence E_n of sets in $\mathbf{R}^d$ converges pointwise to another set E in $\mathbf{R}^d$ if the indicator functions 1_{E_n} converge pointwise to 1_E.*

(i) *If the E_n are all Lebesgue measurable, and converge pointwise to E, then E is Lebesgue measurable also.*

(ii) *(Dominated convergence theorem) Let E_n, E be as in part (i). Suppose that the E_n are all contained in another Lebesgue measurable set F of finite measure. Then $m(E_n)$ converges to $m(E)$.*

Proof. (i) Suppose that E_n converges pointwise to E, then

$$1_E(x) = \limsup_{n\to\infty} 1_{E_n}(x) = \liminf_{n\to\infty} 1_{E_n}(x).$$

For every $x \in E$ we have

$$
\begin{aligned}
E &= \{x \in \mathbf{R}^d : x \in E\} \\
&= \{x \in \mathbf{R}^d : 1_E(x) = \liminf_{n\to\infty} 1_{E_n}(x) = 1\} \\
&= \{x \in \mathbf{R}^d : \sup_{N>0} \inf_{n\ge N} 1_{E_n}(x) = 1\} \\
&= \bigcup_{N>0} \bigcap_{n\ge N} \{x \in \mathbf{R}^d : 1_{E_n}(x) = 1\} \\
&= \bigcup_{N>0} \bigcap_{n\ge N} E_n.
\end{aligned}
$$

This means that E can be represented as the countable union and intersection of E_n, by Theorem 1.2.12(vi, vii), E is Lebesgue measurable.

(ii) Since we have $E = \bigcup_{N>0} \bigcap_{n \geq N} E_n$ and $\bigcap_{n \geq N} E_n \subset \bigcap_{n \geq N+1} E_n$ for every $N \in \mathbf{N}$, by Theorem 1.2.18,

$$\begin{aligned} m(E) &= m\Big(\bigcup_{N>0} \bigcap_{n \geq N} E_n \Big) \\ &= \lim_{N \to \infty} m\Big(\bigcap_{n \geq N} E_n \Big) \\ &\leq \lim_{N \to \infty} m(E_N). \end{aligned}$$

Other hand, we have

$$\begin{aligned} E &= \{x \in \mathbf{R}^d : 1_E(x) = \limsup_{n \to \infty} 1_{E_n}(x) = 1\} \\ &= \{x \in \mathbf{R}^d : \inf_{N>0} \sup_{n \geq N} 1_{E_n}(x) = 1\} \\ &= \bigcap_{N>0} \bigcup_{n \geq N} \{x \in \mathbf{R}^d : 1_{E_n}(x) = 1\} \\ &= \bigcap_{N>0} \bigcup_{n \geq N} E_n. \end{aligned}$$

Since $\bigcup_{n \geq N} E_n \supset \bigcup_{n \geq N+1} E_n$ for every $N \in \mathbf{N}$, we have

$$\begin{aligned} m(E) &= m\Big(\bigcap_{N>0} \bigcup_{n \geq N} E_n \Big) \\ &= \lim_{N \to \infty} m\Big(\bigcup_{n \geq N} E_n \Big) \\ &\geq \lim_{N \to \infty} m(E_N). \end{aligned}$$

Together our conclusions, we have $m(E) = \lim_{n \to \infty} m(E_n)$. □

Remark. In later sections we will generalise the monotone and dominated convergence theorems to measurable functions instead of measurable sets; see Theorem 1.4.51 and Theorem 1.4.55.

Proposition 1.2.21. *Let $E \subset \mathbf{R}^d$. E is contained in a Lebesgue measurable set of measure exactly equal to $m^*(E)$.*

Proof. Let $(U_n)_{n=1}^{\infty}$ be the countable family of open sets containing E, which are Lebesgue measurable. Let $U := \bigcap_{n=1}^{\infty} U_n$, it is also Lebesgue measurable and containing E. Now we want to show that $m(U) = m^*(E)$. By outer regularity, we have $m(U) \geq m^*(E)$, so it suffices to show that

$$m(U) \leq \inf_{E \subset U_n} m(U_n) = m^*(E).$$

Let $\varepsilon > 0$ be arbitrary. There exists an open set U_n such that

$$\begin{aligned} m^*(E) &\geq m(U_n) - \varepsilon \\ &\geq m\Big(\bigcap_{n=1}^{\infty} U_n\Big) - \varepsilon \\ &= m(U) - \varepsilon. \end{aligned}$$

Because ε is arbitrary, we have $m^*(E) \geq m(U)$, as desired. □

Lemma 1.2.22 (Inner regularity)**.** *Let $E \subset \mathbf{R}^d$ be Lebesgue measurable. Then one has*

$$m(E) = \sup_{K \subset E, K \text{ compact}} m(K).$$

Proof. From monotonicity (Proposition 1.2.19) one trivially has

$$m(E) \geq \sup_{K \subset E, K \text{ compact}} m(K)$$

so it suffices to show that

$$m(E) \leq \sup_{K \subset E, K \text{ compact}} m(K).$$

We first suppose that E is bounded. Since E is Lebesgue measurable, by Theorem 1.2.13(iv), there exists a closed set $K \subset E$, so that is compact for E is bounded, such that $m(E \setminus K) \leq \varepsilon$. Thus we have $m(E) = m(K) + m(E \setminus K) \leq m(K) + \varepsilon$, this means that $m(E) = \sup_{K \subset E} m(K)$.

Now we suppose that E is unbounded. Consider the sequence of sets $E_n := E \cap (-n, n)^d$ for $n \in \mathbf{N}$. We can see that E_n is bounded and Lebesgue measurable for all $n \in \mathbf{N}$. Then for every $\varepsilon > 0$ and every $n > 0$ there exists a compact set $K_n \subset E_n$ such that $m(E_n) - \varepsilon \leq m(K_n)$. Since $E_1 \subset E_2 \subset \cdots$, by monotone convergence theorem (Theorem 1.2.18(i)), we have

$$\sup_{K \subset E} m(K) \geq \lim_{n\to\infty} m(K_n) \geq \lim_{n\to\infty} m(E_n) - \varepsilon = m(E) - \varepsilon$$

As ε was arbitrary, we have $\sup_{K \subset E} m(K) \geq m(E)$. This complete the proof. $\square$

Theorem 1.2.23 (Criteria for finite measure)**.** *Let $E \subset \mathbf{R}^d$. Then following are equivalent:*

(i) *E is Lebesgue measurable with finite measure.*

(ii) *(Outer approximation by open) For every $\varepsilon > 0$, there exists an open set $U \supset E$ with finite measure such that $m^*(U \setminus E) \leq \varepsilon$.*

(iii) *(Almost open bounded) For every $\varepsilon > 0$, there exists a bounded open set U such that $m^*(E \triangle U) \leq \varepsilon$. (In other words, E differs from a bounded open set by a set of arbitrarily small Lebesgue outer measure.)*

(iv) *(Inner approximation by compact) For every $\varepsilon > 0$, there exists a compact set $F \subset E$ with $m^*(E \setminus F) \leq \varepsilon$.*

(v) *(Almost compact) For every $\varepsilon > 0$, there exists a compact set F such that $m^*(E \triangle F) \leq \varepsilon$.*

(vi) *(Almost bounded measurable) For every $\varepsilon > 0$, there exists a bounded Lebesgue measurable set F such that $m^*(E \triangle F) \leq \varepsilon$.*

(vii) *(Almost finite measure) For every $\varepsilon > 0$, there exists a Lebesgue measurable set F with finite measure such that*

$m^*(E\triangle F) \leq \varepsilon$.

(viii) *(Almost elementary) For every $\varepsilon > 0$, there exists an elementary set F such that $m^*(E\triangle F) \leq \varepsilon$.*

(ix) *(Almost dyadically elementary) For every $\varepsilon > 0$, there exists an integer n and a finite union F of closed dyadic cubes of sidelength 2^{-n} such that $m^*(E\triangle F) \leq \varepsilon$.*

Proof. Notice that for every bounded Lebesgue measurable set $E \subset \mathbf{R}^d$, we have $m(E) \leq m^{*,(J)}(E) < \infty$, i.e., $m(E)$ is finite. But the converse statement is not always true.

We first show that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii). By Theorem 1.2.13, we only need to show that $m(U)$ is finite. By additivity of Lebesgue measure, we have $m(U) = m(U \setminus E) + m(E) \leq \varepsilon + m(E)$, since $m(E)$ is finite, $m(U)$ is also finite.

(ii) $\Rightarrow$ (iii). For every $\varepsilon > 0$, we can find an open set $U \supset E$ with finite measure such that $m^*(U \setminus E) \leq \varepsilon$. If U is bounded, our proof is done.

While if U is unbounded. Let $U_n := U \cap (-n, n)^d$ for $n \in \mathbf{N}$. Clearly, U_n are bounded and open. Since $U_n \setminus E \subset U_{n+1} \setminus E$ for every $n \in \mathbf{N}$, by Theorem 1.2.18(i), we have

$$m(U\triangle E) = m(U \setminus E) = m\Big(\bigcup_{n=1}^{\infty} U_n \setminus E\Big) = \lim_{n\to\infty} m(U_n \setminus E) \leq \varepsilon.$$

This means that for every $\varepsilon > 0$, there is an $N > 0$ such that

$$m(U_N \setminus E) \leq m(U\triangle E) + \varepsilon \leq 2\varepsilon,$$

where U_N is bounded open set, as desired.

(iii) $\Rightarrow$ (i). Clearly, E is Lebesgue measurable. From $m(U\triangle E) = m(U \setminus E) + m(E \setminus U) \leq \varepsilon$, we have $m(E \setminus U) \leq \varepsilon$. This implies $m(E) = m(E \setminus U) + m(U) \leq \varepsilon + m(U)$. Since U is bounded, so that $m(U)$ is finite. Thus we conclude that $m(E)$ is finite.

Then we show that (i) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (vi) $\Rightarrow$ (vii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (iv). Let $\varepsilon > 0$. By inner regularity, there is a compact set $F \subset E$ such that $m(E) \leq m(F) + \varepsilon$. This implies $m(E \setminus F) = m(E) - m(F) \leq \varepsilon$.

(iv) $\Rightarrow$ (v). This is trivial to prove.

(v) $\Rightarrow$ (vi). This is also trivial.

(vi) $\Rightarrow$ (vii). Bounded set obviously has finite measure, (vi) clearly implies (vii).

(vii) $\Rightarrow$ (i). E is clearly Lebesgue measurable from Theorem 1.2.13(vi). We show that E has finite measure. From $m(F\triangle E) \leq \varepsilon$ we have $m(E) = m(E \setminus F) + m(F) \leq \varepsilon + m(F)$. Since F is bounded, so that has finite measure. Thus $m(E)$ is finite.

Finally, we show that (i) $\Rightarrow$ (ix) $\Rightarrow$ (viii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ix). By definition, there is an open set $U \supset E$ such that $m(U\triangle E) = m(U \setminus E) \leq \varepsilon/2$. Let $\mathcal{Q}_n$ be the collection of n's almost disjoint closed dyadic cubes Q, each of which are contained in U. We can see that $\mathcal{Q} := \lim_{n\to\infty} \mathcal{Q}_n$ is the collection of countable many Q. Then we have $U = \bigcup_{Q\in\mathcal{Q}} Q$ (thanks to Lemma 1.2.10).

We see that $\bigcup_{Q\in\mathcal{Q}_n} Q$ are all contained in U and converges pointwise to $\bigcup_{Q\in\mathcal{Q}} Q$. By Theorem 1.2.20(ii), we have

$$\lim_{n\to\infty} m\Big(\bigcup_{Q\in\mathcal{Q}_n} Q\Big) = m\Big(\bigcup_{Q\in\mathcal{Q}} Q\Big).$$

This means that for every $\varepsilon > 0$, there is an $N > 0$ such that

$$m\Big(\bigcup_{Q\in\mathcal{Q}_N} Q\Big) \geq m\Big(\bigcup_{Q\in\mathcal{Q}} Q\Big) - \varepsilon.$$

Let $F := \bigcup_{Q\in\mathcal{Q}_N} Q$, then we have

$$\begin{aligned} m(F\triangle E) &= m\Big(\bigcup_{Q\in\mathcal{Q}_N} Q \setminus E\Big) + m\Big(E \setminus \bigcup_{Q\in\mathcal{Q}_N} Q\Big) \\ &\leq m\Big(\bigcup_{Q\in\mathcal{Q}} Q \setminus E\Big) + m\Big(E \setminus \bigcup_{Q\in\mathcal{Q}_N} Q\Big) \\ &= m\Big(\bigcup_{Q\in\mathcal{Q}} Q \setminus E\Big) + m(E) - m\Big(\bigcup_{Q\in\mathcal{Q}_N} Q\Big) \\ &\leq m\Big(\bigcup_{Q\in\mathcal{Q}} Q \setminus E\Big) + m(E) - m\Big(\bigcup_{Q\in\mathcal{Q}} Q\Big) + \varepsilon \end{aligned}$$

$$= m\Big(\bigcup_{Q\in\mathcal{Q}} Q \setminus E\Big) + m\Big(E \setminus \bigcup_{Q\in\mathcal{Q}} Q\Big) + \varepsilon$$
$$= m(U \triangle E) + \varepsilon.$$

Thus we have $m(F\triangle E) \le \varepsilon$, as desired.

(ix) $\Rightarrow$ (viii). This immediately comes from the definition of F.

(viii) $\Rightarrow$ (i). Use (vi), we obviously have E is Lebesgue measurable with finite measure. □

Proposition 1.2.24 (Carathéodory criterion, one direction)**.** *Let $E \subset \mathbf{R}^d$. Then following are equivalent:*

(i) *E is Lebesgue measurable.*

(ii) *For every elementary set A, we have $m(A) = m^*(A \cap E) + m^*(A \setminus E)$.*

(iii) *For every box B, we have $|B| = m^*(B \cap E) + m^*(B \setminus E)$.*

Proof. (i) $\Rightarrow$ (ii). Suppose that E is Lebesgue measurable. Then $A \cap E$ and $A \setminus E$ are also Lebesgue measurable. From finite additivity, we have

$$m(A) = m(A \cap E) + m(A \setminus E).$$

(ii) $\Rightarrow$ (iii). Since every box B is elementary set, we have

$$|B| = m(B) = m^*(B \cap E) + m^*(B \setminus E).$$

(iii) $\Rightarrow$ (i). Suppose that $m^*(E)$ is finite. Let $\varepsilon > 0$. By definition of outer measure, there is a countable family $B_1, B_2, \cdots$ of boxes covering E such that

$$\sum_{n=1}^{\infty} |B_n| \le m^*(E) + \varepsilon.$$

We can enlarge each of these boxes B_n to an open box B_n' such that $|B_n'| \le |B_n| + \varepsilon/2^n$ (see the $\varepsilon/2^n$ trick). Then the set $\bigcup_{n=1}^{\infty} B_n'$ is open and contains E, and we have

$$m^*(E) + m^*\Big(\bigcup_{n=1}^{\infty} B_n' \setminus E\Big) = m^*\Big(\bigcup_{n=1}^{\infty} B_n' \cap E\Big) + m^*\Big(\bigcup_{n=1}^{\infty} B_n' \setminus E\Big)$$

$$
\begin{aligned}
&\leq \sum_{n=1}^{\infty} m^*(B'_n \cap E) + m^*(B'_n \setminus E) \\
&= \sum_{n=1}^{\infty} |B'_n| \\
&\leq \sum_{n=1}^{\infty} |B_n| + \varepsilon \\
&\leq m^*(E) + 2\varepsilon.
\end{aligned}
$$

This means that $m^*(\bigcup_{n=1}^{\infty} B'_n \setminus E) \leq 2\varepsilon$ and E is Lebesgue measurable.

While if $m^*(E)$ is infinite. We use a countable family $A_1, A_2, \cdots$ of open boxes covering E. Let $E_n := A_n \cap E$, which has finite outer measure. Use the above conclusion of finite case, E_n is Lebesgue measurable for every $n > 0$. Then for every $n > 0$ we can find an open set $U_n \supset E_n$ such that $m^*(U_n \backslash E_n) \leq \varepsilon/2^n$. From monotonicity and subadditivity of outer measure, we have

$$
m^*\Big(\bigcup_{n=1}^{\infty} U_n \setminus E\Big) \leq \sum_{n=1}^{\infty} m^*(U_n \setminus E) \leq \sum_{n=1}^{\infty} m^*(U_n \setminus E_n) \leq \varepsilon.
$$

Thus we can conclude that E is Lebesgue measurable. $\square$

Proposition 1.2.25. *Let $E \subset \mathbf{R}^d$ be a bounded set. Define the* Lebesgue inner measure $m_*(E)$ *of E by the formula*

$$
m_*(E) := m(A) - m^*(A \setminus E)
$$

for any elementary set A containing E.

(i) *This definition is well-defined, i.e., that if A, A' are two elementary sets containing E, then $m(A) - m^*(A \setminus E)$ is equal to $m(A') - m^*(A' \setminus E)$.*

(ii) *We have $m_*(E) \leq m^*(E)$, and that equality holds if and only if E is Lebesgue measurable.*

Proof. (i) Let A, A' be two elementary sets containing E, then $A \cap A'$ is also elementary contains E. We want to show that

$$m(A) - m^*(A \setminus E) = m(A \cap A') - m^*((A \cap A') \setminus E).$$

From $A = (A \cap A') \cup (A \setminus A')$, Proposition 1.2.24(ii), and the subadditivity of Lebesgue outer measure, we have

$$\begin{aligned}
m(A) - m^*(A \setminus E) &= m(A \cap A') + m(A \setminus A') \\
&\qquad - m^*(((A \cap A') \cup (A \setminus A')) \setminus E) \\
&\geq m(A \cap A') + m(A \setminus A') \\
&\qquad - m^*((A \cap A') \setminus E) - m^*((A \setminus A') \setminus E) \\
&= m(A \cap A') + m(A \setminus A') \\
&\qquad - m^*((A \cap A') \setminus E) - m(A \setminus A') \\
&= m(A \cap A') - m^*((A \cap A') \setminus E).
\end{aligned}$$

For the other hand, we want to show that $m(A) - m^*(A \setminus E) \leq m(A \cap A') - m^*((A \cap A') \setminus E)$. This equivalent to show that $m(A \setminus A') + m^*((A \cap A') \setminus E) \leq m^*(A \setminus E)$ for that $m(A) - m(A \cap A') = m(A \setminus A')$. By definition of outer measure, we can find a sequence of boxes B_n such that $\bigcup_{n=1}^{\infty} B_n \supset A \setminus E$ and

$$m^*(A \setminus E) \geq \sum_{n=1}^{\infty} |B_n| - \varepsilon.$$

Use Venn diagram, we see that $(A \cap A') \setminus E \subset \bigcup_{n=1}^{\infty} B_n \cap (A \cap A')$ and $A \setminus A' \subset \bigcup_{n=1}^{\infty} B_n \setminus (A \cap A')$. Then by Proposition 1.2.24,

$$\begin{aligned}
m(A \setminus A') + m^*((A \cap A') \setminus E) &\leq m\Big(\bigcup_{n=1}^{\infty} B_n \setminus (A \cap A')\Big) \\
&\qquad + m\Big(\bigcup_{n=1}^{\infty} B_n \cap (A \cap A')\Big) \\
&= \sum_{n=1}^{\infty} |B_n| - \varepsilon \\
&\leq m^*(A \setminus E).
\end{aligned}$$

Thus we have $m(A) - m^*(A \setminus E) = m(A \cap A') - m^*((A \cap A') \setminus E)$. A similar argument shows that $m(A') - m^*(A' \setminus E) = m(A \cap A') - m^*((A \cap A') \setminus E)$. Thus $m(A) - m^*(A \setminus E) = m(A') - m^*(A' \setminus E)$, and the Lebesgue inner measure is well-defined.

(ii) Since $A = (A \cap E) \cup (A \setminus E)$, by subadditivity of Lebesgue outer measure, we have

$$m(A) \leq m^*(A \cap E) + m^*(A \setminus E) = m^*(E) + m^*(A \setminus E).$$

Then

$$m_*(E) = m(A) - m^*(A \setminus E) \leq m^*(E).$$

In particular, if E is Lebesgue measurable, we have

$$m(A) = m^*(A \cap E) + m^*(A \setminus E) = m^*(E) + m^*(A \setminus E),$$

so that $m_*(E) = m^*(E)$. □

Definition 1.2.26 (G_δ set and F_σ set)**.** Define a G_δ set to be the countable intersection $\bigcap_{n=1}^{\infty} U_n$ of open sets, and an F_σ set to be the countable union $\bigcup_{n=1}^{\infty} F_n$ of closed sets.[a]

[a]The terminology G_δ comes from German "Gebiete" (means field) and "Durchschnitt" (means cut); F_σ comes from French "fermé" (means closed) and "somme" (means sum).

Proposition 1.2.27. *Let $E \subset \mathbf{R}^d$. Then the following are equivalent:*

(i) *E is Lebesgue measurable.*

(ii) *E is G_δ set with a null set removed, i.e., $E = G \setminus N$ where G is a G_δ set and N is a null set.*

(iii) *E is the union of an F_σ set and a null set, i.e., $E = F \cup N$ where F is an F_σ set and N is a null set.*

Proof. (i) $\Rightarrow$ (ii). Suppose that E is Lebesgue measurable, by outer regularity (Lemma 1.2.11), there are open sets U_n containing E such that

$$m(E) \geq m(G_n) - \frac{1}{n}$$

for every $n \geq 1$. Since E and G_n are all Lebesgue measurable, we have

$$m(G_n) - m(E) = m(G_n \setminus E) \leq \frac{1}{n}.$$

Let $G := \bigcap_{n=1}^{\infty} G_n$, which is a G_δ set and containing E. Then we have

$$m(G \setminus E) \leq m(G_n \setminus E) = m(G_n) - m(E) \leq \frac{1}{n}$$

for every $n \geq 1$. This means that there is a null set $N := G \setminus E$ when $n \to \infty$ such that $E = G \setminus (G \setminus E) = G \setminus N$, as desired.

(ii) $\Rightarrow$ (iii). Let $E = G \setminus N$ be as in above. From Lemma 1.2.12, E is Lebesgue measurable. Then from Lemma 1.2.22, there are compact, and then closed, sets F_n contained in E such that

$$m(E) \leq m(F_n) + \frac{1}{n}$$

for every $n \geq 1$. Since F_n are all Lebesgue measurable, we have

$$m(E) - m(F_n) \leq \frac{1}{n}.$$

Let $F = \bigcup_{n=1}^{\infty} F_n$, which is an F_σ set and contained in E. Then we have

$$m(E \setminus F) = m\Big(\bigcap_{n=1}^{\infty} E \setminus F_n\Big) \leq m(E \setminus F_n) = m(E) - m(F_n) \leq \frac{1}{n}.$$

for every $n \geq 1$. This means that there is a null set $N' := F \setminus F$ when $n \to \infty$ such that $E = F \cup (E \setminus F) = F \cup N'$, as desired.

(iii) $\Rightarrow$ (i). This is immediately comes from Lemma 1.2.12, thus we compete the proof. □

Proposition 1.2.28 (Translation invariance)**.** *If $E \subset \mathbf{R}^d$ is Lebesgue measurable, then $E + x$ is Lebesgue measurable for any $x \in \mathbf{R}^d$, and* *that $m(E+x)=m(E)$.*

Proof. Let E be arbitrary subset $E \subset \mathbf{R}^d$. We first show that Lebesgue outer measure obeys translation invariance. Suppose $B_1, B_2, \cdots$ be a sequence of boxes whose union contains E. Then $B_1+x, B_2, +x, \cdots$ is a sequence of boxes whose union contains $E+x$. Then from the translation invariance of Jordan (or elementary) measure,

$$m^*(E+x) \leq \sum_{n=1}^{\infty} |B_n + x| = \sum_{n=1}^{\infty} |B_n|.$$

Taking the infimum over all sequence $B_1, B_2, \cdots$, we have $m^*(E+x) \leq m^*(E)$.

In the other direction, notice that $E = -x + (E+x)$. Thus from above inequality, we have

$$m^*(E) = m^*(-x + (E+x)) \leq m^*(E+x).$$

Thus we conclude that $m^*(E+x) = m^*(E)$ for every $x \in \mathbf{R}^d$. From Carathéodory criterion, for every Lebesgue measurable set E and every elementary set A, we have

$$\begin{aligned} m(A) &= m(A-x) \\ &= m^*((A-x) \cap E) + m^*((A-x) \setminus E) \\ &= m^*((A-x) \cap E + x) + m^*((A-x) \setminus E + x) \\ &= m^*(A \cap (E+x)) + m^*(A \setminus (E+x)). \end{aligned}$$

Thus $E+x$ is Lebesgue measurable, and we have $m(E+x) = m(E)$ for that $m^*(E+x) = m^*(E)$. □

Lemma 1.2.29. *Let $d, d' \geq 1$ be natural numbers.*

(i) *If $E \subset \mathbf{R}^d$ and $F \subset \mathbf{R}^{d'}$, then*

$$(m^{d+d'})^*(E \times F) \leq (m^d)^*(E)(m^{d'})^*(F),$$

where $(m^d)^*$ denotes d-dimensional Lebesgue outer measure, etc.

(ii) *Let* $E \subset \mathbf{R}^d, F \subset \mathbf{R}^{d'}$ *be Lebesgue measurable sets. Then* $E \times F \subset \mathbf{R}^{d+d'}$ *is Lebesgue measurable, with* $m^{d+d'}(E \times F) = m^d(E) \cdot m^{d'}(F)$.

Proof. (i) The inequality is trivial when E or F has infinity outer measure. We suppose that $(m^d)^*(E), (m^{d'})^*(F) < \infty$. By outer regularity, we have

$$(m^d)^*(E) \geq (m^d)^*(U) - \varepsilon$$

and

$$(m^{d'})^*(F) \geq (m^{d'})^*(V) - \varepsilon,$$

where U and V are open and containing E and F, respectively. Since every open set can be represented by countable union of almost disjoint boxes, thus we have $\bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} A_n \times B_m \supset E \times F$ and

$$\begin{aligned}
(m^d)^*(E)(m^{d'})^*(F) &\geq ((m^d)^*(U) - \varepsilon)((m^{d'})^*(V) - \varepsilon) \\
&= \Big((m^d)^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) - \varepsilon\Big)\Big((m^{d'})^*\Big(\bigcup_{m=1}^{\infty} B_m\Big) - \varepsilon\Big) \\
&\geq (m^d)^*\Big(\bigcup_{n=1}^{\infty} A_n\Big)(m^{d'})^*\Big(\bigcup_{m=1}^{\infty} B_m\Big) - \delta \\
&= \sum_{n=1}^{\infty} |A_n| \sum_{m=1}^{\infty} |B_m| - \delta \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |A_n \times B_m| - \delta \\
&\geq (m^{d+d'})^*(E \times F) - \delta,
\end{aligned}$$

where

$$\delta := \varepsilon\Big((m^d)^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) + (m^{d'})^*\Big(\bigcup_{m=1}^{\infty} B_m\Big)\Big).$$

Because ε is arbitrary, this implies δ is also arbitrary. Thus we have $(m^{d+d'})^*(E \times F) \leq (m^d)^*(E)(m^{d'})^*(F)$.

(ii) Since E, F are Lebesgue measurable, by Proposition 1.2.27, there exist G_δ sets G_1 and G_2 such that $E \subset G_1$ and $F \subset G_2$ such that $m^d(G_1 \setminus E) = m^{d'}(G_2 \setminus F) = 0$. Clearly, $G_1 \times G_2$ is open so that Lebesgue measurable, and we have

$$(G_1 \times G_2) \setminus (E \times F) \subset (G_1 \setminus E) \times G_2 \cup G_1 \times (G_2 \setminus F).$$

Use the inequality in part (i), we have

$$m^{d+d'}((G_1 \times G_2) \setminus (E \times F)) \leq 0.$$

By Proposition 1.2.27 again, $E \times F$ is Lebesgue measurable.

Now we want to show that $m^{d+d'}(E \times F) = m^d(E) \cdot m^{d'}(F)$. This is trivial when $E \times F$ has infinity measure, thus we assume that $E \times F$ has finite measure. By Theorem 1.2.23(ii), there is a bounded open set $U \supset E$ and $V \supset F$ with finite measure. Let $G_1 := \bigcap_{n=1}^\infty U_n$, and $G_2 := \bigcap_{n=1}^\infty V_n$. Let

$$A_k := U \cap \bigcap_{n=1}^k U_n, \qquad B_k := V \cap \bigcap_{n=1}^k V_n.$$

Then we obtain two decreasing sequences $A_1 \supset A_2 \supset \cdots$ and $B_1 \supset B_2 \supset \cdots$ satisfying

$$\bigcap_{k=1}^\infty A_k = \bigcap_{n=1}^\infty U_n, \qquad \bigcap_{k=1}^\infty B_k = \bigcap_{n=1}^\infty V_n.$$

Clearly, we also have $A_1 \times B_1 \supset A_2 \times B_2 \supset \cdots$. Then we have

$$\begin{aligned}
m(E)m(F) &= m\Big(\bigcap_{n=1}^\infty U_n\Big) m\Big(\bigcap_{n=1}^\infty V_n\Big) \\
&= m\Big(\bigcap_{k=1}^\infty A_k\Big) m\Big(\bigcap_{k=1}^\infty B_k\Big) \\
&= \lim_{k\to\infty} m(A_k)m(B_k) \\
&= \lim_{k\to\infty} m(A_k \times B_k)
\end{aligned}$$

$$
\begin{aligned}
&= m\Big(\bigcap_{k=1}^{\infty} A_k \times \bigcap_{k=1}^{\infty} B_k\Big) \\
&= m\Big(\bigcap_{n=1}^{\infty} U_n \times \bigcap_{n=1}^{\infty} V_n\Big) \\
&= m(E \times F),
\end{aligned}
$$

where the first and seventh lines hold from Proposition 1.2.27; the third and fifth lines hold from the monotone convergence theorem (Theorem 1.2.18(ii)); the fourth line holds for that A_k and B_k are bounded, so that can be represented as finitely many boxes, then use Lemma 1.1.18. □

Theorem 1.2.30 (Uniqueness of Lebesgue measure)**.** *Lebesgue measure $E \mapsto m(E)$ is the only map from Lebesgue measurable sets to $[0, +\infty]$ that obeys the following axioms:*

(i) *(Empty set) $m(\emptyset) = 0$.*

(ii) *(Countable additivity) If $E_1, E_2, \cdots \subset \mathbf{R}^d$ is a countable sequence of disjoint Lebesgue measurable sets, then*

$$m\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} m(E_n).$$

(iii) *(Translation invariance) If E is Lebesgue measurable and $x \in \mathbf{R}^d$, then $m(E + x) = m(E)$.*

(iv) *(Normalisation) $m([0, 1]^d) = 1$.*

Proof. Suppose that there is another map m' satisfying above properties. From Corollary 1.2.19, we know that m, m' also obey the monotonicity and countable subadditivity.

By Proposition 1.2.14, every elementary set is Lebesgue measurable, then from the uniqueness of elementary (or Jordan) measure, the map m is unique for every elements of $\mathcal{E}(\mathbf{R}^d)$, i.e., $m(E) = m'(E)$ for every $E \in \mathcal{E}(\mathbf{R}^d)$. In particular, this is hold for every box.

Let $\varepsilon > 0$, for arbitrary $E \subset \mathbf{R}^d$ there is some sequence of boxes such that

$$m^*(E) \geq \sum_{n=1}^{\infty} |B_n| - \varepsilon = m\Big(\bigcup_{n=1}^{\infty} B_n\Big) - \varepsilon \geq m(E) - \varepsilon.$$

As ε was arbitrary, we have $m^*(E) \geq m(E)$. A similar argument shows that $m^*(E) \geq m'(E)$. From this, we implies that for every null set N, $m(N) = m'(N)$.

Now we consider the open set. For every open set which can be represented as the countable union of almost disjoint boxes B_n, we have

$$m'(U) = m'\Big(\bigcup_{n=1}^{\infty} B_n\Big) = \sum_{n=1}^{\infty} m'(B_n) = \sum_{n=1}^{\infty} m(B_n) = m\Big(\bigcup_{n=1}^{\infty} B_n\Big) = m(U)$$

Thus m and m' are also coincide for open sets.

For arbitrary bounded Lebesgue measurable set, there is a bounded open set $U \supset E$ such that $m(U) - m(E) \leq \varepsilon$. Then from monotonicity,

$$m'(E) \leq m'(U) = m(U) \leq m(E) + \varepsilon.$$

This implies $m'(E) \leq m(E)$ for every Lebesgue measurable set E. (One can see that the inequality is trivial when E has infinity measure.) Now we can conclude that for open set $U \supset E$, one has

$$m(E) = m(U) - m(U \setminus E) \leq m'(U) - m'(U \setminus E) = m'(E).$$

This means that m and m' are coincide for bounded Lebesgue measurable set.

For unbounded Lebesgue measurable set, let $E_n := E \cap (-n, n)^d$, then E_n is bounded for every n and $E = \bigcup_{n=1}^{\infty} E_n$. By countable additivity and conclusion above, we have

$$m'(E) = \sum_{n=1}^{\infty} m'(E_n) = \sum_{n=1}^{\infty} m(E_n) = m(E).$$

Thus we conclude that m and m' are coincide for arbitrary Lebesgue measurable set. $\square$

§1.3 The Lebesgue integral

1.3.1 Integration of simple functions

Much as the Riemann integral was set up by first using the integral for piecewise constant functions, the Lebesgue integral is set up using the integral for simple functions.

Definition 1.3.1 (Simple function)**.** A (complex-valued) *simple function* $f : \mathbf{R}^d \to \mathbf{C}$ is a finite linear combination

$$f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$$

of indicator functions 1_{E_i} of Lebesgue measurable sets $E_i \subset \mathbf{R}^d$ for $i = 1, \cdots, k$, where $k \geq 0$ is a natural number and $c_1, \cdots, c_k \in \mathbf{C}$ are complex numbers. An *unsigned simple function*[a] $f : \mathbf{R}^d \to [0, +\infty]$, is defined similarly, but with the c_i taking values in $[0, +\infty]$ rather than $\mathbf{C}$.

[a]The terminology *unsigned* comes from computer programming, means *non-negative.*

Remark. In this definition, we did not require the $E_1, \cdots, E_k$ to be disjoint. However, it is easy enough to arrange this, basically by exploiting Venn diagrams (or finite Boolean algebras). Indeed, any k subsets $E_1, \cdots, E_k$ of $\mathbf{R}^d$ partition $\mathbf{R}^d$ into 2^k disjoint sets, each of which is an intersection of E_i or the complement $\mathbf{R}^d \setminus E_i$ for $i = 1, \cdots, k$ (and in particular, is measurable).

Definition 1.3.2 (Integral of an unsigned simple function)**.** Let

$$f = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$$

be an unsigned simple function, the integral $\operatorname{Simp} \int_{\mathbf{R}^d} f(x) dx$ is defined by the formula

$$\mathrm{Simp} \int_{\mathbf{R}^d} f(x)dx := c_1 m(E_1) + \cdots + c_k m(E_k),$$

thus $\mathrm{Simp} \int_{\mathbf{R}^d} f(x)dx$ will take values in $[0, +\infty]$.

One has to actually check that this definition is well-defined:

Lemma 1.3.3 (Well-definedness of simple integral). *Let $k, k' \geq 0$ be natural numbers, $c_1, \cdots, c_k, c'_1, \cdots, c'_{k'} \in [0, +\infty]$, and let $E_1, \cdots, E_k$, $E'_1, \cdots, E'_{k'} \subset \mathbf{R}^d$ be Lebesgue measurable sets such that the identity*

$$c_1 1_{E_1} + \cdots + c_k 1_{E_k} = c'_1 1_{E'_1} + \cdots + c'_k 1_{E'_{k'}}$$

holds identically on $\mathbf{R}^d$. Then one has

$$c_1 m(E_1) + \cdots + c_k m(E_k) = c'_1 m(E'_1) + \cdots + c'_k m(E'_{k'}).$$

Proof. Proof omitted. □

We now make some important definitions that we will use repeatedly in this text:

Definition 1.3.4 (Almost everywhere and support). A property $P(x)$ of a point $x \in \mathbf{R}^d$ is said to hold *(Lebesgue) almost everywhere* in $\mathbf{R}^d$, or *for (Lebesgue) almost every point* $x \in \mathbf{R}^d$, if the set of $x \in \mathbf{R}^d$ for which $P(x)$ fails has Lebesgue measure zero (i.e., P is true outside of a null set). We usually omit the prefix Lebesgue, and often abbreviate "almost everywhere" or "almost every"as a.e.

Two functions $f, g : \mathbf{R}^d \to Z$ into an arbitrary range Z are said to *agree almost everywhere* if one has $f(x) = g(x)$ for almost every $x \in \mathbf{R}^d$.

The *support* of a function $f : \mathbf{R}^d \to \mathbf{C}$ or $f : \mathbf{R}^d \to [0, +\infty]$ is defined to be the set $\{x \in \mathbf{R}^d : f(x) \neq 0\}$ where f is non-zero.

The following properties of the simple unsigned integral are easily obtained from the definitions:

Lemma 1.3.5 (Basic properties of the simple unsigned integral)**.** *Let $f, g : \mathbf{R}^d \to [0, +\infty]$ be simple unsigned functions.*

(i) *(Unsigned linearity) We have*

$$\begin{aligned} \mathrm{Simp} \int_{\mathbf{R}^d} f(x) + g(x) dx \\ = \mathrm{Simp} \int_{\mathbf{R}^d} f(x) dx + \mathrm{Simp} \int_{\mathbf{R}^d} g(x) dx \end{aligned}$$

and

$$\mathrm{Simp} \int_{\mathbf{R}^d} cf(x) dx = c \times \mathrm{Simp} \int_{\mathbf{R}^d} f(x) dx.$$

for all $c \in [0, +\infty]$.

(ii) *(Finiteness) We have* $\mathrm{Simp} \int_{\mathbf{R}^d} f(x) dx < \infty$ *if and only if f is finite almost everywhere, and its support has finite measure.*

(iii) *(Vanishing) We have* $\mathrm{Simp} \int_{\mathbf{R}^d} f(x) dx = 0$ *if and only if f is zero almost everywhere.*

(iv) *(Equivalence) If f and g agree almost everywhere, then we have* $\mathrm{Simp} \int_{\mathbf{R}^d} f(x) dx = \mathrm{Simp} \int_{\mathbf{R}^d} g(x) dx$.

(v) *(Monotonicity) If $f(x) \leq g(x)$ for almost every $x \in \mathbf{R}^d$, then* $\mathrm{Simp} \int_{\mathbf{R}^d} f(x) dx \leq \mathrm{Simp} \int_{\mathbf{R}^d} g(x) dx$.

(vi) *(Compatibility with Lebesgue measure) For any Lebesgue measurable E, one has* $\mathrm{Simp} \int_{\mathbf{R}^d} 1_E(x) dx = m(E)$.

Furthermore, the simple unsigned integral $f \mapsto \mathrm{Simp} \int_{\mathbf{R}^d} f(x) dx$ is the only map from the space $\mathrm{Simp}^+(\mathbf{R}^d)$ of unsigned simple functions to $[0, +\infty]$ that obeys all of the above properties.

Proof. Let $f = a_1 1_{E_1} + \cdots + a_k 1_{E_n}$ and $g = a'_1 1_{E'_1} + \cdots + a'_{n'} 1_{E'_{n'}}$. We use a Venn diagram argument. The $n + n'$ sets $E_1, \cdots, E_n, E'_1, \cdots, E'_{n'}$ partition $\mathbf{R}^d$ into $2^{n+n'}$ disjoint sets, each of which is an intersection of some of the

$E_1, \cdots, E_n, E'_1, \cdots, E'_{n'}$ and their complements. We throw away any sets that are empty, leaving us with a partition of $\mathbf{R}^d$ into m non-empty disjoint sets $A_1, \cdots, A_k$. As the $E_1, \cdots, E_n, E'_1, \cdots, E'_{n'}$ are Lebesgue measurable, the $A_1, \cdots, A_k$ are too. Then f and g can be rewritten as

$$f = c_1 1_{A_1} + \cdots + c_k 1_{A_k},$$

and

$$g = c'_1 1_{A_1} + \cdots + c'_k 1_{A_k}.$$

(i) Since we have

$$f + g = (c_1 + c'_1) 1_{A_1} + \cdots + (c_k + c'_k) 1_{A_k},$$

thus,

$$\begin{aligned}
\text{Simp} \int_{\mathbf{R}^d} f(x) + g(x) dx &= (c_1 + c'_1) m(A_1) + \cdots + (c_k + c'_k) m(A_k) \\
&= c_1 m(A_1) + \cdots + c_k m(A_k) \\
&\qquad + c'_1 m(A_1) + \cdots + c'_k m(A_k) \\
&= \text{Simp} \int_{\mathbf{R}^d} f(x) dx + \text{Simp} \int_{\mathbf{R}^d} g(x) dx.
\end{aligned}$$

For all $c \in [0, +\infty]$, we have

$$cf = (c \times c_1) 1_{A_1} + \cdots + (c \times c_k) 1_{A_k}.$$

Thus

$$\begin{aligned}
\text{Simp} \int_{\mathbf{R}^d} cf(x) dx &= (c \times c_1) m(A_1) + \cdots + (c \times c_k) m(A_k) \\
&= c(c_1 m(A_1) + \cdots + c_k m(A_k)) \\
&= c \times \text{Simp} \int_{\mathbf{R}^d} f(x) dx.
\end{aligned}$$

(ii) Suppose that $\text{Simp} \int_{\mathbf{R}^d} f(x) dx = c_1 m(E_1) + \cdots + c_k m(E_k) < \infty$. This means that $c_n m(E_n) < \infty$ for every $1 \leq n \leq k$. If there is some $c_i = +\infty$, then we must have $m(E_i) = 0$, otherwise we will contradict our assumption. Then we can see that $f < \infty$ except on null sets E_i. Thus f is finite almost everywhere.

Let $Supp(f)$ be the support of f. Then for every $x \in Supp(f)$ must lies in some E_n. If $m(E_n) = +\infty$, we must have $c_n = 0$ from assumption, a contradiction that $x \notin Supp(f)$. Thus for every $x \in Supp(f)$ and $x \in E_n$, we only have $m(E_n) < \infty$. From the definition of simple function, we have $Supp(f) = \bigcup_n E_n$, thus $m(Supp(f)) = m(\bigcup_n E_n) \leq \sum_n m(E_n) < \infty$.

Conversely, suppose that f is finite a.e. and $m(Supp(f)) < \infty$. Since we either have $c_i = 0$ or $0 < c_i$. Support of f has finite measure means that for all $c_n > 0$, we have $m(E_n) < \infty$; in particular, for $c_m = \infty$, f is finite a.e. requires that E_m be null set, so that $m(E_m) = 0$. And for $c_n = 0$, we have $c_n m(E_n) = 0$. Thus $\text{Simp} \int_{\mathbf{R}^d} = c_1 m(E_1) + \cdots + c_k m(E_k) < \infty$.

(iii) Suppose $\text{Simp} \int_{\mathbf{R}^d} f(x)dx = c_1 m(E_1) + \cdots + c_k m(E_k) = 0$. This means that we either have $c_n = 0$ or $m(E_n) = 0$ for every $1 \leq n \leq k$. Thus except on null sets, we have $f(x) = 0$ a.e. For the other hand, suppose that $f(x) = 0$ a.e. Then except on null sets, we have $c_n = 0$, and on null sets, we have $m(E_n) = 0$. Thus $\text{Simp} \int_{\mathbf{R}^d} f(x)dx = 0$.

(iv) Since f and g agree almost everywhere, we have $c_n \neq c'_n$ only on null sets. Thus $\text{Simp} \int_{\mathbf{R}^d} f(x)dx = \text{Simp} \int_{\mathbf{R}^d} g(x)dx$.

(v) Since $f \leq g$ for a.e. $x \in \mathbf{R}^d$. we have $c_n \leq c'_n$ for non null sets, thus $\text{Simp} \int_{\mathbf{R}^d} f(x)dx \leq \text{Simp} \int_{\mathbf{R}^d} g(x)dx$.

(vi) This is trivial.

Now we prove the uniqueness of the simple unsigned integral. Let $\mathcal{M}$ be the map from $\text{Simp}^+(\mathbf{R}^d)$ to $[0, +\infty]$ satisfying above properties. Let $\mathcal{M}'$ be a different map. Then for arbitrary simple function $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k} \in \text{Simp}^+(\mathbf{R}^d)$, by unsigned linearity and compatibility with Lebesgue measure, we have

$$
\begin{aligned}
\mathcal{M}(f) &= \mathcal{M}\Big(\sum_{n=1}^{k} c_n 1_{E_n}\Big) \\
&= \sum_{n=1}^{k} c_n \mathcal{M}(1_{E_n})
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{n=1}^{k} c_n m(E_n) \\
&= \sum_{n=1}^{k} c_n \mathcal{M}'(1_{E_n}) \\
&= \mathcal{M}'\Big(\sum_{n=1}^{k} c_n 1_{E_n}\Big) \\
&= \mathcal{M}'(f),
\end{aligned}
$$

a contradiction. □

We can now define an absolutely convergent counterpart to the simple unsigned integral. This integral will soon be superceded by the absolutely convergent Lebesgue integral, but we give it here as motivation for the more general notion of integration.

Definition 1.3.6 (Absolutely convergent simple integral)**.** A complex-valued simple function $f : \mathbf{R}^d \to \mathbf{C}$ is said to be *absolutely integrable* if

$$\mathrm{Simp}\int_{\mathbf{R}^d} |f(x)|dx < \infty.$$

If f is absolutely integrable, the integral $\mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx$ is defined for real signed f by the formula

$$\mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx := \mathrm{Simp}\int_{\mathbf{R}^d} f_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} f_-(x)dx$$

where $f_+(x) := \max(f(x), 0)$ and $f_-(x) := \max(-f(x), 0)$ (note that these are unsigned simple functions that are pointwise dominated by $|f|$ and thus have finite integral), and for complex-valued f by the formula[a]

$$\mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx := \mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Re}\, f(x)dx + i\,\mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Im}\, f(x)dx.$$

[a]Strictly speaking, this is an abuse of notation as we have now defined the simple integral $\mathrm{Simp}\int_{\mathbf{R}^d}$ three different times, for unsigned, real signed, and complex-valued simple functions, but one easily verifies that these three definitions agree with each other on their common domains of definition, so it is safe to use a single notation for all three.

The properties of the unsigned simple integral then can be used to deduce analogous properties for the complex-valued integral:

Proposition 1.3.7 (Basic properties of the complex-valued simple integral). *Let $f, g : \mathbf{R}^d \to \mathbf{C}$ be absolutely integrable simple functions.*

(i) *(*-linearity) We have*

$$\mathrm{Simp}\int_{\mathbf{R}^d} f(x) + g(x)dx = \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} g(x)dx$$

and

$$\mathrm{Simp}\int_{\mathbf{R}^d} cf(x)dx = c \times \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx$$

for all $c \in \mathbf{C}$. Also, we have

$$\mathrm{Simp}\int_{\mathbf{R}^d} \overline{f}(x)dx = \overline{\mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx}.$$

(ii) *(Equivalence) If f and g agree almost everywhere, then we have $\mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx = \mathrm{Simp}\int_{\mathbf{R}^d} g(x)dx$.*

(iii) *(Compatibility with Lebesgue measure) For any Lebesgue measurable E, one has $\mathrm{Simp}\int_{\mathbf{R}^d} 1_E(x)dx = m(E)$.*

Furthermore, the complex-valued simple integral $f \mapsto \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx$ is the only map from the space $\mathrm{Simp}^{abs}(\mathbf{R}^d)$ of absolutely integrable simple functions to $\mathbf{C}$ that obeys all of the above properties.

Proof. (i) For the additivity, we first suppose that f and g are real-valued functions. From Definition 1.3.6, we clearly have

$$(f+g)_+ - (f+g)_- = (f_+ - f_-) + (g_+ - g_-).$$

Here we don't prove this identity. Thus

$$(f+g)_+ + f_- + g_- = (f+g)_- + f_+ + g_+.$$

We can see that both sides of the equality above are sums of unsigned simple functions. Thus from Lemma 1.3.5, we have

$$\begin{aligned}&\mathrm{Simp}\int_{\mathbf{R}^d}(f+g)_+(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} f_-(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} g_-(x)dx\\ =\ &\mathrm{Simp}\int_{\mathbf{R}^d}(f+g)_-(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} f_+(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} g_+(x).\end{aligned}$$

Rearranging the equality above, we have

$$\begin{aligned}&\mathrm{Simp}\int_{\mathbf{R}^d}(f+g)_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d}(f+g)_-(x)dx\\ =\ &\mathrm{Simp}\int_{\mathbf{R}^d} f_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} f_-(x)dx\\ &+ \mathrm{Simp}\int_{\mathbf{R}^d} g_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} g_-(x)dx.\end{aligned}$$

From the absolutely integrability of f and g, and the inequalities $(f+g)_+ \le f_+ + g_+$, $(f+g)_- \le f_- + g_-$, the left side of above equality would not taking the form $\infty - \infty$. By Definition 1.3.6, we have

$$\mathrm{Simp}\int_{\mathbf{R}^d} f(x) + g(x)dx = \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} g(x)dx.$$

Above conclusion is hold for real and image parts of complex-valued f, g, i.e., we have

$$\mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Re}\, f(x) + g(x)dx = \mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Re}\, f(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Re}\, g(x)dx,$$

and

$$\mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Im}\, f(x) + g(x)dx = \mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Im}\, f(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Im}\, g(x)dx.$$

These give

$$\mathrm{Simp}\int_{\mathbf{R}^d} f(x)+g(x)dx = \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} g(x)dx$$

for complex-valued f, g, as desired.

For the homogeneity, we also first consider the real-valued f and $c \geq 0$. By Lemma 1.3.5, we have

$$\begin{aligned}\mathrm{Simp}\int_{\mathbf{R}^d} cf(x)dx &= \mathrm{Simp}\int_{\mathbf{R}^d} (cf)_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} (cf)_-(x)dx \\ &= \mathrm{Simp}\int_{\mathbf{R}^d} cf_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} cf_-(x)dx \\ &= c \times \Big(\mathrm{Simp}\int_{\mathbf{R}^d} cf_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} f_-(x)dx\Big) \\ &= c \times \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx.\end{aligned}$$

While if $c < 0$, then

$$\begin{aligned}\mathrm{Simp}\int_{\mathbf{R}^d} cf(x)dx &= \mathrm{Simp}\int_{\mathbf{R}^d} (cf)_+(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} (cf)_-(x)dx \\ &= \mathrm{Simp}\int_{\mathbf{R}^d} (-c)f_-(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} (-c)f_+(x)dx \\ &= -c \times \Big(\mathrm{Simp}\int_{\mathbf{R}^d} f_-(x)dx - \mathrm{Simp}\int_{\mathbf{R}^d} f_+(x)dx\Big) \\ &= c \times \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx.\end{aligned}$$

Thus we have

$$\mathrm{Simp}\int_{\mathbf{R}^d} cf(x)dx = c \times \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx$$

for real-valued function f. Since this is also hold for real and image parts of complex-valued function f, i.e., we have

$$\mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Re}\, cf(x)dx = c \times \mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Re}\, f(x)dx$$

and

$$\mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Im}\, cf(x)dx = c \times \mathrm{Simp}\int_{\mathbf{R}^d} \mathrm{Im}\, f(x)dx.$$

Thus we have

$$\operatorname{Simp}\int_{\mathbf{R}^d} cf(x)dx = c \times \operatorname{Simp}\int_{\mathbf{R}^d} f(x)dx$$

for complex-valued function f, as desired.

Use linearity, we have

$$\begin{aligned}
\operatorname{Simp}\int_{\mathbf{R}^d} \overline{f}(x)dx &= \operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Re}\overline{f}(x)dx + i\operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Im}\overline{f}(x)dx \\
&= \operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Re} f(x)dx + i\operatorname{Simp}\int_{\mathbf{R}^d} (-1)\operatorname{Im} f(x)dx \\
&= \operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Re} f(x)dx - i\operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Im} f(x)dx \\
&= \overline{\operatorname{Simp}\int_{\mathbf{R}^d} f(x)dx}.
\end{aligned}$$

(ii) Since f and g agree a.e., we have $\operatorname{Re} f(x) = \operatorname{Re} g(x)$ and $\operatorname{Im} f(x) = \operatorname{Im} g(x)$ for a.e. $x \in \mathbf{R}^d$. Then by Lemma 1.3.5, we have

$$\operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Re} f(x)dx = \operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Re} g(x)dx$$

and

$$\operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Im} f(x)dx = \operatorname{Simp}\int_{\mathbf{R}^d} \operatorname{Im} g(x)dx.$$

This implies

$$\operatorname{Simp}\int_{\mathbf{R}^d} f(x)dx = \operatorname{Simp}\int_{\mathbf{R}^d} g(x)dx.$$

(iii) This is same with Lemma 1.3.5(vi).

Now we prove the uniqueness of the complex-valued simple integral. Let $\mathcal{M}$ be the map from $\operatorname{Simp}^{abs}(\mathbf{R}^d)$ to $\mathbf{C}$ satisfying above properties. Let $\mathcal{M}'$ be the different map. Then for arbitrary complex-valued simple function $f = c_1 1_{E_1} + \cdots + c_k 1_{E_k} \in \operatorname{Simp}^{abs}(\mathbf{R}^d)$, by *-linearity and compatibility with Lebesgue measure, we have

$$\mathcal{M}(f) = \mathcal{M}\Big(\sum_{n=1}^{k} c_n 1_{E_n}\Big)$$

$$
\begin{aligned}
&= \sum_{n=1}^{k} c_n \mathcal{M}(1_{E_n}) \\
&= \sum_{n=1}^{k} c_n m(E_n) \\
&= \sum_{n=1}^{k} c_n \mathcal{M}'(1_{E_n}) \\
&= \mathcal{M}'\Big(\sum_{n=1}^{k} c_n 1_{E_n}\Big) \\
&= \mathcal{M}'(f),
\end{aligned}
$$

a contradiction. □

1.3.2 Measurable functions

Much as the piecewise constant integral can be completed to the Riemann integral, the unsigned simple integral can be completed to the unsigned Lebesgue integral, by extending the class of unsigned simple functions to the larger class of unsigned Lebesgue measurable functions. One of the shortest ways to define this class is as follows:

Definition 1.3.8 (Unsigned measurable function)**.** An unsigned function $f : \mathbf{R}^d \to [0, +\infty]$ is *unsigned Lebesgue measurable*, or *measurable* for short, if it is the pointwise limit of unsigned simple functions, i.e., if there exists a sequence $f_1, f_2, f_3, \cdots : \mathbf{R}^d \to [0, +\infty]$ of unsigned simple functions such that $f_n(x) \to f(x)$ for every $x \in \mathbf{R}^d$.

This particular definition is not always the most tractable. Fortunately, it has many equivalent forms:

Lemma 1.3.9 (Equivalent notions of measurability)**.** *Let* $f : \mathbf{R}^d \to [0, +\infty]$ *be an unsigned function. Then the following are equivalent:*

(i) *f is unsigned Lebesgue measurable.*

(ii) *f is the pointwise limit of unsigned simple functions f_n (thus the limit $\lim_{n\to\infty} f_n(x)$ exists and is equal to $f(x)$ for all $x \in \mathbf{R}^d$).*

(iii) *f is the pointwise almost everywhere limit of unsigned simple functions f_n (thus the limit $\lim_{n\to\infty} f_n(x)$ exists and is equal to $f(x)$ for almost every $x \in \mathbf{R}^d$).*

(iv) *f is the supremum $f(x) = \sup_n f_n(x)$ of an increasing sequence $0 \le f_1 \le f_2 \le \cdots$ of unsigned simple functions f_n, each of which are bounded with finite measure support.*

(v) *For every $\lambda \in [0, +\infty]$, the set $\{x \in \mathbf{R}^d : f(x) > \lambda\}$ is Lebesgue measurable.*

(vi) *For every $\lambda \in [0, +\infty]$, the set $\{x \in \mathbf{R}^d : f(x) \ge \lambda\}$ is Lebesgue measurable.*

(vii) *For every $\lambda \in [0, +\infty]$, the set $\{x \in \mathbf{R}^d : f(x) < \lambda\}$ is Lebesgue measurable.*

(viii) *For every $\lambda \in [0, +\infty]$, the set $\{x \in \mathbf{R}^d : f(x) \le \lambda\}$ is Lebesgue measurable.*

(ix) *For every interval $I \subset [0, +\infty)$, the set $f^{-1}(I) := \{x \in \mathbf{R}^d : f(x) \in I\}$ is Lebesgue measurable.*

(x) *For every (relatively) open set $U \subset [0, +\infty)$, the set $f^{-1}(U) := \{x \in \mathbf{R}^d : f(x) \in U\}$ is Lebesgue measurable.*

(xi) *For every (relatively) closed set $K \subset [0, +\infty)$, the set $f^{-1}(K) := \{x \in \mathbf{R}^d : f(x) \in K\}$ is Lebesgue measurable.*

Proof. Proof omitted. □

Remark. As we use simple function and Lebesgue measurable sets to define the unsigned simple integral, we require some "integrable function", i.e., unsigned measurable function, to extend the unsigned simple integral. Thus we want to define a function f on a Lebesgue measurable subset E of $\mathbf{R}^d$, if the inverse image of the subset of $\mathbf{R}^d$ is measurable, we say that f is measurable.

With these equivalent formulations, we can now generate plenty of measurable functions:

Proposition 1.3.10 (Unsigned measurable functions)**.**

(i) *Every continuous function* $f : \mathbf{R}^d \to [0, +\infty]$ *is measurable.*

(ii) *Every unsigned simple function is measurable.*

(iii) *The supremum, infimum, limit superior, or limit inferior of at most countable (either countable or finite) family of unsigned measurable functions is unsigned measurable.*

(iv) *An unsigned function that is equal almost everywhere to an unsigned measurable function, is itself measurable.*

(v) *If a sequence* f_n *of unsigned measurable functions converges pointwise almost everywhere to an unsigned limit* f*, then* f *is also measurable.*

(vi) *If* $f : \mathbf{R}^d \to [0, +\infty]$ *is measurable and* $\phi : [0, +\infty] \to [0, +\infty]$ *is continuous, then* $\phi \circ f : \mathbf{R}^d \to [0, +\infty]$ *is measurable.*

(vii) *If* f, g *are unsigned measurable functions, then* $f + g$ *and* fg *are measurable.*

Proof. (i) Since f is continuous, for every open subset $U \subset [0, +\infty)$, $f^{-1}(U)$ is an open set in $\mathbf{R}^d$, so is Lebesgue measurable. Thus by Lemma 1.3.9(x), f is measurable.

(ii) This immediately comes from the definition of unsigned simple function and Lemma 1.3.9.

(iii) For unsigned measurable functions $(f_n)_{n=1}^{\infty}$, we have $\{x \in \mathbf{R}^d : f_n(x) \geq \lambda\}$ are Lebesgue measurable for every $\lambda \in [0, +\infty]$ and $n \geq 1$. Then from

$$\{x \in \mathbf{R}^d : \sup_{n\geq 1} f_n(x) \geq \lambda\} = \bigcup_{n=1}^{\infty} \{x \in \mathbf{R}^d : f_n(x) \geq \lambda\}$$

for every $\lambda \in [0, +\infty]$, which is countable union of Lebesgue measurable sets, hence is also Lebesgue measurable. By Lemma 1.3.9(vi), $\sup_{n\geq 1} f_n$ is unsigned measurable. A similar argument shows that $\inf_n f_n$ is unsigned measurable.

For the limit superior $\limsup_{n\to\infty} f_n$, for every λ we have

$$\begin{aligned}\{x \in \mathbf{R}^d : \limsup_{n\to\infty} f_n(x) \geq \lambda\} &= \{x \in \mathbf{R}^d : \inf_{N>0} \sup_{n\geq N} f_n(x) \geq \lambda\} \\ &= \bigcap_{N>0} \bigcup_{n\geq N} \{x \in \mathbf{R}^d : f_n(x) \geq \lambda\},\end{aligned}$$

such a set is countable union and intersection of Lebesgue measurable sets, so that is also Lebesgue measurable. By Lemma 1.3.9(vi), $\limsup_{n\to\infty} f_n$ is unsigned measurable. A similar argument shows that $\liminf_{n\to\infty} f_n$ is unsigned measurable.

Above conclusions are also hold for finitely many f_n, say $1 \leq n \leq k$. We just leave $\{x \in \mathbf{R}^d : f_m(x) \geq \lambda\}$ be empty for every $m > k$.

(iv) Let N be null set. Since unsigned function f is agree a.e. with unsigned measurable function g, we have $f(x) = g(x)$ for all $x \in \mathbf{R}^d \setminus N$. By Lemma 1.3.9(iii), there exists the unsigned simple functions f_n such that $\lim_{n\to\infty} f_n(x) = g(x)$ for all $x \in \mathbf{R}^d \setminus N$. This implies that $\lim_{n\to\infty} f_n(x) = f(x)$ and f is Lebesgue measurable.

(v) Since f is the pointwise a.e. limit of f_n, then for a.e. $x \in \mathbf{R}^d$, we have

$$f(x) = \lim_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x) = \inf_{N>0} \sup_{n\geq N} f_n(x).$$

This implies that for every λ, we have

$$\{x \in \mathbf{R}^d : f(x) \geq \lambda\} = \bigcap_{N>0} \bigcup_{n\geq N} \{x \in \mathbf{R}^d : f_n(x) \geq \lambda\},$$

which is the union and intersection of Lebesgue measurable sets from the measurability of f_n. Thus f is measurable.

(vi) Let U be an open subset of $[0, +\infty)$, then $\phi^{-1}(U)$ is open in $[0, +\infty)$. Since f is measurable, for every U we have $(\phi \circ f)^{-1}(U) = f^{-1}(\phi^{-1}(U))$ is open, so that is Lebesgue measurable. Thus by Lemma 1.3.9(x), $\phi \circ f$ is measurable.

(vii) Consider $f+g$. We can see that $f+g$ can be written as the composition of direct sum and continuous function $(f \oplus g) \circ h$, where $f \oplus g : \mathbf{R}^d \to [0,+\infty] \times [0,+\infty]$ is the function $f \oplus g(x) = (f(x), g(x))$, and $h : [0,+\infty] \times [0,+\infty] \to [0,+\infty]$ is the function $h(a,b) := a+b$. Since f, g are measurable, for every open subset $U, V \subset \mathbf{R}^d$, we have $f^{-1}(U)$ and $g^{-1}(V)$ are open in $[0,+\infty]$. Then $(f \oplus g)^{-1}(U \times V) = f^{-1}(U) \cap g^{-1}(V)$ is open in $\mathbf{R}^d$. This shows that $f \oplus g$ is measurable. Since h is continuous. By (vi), we have $f+g = (f \oplus g) \circ h$ is measurable.

If we let h be the function $h(a,b) := ab$, which is also continuous, then $fg = (f \oplus g) \circ h$ is measurable. □

Lemma 1.3.11. *Let $f : \mathbf{R}^2 \to [0,+\infty]$. f is a bounded unsigned measurable function if and only if f is the uniform limit of bounded simple functions.*

Proof. We first suppose that f is the uniform limit of bounded simple functions f_n. This implies that f is the pointwise limit of simple functions f_n. By Lemma 1.3.9(ii), f is measurable. Since uniform convergence preserve boundedness, f is bounded.

Conversely, suppose that f is a bounded unsigned measurable function. By Lemma 1.3.9(iv), there is increasing sequence of bounded functions such that $\lim_{n\to\infty} f_n(x) = f(x)$. Then f_n converges uniformly to f. □

Lemma 1.3.12. *An unsigned function $f : \mathbf{R}^d \to [0,+\infty]$ is a simple function if and only if it is measurable and takes on at most finitely many values.*

Proof. Suppose that f is an unsigned simple function. From Theorem 1.3.10(ii), f is measurable. We need to show that f takes on at most finitely many values. For every Lebesgue measurable set E_n, we have $f(E_n) = \{c_n\}$ from definition. Then we have $f(\mathbf{R}^d) = f(\bigcup_{n=1}^k E_n) = \bigcup_{n=1}^k \{c_n\}$, which is countably finite.

Conversely, suppose that f is measurable with countably finite many values. This means that $f(\mathbf{R}^d) = \bigcup_{n=1}^k \{c_n\}$. Since f is measurable, by Theorem 1.3.9(xi), $E_n := f^{-1}(\{c_n\})$ is measurable for all $1 \leq n \leq k$. Then we can define f as $f := c_1 1_{E_1} + \cdots + c_k 1_{E_k}$, which is unsigned simple function. □

Proposition 1.3.13. *Let $f : \mathbf{R}^d \to [0, +\infty]$ be an unsigned measurable function. Then the region $\{(x,t) \in \mathbf{R}^d \times \mathbf{R} : 0 \leq t \leq f(x)\}$ is a measurable subset of $\mathbf{R}^{d+1}$.*

Remark. There is a converse to this statement, but we will wait until Proposition 1.3.24 to prove it, once we have the Fubini-Tonelli theorem (Corollary **??**) available to us.

Proof. By Lemma 1.3.9(iv), there are bounded unsigned simple functions f_n such that $f(x) = \sup_n f_n(x)$. Then

$$\begin{aligned}\{(x,t) : 0 \leq t \leq f(x)\} &= \{(x,t) : 0 \leq t \leq \sup_n f_n(x)\} \\ &= \bigcup_{n=1}^{\infty} \{(x,t) : 0 \leq t \leq f_n(x)\}.\end{aligned}$$

Since every f_n is simple function, by Lemma 1.3.12, for every $n \geq 1$ we have

$$\{(x,t) : 0 \leq t \leq f_n(x)\} = \bigcup_{k=1}^{\infty} E_k \times [0, c_k].$$

By Lemma 1.2.29, $\{(x,t) : 0 \leq t \leq f_n(x)\}$ is Lebesgue measurable for every $n \geq 1$ for that E_m and $[0, c_m]$. Thus $\{(x,t) : 0 \leq t \leq f(x)\}$ is also Lebesgue measurable. □

Now we can define the concept of a complex-valued measurable function. As discussed earlier, it will be convenient to allow for such functions to only be defined *almost everywhere*, rather than *everywhere*, to allow for the possibility that the function becomes singular or otherwise undefined on a null set.

Definition 1.3.14 (Complex measurability)**.** An almost everywhere defined complex-valued function $f : \mathbf{R}^d \to \mathbf{C}$ is *Lebesgue measurable*, or *measurable* for short, if it is the pointwise almost everywhere limit of complex-valued simple functions.

As before, there are several equivalent definitions:

Lemma 1.3.15. *Let $f : \mathbf{R}^d \to \mathbf{C}$ be an almost everywhere defined complex-valued function. Then the following are equivalent:*

(i) *f is measurable.*

(ii) *f is the pointwise almost everywhere limit of complex-valued simple functions.*

(iii) *The (magnitudes of the) positive and negative parts of* $\mathrm{Re}(f)$ *and* $\mathrm{Im}(f)$ *are unsigned measurable functions.*

(iv) *$f^{-1}(U)$ is Lebesgue measurable for every open set $U \subset \mathbf{C}$.*

(v) *$f^{-1}(K)$ is Lebesgue measurable for every closed set $K \subset \mathbf{C}$.*

Proof. (i) and (ii) are equivalent by definition.

(ii) $\Rightarrow$ (iii). Let $(f_n)_{n=1}^{\infty}$ be complex-valued simple functions which converge pointwise for a.e. to f. We show that the positive parts of $\mathrm{Re}(f)$ is unsigned measurable function. Let $\mathrm{Re}\, f_+(x) := \max(\mathrm{Re}\, f(x), 0)$, which is an unsigned function. For every f_n, we have

$$\begin{aligned}
f_n &= c_1 1_{E_1} + \cdots + c_k 1_{E_k} \\
&= (\mathrm{Re}\, c_1 + i \,\mathrm{Im}\, c_1) 1_{E_1} + \cdots + (\mathrm{Re}\, c_k + i \,\mathrm{Im}\, c_k) 1_{E_k} \\
&= \mathrm{Re}\, c_1 1_{E_1} + \cdots + \mathrm{Re}\, c_k 1_{E_1} + i(\mathrm{Im}\, c_1 1_{E_1} + \cdots + \mathrm{Im}\, c_k 1_{E_k}) \\
&= \mathrm{Re}\, f_n + i \,\mathrm{Im}\, f_n
\end{aligned}$$

for a.e. $x \in \mathbf{R}^d$. Then from

$$\begin{aligned}
\lim_{n\to\infty} f_n(x) &= \lim_{n\to\infty} \mathrm{Re}\, f_n(x) + i \lim_{n\to\infty} \mathrm{Im}\, f_n(x) \\
&= \mathrm{Re}\, f(x) + i \,\mathrm{Im}\, f(x),
\end{aligned}$$

we have $\lim_{n\to\infty} \operatorname{Re} f_n(x) = \operatorname{Re} f(x)$ for a.e. x. This means that $\operatorname{Re} f$ is the pointwise a.e. limit of unsigned simple functions $\operatorname{Re} f_n$. From the limit laws for functions, we have

$$\lim_{n\to\infty} \max(\operatorname{Re} f_n(x), 0) = \max(\operatorname{Re} f(x), 0) = \operatorname{Re} f_+(x)$$

for a.e. $x \in \mathbf{R}^d$. Since $(\max(\operatorname{Re} f_n(x), 0))_{n=1}^\infty$ are all unsigned simple functions, by Lemma 1.3.9(iii), $\operatorname{Re} f_+$ is unsigned Lebesgue measurable. This is similar to define $\operatorname{Re} f_-$, $\operatorname{Im} f_+$, $\operatorname{Im} f_-$, and show that these are unsigned Lebesgue measurable.

(iii) $\Rightarrow$ (iv). Suppose that $\operatorname{Re} f_+$, $\operatorname{Re} f_-$, $\operatorname{Im} f_+$, $\operatorname{Im} f_-$ are unsigned Lebesgue measurable. Then $\operatorname{Re} f = \operatorname{Re} f_+ - \operatorname{Im} f_-$ and $\operatorname{Im} f = \operatorname{Im} f_+ - \operatorname{Im} f_-$ are also Lebesgue measurable from Theorem 1.3.10(vi).

Every open subset $U \subset \mathbf{C}$ can be written as $U = U_1 \times U_2$, where U_1, U_2 are open in $\mathbf{R}$. Then for every open set U we have $\operatorname{Re} f^{-1}(U_1)$ and $\operatorname{Im} f^{-1}(U_2)$ are also open. Thus $f^{-1}(U) = \operatorname{Re} f^{-1}(U_1) \cap \operatorname{Im} f^{-1}(U_2)$ is open, so that Lebesgue measurable, for every $U \subset \mathbf{C}$.

(iv) $\Rightarrow$ (v). For every closed set $K \subset \mathbf{C}$, we have open set $\mathbf{C} \setminus K \subset \mathbf{C}$. Then

$$f^{-1}(\mathbf{C} \setminus K) = \mathbf{R}^d \setminus f^{-1}(K)$$

is open. This implies $f^{-1}(K)$ is closed, so that Lebesgue measurable.

(v) $\Rightarrow$ (i). Since (v) implies that f is continuous, then we have f is the the pointwise a.e. limit of complex-valued simple functions. Thus f is measurable. $\square$

Proposition 1.3.16 (Complex-valued measurable functions)**.**

(i) *Every continuous function $f : \mathbf{R}^d \to \mathbf{C}$ is measurable.*

(ii) *A function $f : \mathbf{R}^d \to \mathbf{C}$ is simple if and only if it is measurable and takes on at most finitely many values.*

(iii) *A complex-valued function that is equal almost everywhere to a measurable function, is itself measurable.*

(iv) *If a sequence f_n of complex-valued measurable functions con-*

verges pointwise almost everywhere to a complex-valued limit f, then f is also measurable.

(v) *If $f : \mathbf{R}^d \to \mathbf{C}$ is measurable and $\phi : \mathbf{C} \to \mathbf{C}$ is continuous, then $\phi \circ f : \mathbf{R}^d \to \mathbf{C}$ is measurable.*

(vi) *If f, g are measurable functions, then $f + g$ and fg are measurable.*

Proof. This is similar to prove as Proposition 1.3.10, we just simply replacing "unsigned" with "complex-valued" and "real" with "complex" in the proofs but otherwise leaving all the other details of the proof unchanged. $\square$

Proposition 1.3.17. *Let $f : [a, b] \to \mathbf{R}$ be a Riemann integrable function. If one extends f to all of $\mathbf{R}$ by defining $f(x) = 0$ for $x \notin [a, b]$, then f is measurable.*

Proof. Let $I_1, \cdots, I_n$ be a partition of $[a, b]$. Since f is Riemann integrable, for every $k > 0$, there is a $\delta > 0$ such that

$$\begin{aligned}\int_a^b f(x)dx - \frac{1}{k} &\leq \sum_{i=1}^{n} \inf_{x \in I_i} f(x)|I_i| \\ &\leq \int_a^b f(x)dx \\ &\leq \sum_{i=1}^{n} \sup_{x \in I_i} f(x)|I_i| \\ &\leq \int_a^b f(x)dx + \frac{1}{k}\end{aligned}$$

for $\sup_{1 \leq i \leq n} |I_i| \leq \delta$. Notice that we have $|I_i| = m(I_i)$. Let $\overline{f}_k(x) : [a, b] \to \mathbf{R}$ and $\underline{f}_k(x) : [a, b] \to \mathbf{R}$ be the function

$$\overline{f}_k(x) = \sum_{i=1}^{n} \sup_{x \in I_i} f(x) 1_{I_i}(x)$$

and

$$\underline{f}_k(x) = \sum_{i=1}^{n} \inf_{x \in I_i} f(x) 1_{I_i}(x)$$

satisfying above inequality. We can see that $\overline{f}_k$ is a decreasing sequence and $\underline{f}_k$ is a increasing sequence such that $\underline{f}_k \le f \le \overline{f}_k$ for every $k > 0$. Since every $\underline{f}_k$ and $\overline{f}_k$ are bounded simple function, $\overline{f}(x) = \lim_{k\to\infty} \overline{f}_k(x)$ and $\underline{f}(x) = \lim_{k\to\infty} \underline{f}_k(x)$ are measurable, and we have $\underline{f} \le f \le \overline{f}$.

Because f is Riemann integrable, when $\sup_{1\le i\le n} |I_i| \to 0$, $\overline{f} = \underline{f}$ for a.e. This implies f is equal a.e. to measurable functions $\overline{f}$ and $\underline{f}$. Thus f is also measurable. □

1.3.3 Unsigned Lebesgue integrals

We are now ready to integrate unsigned measurable functions. We begin with the notion of the lower unsigned Lebesgue integral, which can be defined for arbitrary unsigned functions (not necessarily measurable):

Definition 1.3.18 (Lower unsigned Lebesgue integral)**.** Let $\mathbf{R}^d \to [0, +\infty]$ be an unsigned function (not necessarily measurable). We define the *lower unsigned Lebesgue integral* $\underline{\int_{\mathbf{R}^d}} f(x)dx$ to be the quantity

$$\underline{\int_{\mathbf{R}^d}} f(x)dx := \sup_{0\le g\le f; g \text{ simple}} \operatorname{Simp} \int_{\mathbf{R}^d} g(x)dx$$

where g ranges over all unsigned simple functions $g : \mathbf{R}^d \to [0, +\infty]$ that are pointwise bounded by f.

One can also define the *upper unsigned Lebesgue integral*

$$\overline{\int_{\mathbf{R}^d}} f(x)dx := \inf_{h\ge f; h \text{ simple}} \operatorname{Simp} \int_{\mathbf{R}^d} h(x)dx,$$

but we will use this integral much more rarely.

The following properties of the lower Lebesgue integral are easy to establish:

Proposition 1.3.19 (Basic properties of the lower Lebesgue integral)**.** *Let $f, g : \mathbf{R}^d \to [0, +\infty]$ be unsigned functions (not necessarily measur-*

able).

(i) *(Compatibility with the simple integral) If f is simple, then we have*

$$\underline{\int_{\mathbf{R}^d}} f(x)dx = \overline{\int_{\mathbf{R}^d}} f(x)dx = \mathrm{Simp}\int_{\mathbf{R}^d} f(x)dx$$

(ii) *(Monotonicity) If $f \geq g$ pointwise almost everywhere, then we have*

$$\underline{\int_{\mathbf{R}^d}} f(x)dx \leq \underline{\int_{\mathbf{R}^d}} g(x)dx \quad \text{and} \quad \overline{\int_{\mathbf{R}^d}} f(x) \leq \overline{\int_{\mathbf{R}^d}} g(x)dx$$

(iii) *(Homogeneity) If $c \in [0, +\infty)$, then*

$$\underline{\int_{\mathbf{R}^d}} cf(x)dx = c\underline{\int_{\mathbf{R}^d}} f(x)dx.$$

(the claim unfortunately fails for $c = +\infty$, but this is somewhat tricky to show.)

(iv) *(Equivalence) If f, g agree almost everywhere, then*

$$\underline{\int_{\mathbf{R}^d}} f(x)dx = \underline{\int_{\mathbf{R}^d}} g(x)dx \quad \text{and} \quad \overline{\int_{\mathbf{R}^d}} f(x)dx = \overline{\int_{\mathbf{R}^d}} g(x)dx$$

(v) *(Superadditivity)*

$$\underline{\int_{\mathbf{R}^d}} f(x) + g(x)dx \geq \underline{\int_{\mathbf{R}^d}} f(x)dx + \underline{\int_{\mathbf{R}^d}} g(x)dx.$$

(vi) *(Subadditivity of upper integral)*

$$\overline{\int_{\mathbf{R}^d}} f(x) + g(x)dx \leq \overline{\int_{\mathbf{R}^d}} f(x)dx + \overline{\int_{\mathbf{R}^d}} g(x)dx.$$

(vii) *(Divisibility) For any measurable set E, one has*

$$\underline{\int_{\mathbf{R}^d}} f(x)dx = \underline{\int_{\mathbf{R}^d}} f(x)1_E(x)dx + \underline{\int_{\mathbf{R}^d}} f(x)1_{\mathbf{R}^d\setminus E}(x)dx.$$

(viii) *(Horizontal truncation) As $n \to \infty$, $\underline{\int_{\mathbf{R}^d}} \min(f(x), n)dx$ converges to $\underline{\int_{\mathbf{R}^d}} f(x)dx$.*

(ix) *(Vertical truncation) As $n \to \infty$, $\underline{\int_{\mathbf{R}^d}} f(x)1_{|x|\le n}dx$ converges to $\underline{\int_{\mathbf{R}^d}} f(x)dx$.*

(x) *(Reflection) If $f+g$ is a simple function that is bounded with finite measure support (i.e. it is absolutely integrable), then we have*

$$\mathrm{Simp}\int_{\mathbf{R}^d} f(x)+g(x)dx = \underline{\int_{\mathbf{R}^d}} f(x)dx + \overline{\int_{\mathbf{R}^d}} g(x)dx.$$

Proof. (i) When f is simple, the claim is trivial from Definition 1.3.18.

(ii) If $f \ge g$ pointwise almost everywhere, then for every simple function minorizes and majorizes f and g must hold the order for a.e. From Theorem 1.1.24(v), their simple integral also holds the order. Taking the supremum and infimum, we have $\underline{\int_{\mathbf{R}^d}} f(x)dx \le \underline{\int_{\mathbf{R}^d}} g(x)dx$ and $\overline{\int_{\mathbf{R}^d}} f(x) \le \overline{\int_{\mathbf{R}^d}} g(x)dx$.

(iii) This immediately comes from Lemma 1.3.5(i).

(iv) By Lemma 1.3.5(iv), this is trivial.

(v) Let h, h' be two non-negative minorizations of functions f and g, respectively. By Definition 1.3.18, for $\varepsilon > 0$, we have

$$\underline{\int_{\mathbf{R}^d}} f(x)dx \le \mathrm{Simp}\int_{\mathbf{R}^d} h(x)dx + \frac{\varepsilon}{2}$$

and

$$\underline{\int_{\mathbf{R}^d}} g(x)dx \le \mathrm{Simp}\int_{\mathbf{R}^d} h'(x)dx + \frac{\varepsilon}{2}.$$

Then from Lemma 1.3.5(i)

$$\begin{aligned}\underline{\int_{\mathbf{R}^d}} f(x)dx + \underline{\int_{\mathbf{R}^d}} g(x)dx &\le \mathrm{Simp}\int_{\mathbf{R}^d} h(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} h'(x)dx + \varepsilon \\ &= \mathrm{Simp}\int_{\mathbf{R}^d} h(x)+h'(x)dx + \varepsilon\end{aligned}$$

Since ε is arbitrary, we have

$$\underline{\int_{\mathbf{R}^d}} f(x)dx + \underline{\int_{\mathbf{R}^d}} g(x)dx \leq \mathrm{Simp}\int_{\mathbf{R}^d} h(x) + h'(x)dx \leq \underline{\int_{\mathbf{R}^d}} f(x) + g(x)dx.$$

(vi) Use the inequalities $\overline{\int_{\mathbf{R}^d}} f(x)dx \geq \mathrm{Simp}\int_{\mathbf{R}^d} h(x)dx + \varepsilon/2$ and $\overline{\int_{\mathbf{R}^d}} g(x)dx \geq \mathrm{Simp}\int_{\mathbf{R}^d} h'(x)dx + \varepsilon/2$, where h, h' are minorizations of f, g, respectively. Then the proof is similar to (v).

(vii) Since $f(x) = f(x)1_E(x) + f(x)1_{\mathbf{R}^d\setminus E}(x)$ for every measurable set E. From (iv), we have $\underline{\int_{\mathbf{R}^d}} f(x)dx = \underline{\int_{\mathbf{R}^d}} f(x)1_E(x) + f(x)1_{\mathbf{R}^d\setminus E}(x)dx$. From superadditivity, we have

$$\underline{\int_{\mathbf{R}^d}} f(x)dx \geq \underline{\int_{\mathbf{R}^d}} f(x)1_E(x)dx + \underline{\int_{\mathbf{R}^d}} f(x)1_{\mathbf{R}^d\setminus E}(x)dx.$$

Now we prove the remaining half parts. Let h be the simple unsigned function that minorizes f. Then $h\cdot 1_E$ and $h\cdot 1_{\mathbf{R}^d\setminus E}$ are also simple unsigned functions and minorize $f\cdot 1_E$ and $f\cdot 1_{\mathbf{R}^d\setminus E}$. Since we have

$$\begin{aligned}
&\underline{\int_{\mathbf{R}^d}} f(x)1_E(x) + f(x)1_{\mathbf{R}^d\setminus E}(x)dx \\
&\qquad\qquad \leq \mathrm{Simp}\int_{\mathbf{R}^d} h(x)1_E(x) + h(x)1_{\mathbf{R}^d\setminus E}(x)dx + \varepsilon \\
&\qquad\qquad = \mathrm{Simp}\int_{\mathbf{R}^d} h(x)1_E(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} h(x)1_E(x)dx + \varepsilon.
\end{aligned}$$

Because ε is arbitrary, we have

$$\begin{aligned}
&\underline{\int_{\mathbf{R}^d}} f(x)1_E(x) + f(x)1_{\mathbf{R}^d\setminus E}(x)dx \\
&\qquad\qquad \leq \mathrm{Simp}\int_{\mathbf{R}^d} h(x)1_E(x)dx + \mathrm{Simp}\int_{\mathbf{R}^d} h(x)1_E(x)dx \\
&\qquad\qquad \leq \underline{\int_{\mathbf{R}^d}} f(x)1_E(x)dx + \underline{\int_{\mathbf{R}^d}} f(x)1_{\mathbf{R}^d\setminus E}(x)dx,
\end{aligned}$$

as desired.

(viii) Let h be minorization of f for a.e. From (ii), $h(x) \leq \min(f(x), n)$ implies

$$\underline{\int_{\mathbf{R}^d}} h(x)dx \leq \underline{\int_{\mathbf{R}^d}} \min(f(x), n)dx$$

for every h and $n \geq 0$.

Now we want to show that $\underline{\int_{\mathbf{R}^d}} \min(f(x), n)dx \leq \underline{\int_{\mathbf{R}^d}} f(x)dx$ is hold for enough large n. If $f(x) = \infty$ on some Lebesgue measurable set, by Definition 1.3.18, we have $\underline{\int_{\mathbf{R}^d}} f(x)dx = \infty$ and the inequality is hold for every $n \geq 0$. Otherwise, if f is finite for a.e., then there is a natural number N such that $f(x) \leq N = \min(f(x), N)$ for a.e. This implies that for enough large n, $\underline{\int_{\mathbf{R}^d}} \min(f(x), n)dx \leq \underline{\int_{\mathbf{R}^d}} f(x)dx$ is hold. Thus we have

$$\underline{\int_{\mathbf{R}^d}} h(x)dx \leq \underline{\int_{\mathbf{R}^d}} \min(f(x), n)dx \leq \underline{\int_{\mathbf{R}^d}} f(x)dx$$

for enough large n. Taking the supremum and using the squeeze test, we have

$$\lim_{n\to\infty} \underline{\int_{\mathbf{R}^d}} \min(f(x), n)dx = \underline{\int_{\mathbf{R}^d}} f(x)dx.$$

(ix) From Theorem 1.2.18(i), we have

$$\lim_{n\to\infty} m(E \cap \{x \in \mathbf{R}^d : |x| \leq n\}) = m(E)$$

for any measurable set E. From Theorem 1.2.20(i), we have

$$\lim_{n\to\infty} f(x)1_{E_n}(x) = f(x)1_E(x)$$

for every measurable set E. Let h be minorization of f with respect to $E_1, \cdots, E_k$ and $c_1, \cdots, c_k$, then

$$\begin{aligned}
&\lim_{n\to\infty} \mathrm{Simp} \int_{\mathbf{R}^d} h(x)1_{|x|\leq n}(x)dx \\
&= \lim_{n\to\infty} \mathrm{Simp} \int_{\mathbf{R}^d} (c_1 1_{E_1}(x)1_{|x|\leq n}(x) + \cdots + c_k 1_{E_k}(x)1_{|x|\leq n}(x))dx \\
&= \lim_{n\to\infty} \mathrm{Simp} \int_{\mathbf{R}^d} (c_1 1_{E_1\cap\{|x|\leq n\}}(x) + \cdots + c_k 1_{E_k\cap\{|x|\leq n\}}(x))dx \\
&= \lim_{n\to\infty} (c_1 m(E_1 \cap \{|x| \leq n\}) + \cdots + c_k m(E_k \cap \{|x| \leq n\})) \\
&= c_1 m(E_1) + \cdots + c_k m(E_k) \\
&= \mathrm{Simp} \int_{\mathbf{R}^d} h(x).
\end{aligned}$$

Since $h \cdot 1_{|x|\leq n}$ is minorization of $f \cdot 1_{|x|\leq n}$ for every h minorizes f and $f(x)1_{|x|\leq n} \leq f(x)$ for every $n \geq 0$, we have

$$\mathrm{Simp}\int_{\mathbf{R}^d} h(x)1_{|x|\leq n}(x)dx \leq \underline{\int_{\mathbf{R}^d}} f(x)1_{|x|\leq n}(x)dx \leq \underline{\int_{\mathbf{R}^d}} f(x)dx.$$

Use squeeze test and taking the supremum, we have

$$\lim_{n\to\infty} \underline{\int_{\mathbf{R}^d}} f(x)1_{|x|\leq n}(x)dx = \underline{\int_{\mathbf{R}^d}} f(x)dx.$$

(x) This immediately following Lemma 1.3.5(i) and Proposition 1.3.19(i). □

Now we restrict attention to measurable functions.

Definition 1.3.20 (Unsigned Lebesgue integral)**.** If $f : \mathbf{R}^d \to [0, +\infty]$ is measurable, we define the unsigned Lebesgue integral $\int_{\mathbf{R}^d} f(x)dx$ of f to equal the lower unsigned Lebesgue integral $\underline{\int_{\mathbf{R}^d}} f(x)dx$. (For non-measurable functions, we leave the unsigned Lebesgue integral undefined.)

Once nice feature of measurable functions is that the lower and upper Lebesgue integrals can match, if one also assumes some boundedness:

Proposition 1.3.21. *Let $f : \mathbf{R}^d \to [0, +\infty]$ be measurable, bounded, and vanishing outside of a set of finite measure. Then the lower and upper Lebesgue integrals of f agree.*

Proof. Clearly, we have $\underline{\int_{\mathbf{R}^d}} f(x)dx \leq \overline{\int_{\mathbf{R}^d}} f(x)dx$, it suffices to show that $\overline{\int_{\mathbf{R}^d}} f(x)dx \leq \underline{\int_{\mathbf{R}^d}} f(x)dx$. Since f is bounded unsigned measurable function, by Lemma 1.3.11, f is the uniform limit of bounded simple functions f_n. This means that for every $\varepsilon > 0$, there exists an $N > 0$ such that $f_n(x) - \varepsilon \leq f(x) \leq f_n(x) + \varepsilon$ for every $n \geq N$ and $x \in \mathbf{R}^d$. Then

$$\overline{\int_{\mathbf{R}^d}} f(x)dx \leq \mathrm{Simp}\int_{\mathbf{R}^d} f_n(x) + \varepsilon dx$$

$$
\begin{aligned}
&= \text{Simp} \int_{\mathbf{R}^d} f_n(x) - \varepsilon dx + \text{Simp} \int_{\mathbf{R}^d} 2\varepsilon dx \\
&\leq \underline{\int_{\mathbf{R}^d}} f(x)dx + \text{Simp} \int_{\mathbf{R}^d} 2\varepsilon dx.
\end{aligned}
$$

Since f is vanishing outside of a set of finite measure, then we can ignore the set with infinite measure, and as ε converges to zero, we have Simp $\int_{\mathbf{R}^d} 2\varepsilon dx$ converges to zero as well. □

Corollary 1.3.22 (Finite additivity of the Lebesgue integral)**.** *Let $f, g : \mathbf{R}^d \to [0, +\infty]$ be measurable. Then*

$$
\int_{\mathbf{R}^d} f(x) + g(x)dx = \int_{\mathbf{R}^d} f(x)dx + \int_{\mathbf{R}^d} g(x)dx.
$$

Proof. Proof omitted. □

Proposition 1.3.23 (Upper Lebesgue integral and outer Lebesgue measure)**.** *For any set $E \subset \mathbf{R}^d$, $\overline{\int_{\mathbf{R}^d}} 1_E(x)dx = m^*(E)$. Conclude that the upper and lower Lebesgue integrals are not necessarily additive if no measurability hypotheses are assumed.*

Proof. For any $E \subset \mathbf{R}^d$, we can find a sequence of boxes $B_1, B_2, \cdots$ such that

$$
\begin{aligned}
m^*(E) &= \inf_{\bigcup_{n=1}^\infty B_n \supset E} m\Big(\bigcup_{n=1}^\infty B_n\Big) \\
&= \inf_{\bigcup_{n=1}^\infty B_n \supset E} \text{Simp} \int_{\mathbf{R}^d} 1_{\bigcup_{n=1}^\infty B_n} \\
&\geq \overline{\int_{\mathbf{R}^d}} 1_E(x)dx.
\end{aligned}
$$

Now we show that $m^*(E) \leq \overline{\int_{\mathbf{R}^d}} 1_E(x)dx$. From definition,

$$
\overline{\int_{\mathbf{R}^d}} 1_E(x)dx \geq \inf_{g \geq 1_E : g \text{ simple}} \text{Simp} \int_{\mathbf{R}^d} g(x)dx - \varepsilon.
$$

Since 1_E takes the value in $\{0,1\}$, then we can write g as the form $g = 1_{E_1} + \cdots + 1_{E_k}$ where $E_1, \cdots, E_k$ are Lebesgue measurable such that $\bigcup_{n=1}^k E_k \supset E$. Then

$$\begin{aligned}\overline{\int_{\mathbf{R}^d}} 1_E(x)dx &\geq \inf_{1_{E_1}+\cdots+1_{E_k}\geq 1_E} \operatorname{Simp}\int_{\mathbf{R}^d} 1_{E_1} + \cdots 1_{E_k} dx - \varepsilon \\ &= \inf_{E_1\cup\cdots\cup E_k\supset E} \sum_{n=1}^k m(E_k) - \varepsilon \\ &\geq m^*(E) - \varepsilon.\end{aligned}$$

Thus we conclude that $\overline{\int_{\mathbf{R}^d}} 1_E(x)dx = m^*(E)$. □

Proposition 1.3.24 (Area interpretation of integral)**.** *If $f : \mathbf{R}^d \to [0,+\infty]$ is measurable, then $\int_{\mathbf{R}^d} f(x)dx$ is equal to the $d+1$-dimensional Lebesgue measure of the region $\{(x,t) \in \mathbf{R}^d \times \mathbf{R} : 0 \leq t \leq f(x)\}$.*

Remark. This can be used as an alternate, and more geometrically intuitive, definition of the unsigned Lebesgue integral; it is a more convenient formulation for establishing the basic convergence theorems, but not quite as convenient for establishing basic properties such as additivity.

Proof. Denote the area under the graph of f as region(f). We first show that $\underline{\int_{\mathbf{R}^d}} f(x)dx \leq m(\text{region}(f))$. We know that region$(f)$ is measurable. Let $\varepsilon > 0$ be arbitrary, then there exists a countable collection $(B_n)_{n=1}^\infty$ of boxes such that $\bigcup_{n=1}^\infty B_n \supset \text{region}(f)$ and

$$\sum_{n=1}^\infty m(B_n) \leq m(\text{region}(f)) + \varepsilon.$$

Since for every simple function $g = c_1 1_{E_1} + \cdots + c_k 1_{E_k}$ minorizes f, we have $\bigcup_{n=1}^k E_n \times [0, c_n] \subset \bigcup_{n=1}^\infty B_n$. Then

$$\begin{aligned}\underline{\int_{\mathbf{R}^d}} f(x)dx &= \sup_{0\leq g\leq f} \operatorname{Simp}\int_{\mathbf{R}^d} g(x)dx \\ &= \sup_{0\leq g\leq f} \sum_{n=1}^k c_n m(E_n)\end{aligned}$$

$$
\begin{aligned}
&= \sup_{0\le g\le f} m\Big(\bigcup_{n=1}^{k} E_n \times [0, c_n]\Big)\\
&\le m\Big(\bigcup_{n=1}^{\infty} B_n\Big)\\
&\le m(\text{region}(f)) + \varepsilon.
\end{aligned}
$$

The third equality comes from Lemma 1.2.29. As ε was arbitrary, we have $\underline{\int_{\mathbf{R}^d}} f(x)dx \le m(\text{region}(f))$.

For the other hand, we want to show that $\underline{\int_{\mathbf{R}^d}} f(x)dx \ge m(\text{region}(f))$. Since f is measurable, by Lemma 1.3.9(iv), we have

$$
\begin{aligned}
\text{region}(f) &= \{(x,t) \in \mathbf{R}^d \times \mathbf{R} : 0 \le t \le f(x)\}\\
&= \{(x,t) \in \mathbf{R}^d \times \mathbf{R} : 0 \le t \le \sup_{n>0} f_n(x)\}\\
&= \bigcup_{n=1}^{\infty} \{(x,t) \in \mathbf{R}^d \times \mathbf{R} : 0 \le t \le f_n(x)\}.
\end{aligned}
$$

For each function f_n, we have $\{(x,t) \in \mathbf{R}^d \times \mathbf{R} : 0 \le t \le f_n(x)\} = \bigcup_{m=1}^{k} E_{n,m} \times [0, c_{n,m}]$, and it is increasing respect to n. Then from

$$
\text{region}(f) = \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{k} E_{n,i} \times [0, c_{n,i}],
$$

and f_n minorizes f, use the monotone convergence theorem, we have

$$
\begin{aligned}
m(\text{region}(f)) &= m\Big(\bigcup_{n=1}^{\infty} \bigcup_{i=1}^{k} E_{n,i} \times [0, c_{n,i}]\Big)\\
&= \lim_{n\to\infty} m\Big(\bigcup_{i=1}^{k} E_{n,i} \times [0, c_{n,i}]\Big)\\
&= \lim_{n\to\infty} \sum_{i=1}^{k} c_{n,i} m(E_{n,i})\\
&= \lim_{n\to\infty} \text{Simp} \int_{\mathbf{R}^d} f_n(x)dx\\
&\le \underline{\int_{\mathbf{R}^d}} f(x)dx,
\end{aligned}
$$

as desired. $\square$

Theorem 1.3.25 (Uniqueness of the Lebesgue integral). *The Lebesgue integral $f \mapsto \int_{\mathbf{R}^d} f(x)dx$ is the only map from measurable unsigned functions $f : \mathbf{R}^d \to [0, +\infty]$ to $[0, +\infty]$ that obeys the following properties for measurable $f, g : \mathbf{R}^d \to [0, +\infty]$:*

(i) *(Compatibility with the simple integral) If f is simple, then we have $\int_{\mathbf{R}^d} f(x)dx = \mathrm{Simp} \int_{\mathbf{R}^d} f(x)dx$.*

(ii) *(Finite additivity)*

$$\int_{\mathbf{R}^d} f(x) + g(x)dx = \int_{\mathbf{R}^d} f(x)dx + \int_{\mathbf{R}^d} g(x)dx.$$

(iii) *(Horizontal truncation) As $n \to \infty$, $\int_{\mathbf{R}^d} \min(f(x), n)dx$ converges to $\int_{\mathbf{R}^d} f(x)dx$.*

(iv) *(Vertical truncation) As $n \to \infty$, $\int_{\mathbf{R}^d} f(x)1_{|x|\le n}dx$ converges to $\int_{\mathbf{R}^d} f(x)dx$.*

Proof. Let $\mathcal{M}$ and $\mathcal{M}'$ be different maps such that $f \mapsto \int_{\mathbf{R}^d} f(x)dx$ and satisfying the properties. By Lemma 1.3.5, the maps are coincide for every simple function.

We first deal with the case when f is vanishing outside of finite measure, and f is bounded. By Lemma 1.3.11, there is a sequence of bounded simple functions $(f_n)_{n=1}^{\infty}$ converges uniformly to f. We want to show that

$$\mathcal{M}(f) = \int_{\mathbf{R}^d} f(x)dx = \lim_{n\to\infty} \mathrm{Simp} \int_{\mathbf{R}^d} f_n(x)dx.$$

Clearly, we have

$$\int_{\mathbf{R}^d} f(x)dx \geq \lim_{n\to\infty} \mathrm{Simp} \int_{\mathbf{R}^d} f_n(x)dx,$$

we show the other direction. From uniform convergence, the inequalities $f_n(x) - \varepsilon \leq f(x) \leq f_n(x) + \varepsilon$ holds for every $x \in \mathbf{R}^n$ and $n \geq N$ where $N > 0$. Then

$$\int_{\mathbf{R}^d} f(x)dx \leq \mathrm{Simp} \int_{\mathbf{R}^d} f_n(x) + \varepsilon dx$$

$$
\begin{aligned}
&= \operatorname{Simp}\int_{\mathbf{R}^d} f_n(x)dx + \int_{\mathbf{R}^d} \varepsilon dx \\
&\le \lim_{n\to\infty} \operatorname{Simp}\int_{\mathbf{R}^d} f_n(x)dx + \int_{\mathbf{R}^d} \varepsilon dx.
\end{aligned}
$$

Notice that ε is arbitrary follows that $\int_{\mathbf{R}^d} \varepsilon dx$ is also arbitrary. Then we have $\mathcal{M}(f) = \lim_{n\to\infty} \operatorname{Simp}\int_{\mathbf{R}^d} f_n(x)dx$. This is similar to show that $\mathcal{M}'(f) = \lim_{n\to\infty} \operatorname{Simp}\int_{\mathbf{R}^d} f_n(x)dx$. Thus the maps are coincide for bounded measurable function f which is vanishing outside of finite measure.

...[To Be Continued] □

Proposition 1.3.26 (Translation invariance)**.** *Let $f : \mathbf{R}^d \to [0, +\infty]$ be measurable. Then $\int_{\mathbf{R}^d} f(x+y)dx = \int_{\mathbf{R}^d} f(x)dx$ for any $y \in \mathbf{R}^d$.*

Proof. It immediately comes from Proposition 1.2.28, Lemma 1.2.29 and Proposition 1.3.24. □

Proposition 1.3.27 (Compatibility with the Riemann integral)**.** *Let $f : [a,b] \to [0, +\infty]$ be Riemann integrable. If we extend f to $\mathbf{R}$ by declaring f to equal zero outside of $[a,b]$, then $\int_{\mathbf{R}^d} f(x)dx = \int_a^b f(x)dx$.*

Proof. This is immediate consequence of Proposition 1.3.24 and Proposition 1.3.24. □

Lemma 1.3.28 (Markov's inequality)**.** *Let $f : \mathbf{R}^d \to [0, +\infty]$ be measurable. Then for any $0 < \lambda < \infty$, one has*

$$m(\{x \in \mathbf{R}^d : f(x) \ge \lambda\}) \le \frac{1}{\lambda}\int_{\mathbf{R}^d} f(x)dx.$$

Proof. Proof omitted. □

By sending λ to infinity or to zero, we obtain the following important corollary:

Corollary 1.3.29. *Let $f : \mathbf{R}^d \to [0, +\infty]$ be measurable.*

(i) *If $\int_{\mathbf{R}^d} f(x)dx < \infty$, then f is finite almost everywhere.*

(ii) *$\int_{\mathbf{R}^d} f(x)dx = 0$ if and only if f is zero almost everywhere.*

Proof. (i) As $\lambda \to \infty$, we see that $m(\{x \in \mathbf{R}^d : f(x) = \infty\})$ is a subset of null set for that $\int_{\mathbf{R}^d} f(x)dx < \infty$. This means that f taking the infinite value in null set, as desired.

(ii) If $\int_{\mathbf{R}^d} f(x)dx = 0$, as $\lambda \to 0$ we have

$$m(\{x \in \mathbf{R}^d : f(x) = 0\}) \leq m(\{x \in \mathbf{R}^d : f(x) \geq 0\}) \leq \frac{1}{\lambda} \int_{\mathbf{R}^d} f(x)dx = 0.$$

Conversely, f is zero a.e. means that for arbitrary λ we have

$$0 < m(\{x \in \mathbf{R}^d : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int_{\mathbf{R}^d} f(x)dx.$$

As $\lambda \to 0$, we require that $\int_{\mathbf{R}^d} f(x)dx = 0$. □

1.3.4 Absolute integrability

Having set out the theory of the unsigned Lebesgue integral, we can now define the absolutely convergent Lebesgue integral.

Definition 1.3.30 (Absolute integrability)**.** An almost everywhere defined measurable function $f : \mathbf{R}^d \to \mathbf{C}$ is said to be *absolutely integrable* if the unsigned integral

$$\|f\|_{L^1(\mathbf{R}^d)} := \int_{\mathbf{R}^d} |f(x)|dx$$

is finite. We refer to this quantity $\|f\|_{L^1(\mathbf{R}^d)}$ as the $L^1(\mathbf{R}^d)$ *norm* of f, and use $L^1(\mathbf{R}^d)$ or $L^1(\mathbf{R}^d \to \mathbf{C})$ to denote the space of absolutely integrable functions. If f is real-valued and absolutely integrable, we define the Lebesgue integral $\int_{\mathbf{R}^d} f(x)dx$ by the formula

$$\int_{\mathbf{R}^d} f(x)dx := \int_{\mathbf{R}^d} f_+(x)dx - \int_{\mathbf{R}^d} f_-(x)dx$$

where $f_+ := \max(f, 0), f_- := \max(-f, 0)$ are the magnitudes of the positive and negative components of f (note that the two unsigned integrals on the right-hand side are finite, as f_+, f_- are pointwise dominated by $|f|$). If f is complex-valued and absolutely integrable, we define the Lebesgue integral $\int_{\mathbf{R}^d} f(x)dx$ by the formula

$$\int_{\mathbf{R}^d} f(x)dx := \int_{\mathbf{R}^d} \operatorname{Re} f(x)dx + i \int_{\mathbf{R}^d} \operatorname{Im} f(x)dx$$

where the two integrals on the right are interpreted as real-valued absolutely integrable Lebesgue integrals. It is easy to see that the unsigned, real-valued, and complex-valued Lebesgue integrals defined in this manner are compatible on their common domains of definition.

The linearity properties of the unsigned integral induce analogous linearity properties of the absolutely convergent Lebesgue integral:

Proposition 1.3.31 (Integration is linear)**.** *The integration $f \mapsto \int_{\mathbf{R}^d} f(x)dx$ is a (complex) linear operation from $L^1(\mathbf{R}^d)$ to $\mathbf{C}$. In other words,*

$$\int_{\mathbf{R}^d} f(x) + g(x)dx = \int_{\mathbf{R}^d} f(x)dx + \int_{\mathbf{R}^d} g(x)dx$$

and

$$\int_{\mathbf{R}^d} cf(x)dx = c\int_{\mathbf{R}^d} f(x)dx$$

for all absolutely integrable $f, g : \mathbf{R}^d \to \mathbf{C}$ and complex numbers c. Also, we have

$$\int_{\mathbf{R}^d} \overline{f(x)}dx = \overline{\int_{\mathbf{R}^d} f(x)dx},$$

which makes integration not just a linear operation, but a *-linear *operation.*

Proof. First, by Corollary 1.3.22, the assertion is hold for real-valued functions f and g that

$$\begin{aligned}\int_{\mathbf{R}^d} f(x)+g(x)dx &= \int_{\mathbf{R}^d} f_+(x)+g_+(x)dx - \int_{\mathbf{R}^d} f_-(x)+g_-(x)dx \\ &= \int_{\mathbf{R}^d} f_+(x)dx + \int_{\mathbf{R}^d} g_+(x)dx \\ &\quad - \int_{\mathbf{R}^d} f_-(x)dx - \int_{\mathbf{R}^d} g_-(x)dx \\ &= \int_{\mathbf{R}^d} f(x)dx + \int_{\mathbf{R}^d} g(x)dx.\end{aligned}$$

Then it follows that for complex-valued functions f, g, we have

$$\begin{aligned}\int_{\mathbf{R}^d} f(x)+g(x)dx &= \int_{\mathbf{R}^d} \operatorname{Re} f(x) + \operatorname{Re} g(x)dx + i\int_{\mathbf{R}^d} \operatorname{Im} f(x) + \operatorname{Im} g(x)dx \\ &= \int_{\mathbf{R}^d} \operatorname{Re} f(x)dx + \int_{\mathbf{R}^d} \operatorname{Re} g(x)dx \\ &\quad + i\int_{\mathbf{R}^d} \operatorname{Im} f(x)dx + i\int_{\mathbf{R}^d} \operatorname{Im} g(x)dx \\ &= \int_{\mathbf{R}^d} f(x)dx + \int_{\mathbf{R}^d} g(x)dx.\end{aligned}$$

Now we prove the homogeneity. We first consider the real-valued function f and $c \in \mathbf{R}$. When $c \geq 0$, the equality follows Proposition 1.3.19. While if $c < 0$, then

$$\begin{aligned}\int_{\mathbf{R}^d} cf(x)dx &= \int_{\mathbf{R}^d} (cf)_+(x)dx - \int_{\mathbf{R}^d} (cf)_-(x)dx \\ &= \int_{\mathbf{R}^d} (-c)f_-(x)dx - \int_{\mathbf{R}^d} (-c)f_+(x)dx \\ &= -c\Big(\int_{\mathbf{R}^d} f_-(x)dx - \int_{\mathbf{R}^d} f_+(x)dx\Big) \\ &= c\int_{\mathbf{R}^d} f(x)dx.\end{aligned}$$

Thus for real-valued function f and $c \in \mathbf{R}$, we have

$$\int_{\mathbf{R}^d} cf(x)dx = c\int_{\mathbf{R}^d} f(x)dx.$$

Then we consider the complex-valued function f and $c \in \mathbf{C}$. This is easy to see that

$$\begin{aligned}
\int_{\mathbf{R}^d} cf(x)dx &= \int_{\mathbf{R}^d} \operatorname{Re}(cf)(x)dx + i\int_{\mathbf{R}^d} \operatorname{Im}(cf)(x)dx \\
&= \int_{\mathbf{R}^d} \operatorname{Re}(c)\operatorname{Re} f(x) - \operatorname{Im}(c)\operatorname{Im} f(x)dx \\
&\quad + i\int_{\mathbf{R}^d} \operatorname{Re}(c)\operatorname{Im} f(x) + \operatorname{Im}(c)\operatorname{Re} f(x)dx \\
&= \operatorname{Re}(c)\int_{\mathbf{R}^d} \operatorname{Re} f(x)dx - \operatorname{Im}(c)\int_{\mathbf{R}^d} \operatorname{Im} f(x)dx \\
&\quad + i\operatorname{Re}(c)\int_{\mathbf{R}^d} \operatorname{Im} f(x)dx + i\operatorname{Im}(c)\int_{\mathbf{R}^d} \operatorname{Re} f(x)dx \\
&= c\int_{\mathbf{R}^d} f(x)dx.
\end{aligned}$$

Use linearity, we have

$$\begin{aligned}
\int_{\mathbf{R}^d} \overline{f(x)}dx &= \int_{\mathbf{R}^d} \operatorname{Re}\overline{f(x)}dx + i\int_{\mathbf{R}^d} \operatorname{Im}\overline{f(x)} \\
&= \int_{\mathbf{R}^d} \operatorname{Re} f(x)dx + i\int_{\mathbf{R}^d} (-1)\operatorname{Im} f(x) \\
&= \int_{\mathbf{R}^d} \operatorname{Re} f(x)dx - i\int_{\mathbf{R}^d} \operatorname{Im} f(x) \\
&= \overline{\int_{\mathbf{R}^d} f(x)dx}.
\end{aligned}$$

$\square$

Proposition 1.3.32 (Absolute summability is a special case of absolute integrability)**.** *Let $(c_n)_{n\in\mathbf{Z}}$ be a doubly infinite sequence of complex numbers, and let $f : \mathbf{R} \to \mathbf{C}$ be the function*

$$f(x) := \sum_{n\in\mathbf{Z}} c_n 1_{[n,n+1)}(x) = c_{\lfloor x\rfloor}$$

where $\lfloor x\rfloor$ is the greatest integer less than or equal to x. Then f is absolutely integrable if and only if the series $\sum_{n\in\mathbf{Z}} c_n$ is absolutely *convergent, in which case one has $\int_{\mathbf{R}^d} f(x)dx = \sum_{n\in\mathbf{Z}} c_n$.*

Proof. Suppose that f is absolutely integrable, we have

$$\int_{\mathbf{R}^d} |f(x)|dx = \int_{\mathbf{R}^d} \Big|\sum_{n\in\mathbf{Z}} c_n 1_{[n,n+1)}(x)\Big| dx < \infty.$$

For every $x \in \mathbf{R}$, we have

$$\Big|\sum_{n\in\mathbf{Z}} c_n 1_{[n,n+1)}(x)\Big| = |c_{\lfloor x\rfloor}| = \sum_{n\in\mathbf{Z}} |c_n| 1_{[n,n+1)}(x),$$

thus

$$\int_{\mathbf{R}^d} \sum_{n\in\mathbf{Z}} |c_n| 1_{[n,n+1)}(x) dx < \infty.$$

Since $\sum_{n\in\mathbf{Z}} |c_n| 1_{[n,n+1)}(x)$ is unsigned, consider the vertical truncation,

$$\begin{aligned}\int_{\mathbf{R}^d} \Big(\sum_{n\in\mathbf{Z}} |c_n| 1_{[n,n+1)}(x)\Big) 1_{|x|\le n} dx &= \int_{\mathbf{R}^d} \Big(\sum_{i=1}^{n} |c_i| 1_{[i,i+1)}(x)\Big) 1_{|x|\le n} dx \\ &= \sum_{i=1}^{n} |c_i|.\end{aligned}$$

By Proposition 1.3.19(ix), we have

$$\begin{aligned}\int_{\mathbf{R}^d} \sum_{n\in\mathbf{Z}} |c_n| 1_{[n,n+1)}(x) dx &= \lim_{n\to\infty} \int_{\mathbf{R}^d} \Big(\sum_{n\in\mathbf{Z}} |c_n| 1_{[n,n+1)}(x)\Big) 1_{|x|\le n} dx \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} |c_i| \\ &= \sum_{n=1}^{\infty} |c_n| < \infty.\end{aligned}$$

Thus $\sum_{n\in\mathbf{Z}} c_n$ is absolutely convergent. Since $\int_{\mathbf{R}^d} f(x)dx$ can be represented as form

$$\begin{aligned}\int_{\mathbf{R}^d} f(x)dx = &\Big(\int_{\mathbf{R}^d} \operatorname{Re} f_+(x)dx - \int_{\mathbf{R}^d} \operatorname{Re} f_-(x)dx\Big) \\ &+ i\Big(\int_{\mathbf{R}^d} \operatorname{Im} f_+(x)dx - \int_{\mathbf{R}^d} \operatorname{Im} f_-(x)dx\Big),\end{aligned}$$

each of which are unsigned. A similar argument shows that

$$\begin{aligned}\int_{\mathbf{R}^d} f(x)dx &= \Big(\sum_{n=1}^{\infty}\mathrm{Re}(c_n)_+ - \sum_{n=1}^{\infty}\mathrm{Re}(c_n)_-\Big) \\ &\qquad + i\Big(\sum_{n=1}^{\infty}\mathrm{Im}(c_n)_+ - \sum_{n=1}^{\infty}\mathrm{Im}(c_n)_-\Big) \\ &= \sum_{n=1}^{\infty} c_n.\end{aligned}$$

For the converse, the process is similar. □

We define

$$\int_E f(x)dx := \int_{\mathbf{R}^d} f(x)1_E(x)dx,$$

then we have following proposition:

Proposition 1.3.33. *If E, F are disjoint measurable subsets of $\mathbf{R}^d$, and $f : E \cup F \to \mathbf{C}$ is absolutely integrable, then*

$$\int_E f(x)dx = \int_{E\cup F} f(x)1_E(x)dx$$

and

$$\int_E f(x)dx + \int_F f(x)dx = \int_{E\cup F} f(x)dx.$$

Proof. Since $1_E = 1_E \cdot 1_{E\cup F}$, we have

$$\begin{aligned}\int_E f(x)dx &= \int_{\mathbf{R}^d} f(x)1_E(x)dx \\ &= \int_{\mathbf{R}^d} f(x)1_E(x)1_{E\cup F}(x)dx \\ &= \int_{E\cup F} f(x)1_E(x)dx.\end{aligned}$$

From $1_E + 1_F = 1_{E\cup F}$ and the linearity of integration, we have

$$\int_E f(x)dx + \int_F f(x)dx = \int_{\mathbf{R}^d} f(x)1_E(x)dx + \int_{\mathbf{R}^d} f(x)1_F(x)dx$$

$$= \int_{\mathbf{R}^d} f(x)(1_E(x) + 1_F(x))dx$$
$$= \int_{\mathbf{R}^d} f(x)1_{E\cup F}(x)dx$$
$$= \int_{E\cup F} f(x)dx.$$

□

We will study the properties of the absolutely convergent Lebesgue integral in more detail in later notes, as a special case of the more general Lebesgue integration theory on abstract measure spaces. For now, we record one very basic inequality:

Lemma 1.3.34 (Triangle inequality)**.** *Let* $f \in L^1(\mathbf{R}^d \to \mathbf{C})$*. Then*

$$\Big| \int_{\mathbf{R}^d} f(x)dx \Big| \le \int_{\mathbf{R}^d} |f(x)|dx.$$

Proof. Proof omitted. □

1.3.5 Littlewood's three principles

Littlewood's three principles are informal heuristics that convey much of the basic intuition behind the measure theory of Lebesgue. Briefly, the three principles are as follows:

(i) Every (measurable) set is nearly a finite sum of intervals;

(ii) Every (absolutely integrable) function is nearly continuous; and

(iii) Every (pointwise) convergent sequence of functions is nearly uniformly convergent.

Various manifestations of the first principle were given in Theorem 1.2.13 and Theorem 1.2.23. Now we turn to the second principle. Define a *step function* to be a finite linear combination of indicator functions 1_B of boxes B.

Theorem 1.3.35 (Approximation of L^1 functions)**.** *Let $f \in L^1(\mathbf{R}^d)$ and $\varepsilon > 0$.*

(i) *There exists an absolutely integrable simple function g such that $\|f - g\|_{L^1(\mathbf{R}^d)} \leq \varepsilon$.*

(ii) *There exists a step function g such that $\|f - g\|_{L^1(\mathbf{R}^d)} \leq \varepsilon$.*

(iii) *There exists a continuous, compactly supported*[a] *g such that $\|f - g\|_{L^1(\mathbf{R}^d)} \leq \varepsilon$.*

[a]We call a function *compactly supported* if its support is contained in a compact set.

Proof. Proof omitted. □

This is note the only way to make Littlewood's second principle manifest; we return to this point shortly. For now, we turn to Littlewood's third principle. We recall three basic ways in which a sequence $f_n : \mathbf{R}^d \to \mathbf{C}$ of functions can converge to a limit $f : \mathbf{R}^d \to \mathbf{C}$:

(i) (Pointwise convergence) $f_n(x) \to f(x)$ for every $x \in \mathbf{R}^d$.

(ii) (Pointwise almost everywhere convergence) $f_n(x) \to f(x)$ for almost every $x \in \mathbf{R}^d$.

(iii) (Uniform convergence) For every $\varepsilon > 0$, there exists N such that $|f_n(x) - f(x)| \leq \varepsilon$ for all $n \geq N$ and all $x \in \mathbf{R}^d$.

Uniform convergence implies pointwise convergence, which in turn implies pointwise almost everywhere convergence.

We now add a fourth mode of convergence, that is weaker than uniform convergence but stronger than pointwise convergence:

Definition 1.3.36 (Locally uniform convergence)**.** A sequence of functions $f_n : \mathbf{R}^d \to \mathbf{C}$ converges *locally uniformly* to a limit $f : \mathbf{R}^d \to \mathbf{C}$ if, for every bounded subset E of $\mathbf{R}^d$, f_n converges uniformly to f on E. In other words, for every bounded $E \subset \mathbf{R}^d$ and every $\varepsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \varepsilon$ for all $n \geq N$ and $x \in E$.

Remark. A property P is said to hold *locally* on some domain X if, for every point x_0 in that domain, there is an open neighbourhood of x_0 in X on which P holds.

Following example shows that pointwise convergence is a weaker concept than local uniform convergence:

Example 1.3.37. The functions $f_n(x) := \frac{1}{nx} 1_{x>0}$ for $n = 1, 2, \cdots$ (with the convention that $f_n(0) = 0$) converge pointwise everywhere to zero, but do not converge locally uniformly.

A remarkable theorem of Egorov, which demonstrates Littlewood's third principle, asserts that one can recover local uniform convergence as long as one is willing to delete a set of small measure:

Theorem 1.3.38 (Egorov's theorem)**.** *Let $f_n : \mathbf{R}^d \to \mathbf{C}$ be a sequence of measurable functions that converge pointwise almost everywhere to another function $f : \mathbf{R}^d \to \mathbf{C}$, and let $\varepsilon > 0$. Then there exists a Lebesgue measurable set A of measure at most ε, such that f_n converges locally uniformly to f outside of A.*

Proof. Proof omitted. □

Unfortunately, one cannot in general upgrade local uniform convergence to uniform convergence in Egorov's theorem:

Example 1.3.39. Let $f_n := 1_{[n,n+1]}$ on $\mathbf{R}$. Then f_n converges pointwise (and locally uniformly) to the zero function $f \equiv 0$. However, for any $0 < \varepsilon < 1$ and any n, we have $|f_n(x) - f(x)| > \varepsilon$ on the interval $[n, n+1]$ with measure 1. Thus, if one wanted f_n to converge uniformly to f outside of a set A, then that set A has to contain a set of measure 1.

However, if all the f_n and f were supported on a fixed set E of finite measure (e.g. on a ball $B(0, R)$), then the above "escape to horizontal infinity" cannot occur, it is easy to see from the above argument that one can recover uniform convergence (and not just locally uniform convergence) outside of a set of arbitrarily small measure.

We now use Theorem 1.3.38 to give another version of Littlewood's second principle, known as *Lusin's theorem*:

Theorem 1.3.40 (Lusin's theorem)**.** *Let $f : \mathbf{R}^d \to \mathbf{C}$ be absolutely integrable, and let $\varepsilon > 0$. Then there exists a Lebesgue measurable set $E \subset \mathbf{R}^d$ of measure at most ε such that the restriction of f to the complementary set $\mathbf{R}^d \setminus E$ is continuous on that set.*

Proof. Proof omitted. □

Proposition 1.3.41 (Relaxed Lusin's theorem)**.** *The hypothesis that f is absolutely integrable in Lusin's theorem can be relaxed to being locally absolutely integrable (i.e. absolutely integrable on every bounded set), and then relaxed further to that of being measurable (but still finite everywhere or almost everywhere).*

Proof. (i) Let f be locally absolutely integrable. By Theorem 1.3.35, for any $n \geq 1$ and $R > 0$ one can find a continuous, compactly supported function f_n such that $\|f - f_n\|_{L^1(\mathbf{R}^d)} \leq \varepsilon/4^n$ for every $x \in B(0, R)$. By Markov's inequality, that implies that $|f(x) - f_n(x)| \leq 1/2^n$ for all $x \in B(0, R) \setminus E_n$ where E_n is a Lebesgue measurable set of measure at most $\varepsilon/2^n$. Letting $E := \bigcup_{n=1}^{\infty} E_n$. Then we have $f_n|_{B(0,R)}$ converges uniformly to $f|_{B(0,R)}$ outside of E which is a Lebesgue measurable with measure at most ε, and this implies that $f|_{B(0,R)}$ is continuous for all $R > 0$. Since $f|_{B(0,R)}$ converge pointwise to f respect to R, we conclude that f is also continuous. We conclude that f is continuous on $\mathbf{R}^d \setminus E$.

(ii) Furthermore, we suppose that f to be measurable. Consider a bounded measurable function $f1_{|f|\leq n}$, which is clearly absolutely integrable. By Lusin's theorem, we can find a Lebesgue measurable set E_n for every

$n > 0$ with measure at most $\varepsilon/2^{n+1}$, such that $f1_{|f|\leq n}$ be continuous on $\mathbf{R}^d \setminus E_n$. Letting $E := \bigcup_{n=1}^{\infty} E_n$. Since $f1_{|f|\leq n}$ converge to f pointwise a.e., by Egorov's theorem, we can find a measurable set A of measure at most $\varepsilon/2$ such that f_n converges uniformly to f outside of $A \cup E$ with measure at most ε. We conclude that the restriction f to $\mathbf{R}^d \setminus (A \cup E)$ is continuous, as required. □

Lemma 1.3.42. *A function $f : \mathbf{R}^d \to \mathbf{C}$ is measurable if and only if it is the pointwise almost everywhere limit of continuous functions $f_n : \mathbf{R}^d \to \mathbf{C}$.*

Proof. Suppose that f is measurable and $n \geq 1$. By Theorem 1.3.44(ii), for every $n \geq 1$ there exists a measurable set $E \subset \mathbf{R}^d$ of measure at most $\varepsilon/2^n$ and $M < \infty$ such that $|f(x)| \leq M$ for all $x \in B(0,n) \setminus E$. Then we see that f is absolutely integrable on $B(0,n) \setminus E$. By Theorem 1.3.35, there is a sequence of continuous, compactly supported function f_n such that $\|f - f_n\|_{L^1(\mathbf{R}^d)} \leq 1/n2^n$. Then by Markov's inequality, we have

$$m(\{x \in B(0,n) \setminus E : |f(x) - f_n(x)| \geq \frac{1}{n}\}) \leq \frac{1}{2^n}.$$

As n converges to infinity, we see that f_n converges pointwise to f for all x outside of a Lebesgue measurable set with measure zero.

Conversely, suppose that f_n converges pointwise to f, where f_n are continuous. By Proposition 1.3.16(i), f_n are measurable. Then by Proposition 1.3.16(iv), f is measurable. □

We can use Lemma 1.3.42 and Egorov's theorem (Theorem 1.3.38) to give an alternate proof of Lusin's theorem for arbitrary measurable functions:

Theorem 1.3.43 (Lusin's theorem, again). *Let $f : \mathbf{R}^d \to \mathbf{C}$ be measurable, and let $\varepsilon > 0$. Then there exists a Lebesgue measurable set $E \subset \mathbf{R}^d$ of measure at most ε such that the restriction of f to the* *complementary set $\mathbf{R}^d \setminus E$ is continuous on that set.*

Proof.

The following facts are not, strictly speaking, instances of any of Littlewood's three principles, but are in a similar spirit.

Theorem 1.3.44 (Littlewood-like principles)**.**

(i) *(Absolutely integrable functions almost have bounded support) Let $f : \mathbf{R}^d \to \mathbf{C}$ be an absolutely integrable function, and let $\varepsilon > 0$. Then there exists a ball $B(0, R)$ outside of which f has an L^1 norm of at most ε, or in other words, that $\int_{\mathbf{R}^d \setminus B(0,R)} |f(x)|dx \leq \varepsilon$.*

(ii) *(Measurable functions are almost locally bounded) Let $f : \mathbf{R}^d \to \mathbf{C}$ be a measurable function, and let $\varepsilon > 0$. Then there exists a measurable set $E \subset \mathbf{R}^d$ of measure at most ε outside of which f is locally bounded, or in other words, that for every $R > 0$ there exists $M < \infty$ such that $|f(x)| \leq M$ for all $x \in B(0, R) \setminus E$.*

Proof. (i) By Theorem 1.3.35(iii), there exists a continuous, compactly supported g such that $\|f - g\|_{L^1(\mathbf{R}^d)} \leq \varepsilon$. Since g is compactly supported, there is a ball $B(0, R)$ such that for all $x \in \mathbf{R}^d$ such that $g(x) \neq 0$, we have $x \in B(0, R)$. Then

$$\begin{aligned}
\int_{\mathbf{R}^d \setminus B(0,R)} |f(x)|dx &\leq \int_{\mathbf{R}^d \setminus B(0,R)} |f(x)|dx + \int_{B(0,R)} |f(x) - g(x)|dx \\
&= \int_{\mathbf{R}^d \setminus B(0,R)} |f(x) - g(x)|dx + \int_{B(0,R)} |f(x) - g(x)|dx \\
&= \int_{\mathbf{R}^d} |f(x) - g(x)|dx \\
&\leq \varepsilon.
\end{aligned}$$

(ii) f is measurable follows that we can find a sequence of simple measurable functions f_n converges pointwise to f. Then by Theorem 1.3.38, there exists a Lebesgue measurable set $E \subset \mathbf{R}^d$ of measure at most ε, such that f_n converges uniformly to f on $B(0,R) \setminus E$ for all $R > 0$.

§1.4 Abstract measure spaces

1.4.1 Boolean algebras

We begin by recalling the concept of a *Boolean algebra*.

Definition 1.4.1 (Boolean algebras)**.** Let X be a set. A (concrete) *Boolean algebra* on X is a collection $\mathcal{B}$ of subsets of X which obeys the following properties:

(i) (Empty set) $\emptyset \in \mathcal{B}$.

(ii) (Complement) If $E \in \mathcal{B}$, then the complement $E^c := X \setminus E$ also lies in $\mathcal{B}$.

(iii) (Finite unions) If $E, F \in \mathcal{B}$, then $E \cup F \in \mathcal{B}$.

We sometimes say that E is *$\mathcal{B}$-measurable*, or *measurable with respect to $\mathcal{B}$*, if $E \in \mathcal{B}$.

Given two Boolean algebras $\mathcal{B}, \mathcal{B}'$ on X, we say that $\mathcal{B}'$ is *finer than*, or *a refinement of* $\mathcal{B}$, or that $\mathcal{B}$ is *coarser than*, *a sub-algebra of*, or *a coasrsening of* $\mathcal{B}'$, if $\mathcal{B} \subset \mathcal{B}'$.

Remark. The Boolean algebra is closed under *complement* and *finite union*, it is easy to see that a Boolean algebra is also closed under the *intersection* $E \cap F$, *set difference* $E \setminus F$, *symmetric difference* $E \triangle F$.

Example 1.4.2 (Elementary algebra)**.** Let $\overline{\mathcal{E}[\mathbf{R}^d]}$ be the collection of those sets $E \subset \mathbf{R}^d$ that are either elementary sets, or co-elementary sets (i.e. the complement of an elementary set). Then $\overline{\mathcal{E}[\mathbf{R}^d]}$ is a Boolean algebra. We will call this algebra the *elementary Boolean algebra* of $\mathbf{R}^d$.

Proof. Clearly, $\emptyset \in \overline{\mathcal{E}[\mathbf{R}^d]}$, then X, which is the co-elementary set of $\emptyset$, is also contained in $\overline{\mathcal{E}[\mathbf{R}^d]}$. Thus by Proposition 1.1.2, $\overline{\mathcal{E}[\mathbf{R}^d]}$ is a Boolean algebra. $\square$

Example 1.4.3. (Jordan algebra) Let $\overline{\mathcal{J}[\mathbf{R}^d]}$ be the collection of subsets of $\mathbf{R}^d$ that are either Jordan measurable or co-Jordan measurable (i.e. the complement of a Jordan measurable set). Then $\overline{\mathcal{J}[\mathbf{R}^d]}$ is a Boolean algebra that is finer than the elementary algebra. We refer to this algebra as the *Jordan algebra* on $\mathbf{R}^d$ (but caution that there is a completely different concept of a Jordan algebra in abstract algebra).

Proof. This is immediate consequence of Proposition 1.1.9(i). $\square$

Example 1.4.4. (Lebesgue algebra) Let $\overline{\mathcal{L}[\mathbf{R}^d]}$ be the collection of Lebesgue measurable subsets of $\mathbf{R}^d$. Then $\overline{\mathcal{L}[\mathbf{R}^d]}$ is a Boolean algebra that is finer than the Jordan algebra; we refer to this algebra as the *Lebesgue algebra* on $\mathbf{R}^d$.

Proof. This immediately comes from Lemma 1.2.12(iv), (v), and (vii). $\square$

Example 1.4.5 (Null algebra)**.** Let $\mathcal{N}[\mathbf{R}^d]$ be the collection of subsets of $\mathbf{R}^d$ that are either Lebesgue null sets or Lebesgue co-null sets (the complement of null sets). Then $\mathcal{N}[\mathbf{R}^d]$ is a Boolean algebra that is coarser than the Lebesgue algebra; we refer to it as the *null algebra* on $\mathbf{R}^d$.

Proof. Clearly, $\emptyset$ and X are contained in $\mathcal{N}[\mathbf{R}^d]$. The second condition is trivial. Let $E, F \in \mathcal{N}[\mathbf{R}^d]$. If E, F are null sets, then $E \cup F$ still null sets and is contained in $\mathcal{N}[\mathbf{R}^d]$. If E, F are co-null sets, then $\mathbf{R}^d \setminus E$ and $\mathbf{R}^d \setminus F$ are null sets, then $\mathbf{R}^d \setminus (E \cup F) = (\mathbf{R}^d \setminus E) \cap (\mathbf{R}^d \setminus F)$ which is also null set. Thus $E \cup F$ is co-null set and contained in $\mathcal{N}[\mathbf{R}^d]$. While if E is null set, and F is co-null set. Clearly, $\mathbf{R}^d \setminus (E \cup F) = (\mathbf{R}^d \setminus E) \cap (\mathbf{R}^d \setminus F)$ is null set, thus $E \cup F$ is co-null set and contained in $\mathcal{N}[\mathbf{R}^d]$. $\square$

Example 1.4.6 (Restriction)**.** Let $\mathcal{B}$ be a Boolean algebra on a set X, and let Y be a subset of X (not necessarily $\mathcal{B}$-measurable). Then the *restriction* $\mathcal{B} \downharpoonright_Y := \{E \cap Y : E \in \mathcal{B}\}$ of $\mathcal{B}$ to Y is a Boolean algebra on Y. If Y is $\mathcal{B}$-measurable, then

$$\mathcal{B} \downharpoonright_Y = \mathcal{B} \cap 2^Y = \{E \subset Y : E \in \mathcal{B}\}.$$

Proof. Obviously, $\emptyset \in \mathcal{B} \downharpoonright_Y$. If $E \cap Y \in \mathcal{B} \downharpoonright_Y$, we want to show that $(E \cap Y)^c = Y \setminus (E \cap Y) \in \mathcal{B} \downharpoonright_Y$. Since $E^c \in \mathcal{B}$, we have $E^c \cap Y \in \mathcal{B} \downharpoonright_Y$. Then

$$E^c \cap Y = (X \cap Y) \setminus (E \cap Y) = Y \setminus (E \cap Y) = (E \cap Y)^c.$$

Thus $(E \cap Y)^c \in \mathcal{B} \downharpoonright_Y$. Finally, if $E \cap Y, F \cap Y \in \mathcal{B} \downharpoonright_Y$, this implies that $E \cup F \in \mathcal{B}$ so that $(E \cup F) \cap Y \in \mathcal{B} \downharpoonright_Y$. From

$$(E \cup F) \cap Y = (E \cap Y) \cup (F \cap Y)$$

we obtained that $(E \cap Y) \cup (F \cap Y) \in \mathcal{B} \downharpoonright_Y$. Together our conclusions, $\mathcal{B} \downharpoonright_Y$ is a Boolean algebra.

If Y is $\mathcal{B}$-measurable, for every $E \cap Y \in \mathcal{B} \downharpoonright_Y$ we have $E \cap Y \in \mathcal{B}$. Thus $E \cap Y \in \mathcal{B} \cap 2^Y$. This means that $\mathcal{B} \downharpoonright_Y \subset \mathcal{B} \cap 2^Y$. For the other hand, for every $E \in \mathcal{B} \cap 2^Y$, we have $E = E \cap Y \in \mathcal{B} \downharpoonright_Y$. Thus $\mathcal{B} \cap 2^Y \subset \mathcal{B} \downharpoonright_Y$. $\square$

Example 1.4.7 (Atomic algebra)**.** Let X be partitioned into a union $X = \bigcup_{\alpha \in I} A_\alpha$ of disjoint sets A_α, which we refer to as *atoms*. Then this partition generates a Boolean algebra $\mathcal{A}((A_\alpha)_{\alpha \in I})$, defined as the collection of all the sets E of the form $E = \bigcup_{\alpha \in J} A_\alpha$ for some $J \subset I$, i.e., $\mathcal{A}((A_\alpha)_{\alpha \in I})$ is the collection of all sets that can be represented as the union of one or more atoms. This is easily verified to be a Boolean algebra, and we refer to it as the *atomic algebra* with atoms $(A_\alpha)_{\alpha \in I}$.

Lemma 1.4.8. *The non-empty atoms of an atomic algebra are determined up to relabeling. More precisely, if $X = \bigcup_{\alpha\in I} A_\alpha = \bigcup_{\alpha'\in I'} A'_{\alpha'}$ are two partitions of X into non-empty atoms $A_\alpha, A'_{\alpha'}$, then*

$$\mathcal{A}((A_\alpha)_{\alpha\in I}) = \mathcal{A}((A'_{\alpha'})_{\alpha'\in I'})$$

if and only if there exists a bijection $\phi : I \to I'$ such that $A'_{\phi(\alpha)} = A_\alpha$ for all $\alpha \in I$.

Proof. Suppose that $\mathcal{A}((A_\alpha)_{\alpha\in I}) = \mathcal{A}((A'_{\alpha'})_{\alpha'\in I'})$. We can see that for every $A_\alpha \in \mathcal{A}((A_\alpha)_{\alpha\in I})$ is also contained in $\mathcal{A}((A'_{\alpha'})_{\alpha'\in I'})$. Then we can find a set $E' = \bigcup_{\alpha'\in J'} A'_{\alpha'}$ in $\mathcal{A}((A'_{\alpha'})_{\alpha'\in I'})$ for some $J' \subset I'$ such that $A_\alpha = E'$. This assertion also holds for every $A'_{\alpha'}$. We need to show that all of J' are singleton. Suppose for sake of contradiction that there is at least one J' are not singleton, then we can find an atom A_α such that $A_\alpha = \bigcup_{\alpha'\in J'} A'_{\alpha'}$. Since A_α are disjoint for $\alpha \in I$, we see that there is no subset $J \subset I$ such that $A'_{\alpha'} = \bigcup_{\alpha\in J} A_\alpha$ for $\alpha' \in J'$, a contradiction. Thus for every A_α there is a singleton $\{\alpha'\} \subset I'$ such that $A_\alpha = A'_{\alpha'}$, and and this holds for every $A'_{\alpha'}$. This means that there exists a bijection $\phi : I \to I'$ such that $\phi(\alpha) = \alpha'$ such that $A'_{\phi(\alpha)} = A_\alpha$ for all $\alpha \in I$.

Conversely, suppose that there exists a bijection $\phi : I \to I'$ such that $A'_{\phi(\alpha)} = A_\alpha$ for all $\alpha \in I$. Then for every $\bigcup_{\alpha\in J} A_\alpha \in \mathcal{A}((A_\alpha)_{\alpha\in I})$ we have $\bigcup_{\alpha\in J} A_\alpha = \bigcup_{\phi(\alpha)\in J'} A'_{\phi(\alpha)} \in \mathcal{A}((A'_{\alpha'})_{\alpha'\in I'})$. Thus $\mathcal{A}((A_\alpha)_{\alpha\in I}) \subset \mathcal{A}((A'_{\alpha'})_{\alpha'\in I'})$. A similar argument shows that $\mathcal{A}((A'_{\alpha'})_{\alpha'\in I'}) \subset \mathcal{A}((A_\alpha)_{\alpha\in I})$, as desired. $\square$

While many Boolean algebras are atomic, many are not, as the following two lemmas indicate.

Lemma 1.4.9. *Every finite Boolean algebra is an atomic algebra. (A Boolean algebra $\mathcal{B}$ is* finite *if its cardinality is finite, i.e., there are only finitely many measurable sets.) Every finite Boolean algebra has a cardinality of the form 2^n for some natural number n. From this lemma* *and Lemma 1.4.8 we see that there is a one-to-one correspondence between finite Boolean algebras on X and finite partitions of X into non-empty sets (up to relabeling).*

Proof. Let $E_1, \cdots, E_k \in \mathcal{B}$. Clearly, we have $E_n \subset X$ for every $1 \leq n \leq k$, and $\bigcup_{n=1}^k E_n = X$. We use a Venn diagram argument. The k sets $E_1, \cdots, E_k$ partition X into 2^k disjoint sets, each of which is an intersection of some of the $E_1, \cdots, E_k$ and their complements. Throw away any sets that are empty, leaving us with a partition of X into m non-empty disjoint sets $A_1, \cdots, A_m$ for some $0 \leq m \leq 2^k$. Then we can see that every element $E \in \mathcal{B}$ can be represented as the union of some of $A_1, \cdots, A_m$, hence $\mathcal{B}$ is an atomic algebra. □

Lemma 1.4.10. *The elementary, Jordan, Lebesgue, and null algebras are* not *atomic algebras.*

Proof. We suppose for sake of contradiction that all of algebras above are atomic algebras.

Elementary algebra is not atomic algebra. Since every elementary set can be represented as the union of finitely many boxes, and every singleton $\{x\}$ for $x \in \mathbf{R}^d$ is also box. Then we require that $\{x\}_{x\in\mathbf{R}^d}$ be the atoms which such that $X = \bigcup_{x\in\mathbf{R}^d}\{x\}$; otherwise, we can find an $x^* \in \mathbf{R}^d$ such that $\{x^*\}$ can not be presented as the union of atoms. From the definition of atomic algebra, we have $[0,1] \cap \mathbf{Q} = \bigcup_{x\in[0,1]\cap\mathbf{Q}}\{x\}$ is not elementary. This means that elementary algebra is a proper subcollection of such atomic algebra.

Jordan algebra is not atomic algebra. Let $(A_\alpha)_{\alpha\in I}$ be atoms of Jordan algebra. Clearly, every Jordan measurable set can be represented as the for $E_n = \bigcup_{\alpha\in J_n} A_\alpha$ where $J_n \subset I$. Then by definition of atomic algebra, $[0,1] \setminus \mathbf{Q} = \bigcup_{n=1}^\infty E_n = \bigcup_{\alpha\in J} A_\alpha$ where $J := \bigcup_{n=1}^\infty J_n \subset I$ is also contained in Jordan algebra, a contradiction.

Lebesgue algebra is not atomic algebra. Notice that all of elementary sets are Lebesgue measurable. Thus we still use singleton $\{x\}_{x\in\mathbf{R}^d}$

as the atoms. There exists $E \subset [0,1]$ is not Lebesgue measurable, but $E = \bigcup_{x \in E} \{x\}$ lies in atomic algebra, a contradiction.

Null algebra is not atomic algebra. Since every singleton set is null set. Then $\{x\}_{x \in \mathbf{R}^d}$ is the atoms with same reason. Then we have $[0,1] = \bigcup_{x \in [0,1]} \{x\}$ lies in atomic algebra, a contradiction. $\square$

Now we describe some further ways to generate Boolean algebras.

Lemma 1.4.11 (Intersection of algebras)**.** *Let $(\mathcal{B}_\alpha)_{\alpha \in I}$ be a family of Boolean algebras on a set X, indexed by a (possibly infinite or uncountable) label set I. Then the intersection $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha := \bigcap_{\alpha \in I} \mathcal{B}_\alpha$ of these algebras is still a Boolean algebra, and is the finest Boolean algebra that is coarser than all $\mathcal{B}_\alpha$. (If I is empty, we adopt the convention that $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha$ is the discrete algebra.)*

Proof. Clearly, we have $\emptyset \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$ for that $\emptyset \in \mathcal{B}_\alpha$ for every $\alpha \in I$. If $E \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$, this means that $E \in \mathcal{B}_\alpha$ for every $\alpha \in I$. Then we have $E^c \in \mathcal{B}_\alpha$ for every $\alpha \in I$, so that $E^c \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$. If $E, F \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$, then $E, F \in \mathcal{B}_\alpha$ for every $\alpha \in I$, this implies $E \cup F \in \mathcal{B}_\alpha$ for every $\alpha \in I$. Thus we have $E \cup F \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$. By definition, $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha$ is a Boolean algebra.

Let $\mathcal{B}$ be an arbitrary Boolean algebra that is coarser than all of $\mathcal{B}_\alpha$, i.e., $\mathcal{B} \subset \mathcal{B}_\alpha$ for all $\alpha \in I$. Then we have $\mathcal{B} \subset \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$, this means that $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha$ is the finest Boolean algebra that is coarser than all of $\mathcal{B}_\alpha$. $\square$

Definition 1.4.12 (Generation of algebras)**.** Let $\mathcal{F}$ be any family of sets in X. We define $\langle \mathcal{F} \rangle_{\text{bool}}$ to be the intersection of all the Boolean algebras that contain $\mathcal{F}$, which is again a Boolean algebra by Lemma 1.4.11. Equivalently, $\langle \mathcal{F} \rangle_{\text{bool}}$ is the coarsest Boolean algebra that contains $\mathcal{F}$. We say that $\langle \mathcal{F} \rangle_{\text{bool}}$ is the Boolean algebra *generated* by $\mathcal{F}$.

Proposition 1.4.13. *The elementary algebra $\overline{\mathcal{E}[\mathbf{R}^d]}$ is generated by the collection of boxes in $\mathbf{R}^d$.*

Proof. Let $\mathcal{B}$ be the collection of boxes in $\mathbf{R}^d$. We need to show that $\langle\mathcal{B}\rangle_{\text{bool}} = \overline{\mathcal{E}[\mathbf{R}^d]}$. Clearly, we have $\langle\mathcal{B}\rangle_{\text{bool}} \subset \overline{\mathcal{E}[\mathbf{R}^d]}$, it is sufficient to show that $\langle\mathcal{B}\rangle_{\text{bool}} \supset \overline{\mathcal{E}[\mathbf{R}^d]}$. For every element $E \in \overline{\mathcal{E}[\mathbf{R}^d]}$, it is either elementary or co-elementary set. If E is elementary set, then there is finitely many boxes $B_1, \cdots, B_k$ such that $E = \bigcup_{n=1}^k B_n$, so that $E \in \langle\mathcal{B}\rangle_{\text{bool}}$. While if E is co-elementary set, then E^c is elementary, and can be represented as the union of finitely many boxes $B_1, \cdots, B_k$. Then $E = (\bigcup_{n=1}^k B_k)^c$. Obviously, we have $E \in \langle\mathcal{B}\rangle_{\text{bool}}$. This complete the proof. □

1.4.2 σ-algebras and measurable spaces

In order to obtain a measure and integration theory that can cope well with limits, the finite union axiom of a Boolean algebra is insufficient, and must be improved to a countable union axiom:

Definition 1.4.14 (σ-algebras)**.** Let X be a set. A *σ-algebra* on X is a collection $\mathcal{B}$ of X which obeys the following properties:

(i) (Empty set) $\emptyset \in \mathcal{B}$.

(ii) (Complement) If $E \in \mathcal{B}$, then the complement $E^c := X \setminus E$ also lies in $\mathcal{B}$.

(iii) (Countable unions) If $E_1, E_2, \cdots \in \mathcal{B}$, then $\bigcup_{n=1}^\infty E_n \in \mathcal{B}$.

We refer to the pair $(X, \mathcal{B})$ of a set X together with a σ-algebra on that set as a *measurable space*.

Remark. The σ-algebra are closed under countable intersections as well as countable unions.

Example 1.4.15. All atomic algebras are σ-algebras. In particular, the discrete algebra and trivial algebra are σ-algebras, as are the finite algebras and the dyadic algebras on Euclidean spaces.

Proof. Since atomic algebras are Boolean algebra, we only need to show that the third condition is hold. Let $E_1, E_2, \cdots \in \mathcal{A}((A_\alpha)_\alpha)$, this means that there is subsets $J_n \subset I$ such that $E_n = \bigcup_{\alpha \in J_n} A_\alpha$ for all $n \geq 1$. Then $\bigcup_{n=1}^\infty E_n = \bigcup_{\alpha \in J} A_\alpha$, where $J := \bigcup_{n=1}^\infty J_n \subset I$, is contained in $\mathcal{A}((A_\alpha)_\alpha)$. Thus all atomic algebras are σ-algebras. $\square$

Example 1.4.16. The Lebesgue and null algebras are σ-algebras, but the elementary and Jordan algebras are not.

Proof. By Lemma 1.2.12, Lebesgue algebras are σ-algebras. From the countable additivity of Lebesgue measure, null algebras are also σ-algebras. For the elementary and Jordan algebras, notice that we can write $[0,1] \cap \mathbf{Q}$ and $[0,1] \setminus \mathbf{Q}$ as the countable union of elementary sets and Jordan measurable sets, respectively. These give the counterexamples, and elementary and Jordan algebras are not σ-algebras. $\square$

Example 1.4.17. Any restriction $\mathcal{B} \downharpoonright_Y$ of a σ-algebra $\mathcal{B}$ to a subspace Y of X (as defined in Example 1.4.6) is again a σ-algebra on the subspace Y.

Proof. We only need to show the third condition. Suppose that $E_1, E_2, \cdots \in \mathcal{B}$, then $E_1 \cap Y, E_2 \cap Y, \cdots \in \mathcal{B} \downharpoonright_Y$. Since $\bigcup_{n=1}^\infty E_n \in \mathcal{B}$, we have $\bigcup_{n=1}^\infty E_n \cap Y \in \mathcal{B} \downharpoonright_Y$. Then $\bigcup_{n=1}^\infty (E_n \cap Y) = \bigcup_{n=1}^\infty E_n \cap Y$ is contained in $\mathcal{B} \downharpoonright_Y$, as desired. $\square$

Proposition 1.4.18 (Intersection of σ-algebras)**.** *The intersection $\bigwedge_{\alpha \in I} \mathcal{B}_\alpha := \bigcap_{\alpha \in I} \mathcal{B}_\alpha$ of an arbitrary (and possibly infinite or uncountable) number of σ-algebras $\mathcal{B}_\alpha$ is again a σ-algebra, and is the finest σ-algebra that is coarser than all of the $\mathcal{B}_\alpha$.*

Proof. Clearly, we have $\emptyset \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$ for that $\emptyset \in \mathcal{B}_\alpha$ for every $\alpha \in I$. If $E \in \bigwedge_{\alpha \in I} \mathcal{B}_\alpha$, this means that $E \in \mathcal{B}_\alpha$ for every $\alpha \in I$. Then we have

$E^c \in \mathcal{B}_\alpha$ for every $\alpha \in I$, so that $E^c \in \bigwedge_{\alpha\in I} \mathcal{B}_\alpha$. If $E_1, E_2, \cdots \in \bigwedge_{\alpha\in I} \mathcal{B}_\alpha$, then $E_1, E_2, \cdots \in \mathcal{B}_\alpha$ for every $\alpha \in I$, this implies $\bigcup_{n=1}^\infty E_n \in \mathcal{B}_\alpha$ for every $\alpha \in I$. Thus we have $\bigcup_{n=1}^\infty E_n \in \bigwedge_{\alpha\in I} \mathcal{B}_\alpha$. By Definition 1.4.14, $\bigwedge_{\alpha\in I} \mathcal{B}_\alpha$ is a σ-algebra.

Let $\mathcal{B}$ be an arbitrary σ-algebra that is coarser than all of $\mathcal{B}_\alpha$, i.e., $\mathcal{B} \subset \mathcal{B}_\alpha$ for all $\alpha \in I$. Then we have $\mathcal{B} \subset \bigwedge_{\alpha\in I} \mathcal{B}_\alpha$, this means that $\bigwedge_{\alpha\in I} \mathcal{B}_\alpha$ is the finest σ-algebra that is coarser than all of $\mathcal{B}_\alpha$. □

Definition 1.4.19 (Generation of σ-algebras)**.** Let $\mathcal{F}$ be any family of sets in X. We define $\langle \mathcal{F} \rangle$ to be the intersection of all the σ-algebras that contain $\mathcal{F}$, which is again a σ-algebra by Proposition 1.4.18. Equivalently, $\langle \mathcal{F} \rangle$ is the coarsest σ-algebra that contains $\mathcal{F}$. We say that $\langle \mathcal{F} \rangle$ is the σ-algebra *generated* by $\mathcal{F}$.

Remark. From the definitions, it is clear that we have the following principle, somewhat analogous to the principle of mathematical induction: if $\mathcal{F}$ is a family of sets in X, and $P(E)$ is a property of sets $E \subset X$ which obeys the following axioms:

(i) $P(\emptyset)$ is true.

(ii) $P(E)$ is true for all $E \in \mathcal{F}$.

(iii) If $P(E)$ is true for some $E \subset X$, then $P(X \setminus E)$ is true also.

(iv) If $E_1, E_2, \cdots \subset X$ are such that $P(E_n)$ is true for all n, then $P(\bigcup_{n=1}^\infty E_n)$ is true also.

Then one can conclude that $P(E)$ is true for all $E \in \langle \mathcal{F} \rangle$. Indeed, the set of all E for which $P(E)$ holds is a σ-algebra that contains $\mathcal{F}$, whence the claim.

Now we turn to an important example of a σ-algebra:

Definition 1.4.20 (Borel σ-algebra)**.** Let X be a metric space, or more generally a topological space. The *Borel σ-algebra* $\mathcal{B}[X]$ of X is defined to be the σ-algebra generated by the open subsets of X. Elements of

$\mathcal{B}[X]$ will be called *Borel measurable.*

We define the Borel σ-algebra to be generated by the open sets. However, they are also generated by several other sets:

Proposition 1.4.21. *The Borel σ-algebra $\mathcal{B}[\mathbf{R}^d]$ of a Euclidean set is generated by any of the following collections of sets:*

(i) *The open subsets of $\mathbf{R}^d$.*

(ii) *The closed subsets of $\mathbf{R}^d$.*

(iii) *The compact subsets of $\mathbf{R}^d$.*

(iv) *The open balls of $\mathbf{R}^d$.*

(v) *The boxes in $\mathbf{R}^d$.*

(vi) *The elementary sets in $\mathbf{R}^d$.*

Proof. We first show that (i) and (ii) generate the same σ-algebra. Let $\mathcal{F}$ be the collection of open subsets of $\mathbf{R}^d$, and let $\mathcal{F}'$ be the collection of closed subsets of $\mathbf{R}^d$. For every closed set $F \in \mathcal{F}'$, we have F^c is open. Then $F^c \in \mathcal{F}$. Then for every σ-algebra $\mathcal{S}$ contains $\mathcal{F}$, we have $F^c \in \mathcal{S}$. Since $\mathcal{S}$ is a σ-algebra, we have $(F^c)^c = F \in \mathcal{S}$. This means that $\mathcal{S}$ also contains $\mathcal{F}'$. A similar argument shows that $\mathcal{F}' \subset \mathcal{S}$ implies $\mathcal{F} \subset \mathcal{S}$. Thus $\mathcal{F}$ and $\mathcal{F}'$ generate the same σ-algebra.

To show that (ii) and (iii) generate the same σ-algebra. Let $\mathcal{F}$ be the collection of closed subsets of $\mathbf{R}^d$, and let $\mathcal{F}'$ be the collection of compact subsets of $\mathbf{R}^d$. This is trivial to see that for every σ-algebra $\mathcal{S}$ such that $\mathcal{F} \subset \mathcal{S}$ implies $\mathcal{F}' \subset \mathcal{S}$. For the converse, since every closed set can be represented as $F = \bigcup_{n=1}^{\infty} F \cap \overline{B(0,n)}$, where $\overline{B(0,n)}$ are closed balls with radius n and centered at origin so that are compact, then we can see that $F \cap \overline{B(0,n)}$ are all compact and lie in $\mathcal{F}' \subset \mathcal{S}$. Since $\mathcal{S}$ is a σ-algebra, we have $F = \bigcup_{n=1}^{\infty} F \cap \overline{B(0,n)} \in \mathcal{S}$. This means that $\mathcal{F} \subset \mathcal{S}$. Thus $\mathcal{F}$ and $\mathcal{F}'$ generate the same σ-algebra.

Then we show that (i) and (iv) generate the same σ-algebra. Let $\mathcal{F}$ be the collection of open subsets of $\mathbf{R}^d$, and let $\mathcal{F}'$ be the collection of open balls in $\mathbf{R}^d$. This is easy to see that for every σ-algebra $\mathcal{S}$ such that $\mathcal{F} \subset \mathcal{S}$ implies $\mathcal{F}' \subset \mathcal{S}$. For the other hand, suppose that $\mathcal{F}' \subset \mathcal{S}$. Since every open set can be represented as the countable union of open balls, then for every open set there are some open balls $(B_\alpha)_{\alpha \in I}$ all lie in $\mathcal{S}$ such that $U = \bigcup_{\alpha \in I} B_\alpha$. Because $\mathcal{S}$ is a σ-algebra, we have $U \in \mathcal{S}$. This means that $\mathcal{F} \subset \mathcal{S}$. Thus $\mathcal{F}$ and $\mathcal{F}'$ generate the same σ-algebra.

To show that (v) and (vi) generate the same σ-algebra. Let $\mathcal{F}$ be the collection of boxes in $\mathbf{R}^d$, and let $\mathcal{F}'$ be the collection of elementary sets in $\mathbf{R}^d$. This is trivial that for every σ-algebra $\mathcal{S}$ contains $\mathcal{F}'$ implies $\mathcal{S}$ contains $\mathcal{F}$. For the other direction, suppose that $\mathcal{F} \subset \mathcal{S}$. Since every elementary set E can be represented as the union of finitely many boxes, thus E lies in $\mathcal{S}$. Thus $\mathcal{F}$ and $\mathcal{F}'$ generate the same σ-algebra.

Finally, we want to show that (i) and (v) generate the same σ-algebra. Let $\mathcal{F}$ be the collection of open subsets of $\mathbf{R}^d$, and let $\mathcal{F}'$ be the collection of boxes in $\mathbf{R}^d$. Since every open set can be represented as the countable union of almost disjoin boxes, for every σ-algebra $\mathcal{S} \supset \mathcal{F}'$, we have $\mathcal{F} \subset \mathcal{S}$. Now we suppose that $\mathcal{S}$ contains all open subsets of $\mathbf{R}^d$, i.e., $\mathcal{F} \subset \mathcal{S}$. We only show that $\mathcal{S}$ contains all of intervals, and this is easy to extend to d-dimension (i.e. boxes) with Cartesian product. Clearly, $\mathcal{S}$ contains the interval with the form (a, b) where $a < b$. We can see that for singleton $\{a\}$ we have

$$\{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a + \frac{1}{n} \right),$$

where $(a - 1/n, a + 1/n)$ are open. Since $\mathcal{S}$ is closed under countable intersections, we have $\{a\} \in \mathcal{S}$, so that $[a, b) = (a, b) \cup \{a\}$ also lies in $\mathcal{S}$. Similarly, we have $(a, b], [a, b] \in \mathcal{S}$. Thus $\mathcal{S}$ contains all of intervals in 1-dimension. Notice that the Cartesian product of open sets still open, we obtain that $\mathcal{S}$ contains all boxes in d-dimensions. This complete the proof. $\square$

Lemma 1.4.22. *Let E, F be Borel measurable subsets of $\mathbf{R}^{d_1}, \mathbf{R}^{d_2}$, respectively. Then $E \times F$ is a Borel measurable subset of $\mathbf{R}^{d_1+d_2}$.*

Proof. Suppose that $F \subset \mathbf{R}^{d_2}$ is a box, and $E \subset \mathbf{R}^{d_1}$ is a Borel measurable set. We first show that $E \times F$ is Borel measurable for every Borel measurable set E. When $E = \emptyset$, we have $E \times F = \emptyset$ is Borel measurable. For the complement property: If there is Borel measurable set E such that $E \times F$ be Borel measurable, then for Borel measurable set $\mathbf{R}^{d_1} \setminus E$ we have

$$(\mathbf{R}^{d_1} \setminus E) \times F = (\mathbf{R}^{d_1} \times F) \setminus (E \times F)$$

is Borel measurable if $\mathbf{R}^{d_1} \times F$ is Borel measurable. Consider the box $[-n, n]^{d_1} \times F$, which is Borel measurable for every $n \geq 1$. Then

$$\mathbf{R}^{d_1} \times F = \bigcup_{n=1}^{\infty} [-n, n]^d \times F = \bigcup_{n=1}^{\infty} ([-n, n]^d \times F)$$

is Borel measurable. For the countable union property: If there is a sequence of Borel measurable sets $E_1, E_2, \cdots$ such that $E_n \times F$ be Borel measurable for every $n \geq 1$. Then

$$\bigcup_{n=1}^{\infty} E_n \times F = \bigcup_{n=1}^{\infty} (E_n \times F)$$

is Borel measurable. Then $E \times F$ is Borel measurable for every Borel measurable set E and box F.

Now we taking the induction on F, and fix E to be a Borel measurable set. When $F = \emptyset$, we have $E \times F = \emptyset$ is Borel measurable. If there is Borel measurable set F such that $E \times F$ be Borel measurable, then for $\mathbf{R}^{d_2} \setminus F$ we have

$$E \times (\mathbf{R}^{d_2} \setminus F) = (E \times \mathbf{R}^{d_2}) \setminus (E \times F)$$

is Borel measurable for that $E \times \mathbf{R}^{d_2}$ is Borel measurable. If there is a sequence of Borel measurable sets $F_1, F_2, \cdots$ such that $E \times F_n$ be Borel measurable for every $n \geq 1$. Then

$$E \times \bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} (E \times F)$$

is Borel measurable. Thus we conclude that for every Borel measurable sets $E \subset \mathbf{R}^{d_1}$ and $F \subset \mathbf{R}^{d_2}$, $E \times F \subset \mathbf{R}^{d_1+d_2}$ is also Borel measurable. $\square$

Lemma 1.4.23. *Let E be a Borel measurable subset of $\mathbf{R}^{d_1+d_2}$. Then for any $x_1 \in \mathbf{R}^{d_1}$, the slice $\{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in E\}$ is a Borel measurable subset of $\mathbf{R}^{d_2}$. Similarly, for every $x_2 \in \mathbf{R}^{d_2}$, the slice $\{x_1 \in \mathbf{R}^{d_1} : (x_1, x_2) \in E\}$ is a Borel measurable subset of $\mathbf{R}^{d_1}$.*

Proof. When $E = \emptyset$, the slice $\{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in E\} = \emptyset$ is Borel measurable. If there is E such that $\{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in E\}$ be Borel measurable, then

$$\{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in \mathbf{R}^{d_1+d_2} \setminus E\} = \mathbf{R}^{d_2} \setminus \{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in E\}$$

is Borel measurable. For Borel measurable sets $E_1, E_2, \cdots$ such that $\{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in E_n\}$ be Borel measurable for every $n \geq 1$, we have

$$\{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in \bigcup_{n=1}^{\infty} E_n\} = \bigcup_{n=1}^{\infty} \{x_2 \in \mathbf{R}^{d_2} : (x_1, x_2) \in E_n\}$$

is Borel measurable. Thus for every Borel measurable set E, the slice respect to x_1 is Borel measurable. A similar argument shows that the slice respect to x_2 is also Borel measurable for every E. $\square$

Proposition 1.4.24. *The Lebesgue σ-algebra on $\mathbf{R}^d$ is generated by the union of the Borel σ-algebra and the null σ-algebra.*

Proof.

1.4.3 Countably additive measures and measure spaces

Having set out the concept of a measurable space, we now endow these structures with a measure.

We begin with the finitely additive theory, although this theory is too weak for our purposes and will soon be supplanted by the countably additive theory.

Definition 1.4.25 (Finitely additive measure)**.** Let $\mathcal{B}$ be a Boolean algebra on a space X. An (unsigned) *finitely additive measure* μ on $\mathcal{B}$ is a map $\mu : \mathcal{B} \to [0, +\infty]$ that obeys the following axioms:

(i) (Empty set) $\mu(\emptyset) = 0$.

(ii) (Finitely additivity) Whenever $E, F \in \mathcal{B}$ are disjoint, then $\mu(E \cup F) = \mu(E) + \mu(F)$.

Example 1.4.26. Lebesgue measure m is a finitely additive measure on the Lebesgue σ-algebra,a nd hence on all sub-algebras (such as the null algebra, the Jordan algebra, or the elementary algebra). In particular, Jordan measure and elementary measure are finitely additive (adopting the convention that co-Jordan measurable sets have infinite Jordan measure, and co-elementary sets have infinite elementary measure).

Example 1.4.27 (Dirac measure)**.** Let $x \in X$ and $\mathcal{B}$ be an arbitrary Boolean algebra on X. Then the *Dirac measure* δ_x at x, defined by setting $\delta_x(E) := 1_E(x)$, is finitely additive.

Proposition 1.4.28. *Let $\mu : \mathcal{B} \to [0, +\infty]$ be a finitely additive measure on a Boolean algebra $\mathcal{B}$. One has following properties:*

(i) *(Monotonicity) If E, F are $\mathcal{B}$-measurable and $E \subset F$, then $\mu(E) \leq \mu(F)$.*

(ii) *(Finite additivity) If k is a natural number, and $E_1, \cdots, E_k$ are $\mathcal{B}$-measurable and disjoint, then $\mu(E_1 \cup \cdots \cup E_k) = \mu(E_1) + \cdots + \mu(E_k)$.*

(iii) *(Finite subadditivity) If k is a natural number, and $E_1, \cdots, E_k$ are $\mathcal{B}$-measurable and disjoint, then $\mu(E_1 \cup \cdots \cup E_k) \leq \mu(E_1) + \cdots + \mu(E_k)$.*

(iv) *(Inclusion-exclusion for two sets) If E, F are $\mathcal{B}$-measurable, then $\mu(E \cup F) + \mu(E \cap F) = \mu(E) + \mu(F)$.*

Proof. (i) Since $F \setminus E, E$ are disjoint, by Definition 1.4.25, we have $\mu(F) = \mu(F \setminus E) + \mu(E) \geq \mu(E)$.

(ii) This is easy to prove with induction.

(iii) We use induction on k. When $k = 2$, we have

$$\begin{aligned}\mu(E_1 \cup E_2) &= \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1) + \mu(E_1 \cap E_2)\\ &\leq \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1) + \mu(E_1 \cap E_2) + \mu(E_2 \cap E_1)\\ &= \mu(E_1) + \mu(E_2).\end{aligned}$$

Suppose inductively that the inequality is hold for $k-1$. Then

$$\begin{aligned}\mu\Big(\bigcup_{n=1}^{k} E_n\Big) &= \mu\Big(\bigcup_{n=1}^{k-1} E_n \cup E_k\Big)\\ &= \mu\Big(\bigcup_{n=1}^{k-1} E_n \setminus E_k\Big) + \mu\Big(E_k \setminus \bigcup_{n=1}^{k-1} E_n\Big) + \mu\Big(\bigcup_{n=1}^{k-1} E_n \cap E_k\Big)\\ &\leq \mu\Big(\bigcup_{n=1}^{k-1} E_n\Big) + \mu(E_k)\\ &\leq \sum_{n=1}^{k-1} \mu(E_n) + \mu(E_k)\\ &= \sum_{n=1}^{k} \mu(E_n).\end{aligned}$$

This close the induction.

(iv) This is easy to see that

$$\begin{aligned}\mu(E \cup F) + \mu(E \cap F) &= \mu(E \setminus F) + \mu(F \setminus E) + 2\mu(E \cap F)\\ &= \mu(E) + \mu(F).\end{aligned}$$

$\square$

Proposition 1.4.29. *Let $\mathcal{B}$ be a finite Boolean algebra, generated by a finite family $A_1, \cdots, A_k$ of non-empty atoms. Then for every finitely additive measure μ on $\mathcal{B}$ there exists $c_1, \cdots, c_k \in [0, +\infty]$ such that*

$$\mu(E) = \sum_{1 \leq j \leq k : A_j \subset E} c_j.$$

Equivalently, if x_j is a point in A_j for each $1 \leq j \leq k$, then

$$\mu = \sum_{j=1}^{k} c_j \delta_{x_j}.$$

Furthermore, the $c_1, \cdots, c_k$ are uniquely determined by μ.

Proof. By definition, for every $E \in \mathcal{B}$ there exists $1 \leq n \leq k$ such that $E = \bigcup_{j=1}^{n} A_j$. Because atoms are disjoint, for every A_j where $n+1 \leq j \leq k$ we must have $A_i \cap E = \emptyset$. Thus we can rewrite the set as $E = \bigcup_{1 \leq j \leq k : A_i \subset E} A_i$. From finitely additivity, we have

$$\mu(E) = \mu\Big(\bigcup_{1 \leq j \leq k : A_j \subset E} A_j \Big) = \sum_{1 \leq j \leq k : A_i \subset E} \mu(A_j).$$

Since $\mu(A_j)$ takes the value in $[0, +\infty]$, this means that for every $1 \leq j \leq k$ there exists $c_j \in [0, +\infty]$ such that $\mu(A_j) = c_j$. Thus

$$\mu(E) = \sum_{1 \leq j \leq k : A_j \subset E} c_j.$$

Suppose that there is $c_1', \cdots, c_k' \in [0, +\infty]$ different from $c_1, \cdots, c_k$. Then from above conclusion, we have $\mu(A_j) = c_j$ and $\mu(A_j) = c_j'$ for every $1 \leq j \leq k$. Then we have $c_j = c_j'$. □

We now specialise to the *countably additive measures* on σ-algebras.

Definition 1.4.30 (Countably additive measure)**.** Let $(X, \mathcal{B})$ be a measurable space. An (unsigned) *countably additive measure* μ on $\mathcal{B}$, or *measure* for short, is a map $\mu : \mathcal{B} \to [0, +\infty]$ that obeys the following axioms:

(i) (Empty set) $\mu(\emptyset) = 0$.

(ii) (Countable additivity) Whenever $E_1, E_2, \dots \in \mathcal{B}$ are a countable sequence of disjoint measurable sets, then $\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$.

A triplet $(X, \mathcal{B}, \mu)$, where $(X, \mathcal{B})$ is a measurable space and $\mu : \mathcal{B} \to [0, +\infty]$ is a countably additive measure, is known as a *measure space*.

Lemma 1.4.31 (Countable combinations of measures)**.** *Let $(X, \mathcal{B})$ be a measurable space.*

(i) *If μ is a countably additive measure on $\mathcal{B}$, and $c \in [0, +\infty]$, then $c\mu$ is also countably additive.*

(ii) *If $\mu_1, \mu_2, \cdots$ are a sequence of countably additive measures on $\mathcal{B}$, then the sum $\sum_{n=1}^{\infty} \mu_n : E \to \sum_{n=1}^{\infty} \mu_n(E)$ is also a countably additive measure.*

Proof. (i) This is easy to see that $c\mu(\emptyset) = 0$, and

$$c\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = c\sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} c\mu(E_n).$$

Thus $c\mu$ is also countably additive.

(ii) For empty set, we have

$$\sum_{n=1}^{\infty} \mu_n(\emptyset) = \sum_{n=1}^{\infty} 0 = 0.$$

For countable additivity, we have

$$\sum_{n=1}^{\infty} \mu_n\Big(\bigcup_{k=1}^{\infty} E_k\Big) = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mu_n(E_k) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} \mu_n(E_k).$$

Thus $\sum_{n=1}^{\infty} \mu_n$ is also a countably additive measure. □

Note that countable additivity measures are necessarily finitely additive (by padding out a finite union into a countable union using the empty set), and so countably additive measures inherit all the properties of finitely additive properties, such as monotonicity and finite subadditivity. But one also has additional properties:

Theorem 1.4.32. *Let $(X, \mathcal{B}, \mu)$ be a measure space.*

(i) *(Countable subadditivity) If $E_1, E_2, \cdots$ are $\mathcal{B}$-measurable, then we have $\mu(\bigcup_{n=1}^{\infty} E_n) \le \sum_{n=1}^{\infty} \mu(E_n)$.*

(ii) *(Upwards monotone convergence) If $E_1 \subset E_2 \subset \cdots$ are $\mathcal{B}$-measurable, then*

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} \mu(E_n) = \sup_n \mu(E_n).$$

(iii) *(Downwards monotone convergence) If $E_1 \supset E_2 \supset \cdots$ are $\mathcal{B}$-measurable, and $\mu(E_n) < \infty$ for at least one n, then*

$$\mu\Big(\bigcap_{n=1}^{\infty} E_n\Big) = \lim_{n\to\infty} \mu(E_n) = \inf_n \mu(E_n).$$

Proof. (i) is immediately comes from Corollary 1.2.19. The proofs of (ii) and (iii) are same with Theorem 1.2.18. □

Theorem 1.4.33 (Dominated convergence for sets)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $E_1, E_2, \cdots$ be a sequence of $\mathcal{B}$-measurable sets that converge to another set E, in the sense that 1_{E_n} converges pointwise to 1_E.*

(i) *E is also $\mathcal{B}$-measurable.*

(ii) *If there exists a $\mathcal{B}$-measurable set F of finite measure (i.e. $\mu(F) < \infty$) that contains all of the E_n, then $\lim_{n\to\infty} \mu(E_n) = \mu(E)$.*

Proof. (i) As we shown in Theorem 1.2.20, we have $E = \bigcup_{N>0} \bigcap_{n\ge N} E_n = \bigcap_{N>0} \bigcup_{n\ge N} E_n$. Since $\mathcal{B}$ is closed under countable union and intersection, we have E is also $\mathcal{B}$-measurable set.

(ii) Since $\bigcup_{n>N}(E_n \triangle E)$ is a decreasing sequence respect to N, then by downwards monotone convergence, we have

$$\mu\Big(\bigcap_{N>0}\bigcup_{n>N} E_n \triangle E\Big) = \lim_{N\to\infty} \mu\Big(\bigcup_{n>N} E_n \triangle E\Big).$$

This is easy to see that we have

$$\bigcup_{n>N} E_n \triangle E = \bigcup_{n>N} E_n \setminus E \cup \bigcup_{n>N} (E \setminus E_n) \supset \bigcup_{n>N} E_n \setminus E \supset E_N \setminus E,$$

and

$$\bigcup_{n>N} E_n \triangle E \supset E \setminus \bigcap_{n>N} E_n \supset E \setminus E_N.$$

Hence

$$\mu\Big(\bigcap_{N>0}\bigcup_{n>N} E_n \triangle E\Big) \geq \lim_{N\to\infty} |\mu(E_N) - \mu(E)|$$

Since

$$\begin{aligned}\bigcap_{N>0}\bigcup_{n>N} E_n \triangle E &= \bigcap_{N>0}\bigcup_{n>N} E_n \setminus E \cup \bigcap_{N>0}\bigcup_{n>N} (E \setminus E_n) \\ &= \bigcap_{N>0}\bigcup_{n>N} E_n \setminus E \cup E \setminus \bigcup_{N>0}\bigcap_{n>N} E_n \\ &= \emptyset,\end{aligned}$$

we have $\lim_{N\to\infty} |\mu(E_N) - \mu(E)| \leq \mu(\emptyset) = 0$. Thus $\lim_{n\to\infty} \mu(E_n) = \mu(E)$ is hold. □

A useful technical property, enjoyed by some measure spaces, is that of *completeness*:

Definition 1.4.34 (Completeness)**.** A *null set* of a measure space $(X, \mathcal{B}, \mu)$ is defined to be a $\mathcal{B}$-measurable set of measure zero. A *sub-null set* is any subset of a null set. A measure space is said to be *complete* if every sub-null set is a null set.

Completion is a convenient property to have in some cases, particularly when dealing with properties that hold almost everywhere. Fortunately, it is fairly easy to modify any measure space to be complete:

Proposition 1.4.35 (Completion)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space. Then there exists a unique refinement $(X, \overline{\mathcal{B}}, \overline{\mu})$, known as the* completion *of $(X, \mathcal{B}, \mu)$, which is the coarsest refinement of $(X, \mathcal{B}, \mu)$ that is complete. Furthermore, $\overline{\mathcal{B}}$ consists precisely of those sets that differ from a $\mathcal{B}$-measurable set by a $\mathcal{B}$-subnull set.*

Proof. Let $E \in \mathcal{B}$, and let $N \subset F$ with $F \in \mathcal{B}$ and $\mu(F) = 0$ (i.e. F is a null set and N is the sub-null set of F). Define $\overline{\mathcal{B}}$ be the collection of $E \triangle N$. We assert that (i) $\overline{\mathcal{B}}$ is exactly a σ-algebra, and (ii) define $\overline{\mu}(E \triangle N) := \mu(E)$, then $\overline{\mu}$ is well-defined, and is a countably additive measure on $\overline{\mathcal{B}}$.

(i) This is trivial that $\emptyset \in \overline{\mathcal{B}}$. From the relation

$$E \triangle N = (E \setminus F) \cup [F \cap (E \triangle N)]$$

and

$$E \cup N = (E \setminus F) \triangle (F \cap (E \cup N))$$

shows that $\overline{\mathcal{B}}$ may also be described as the collection of all sets of the form $E \cup N$. Then

$$(E \cup N)^c = (E \cup F)^c \cup (F \setminus N)$$

shows that $\overline{\mathcal{B}}$ is closed under the complements. And this is easy to see that $\overline{\mathcal{B}}$ is closed under the countable unions. Thus $\overline{\mathcal{B}}$ is a σ-algebra.

(ii) If $E_1 \triangle N_1 = E_2 \triangle N_2$, then $E_1 \triangle E_2 = N_1 \triangle N_2$.[1] Since $N_1 \triangle N_2$ is

[1]Following results are useful: Let E, F, G be arbitrary subsets of X, then

$$\begin{aligned} E \triangle F &= F \triangle E, \\ E \triangle (F \triangle G) &= (E \triangle F) \triangle G, \\ E \triangle \emptyset &= E, \end{aligned}$$

the subset of null set, we have $\mu(E_1 \triangle E_2) = 0$. Then we have

$$\mu(E_1) = \mu(E_1 \setminus E_2) + \mu(E_1 \cap E_2) = \mu(E_2 \setminus E_1) + \mu(E_1 \cap E_2) = \mu(E_2).$$

Thus $\overline{\mu}$ is well-defined, in other words, for every $E_1 \triangle N_1 = E_2 \triangle N_2$ such that $\overline{\mu}(E_1 \triangle N_1) = \overline{\mu}(E_2 \triangle N_2)$, we have $\mu(E_1) = \mu(E_2)$.

Now we show that $\overline{\mu}$ is a countably additive measure. Clearly, we have $\overline{\mu}(\emptyset) = \mu(\emptyset) = 0$. For $E_1 \cup N_1, E_2 \cup N_2, \cdots \in \overline{\mathcal{B}}$, we have

$$\begin{aligned}
\overline{\mu}\Big(\bigcup_{n=1}^{\infty} E_n \cup N_n\Big) &= \overline{\mu}\Big(\bigcup_{n=1}^{\infty} E_n \cup \bigcup_{n=1}^{\infty} N_n\Big) \\
&= \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \\
&= \sum_{n=1}^{\infty} \mu(E_n) \\
&= \sum_{n=1}^{\infty} \overline{\mu}(E_n \cup N_n).
\end{aligned}$$

Thus $\overline{\mu}$ is a countably additive measure on $\overline{\mathcal{B}}$

Finally, we show that $\overline{\mathcal{B}}$ is the coarsest refinement of $\mathcal{B}$. This is easy to see that $\mathcal{B} \subset \overline{\mathcal{B}}$. Let $(X, \mathcal{B}', \mu')$ be another refinement of $(X, \mathcal{B}, \mu)$ which is complete. Since N is a sub-null set, then by definition of completeness, $N \in \mathcal{B}'$, and $E \in \mathcal{B} \subset \mathcal{B}'$. Thus for every $E \triangle N \in \overline{\mathcal{B}}$, we have $E \triangle N \in \mathcal{B}'$. This implies $\overline{\mathcal{B}} \subset \mathcal{B}'$. Thus $\overline{\mathcal{B}}$ is the coarsest refinement of $\mathcal{B}$. $\square$

$$\begin{aligned}
E \triangle X &= E^c, \\
E \triangle E &= \emptyset, \\
E \triangle E^c &= X, \\
E \triangle F &= (E \cup F) \setminus (E \cap F) \\
E \cap (F \triangle G) &= (E \cap F) \triangle (E \cap G).
\end{aligned}$$

Lemma 1.4.36 (Approximation by an algebra)**.** *Let $\mathcal{A}$ be a Boolean algebra on X, and let μ be a measure on $\langle\mathcal{A}\rangle$.*

(i) *If $\mu(X) < \infty$, then for every $E \in \langle\mathcal{A}\rangle$ and $\varepsilon > 0$ there exists $F \in \mathcal{A}$ such that $\mu(E \triangle F) < \varepsilon$.*

(ii) *More generally, if $X = \bigcup_{n=1}^\infty A_n$ for some $A_1, A_2, \cdots \in \mathcal{A}$ with $\mu(A_n) < \infty$ for all n, $E \in \langle\mathcal{A}\rangle$ has finite measure, and $\varepsilon > 0$, then there exists $F \in \mathcal{A}$ such that $\mu(E \triangle F) < \varepsilon$.*

Proof. (i) Let

$$\mathcal{B} := \{E \subset X : \forall \varepsilon > 0, \exists F \in \mathcal{A} \text{ such that } \mu(E \triangle F) < \varepsilon\}.$$

Clearly, $\mathcal{B}$ is finer than $\mathcal{A}$ for that $\mu(F \triangle F) = 0$. If we can show that $\mathcal{B}$ is a σ-algebra, then $\langle\mathcal{A}\rangle \subset \mathcal{B}$ follows that for every element $E \in \langle\mathcal{A}\rangle$ there exists $F \in \mathcal{A}$ such that $\mu(E \triangle F) < \varepsilon$.

This is trivial to see that $\emptyset \in \mathcal{B}$. To show that $\mathcal{B}$ is closed under the complements: Let $E \in \mathcal{B}$, and $\varepsilon > 0$. Then there is $F \in \mathcal{A}$ such that $\mu(E \triangle F) < \varepsilon$. Since $\mathcal{A}$ is a Boolean algebra, we have $F^c \in \mathcal{A}$. Then from

$$E^c \triangle F^c = (E \triangle X) \triangle (F \triangle X) = E \triangle F \triangle \emptyset = E \triangle F,$$

we have $\mu(E^c \triangle F^c) \leq \varepsilon$. Thus $E^c \in \mathcal{B}$. To show that $\mathcal{B}$ is closed under the countable unions: Let $E_1, E_2, \cdots \in \mathcal{B}$, then there are $F_1, F_2, \cdots \in \mathcal{A}$ such that $\mu(E_n \triangle F_n) \leq \varepsilon/2^n$ for every n. Then from

$$\Big(\bigcup_{n=1}^\infty E_n\Big) \triangle \Big(\bigcup_{n=1}^\infty F_n\Big) \subset \bigcup_{n=1}^\infty E_n \triangle F_n$$

we have $\mu((\bigcup_{n=1}^\infty E_n) \triangle (\bigcup_{n=1}^\infty F_n)) \leq \varepsilon$, so that $\bigcup_{n=1}^\infty E_n \in \mathcal{B}$. Thus $\mathcal{B}$ is a σ-algebra.

(ii) Notice that the restriction $\mathcal{A} \downharpoonleft_{A_n}$ of $\mathcal{A}$ to A_n is a Boolean algebra on A_n. Since $\mu(A_n) < \infty$, use the conclusion of (i), for every $E \cap A_n \in \langle \mathcal{A} \downharpoonleft_{A_n} \rangle$ there exists $F_n \in \mathcal{A} \downharpoonleft_{A_n}$ such that $\mu(E \cap A_n \triangle F_n) \leq \varepsilon/2^n$. Let $F := \bigcup_{n=1}^\infty F_n$, we have $\mu(E \triangle F) \leq \varepsilon$. $\square$

1.4.4 Measurable functions, and integration on a measure space

Now we are ready to define integration on measure spaces. We first need the notion of a *measurable function*, which is analogous to that of a continuous function in topology. Recall that a function $f : X \to Y$ between two topological spaces X, Y is continuous if the inverse image $f^{-1}(U)$ of any open set is open. In a similar spirit, we have

Definition 1.4.37 (Measurable functions)**.** Let $(X, \mathcal{B})$ be a measurable space, and let $f : X \to [0, +\infty]$ or $f : X \to \mathbf{C}$ be an unsigned or complex-valued function. We say that f is *measurable* if $f^{-1}(U)$ is $\mathcal{B}$-measurable for every open subset U of $[0, +\infty]$ or $\mathbf{C}$.

From Lemma 1.3.9, we see that this generalises the notion of a Lebesgue measurable function.

Proposition 1.4.38. *Let $(X, \mathcal{B})$ be a measurable space.*

(i) *A function $f : X \to [0, +\infty]$ is measurable if and only if the level sets $\{x \in X : f(x) > \lambda\}$ are $\mathcal{B}$-measurable.*

(ii) *An indicator function 1_E of a set $E \subset X$ is measurable if and only if E itself is $\mathcal{B}$-measurable.*

(iii) *A function $f : X \to [0, +\infty]$ or $f : X \to \mathbf{C}$ is measurable if and only if $f^{-1}(E)$ is $\mathcal{B}$-measurable for every Borel-measurable subset E of $[0, +\infty]$ or $\mathbf{C}$.*

(iv) *A function $f : X \to \mathbf{C}$ is measurable if and only if its real and imaginary parts are measurable.*

(v) *A function $f : X \to \mathbf{R}$ is measurable if and only if the magnitudes $f_+ := \max(f, 0), f_- := \max(-f, 0)$ of its positive and negative parts are measurable.*

(vi) *If $f_n : X \to [0, +\infty]$ or $f_n : X \to \mathbf{C}$ are a sequence of measurable functions that converge pointwise to a limit $f :$*

$X \to [0,+\infty]$ or $f : X \to \mathbf{C}$, then f is also measurable.

(vii) *If $f : X \to [0,+\infty]$ or $f : X \to \mathbf{C}$ is measurable and $\phi : [0,+\infty] \to [0,+\infty]$ or $\phi : \mathbf{C} \to \mathbf{C}$ is continuous, then $\phi \circ f$ is measurable.*

(viii) *The sum or product of two measurable functions in $[0,+\infty]$ or $\mathbf{C}$ is still measurable.*

Proof. (i) If f is measurable, for every open subset $U \subset [0,+\infty]$ we have $f^{-1}(U) \in \mathcal{B}$. Since $(\lambda,+\infty]$ are open in $[0,+\infty]$ for every $\lambda \in [0,+\infty]$, we have

$$\{x \in X : f(x) > \lambda\} = f^{-1}((\lambda,+\infty])$$

are $\mathcal{B}$-measurable.

For the converse, we first show that $f^{-1}((a,b))$ is $\mathcal{B}$-measurable, where $a,b \in [0,+\infty]$ and $a < b$. From

$$\begin{aligned} f^{-1}((a,b)) &= f^{-1}((a,+\infty] \setminus [b,+\infty]) \\ &= f^{-1}((a,+\infty]) \setminus f^{-1}([b,+\infty]) \\ &= f^{-1}((a,+\infty]) \setminus \bigcap_{b' \in \mathbf{Q}\cap[0,+\infty]:b'<b} f^{-1}((b',+\infty])) \end{aligned}$$

for $b \in (0,+\infty]$, we have $f^{-1}((a,b))$ is also $\mathcal{B}$-measurable. Obviously, $f^{-1}([0,+\infty]) = X$ is $\mathcal{B}$-measurable, then we have

$$\begin{aligned} f^{-1}([0,a)) &= f^{-1}([0,+\infty] \setminus [a,+\infty]) \\ &= X \setminus f^{-1}([a,+\infty]) \\ &= X \setminus \bigcap_{a' \in \mathbf{Q}\cap[0,+\infty]:a'<a} f^{-1}((a,+\infty]) \end{aligned}$$

for $a \in (0,+\infty]$ is also $\mathcal{B}$-measurable. Since every open subset $U \subset [0,+\infty]$ can be represented as the countable union of open intervals, thus $f^{-1}(U)$ is $\mathcal{B}$-measurable.

(ii) If 1_E is measurable, for every open subset $U \subset [0,+\infty]$, we have $1_E^{-1}(U) = E, X \setminus E, X$ or $\emptyset$ are all $\mathcal{B}$-measurable. Thus E is $\mathcal{B}$-measurable.

Conversely, if E is $\mathcal{B}$-measurable, it follows that $X \setminus E$ is also $\mathcal{B}$-measurable. Thus for every open subset $U \subset [0, +\infty]$, $1_E^{-1}(U)$ are $\mathcal{B}$-measurable.

(iii) Let f be measurable. We see that $f^{-1}(\emptyset)$ is $\mathcal{B}$-measurable, and $f^{-1}(E)$ is $\mathcal{B}$-measurable for every open subset E of $[0, +\infty]$ or $\mathbf{C}$. If E is open in $[0, +\infty]$ or $\mathbf{C}$, then $f^{-1}([0, +\infty] \setminus E) = X \setminus f^{-1}(E)$ or $f^{-1}(\mathbf{C} \setminus E) = X \setminus f^{-1}(E)$ is also $\mathcal{B}$-measurable. If $E_1, E_2, \cdots$ are open in $[0, +\infty]$ or $\mathbf{C}$, then $f^{-1}(\bigcup_{n=1}^{\infty} E_n) = \bigcup_{n=1}^{\infty} f^{-1}(E_n)$ is also $\mathcal{B}$-measurable. Thus is $\mathcal{B}$-measurable for every Borel-measurable subset E of $[0, +\infty]$ or $\mathbf{C}$. This proves the "only if" part; the "if" part is trivial.

(v) Suppose that f is measurable. Use (i), consider the level set, for every $\lambda \in \mathbf{R}$, we have

$$\{x \in X : f_+(x) > \lambda\} = \{x \in X : \max(f(x), 0) > \lambda\}.$$

We see that such a set equals to $\{x \in X : f(x) > \lambda\}$, X or $\emptyset$, which are all measurable. Thus f_+ is measurable. This is similar to show that f_- is also measurable.

Conversely, since f_+ and f_- are measurable, for every open subset $U \subset \mathbf{R}$ we have

$$f^{-1}(U) = (f_+)^{-1}([0, +\infty) \cap U) \cup (f_-)^{-1}((-\infty, 0] \cap U)$$

is also measurable. Thus f is measurable.

Now we show (vii) and introduce a lemma, and use these to prove (iv).

(vii) By definition, for every open subset of $[0, +\infty]$ or $\mathbf{C}$, $\phi^{-1}(U)$ is also open. Then $(\phi \circ f)^{-1}(U) = f^{-1}(\phi^{-1}(U))$ is open in X for that f is measurable. Thus $\phi \circ f$ is measurable.

Lemma. If f and g are measurable function from X to $\mathbf{R}$, then direct sum $f \oplus g := (f, g)$ is also measurable.

Proof of Lemma. Since f and g are measurable, for every open subset $U, V \subset X$ we have $f^{-1}(U)$ and $f^{-1}(V)$ are $\mathcal{B}$-measurable. Then $(f \oplus g)^{-1}(U \times V) = f^{-1}(U) \cap g^{-1}(V)$ is also $\mathcal{B}$-measurable. Thus $f \oplus g$ is measurable.

(iv) Suppose that f is measurable. Since the projections π_1 and π_2 are continuous which defined as $\pi_1(x, y) := x$ and $\pi_2(x, y) := y$. We have

$\operatorname{Re} f = \pi_1 \circ f$ and $\operatorname{Im} f = \pi_2 \circ f$ are measurable from (vii). For the other hand, if $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable, we have $f = \operatorname{Re} f \oplus \operatorname{Im} f$ is also measurable.

(vi) We first consider the unsigned case. Since f_n converges pointwise to f, we have $f = \limsup_{n\to\infty} f_n$. From (i), f_n are measurable follows that

$$\{x \in X : f_n(x) > \lambda\}$$

are $\mathcal{B}$-measurable. Then

$$\begin{aligned}\{x \in X : f(x) > \lambda\} &= \{x \in X : \inf_{N>0} \sup_{n\geq N} f_n(x) > \lambda\} \\ &= \bigcap_{N>0} \bigcup_{n\geq N} \{x \in X : f_n(x) > \lambda\}\end{aligned}$$

is also $\mathcal{B}$-measurable. This implies that f is measurable. Similarly, we can show that $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable, then use (iv), f is measurable on $\mathbf{C}$. □

Definition 1.4.39 (μ-almost everywhere)**.** A property $P(x)$ of an element $x \in X$ of a measure space $(X, \mathcal{B}, \mu)$ holds *μ-almost everywhere* if it holds outside of a sub-null set.

Theorem 1.4.40 (Egorov's theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a finite measure space (so $\mu(X) < \infty$), and let $f_n : X \to \mathbf{C}$ be a sequence of measurable functions that converge pointwise almost everywhere to a limit $f : X \to \mathbf{C}$, and let $\varepsilon > 0$. Then there exists a measurable set E of measure at most ε such that f_n converges uniformly to f outside of E.*

Proof. The definition of pointwise convergence implies that

$$\bigcup_{N>0} \bigcap_{n\geq N} \{x \in X : |f_n(x) - f(x)| \leq \frac{1}{m}\} = X.$$

Let

$$E_{N,m} := \bigcap_{n\geq N} \{x \in X : |f_n(x) - f(x)| \leq \frac{1}{m}\}.$$

Clearly, we have $E_{1,m} \subset E_{2,m} \subset \cdots$ is increasing, and

$$\bigcup_{N>0} E_{N,m} = X.$$

By monotone convergence theorem, we have

$$\lim_{N\to\infty} \mu(E_{N,m}) = \mu(X),$$

it follows that there exists an N_m such that

$$\mu(X) - \mu(E_{N_m,m}) \le \frac{\varepsilon}{2^m}.$$

Let $E := \bigcap_{m>0} E_{N_m,m}$. Then

$$\begin{aligned}
\mu(X \setminus E) &= \mu\Big(X \setminus \bigcap_{m>0} E_{N_m,m}\Big) \\
&= \mu\Big(\bigcup_{m>0} X \setminus E_{N_m,m}\Big) \\
&= \sum_{m>0} \mu(X \setminus E_{E_m,m}) \\
&\le \varepsilon.
\end{aligned}$$

To show that f_n converges uniformly to f on E. Let $\delta > 0$. We can find an enough large m such that $\frac{1}{m} < \delta$. Then $E \subset E_{N_m,m}$, this means that

$$|f_n(x) - f(x)| \le \frac{1}{m} < \delta$$

for every $n \ge N_m$ and $x \in E$. Thus f_n converges uniformly to f on E. □

In Section 1.3 we defined first a simple integral, then an unsigned integral, and then finally an absolutely convergent integral. We perform the same three stages here. We begin with the simple integral:

Definition 1.4.41 (Integral of simple functions)**.** An (unsigned) *simple function* $f : X \to [0, +\infty]$ on a measure space $(X, \mathcal{B}, \mu)$ is a measurable function that takes on finitely many values $a_1, \cdots, a_k$. We then define the simple integral $\mathrm{Simp}\int_X f d\mu$ by the formula

$$\mathrm{Simp}\int_X f d\mu := \sum_{j=1}^{k} a_j \mu(f^{-1}(\{a_j\})).$$

In order to set out the basic properties of the simple integral, the following preliminary result is handy:

Lemma 1.4.42. *Let $f : X \to [0, +\infty]$ be a simple function on a measure space $(X, \mathcal{B}, \mu)$, and suppose that there are disjoint measurable sets $E_1, \cdots, E_m$ such that f is supported on $E_1 \cup \cdots \cup E_m$ and equals c_i on each E_i for some $c_i \in [0, +\infty]$. Then*

$$\mathrm{Simp}\int_X f d\mu = \sum_{j=1}^{m} c_j \mu(E_j).$$

Proof. Proof omitted. □

Proposition 1.4.43 (Basic properties of the simple integral)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to [0, +\infty]$ be simple functions.*

(i) *(Monotonicity) If $f \le g$ pointwise, then*

$$\mathrm{Simp}\int_X f d\mu \le \mathrm{Simp}\int_X g d\mu.$$

(ii) *(Compatibility with measure) For every $\mathcal{B}$-measurable set E, we have $\mathrm{Simp}\int_X 1_E d\mu = \mu(E)$.*

(iii) *(Homogeneity) For every $c \in [0, +\infty]$, on has $\mathrm{Simp}\int_X cf d\mu = c \times \mathrm{Simp}\int_X f d\mu$.*

(iv) *(Finite additivity)*

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

(v) *(Insensitivity to refinement) If* $(X, \mathcal{B}', \mu')$ *is a refinement of* $(X, \mathcal{B}, \mu)$*, then we have* $\operatorname{Simp} \int_X f d\mu = \operatorname{Simp} \int_X f d\mu'$*.*

(vi) *(Almost everywhere equivalent) If* $f(x) = g(x)$ *for* μ*-almost every* $x \in X$*, then* $\operatorname{Simp} \int_X f d\mu = \operatorname{Simp} \int_X g d\mu$*.*

(vii) *(Finiteness)* $\operatorname{Simp} \int_X f d\mu < \infty$ *if and only if* f *is finite almost everywhere, and is supported on a set of finite measure.*

(viii) *(Vanishing)* $\operatorname{Simp} \int_X f d\mu = 0$ *if and only if* f *is zero alost everywhere.*

Proof. Suppose that f takes on finitely many values $a_1, \cdots, a_k$, and g takes on finitely many values $b_1, \cdots, b_{k'}$. Since $[0, +\infty] = \bigcup_{i=1}^k f^{-1}(\{a_i\}) = \bigcup_{j=1}^{k'} g^{-1}(\{b_j\})$, we have disjoint sets $f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})$ such that

$$f^{-1}(\{a_i\}) = \bigcup_{j=1}^{k'} f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})$$

and

$$g^{-1}(\{b_j\}) = \bigcup_{i=1}^{k} f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\}).$$

(i) If $f \leq g$, we have $a_i = a_{ij}$ and $b_j = b_{ij}$ on every $f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})$ such that $a_{ij} \leq b_{ij}$. Then by Lemma 1.4.42,

$$\begin{aligned}
\operatorname{Simp} \int_X f d\mu &= \sum_{i=1}^k a_i \mu(f^{-1}(\{a_i\})) \\
&= \sum_{i=1}^k a_i \mu\Big(\bigcup_{j=1}^{k'} f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})\Big) \\
&= \sum_{i=1}^k \sum_{j=1}^{k'} a_{ij} \mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})) \\
&\leq \sum_{j=1}^{k'} \sum_{i=1}^{k} b_{ij} \mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\}))
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{k'} b_i \mu\Big(\bigcup_{i=1}^{k} f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})\Big) \\
&= \sum_{j=1}^{k'} b_i \mu(g^{-1}(\{b_j\})) \\
&= \operatorname{Simp} \int_X g d\mu.
\end{aligned}
$$

(ii) From Lemma 1.4.42, we have

$$
\begin{aligned}
\operatorname{Simp} \int_X 1_E d\mu &= 1 \times \mu(1_E^{-1}(\{1\})) + 0 \times \mu(1_E^{-1}(\{0\})) \\
&= \mu(1_E^{-1}(\{1\})) \\
&= \mu(E).
\end{aligned}
$$

(iii) Notice that $(cf)^{-1}(\{ca_i\}) = f^{-1}(\{a_i\})$, then

$$
\begin{aligned}
\operatorname{Simp} \int_X cf d\mu &= \sum_{i=1}^{k} ca_i \mu((cf)^{-1}(\{a_i\})) \\
&= c \times \sum_{i=1}^{k} a_i \mu(f^{-1}(\{a_i\})) \\
&= c \times \operatorname{Simp} \int_X f d\mu.
\end{aligned}
$$

(iv) Since $(f+g)^{-1}(\{a_i + b_j\}) = f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})$, we have

$$
\begin{aligned}
\operatorname{Simp} \int_X f d\mu + \operatorname{Simp} \int_X g d\mu &= \sum_{i=1}^{k} \sum_{j=1}^{k'} a_i \mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})) \\
&\quad + \sum_{j=1}^{k'} \sum_{i=1}^{k} b_j \mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k'} (a_i + b_j) \mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k'} (a_i + b_j) \mu((f+g)^{-1}(\{a_i + b_j\}))
\end{aligned}
$$

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

(v) We only need to show that $\mu(f^{-1}(\{a_i\})) = \mu'(f^{-1}(\{a_i\}))$ for every $1 \leq i \leq k$. Let $Y := \bigcup_{i=1}^k f^{-1}(\{a_i\})$. Since every $f^{-1}(\{a_i\})$ is contained in $\mathcal{B}$ and $\mathcal{B}'$ at the time, this is easy to see that $\mathcal{B}\downharpoonright_Y = \mathcal{B}'\downharpoonright_Y$. Then we have

$$\mu(f^{-1}(\{a_i\})) = \mu\downharpoonright_Y (f^{-1}(\{a_i\})) = \mu'\downharpoonright_Y (f^{-1}(\{a_i\})) = \mu'(f^{-1}(\{a_i\}))$$

for every $1 \leq i \leq k$.

(vi) If f, g taking the infinite value in non null set, the assertion is trivial. Suppose that f, g are finite. If $f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})$ is non sub-null set, then $a_i = b_j$, so that

$$a_i\mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})) = b_j\mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})).$$

While if $f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})$ is a sub-null set with $a_i \neq b_j$, then we have

$$a_i\mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})) = b_j\mu(f^{-1}(\{a_i\}) \cap g^{-1}(\{b_j\})) = 0.$$

Thus for arbitrary i, j, we have $\mathrm{Simp}\int_X f d\mu = \mathrm{Simp}\int_X g d\mu$.

The proofs of (vii) and (viii) are same with Lemma 1.3.5. □

Proposition 1.4.44 (Inclusion-exclusion principle)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $A_1, \cdots, A_n$ be $\mathcal{B}$-measurable sets of finite measure. Then*

$$\mu\Big(\bigcup_{i=1}^n A_i\Big) = \sum_{J\subset\{1,\cdots,n\}: J\neq\emptyset} (-1)^{|J|-1}\mu\Big(\bigcap_{i\in J} A_i\Big).$$

Proof. Notice that

$$1_{\bigcup_{i=1}^n A_i} = 1 - \prod_{i=1}^n (1 - 1_{A_i}).$$

Then consider the simple integral $\mathrm{Simp}\int_X (1 - \prod_{i=1}^n (1 - 1_{A_i}))d\mu$, we have

$$\mathrm{Simp}\int_X \Big(1 - \prod_{i=1}^n (1 - 1_{A_i})\Big) d\mu = \mathrm{Simp}\int_X \bigcup_{i=1}^n A_i d\mu = \mu\Big(\bigcup_{i=1}^n A_i\Big).$$

For the remaining parts, we use induction to show that

$$1-\prod_{i=1}^{n}(1-1_{A_i})=\sum_{J\subset\{1,\cdots,n\}:J\neq\emptyset}(-1)^{|J|-1}\prod_{i\in J}1_{A_i}$$

When $n=2$, we have

$$1-(1-1_{A_1})(1-1_{A_2})=1_{A_1}+1_{A_2}-1_{A_1\cap A_2}.$$

Now we suppose inductively that

$$1-\prod_{i=1}^{n-1}(1-1_{A_i})=\sum_{J\subset\{1,\cdots,n-1\}:J\neq\emptyset}(-1)^{|J|-1}\prod_{i\in J}1_{A_i}.$$

Then

$$\begin{aligned}1-\prod_{i=1}^{n}(1-1_{A_i})&=1-(1-1_{A_n})\prod_{i=1}^{n-1}(1-1_{A_i})\\&=\Big(1-\prod_{i=1}^{n-1}(1-1_{A_i})\Big)-1_{A_n}\Big(1-\prod_{i=1}^{n-1}(1-1_{A_i})\Big)+1_{A_n}\\&=(1-1_{A_n})\Big(1-\prod_{i=1}^{n-1}(1-1_{A_i})\Big)+1_{A_n}\\&=(1-1_{A_n})\Big(\sum_{J\subset\{1,\cdots,n-1\}:J\neq\emptyset}(-1)^{|J|-1}\prod_{i\in J}1_{A_i}\Big)+1_{A_n}\\&=\sum_{J\subset\{1,\cdots,n-1\}:J\neq\emptyset}(-1)^{|J|-1}\prod_{i\in J}1_{A_i}+1_{A_n}\\&\quad-\sum_{J\subset\{1,\cdots,n-1\}:J\neq\emptyset}(-1)^{|J|-1}\prod_{i\in J\cup\{n\}}1_{A_i}\\&=\sum_{J\subset\{1,\cdots,n\}:J\neq\emptyset}(-1)^{|J|-1}\prod_{i\in J}1_{A_i}.\end{aligned}$$

This closed the induction. Then we have

$$\mathrm{Simp}\int_X\Big(1-\prod_{i=1}^{n}(1-1_{A_i})\Big)d\mu$$

$$= \operatorname{Simp} \int_X \Big(\sum_{J \subset \{1,\cdots,n\}: J \neq \emptyset} (-1)^{|J|-1} \prod_{i \in J} 1_{A_i} \Big) d\mu$$
$$= \operatorname{Simp} \int_X \Big(\sum_{J \subset \{1,\cdots,n\}: J \neq \emptyset} (-1)^{|J|-1} 1_{\bigcap_{i \in J} A_i} \Big) d\mu$$
$$= \sum_{J \subset \{1,\cdots,n\}: J \neq \emptyset} (-1)^{|J|-1} \mu\Big(\bigcap_{i \in J} A_i \Big),$$

as desired. □

From the simple integral, we can now define the unsigned integral, in analogy to the way the unsigned Lebesgue integral was constructed in Section 1.3.3.

Definition 1.4.45. Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f : X \to [0, +\infty]$ be measurable. Then we define the unsigned integral $\int_X f d\mu$ of f by the formula

$$\int_X f d\mu := \sup_{0 \leq g \leq f : g \text{ simple}} \operatorname{Simp} \int_X g d\mu.$$

Clearly, this definition generalises Definition 1.3.20. Indeed, if $f : \mathbf{R}^d \to [0, +\infty]$ is Lebesgue measurable, then $\int_{\mathbf{R}^d} f(x) dx = \int_{\mathbf{R}^d} f dm$.

We record some easy properties of this integral:

Lemma 1.4.46 (Easy properties of the unsigned integral)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f, g : X \to [0, +\infty]$ be measurable.*

(i) *(Almost everywhere equivalence) If $f = g$ μ-almost everywhere, then $\int_X f d\mu = \int_X g d\mu$.*

(ii) *(Monotonicity) If $f \leq g$ μ-almost everywhere, then $\int_X f d\mu \leq \int_X g d\mu$.*

(iii) *(Homogeneity) We have $\int_X c f d\mu = c \int_X f d\mu$ for every $c \in [0, +\infty]$.*

(iv) *(Superadditivity) We have $\int_X (f + g) d\mu \geq \int_X f d\mu + \int_X g d\mu$.*

(v) *(Compatibility with the simple integral) If f is simple, then we have $\int_X f d\mu = \mathrm{Simp} \int_X f d\mu$.*

(vi) *(Markov's inequality) For any $0 < \lambda < \infty$, one has*

$$\mu(\{x \in X : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int_X f d\mu.$$

In particular, if $\int_X f d\mu < \infty$, then the sets $\{x \in X : f(x) \geq \lambda\}$ have finite measure for each $\lambda > 0$.

(vii) *(Finiteness) If $\int_X f d\mu < \infty$, then $f(x)$ is finite for μ-almost every x.*

(viii) *(Vanishing) If $\int_X f d\mu = 0$, then $f(x)$ is zero for μ-almost every x.*

(ix) *(Horizontal truncation) We have $\lim_{n\to\infty} \int_X \min(f, n) d\mu = \int_X f d\mu$.*

(x) *(Vertical truncation) If $E_1 \subset E_2 \subset \cdots$ is an increasing sequence of $\mathcal{B}$-measurable sets, then*

$$\lim_{n\to\infty} \int_X f 1_{E_n} d\mu = \int_X f 1_{\bigcup_{n=1}^{\infty} E_n} d\mu.$$

(xi) *(Restriction) If Y is a measurable subset of X, then*

$$\int_X f 1_Y d\mu = \int_Y f \downharpoonright_Y d\mu \downharpoonright_Y,$$

where $f \downharpoonright_Y : Y \to [0, +\infty]$ is the restriction of $f : X \to [0, +\infty]$ to Y, and the restriction $\mu \downharpoonright_Y$ was defined in Example 1.4.25. We will often abbreviate $\int_Y f \downharpoonright_Y d\mu \downharpoonright_Y$ (by slight abuse of notation) as $\int_Y f d\mu$.

Proof. (i) - (v) are similar to Proposition 1.3.19

1.4.5 The convergence theorems

Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_1, f_2, \cdots : X \to [0, +\infty]$ be a sequence of countable functions. Suppose that as $n \to \infty$, $f_n(x)$ converges pointwise either everywhere, or μ-almost everywhere, to a measurable limit f. A basic question in the subject is to determine the conditions under which such pointwise convergence would imply convergence of the integral:

$$\int_X f_n d\mu \xrightarrow{?} \int_X f d\mu.$$

To put it another way: When can we ensure that one can interchange integrals and limits,

$$\lim_{n\to\infty}\int_X f_n d\mu \stackrel{?}{=} \int_X \lim_{n\to\infty} f_n d\mu?$$

There are certainly some cases in which one can safely do this:

Proposition 1.4.47 (Uniform convergence on a finite measure space)**.** *Suppose that $(X, \mathcal{B}, \mu)$ is a finite measure space (so $\mu(X) < \infty$), and $f_n : X \to [0, +\infty]$ (resp. $f_n : X \to \mathbf{C}$) are a sequence of unsigned measurable functions (resp. absolutely integrable functions) that converge uniformly to a limit f. Then*

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

Proof. Since f_n converge uniformly to f, for every x and $\varepsilon > 0$ there is $N > 0$ such that $f - f_n \le \varepsilon/\mu(X)$ for every $n \ge N$. Then for every $n \ge N$ we have

$$\left|\int_X f d\mu - \int_X f_n d\mu\right| = \left|\int_X (f - f_n) d\mu\right| \le \left|\int_X \frac{\varepsilon}{\mu(X)} d\mu\right| \le \varepsilon.$$

Thus

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

This conclusion is easy to extend to real-valued function f (which equals to $f_+ - f_-$), and then complex-valued function. $\square$

However, there are also cases in which one cannot interchange limits and integrals, even when the f_n are unsigned. We give the three classic examples, all of "moving bump" type, though the way in which the bump moves varies from example to example:

Example 1.4.48 (Escape to horizontal infinity)**.** Let X be the real line with Lebesgue measure, and let $f_n := 1_{[n,n+1]}$. Then f_n converges pointwise to $f := 0$, but $\int_{\mathbf{R}} f_n(x)dx = 1$ does not converge to $\int_{\mathbf{R}} f(x)dx = 0$. Somehow, all the mass in the f_n has escaped by moving off to infinity in a horizontal direction, leaving none behind for the pointwise limit f.

Example 1.4.49 (Escape to width infinity)**.** Let X be the real line with Lebesgue measure, and let $f_n := \frac{1}{n}1_{[0,n]}$. Then f_n converges *uniformly* to $f := 0$, but $\int_{\mathbf{R}} f_n(x)dx = 1$ still does not converge to $\int_{\mathbf{R}} f(x)dx = 0$. Proposition 1.4.47 would prevent this from happening if all the f_n where supported in a single set of finite measure, but the support of the f_n are becoming increasingly wide, and so Proposition 1.4.47 does not apply.

Example 1.4.50 (Escape to vertical infinity)**.** Let X be the unite interval $[0,1]$ with Lebesgue measure (restricted from $\mathbf{R}$), and let $f_n := n1_{[\frac{1}{n},\frac{2}{n}]}$. Now, we have finite measure, and f_n converges pointwise to f, but no uniform convergence. And again, $\int_{\mathbf{R}} f_n(x)dx = 1$ is not converging to $\int_{\mathbf{R}} f(x)dx = 0$. This time, the mass has escaped vertically, through the increasingly large values of f_n.

However, once one shuts down these avenues of escape to infinity, it turns out that one can recover convergence of the integral. There are two major ways to accomplish this. One is to enforce *monotonicity*, which prevents each f_n from abandoning the location where the mass of the preceding $f_1, \cdots, f_{n-1}$ was concentrated and which thus shuts down the above three escape scenarios. More precisely, we have the *monotone convergence theo-*

rem:

Theorem 1.4.51 (Monotone convergence theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $0 \le f_1 \le f_2 \le \cdots$ be a monotone non-decreasing sequence of unsigned measurable functions on X. Then we have*

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

Proof. Proof omitted. □

This has a number of important corollaries. First, we can generalise (part of) Tonelli's theorem for exchanging sums:

Corollary 1.4.52 (Tonelli's thm for sums and integrals)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_1, f_2, \cdots, : X \to [0, +\infty]$ be a sequence of unsigned measurable functions. Then one has*

$$\int_X \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu.$$

Proof. Proof omitted. □

Lemma 1.4.53 (Borel-Cantelli lemma)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $E_1, E_2, E_3, \cdots$ be a sequence of $\mathcal{B}$-measurable sets such that*

$$\sum_{n=1}^{\infty} \mu(E_n) < \infty.$$

Then almost every $x \in X$ is contained in at most finitely many of the E_n. In other words, $\{n \in \mathbf{N} : x \in E_n\}$ is finite for almost every $x \in X$.

Proof. By Corollary 1.4.52, for the sequence of unsigned measurable functions $1_{E_1}, 1_{E_2}, 1_{E_3}, \cdots$, we have

$$\sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} \int_X 1_{E_n} d\mu = \int_X \sum_{n=1}^{\infty} 1_{E_n} d\mu < \infty.$$

It follows that $\sum_{n=1}^{\infty} 1_{E_n} d\mu < \infty$, so that almost every $x \in X$ is contained in at most finitely many of E_n. $\square$

Second, when one does not have monotonicity, one can at least obtain an important inequality, known as *Fatou's lemma*:

Corollary 1.4.54 (Fatou's lemma)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_1, f_2, \cdots : X \to [0, +\infty]$ be a sequence of unsigned measurable functions. Then*

$$\int_X \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

Proof. Proof omitted. $\square$

Finally, we give the other major way to shut down loss of mass via escape to infinity, which is to *dominate* all of the functions involved by an absolutely convergent one. This result is known as the *dominated convergence theorem*:

Theorem 1.4.55 (Dominated convergence theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_1, f_2, \cdots : X \to \mathbf{C}$ be a sequence of measurable functions that converge pointwise μ-almost everywhere to a measurable limit $f : X \to \mathbf{C}$. Suppose that there is an unsigned absolutely integrable function $G : X \to [0, +\infty]$ such that $|f_n|$ are pointwise μ-almost everywhere bounded by G for each n. Then we have*

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

Proof. Proof omitted. $\square$

Remark. We deduce the dominated convergence theorem from Fatou's lemma, and Fatou's lemma from the monotone convergence theorem. However, one can obtain these theorems in a different order, depending on one's taste, as they are so closely related. For instance, one can obtains the slightly simpler *bounded convergence theorem*, which is the dominated convergence theorem under the assumption that the functions are uniformly bounded and all supported on a single set of finite measure, and then uses that to deduce Fatou's lemma, which in turn is used to deduce the monotone convergence theorem; and then the horizontal and vertical truncation properties are used to extend the bounded convergence theorem to the dominated convergence theorem. It is instructive to view a couple different derivations of these key theorems to get more of an intuitive understanding as to how they work.

§1.5 Modes of convergence

As an undergraduate, one learns of the following two basic *modes of convergence*:

(i) (Pointwise convergence) We say that f_n converges to f *pointwise* if, for every $x \in X$, $f_n(x)$ converges to $f(x)$. In other words, for every $\varepsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \varepsilon$ whenever $n \geq N$.

(ii) (Uniform convergence) We say that f_n converges to f *uniformly* if, for every $\varepsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \varepsilon$ for every $n \geq N$ and every $x \in X$.

The difference between uniform convergence and pointwise convergence is that with the former, the time N at which $f_n(x)$ must be permanently ε-close to $f(x)$ is not permitted to depend on x, but must instead be chosen uniformly in x. Uniform convergence implies pointwise convergence, but not conversely.

However, pointwise and uniform convergence are only two of dozens of many other modes of convergence that are of importance in analysis.

We will discuss some of the modes of convergence that arise from measure theory, when the domain X is equipped with the structure of a measure space $(X, \mathcal{B}, \mu)$, and the functions f_n (and their limit f) are measurable with respect to this space. In this context, we have some additional modes of convergence:

(iii) (Pointwise a.e. convergence) We say that f_n converges to f *pointwise almost everywhere* if, for (μ-)almost everywhere $x \in X$, $f_n(x)$ converges to $f(x)$.

(iv) (L^∞ convergence) We say that f_n converges to f *uniformly almost everywhere*, *essentially uniformly*, or *in L^∞ norm* if, for every $\varepsilon > 0$, there exists N such that for every $n \geq N$, $|f_n(x) - f(x)| \leq \varepsilon$ for μ-almost every $x \in X$.

(v) (Almost uniform convergence) We say that f_n converges to f *almost uniformly* if, for every $\varepsilon > 0$, there exists an exceptional set $E \in \mathcal{B}$ of measure $\mu(E) \leq \varepsilon$ such that f_n converges uniformly to f on the complement of E.

(vi) (L^1 convergence) We say that f_n converges to f *in L^1 norm* if the quantity $\|f_n - f\|_{L^1(\mu)} = \int_X |f_n(x) - f(x)| d\mu$ converges to zero as $n \to \infty$.

(vii) (Convergence in measure) We say that f_n converges to f *in measure* if, for every $\varepsilon > 0$, the measure $\mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\})$ converges to zero as $n \to \infty$.

Proposition 1.5.1 (Linearity of convergence)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, let $f_n, g_n : X \to \mathbf{C}$ be sequences of measurable functions, and let $f, g : X \to \mathbf{C}$ be measurable functions.*

(i) *f_n converges to f along one of the above seven modes of convergence if and only if $|f_n - f|$ converges to 0 along the same mode.*

(ii) *If f_n converges to f along one of the above seven modes of convergence, and g_n converges to g along the same mode.*

Then $f_n + g_n$ converges to $f + g$ along the same mode, and that cf_n converges to cf along the same mode for any $c \in \mathbf{C}$.

(iii) *(Squeeze test) If f_n converges to 0 along one of the above seven modes, and $|g_n| \le |f_n|$ pointwise for each n. Then g_n converges to 0 along the same mode.*

Proof. This is easy to prove that the assertions are hold for (i) and (ii). Then use (i) to prove (ii), and use (ii) to prove (iv) and (v) that the assertions are hold. Here we only show the cases of (vi) and (vii).

(i) For the convergence in L^1 norm. Since $||f_n(x)-f(x)|-0| = |f_n(x)-f(x)|$, we have

$$\int_X ||f_n(x) - f(x)| - 0| d\mu = \int_X |f_n(x) - f(x)| d\mu.$$

Thus f_n converges to f in L^1 norm if and only if $|f_n - f|$ converges to 0 in L^1 norm.

Similarly, for the convergence in measure, we have

$$\mu(\{x \in X : ||f_n(x) - f(x)| - 0| \ge \varepsilon\}) = \mu(\{x \in X : |f_n(x) - f(x)| \ge \varepsilon\}).$$

Thus f_n converges to f in measure if and only if $|f_n - f|$ converges to 0 in measure.

(ii) From the triangle inequality, we have

$$\begin{aligned} 0 \le \|(f_n + g_n) - (f + g)\|_{L^1(\mu)} &= \|(f_n - f) + (g_n - g)\|_{L^1(\mu)} \\ &\le \|f_n - f\|_{L^1(\mu)} + \|g_n - g\|_{L^1(\mu)}. \end{aligned}$$

Thus when $n \to \infty$, we have $\|(f_n + g_n) - (f + g)\|_{L^1(\mu)}$ converges to 0. For $c \in \mathbf{C}$, we have

$$\|cf_n - cf\|_{L^1(\mu)} = \|c(f_n - f)\|_{L^1(\mu)} = |c|\|f_n - f\|_{L^1(\mu)}.$$

Thus when $n \to \infty$, we have $\|cf_n - cf\|_{L^1(\mu)}$ converges to 0.

For the convergence in measure: Use triangle inequality, we have

$$\{x \in X : |f_n(x) - f(x)| \ge \varepsilon\} \cup \{x \in X : |g_n(x) - g(x)| \ge \varepsilon\}$$

$$\supset \{x \in X : |(f_n(x) + g_n(x)) - (f(x) + g(x))| \geq \varepsilon\}.$$

Then

$$\begin{aligned} &\mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}) + \mu(\{x \in X : |g_n(x) - g(x)| \geq \varepsilon\}) \\ &\geq \mu(\{x \in X : |(f_n(x) + g_n(x)) - (f(x) + g(x))| \geq \varepsilon\}) \end{aligned}$$

Thus when $n \to \infty$, we have $\mu(\{x \in X : |(f_n(x)+g_n(x)) - (f(x)+g(x))| \geq \varepsilon\})$ converges to 0. For $c \in \mathbf{C}$, we have

$$\{x \in X : |cf_n(x) - cf(x)| \geq \varepsilon\} = \{x \in X : |f_n(x) - f(x)| \geq \varepsilon/|c|\}.$$

Since ε is arbitrary, f_n converges to f in measure follows that $\mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon/|c|\})$ converges to 0 as $n \to \infty$. It follows that $\mu(\{x \in X : |cf_n(x) - cf(x)| \geq \varepsilon\})$ converges to 0 as well.

(iii) For the convergence in L^1 norm, we have

$$\|g_n\|_{L^1(\mu)} \leq \|f_n\|_{L^1(\mu)}$$

converges to 0 as $n \to \infty$.

For the convergence in measure, we have

$$\mu(\{x \in X : |g_n| \geq \varepsilon\}) \leq \mu(\{x \in X : |f_n| \geq \varepsilon\})$$

converges to 0 as $n \to \infty$. □

Proposition 1.5.2 (Easy implications)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space, and let $f_n : X \to \mathbf{C}$ and $f : X \to \mathbf{C}$ be measurable functions.*

(i) *If f_n converges to f uniformly, then f_n converges to f pointwise.*

(ii) *If f_n converges to f uniformly, then f_n converges to f in L^∞ norm. Conversely, if f_n converges to f in L^∞ norm, then f_n converges to f uniformly outside of a null set (i.e. there exists a null set E such that the restriction $f_n \downharpoonright_{X\setminus E}$ of f_n to the complement of E converges to the restriction $f \downharpoonright_{X\setminus E}$ of f).*

(iii) *If f_n converges to f in L^∞ norm, then f_n converges to f*

almost uniformly.

(iv) *If f_n converges to f almost uniformly, then f_n converges to f pointwise almost everywhere.*

(v) *If f_n converges to f in L^1 norm, then f_n converges to f in measure.*

(vi) *If f_n converges to f almost uniformly, then f_n converges to f in measure.*

Proof. These are easy to prove through compare the definitions. □ We

give four key examples that distinguish between these modes, in the case when X is the real line $\mathbf{R}$ with Lebesgue measure. The first three of these examples were already introduced in Section 1.4, but the forth is new, and also important.

Example 1.5.3 (Escape to horizontal infinity)**.** Let $f_n := 1_{[n,n+1]}$. Then f_n converges to zero *pointwise* (and thus, *pointwise almost everywhere*), but not *uniformly*, *in L^∞ norm*, *almost uniformly*, *in L^1 norm*, or *in measure*.

Example 1.5.4. Let $f_n := \frac{1}{n} 1_{[0,n]}$. Then f_n converges to zero *uniformly* (and thus, *pointwise*, *pointwise almost everywhere*, *in L^∞*, *almost uniformly*, and *in measure*), but not *in L^1 norm*.

Example 1.5.5 (Escape to vertical infinity)**.** Let $f_n := n 1_{[\frac{1}{n}, \frac{2}{n}]}$. Then f_n converges to zero *pointwise* (and thus, *pointwise almost everywhere*) and *almost uniformly* (and hence *in measure*), but not *uniformly*, *in L^∞ norm*, or *in L^1 norm*.

Example 1.5.6 (Typewriter sequence)**.** Let f_n be defined by the formula

$$f_n := 1_{[\frac{n-2^k}{2^k}, \frac{n-2^k+1}{2^k}]}$$

whenever $k \geq 0$ and $2^k \leq n \leq 2^{k+1}$. This is a sequence of indicator functions of intervals of decreasing length, marching across the unit interval $[0, 1]$ over and over again. Then f_n converges to zero *in measure* and *in L^1 norm*, but not *pointwise almost everywhere* (and hence also not *pointwise*, not *almost uniformly*, nor *in L^∞ norm*, nor *uniformly*).

1.5.1 Uniqueness

Let $(X, \mathcal{B}, \mu)$ be a measure space. We abbreviate "μ-almost everywhere" as "almost everywhere" throughout.

Even though the modes of convergence all differ from each other, they are all *compatible* in the sense that they never disagree about *which* function f a sequence of functions f_n converges to, outside of a set of measure zero. More precisely:

Proposition 1.5.7. *Let $f_n : X \to \mathbf{C}$ be a sequence of measurable functions, and let $f, g : X \to \mathbf{C}$ be two additional measurable functions. Suppose that f_n converges to f along one of the seven modes of convergence defined above, and f_n converges to g along another of the seven modes of convergence (or perhaps the same mode of convergence as for f). Then f and g agree almost everywhere.*

Proof. Proof omitted. □

1.5.2 The case of a step function

We define each of *step function* $f_n = A_n 1_{E_n}$ to be a finite linear combination of indicator functions 1_{E_n} of a measurable set E_n. For simplicity we assume that the $A_n > 0$ are positive reals, and that the E_n have a positive measure

$\mu(E_n) > 0$. We also assume the A_n exhibit one of two modes of behaviour: either the A_n converge to zero, or else they are bounded away from zero (i.e. there exists $c > 0$ such that $A_n \geq c$ for eery n).

Given such a sequence $f_n = A_n 1_{E_n}$ of step functions, we now ask, for each of the seven modes of convergence, what it means for this sequence to converge to zero along that mode. It turns out that the answer to the question is controlled more or less completely by the following three quantities:

(i) The *height* A_n of the n^{th} function f_n;

(ii) The *width* $\mu(E_n)$ of the n^{th} function f_n; and

(iii) The N^{th} *tail support* $E_N^* := \bigcup_{n \geq N} E_n$ of the sequence $f_1, f_2, f_3, \cdots$.

Proposition 1.5.8 (Convergence for step functions). *Let the notation and assumptions be as above. Then*

(i) *f_n converges uniformly to zero if and only if $A_n \to 0$ as $n \to \infty$.*

(ii) *f_n converges in L^∞ norm to zero if and only if $A_n \to 0$ as $n \to \infty$.*

(iii) *f_n converges almost uniformly to zero if and only if*

$$\min(A_n, \mu(E_n^*)) \to 0$$

as $n \to \infty$.

(iv) *f_n converges pointwise to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\bigcap_{N=1}^\infty E_N^* = \emptyset$.*

(v) *f_n converges pointwise almost everywhere to zero if and only if $A_n \to 0$ as $n \to \infty$, or $\bigcap_{N=1}^\infty E_N^*$ is a null set.*

(vi) *f_n converges in measure to zero if and only if*

$$\min(A_n, \mu(E_n)) \to 0$$

as $n \to \infty$.

(vii) *f_n converges in L^1 norm to zero if and only if $A_n\mu(E_n) \to 0$ as $n \to \infty$.*

Proof. (i) Suppose that f_n converges uniformly to zero. For every $\varepsilon > 0$ there exists an $N > 0$ such that $|A_n 1_{E_n}| \le \varepsilon$ for every $n \ge N$ and $x \in X$. This means that A_n cannot be bounded away from zero. Thus we have $A_n \to 0$ as $n \to \infty$. This provers the "only if" part; the "if" part is trivial.

(ii) Since f converges uniformly outside of a null set is equivalent to converges in L^∞, the assertion is immediately comes from (i).

(iii) Suppose that f_n converges almost uniformly to zero. For every $\varepsilon > 0$, there exists a measurable set E with the measure $\mu(E) \le \varepsilon$ such that f_n converges uniformly to zero on E^c. Notice that for every E_n, we have $E_n \cap E^c \neq \emptyset$, $E_n \subset E^c$, or $E_n \subset E$. When $E_n \cap E^c \neq \emptyset$, for every $x \in E_n \cap E^c$ we have $A_n \to 0$ as $n \to \infty$ from (i). Thus $\min(A_n, \mu(E_n^*)) = A_n \to 0$ as $n \to \infty$. When $E_n \subset E^c$, then we still have $\min(A_n, \mu(E_n^*)) = A_n \to 0$ as $n \to \infty$. (In above two cases, $\mu(E_n)$ is not necessary converges to 0 as $n \to \infty$ for that they intersect with E^c.) While if $E_k \subset E$. Since there is an $N' > 0$ such that $|f_n| \le \varepsilon$ for all $n \ge N$ and $x \in E^c$, then for every $x \in E_k \subset E$ we have

$$\mu(E_n^*) \le \mu\Big(\bigcup_{k\ge n} E_k\Big) \le \mu(E) \le \varepsilon$$

for every $n \ge N$. This means that $\mu(E_n^*) \to 0$ as $n \to \infty$. Thus we conclude that we have $\min(A_n, \mu(E_n^*)) \to 0$ as $n \to \infty$.

For the other direction, suppose that $\min(A_n, \mu(E_n^*)) \to 0$ as $n \to \infty$. If $\min(A_n, \mu(E_n^*)) = A_n$, then we have f_n converges uniformly, and thus almost uniformly to zero. While if $\min(A_n, \mu(E_n^*)) = \mu(E_n^*)$, there is an $N > 0$ such that $\mu(E_n^*) \le \varepsilon$ for every $n \ge N$. Let $E := E_N^*$. Then for every $x \in E^c$ and $n \ge N$, we have $1_{E_n} = 0$ for that $x \notin E_n \subset E$ for every $n \ge N$. This follows that we have a measurable set E with $\mu(E) \le \varepsilon$ such that $|f_n| = 0 \le \varepsilon$ for every $n \ge N$ and $x \in E^c$, thus f_n converges almost uniformly to zero.

(iv) Suppose that f_n converges pointwise to zero. For every $\varepsilon > 0$ and every $x \in X$ there exists an $N > 0$ such that $|f_n| \le \varepsilon$ whenever $n \ge N$.

This means that we either have A_n converges to zero or 1_{E_n} converges to zero as $n \to \infty$. In the latter case, we have $\lim_{n\to\infty} 1_{E_n} = 0$, write this as set-theoretic form, we have

$$\begin{aligned} x \notin \{x \in X : \lim_{n\to\infty} 1_{E_n} = 1\} &= \{x \in X : \limsup_{n\to\infty} 1_{E_n} = 1\} \\ &= \bigcap_{N\geq 1} \bigcup_{n\geq N} E_n \\ &= \bigcap_{N\geq 1} E_N^*. \end{aligned}$$

Since this is hold for all $x \in X$, we have $\bigcap_{N\geq 1} E_N^* = \emptyset$.

Conversely, suppose that $A_n \to 0$ as $n \to \infty$, or $\bigcap_{N=1}^{\infty} E_N^* = \emptyset$. If $A_n \to 0$ as $n \to \infty$. Then f_n converges uniformly to zero, and thus converges pointwise to zero. While if $\bigcap_{N=1}^{\infty} E_N^* = \emptyset$, this means that none of $x \in X$ such that $\lim_{n\to\infty} 1_{E_n} = 1$, thus we have $\lim_{n\to\infty} 1_{E_n} = 0$ for all $x \in X$ and f_n converges pointwise to zero.

(v) Since f converges pointwise outside of a null set is equivalent to converges pointwise a.e., the assertion is immediately comes from (iv).

(vi) Suppose that f_n converges to zero in measure, then for every $m \geq 1$, $\{x \in X : |f_n(x)| \geq 1/m\}$ converges to zero as $n \to \infty$. Let

$$E_{n,m} := \{x \in X : |f_n(x)| \geq \frac{1}{m}\}.$$

Then for every $m \geq 1$, there exists $N > 0$ such that $\mu(E_{n,m}) \leq \varepsilon$ for all $n \geq N$. For every E_n, we have $E_n \cap E_{n,m}^c \neq \emptyset$, $E_n \subset E_{n,m}^c$, or $E_n \subset E_{n,m}$. When $E_n \cap E_{n,m}^c \neq \emptyset$, for every $x \in E_n \cap E_{n,m}^c \neq \emptyset$ we have $|f_n| = |A_n| \leq \varepsilon$ whenever $n \geq N$. Thus $\min(A_n, \mu(E_n)) = A_n \to 0$ as $n \to \infty$. When $E_n \subset E_{n,m}^c$, for every $x \in E_n \subset E_{n,m}^c$ we still have $\min(A_n, \mu(E_n)) = A_n \to 0$ as $n \to \infty$. When $E_n \subset E_{n,m}$, we have $\mu(E_n) \leq m(E_{n,m}) \leq \varepsilon$ whenever $n \geq N$. Thus $\min(A_n, \mu(E_n)) = \mu(E_n) \to 0$ as $n \to \infty$. Thus we conclude that $\min(A_n, \mu(E_n)) \to 0$ as $n \to \infty$.

For the converse, suppose that $\min(A_n, \mu(E_n)) \to 0$ as $n \to \infty$. If $\min(A_n, \mu(E_n)) = A_n$, this implies that f_n converges uniformly to zero, and thus converges to zero in measure. While if $\min(A_n, \mu(E_n)) = \mu(E_n)$, and we suppose that A_n are bounded away from zero, then for every $x \in E_{n,m}$ we have $x \in E_n$, thus $E_{n,m} \subset E_n$. This implies that for every m, there exists $N > 0$ such that $\mu(E_{n,m}) \leq \mu(E_n)$ for all $n \geq N$. Thus f_n converges to zero in measure.

(vii) From following equality

$$\|f_n\|_{L^1(\mu)} = \int_X A_n 1_{E_n} d\mu = A_n \mu(E_n),$$

we see that f_n converges to zero in L^1 norm if and only if $A_n \mu(E_n) \to 0$ as $n \to \infty$. $\square$

Remark. To put it more informally: When the height goes to zero, then one has convergence to zero in all modes except possibly for L^1 convergence, which requires that the product of the height and the width goes to zero. If instead, the height is bounded away from zero and the width is positive, then we never have uniform or L^∞ convergence, but we have convergence in measure if the width goes to zero, we have almost uniform convergence if the tail support (which has larger measure than the width) has measure that goes to zero, we have pointwise almost everywhere convergence if the tail support shrinks to a null set, and pointwise convergence if the tail support shrinks to the emptyset.

1.5.3 Finite measure spaces

The situation simplifies somewhat if the space X has finite measure (and in particular, in the case when $(X, \mathcal{B}, \mu)$ is a *probability space*). This shuts down two of the four examples (namely, escape to horizontal infinity or width infinity) and creates a few more equivalences. Indeed, from Egorov's theorem (Theorem 1.3.38), we now have

Theorem 1.5.9 (Egorov's theorem, again)**.** *Let X have finite measure, and let $f_n : X \to \mathbf{C}$ and $f : X \to \mathbf{C}$ be measurable functions. Then f_n converges to f pointwise almost everywhere if and only if f_n converges to f almost uniformly.*

Another nice feature of the finite measure case is that L^∞ convergence implies L^1 convergence:

Proposition 1.5.10. *Let X have finite measure, and let $f_n : X \to \mathbf{C}$ and $f : X \to \mathbf{C}$ be measurable functions. If f_n converges to f in L^∞ norm, then f_n also converges to f in L^1 norm.*

Proof. Suppose that f_n converges to f in L^∞ norm, for every $\varepsilon > 0$, there exists $N > 0$ such that $|f_n(x) - f(x)| \leq \varepsilon/\mu(X)$ for every $n \geq N$ and for μ-almost every $x \in X$. This means that we have sub-null set

$$E := \{x \in X : |f_n(x) - f(x)| \geq \frac{\varepsilon}{\mu(X)}\}.$$

Then

$$\begin{aligned}
\|f_n - f\|_{L^1(\mu)} &= \int_X |f_n(x) - f(x)| d\mu \\
&= \int_E |f_n(x) - f(x)| d\mu + \int_{X \setminus E} |f_n(x) - f(x)| d\mu \\
&\leq \mu(X) \frac{\varepsilon}{\mu(X)} \\
&\leq \varepsilon
\end{aligned}$$

for all $n \geq N$. Thus f_n converges to f in L^1 norm. □

1.5.4 Fast convergence

The typewriter example shows that L^1 convergence is not strong enough to force almost uniform or pointwise almost everywhere convergence. However, this can be rectified if one assumes that the L^1 convergence is sufficiently fast:

Proposition 1.5.11 (Fast L^1 convergence)**.** *Suppose that $f_n, f : X \to \mathbf{C}$ are measurable functions such that $\sum_{n=1}^\infty \|f_n - f\|_{L^1(\mu)} < \infty$; thus, not only do the quantities $\|f_n - f\|_{L^1(\mu)}$ go to zero (which would mean L^1 convergence), but they converge in an absolutely summable fashion.*

Then

(i) f_n *converges pointwise almost everywhere to* f.

(ii) f_n *converges almost uniformly to* f.

Proof. (i) We first consider the step function $f_n = A_n 1_{E_n}$ and let $f = 0$. f_n converges to zero in L^1 norm implies, by Proposition 1.5.8(vii), that $A_n \mu(E_n)$ converges to zero as $n \to \infty$. If E_n is not null set, then it requires that A_n converges to zero. Thus except all of null sets, there exists N such that $f_n = A_n 1_{E_n} \leq \varepsilon$ for μ-almost every $x \in X$ and every $n \geq N$. Thus f_n converges pointwise a.e. to zero.

In general, we show that f_n converges pointwise a.e. to f. From Markov's inequality, for every $0 < \lambda < \infty$ we have

$$\{x \in X : |f_n(x) - f(x)| \geq \lambda\} \leq \frac{1}{\lambda} \int_X |f_n(x) - f(x)| d\mu$$

§1.6 Differentiation theorems

Throughout this section, the notions of measurability and "almost everywhere" are understood to be with respect to Lebesgue measure.

Definition 1.6.1 (Some concepts of differentiation)**.** Let $[a, b]$ be a compact interval of positive length ($-\infty < a < b < +\infty$). A function $F : [a, b] \to \mathbf{R}$ is said to be *differentiable* at a point $x \in [a, b]$ if the limit

$$F'(x) := \lim_{y \to x; y \in [a,b] \setminus \{x\}} \frac{F(y) - F(x)}{y - x}$$

exists. In that case, we call $F'(x)$ the *strong derivative*, *classical derivative*, or just *derivative* for short, of F at x. We say that F is *everywhere differentiable*, or *differentiable* for short, if it is differentiable at all points $x \in [a, b]$, and *differentiable almost everywhere* if it is differentiable at almost every point $x \in [a, b]$. If F is differentiable everywhere and its derivative F' is continuous, then we say that F is *continuously* *differentiable.*

Proposition 1.6.2. *If $F : [a, b] \to \mathbf{R}$ is everywhere differentiable, then F is continuous and F' is measurable. If F is almost everywhere differentiable, then the (almost everywhere defined) function F' is measurable.*

Proof. By definition, for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\Big|\frac{F(y) - F(x)}{y - x} - F'(x)\Big| \leq \varepsilon$$

whenever $y \in [a, b] \setminus \{x\}$ such that $|y - x| \leq \delta$. Following inequality

$$|F(y) - (F(x) + F'(x)(y - x))| \leq \varepsilon|y - x|$$

can be derived from above definition when $y \neq x$, and it is clearly true when $y = x$. Then we conclude that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|F(y) - (F(x) + F'(x)(y - x))| \leq \varepsilon|y - x|,$$

and hence

$$|F(y) - F(x)| \leq (\varepsilon + |F'(x)|)|y - x|$$

for all $y \in [a, b]$ such that $|y - x| \leq \delta$. Let $\zeta = (\varepsilon + |F'(x)|)|y - x| > 0$. Then for every $\zeta > 0$ there exists $\delta > 0$ such that $|F(y) - F(x)| \leq \zeta$ whenever $y \in [a, b]$ such that $|y - x| \leq \delta$. Since this is hold for every $x \in [a, b]$, thus F is continuous on $[a, b]$.

Since f is continuous, for every open subset $U \subset \mathbf{R}$, $f^{-1}(U)$ is an open set in $\mathbf{R}$, so is Lebesgue measurable. Thus by Definition 1.4.37, f is measurable.

The second assertion is immediate consequence of the first. □

In single-variable calculus, the operations of integration and differentiation are connected by a number of basic theorems, starting with *Rolle's theorem*.

Theorem 1.6.3 (Rolle's theorem)**.** *Let $[a,b]$ be a compact interval of positive length, and let $F : [a,b] \to \mathbf{R}$ be a differentiable function such that $F(a) = F(b)$. Then there exists $x \in (a,b)$ such that $F'(x) = 0$.*

Proof. Proof omitted. □

One can easily amplify Rolle's theorem to the *mean value theorem*:

Corollary 1.6.4 (Mean value theorem)**.** *Let $[a,b]$ be a compact interval of positive length, and let $F : [a,b] \to \mathbf{R}$ be a differentiable function. Then there exists $x \in (a,b)$ such that $F'(x) = \frac{F(b)-F(a)}{b-a}$.*

Proof. Proof omitted.

Lemma 1.6.5 (Uniqueness of antiderivatives up to constants)**.** *Let $[a,b]$ be a compact interval of positive length, and let $F : [a,b] \to \mathbf{R}$ and $G : [a,b] \to \mathbf{R}$ be differentiable functions. Then $F'(x) = G'(x)$ for every $x \in [a,b]$ if and only if $F(x) = G(x) + C$ for some constant $C \in \mathbf{R}$ and all $x \in [a,b]$.*

Proof. Suppose that $F'(x) = G'(x)$, by mean value theorem, for every $x \in [a,b]$, there exists $y \in (a,b)$ (either $y \in (a,x)$ or $y \in (x,b)$) where $x \neq y$ such that

$$F'(x) - G'(x) = \frac{F(x) - G(x) - (F(y) - G(y))}{x - y} = 0.$$

This means that $F(x) - G(x) = F(y) - G(y)$. Let $C := F(y) - G(y)$, we have $F(x) = G(x) + C$ for all $x \in [a,b]$. The converse is immediately comes from definition. □

We can use the mean value theorem to deduce one of the fundamental theorems of calculus:

Theorem 1.6.6 (Second fundamental theorem of calculus)**.** *Let $F : [a, b] \to \mathbf{R}$ be a differentiable function, such that F' is Riemann integrable. Then the Riemann integral $\int_a^b F'(x)dx$ of F' is equal to $F(b)-F(a)$. In particular, we have $\int_a^b F'(x)dx = F(b)-F(a)$ whenever F is continuously differentiable.*

Proof. Proof omitted. □

Of course, we also have the other half of th efundamental theorem of calculus:

Theorem 1.6.7 (First fundamental theorem of calculus)**.** *Let $[a, b]$ be a compact interval of positive length. Let $f : [a, b] \to \mathbf{C}$ be a continuous function, and let $F : [a, b] \to \mathbf{C}$ be the indefinite integral $F(x) := \int_a^x f(t)dt$. Then F is differentiable on $[a, b]$, with derivative $F'(x) = f(x)$ for all $x \in [a, b]$. In particular, F is continuously differentiable.*

Proof. Proof omitted. □

Corollary 1.6.8 (Differentiation theorem for continuous functions)**.** *Let $f : [a, b] \to \mathbf{C}$ be a continuous function on a compact interval. Then we have*

$$\lim_{h \to 0^+} \frac{1}{h} \int_{[x,x+h]} f(t)dt = f(x)$$

for all $x \in [a, b)$,

$$\lim_{h \to 0^+} \frac{1}{h} \int_{[x-h,x]} f(t)dt = f(x)$$

for all $x \in (a, b]$,

$$\lim_{h \to 0^+} \frac{1}{2h} \int_{[x-h,x+h]} f(t)dt = f(x)$$

for all $x \in (a,b)$.

1.6.1 The Lebesgue differentiation theorem in one dimension

The main objective of this section is to show

Theorem 1.6.9 (Lebesgue differentiation theorem, one dimensional case)**.** *Let $f : \mathbf{R} \to \mathbf{C}$ be an absolutely integrable function, and let $F : \mathbf{R} \to \mathbf{C}$ be the indefinite integral $F(x) := \int_{[-\infty,x]} f(t)dt$. Then F is continuous and almost everywhere differentiable, and $F'(x) = f(x)$ for almost every $x \in \mathbf{R}$.*

The continuity is easy:

Theorem 1.6.10. *Let $f : \mathbf{R} \to \mathbf{C}$ be an absolutely integrable function, and let $F : \mathbf{R} \to \mathbf{C}$ be the indefinite integral $F(x) := \int_{[-\infty,x]} f(t)dt$. Then F is continuous.*

Proof. For every $x, y \in \mathbf{R}$, we have

$$\begin{aligned} |F(x) - F(y)| &= \Big| \int_{[-\infty,x]} f(t)dt - \int_{[-\infty,y]} f(t)dt \Big| \\ &= \Big| \int_{[x,y]\cup[y,x]} f(t)dt \Big| \\ &\le \int_{[x,y]\cup[y,x]} |f(t)|dt < \infty. \end{aligned}$$

This implies that $|f(t)|$ is finite. Then we can find a real number $M > 0$ such that $|f(t)| \le M$ for all $t \in [x,y] \cup [y,x]$. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|F(x) - F(y)| \le \int_{[x,y]\cup[y,x]} |f(t)|dt \le M|x-y| \le \varepsilon$$

whenever $|x - y| \le \delta$. Thus F is continuous. □

The main difficulty is to show $F'(x) = f(x)$ for almost every $x \in \mathbf{R}$. This will follow from

Theorem 1.6.11 (Lebesgue differentiation theorem, second formulation)**.** *Let $f : \mathbf{R} \to \mathbf{C}$ be an absolutely integrable function. Then*

$$\lim_{h \to 0^+} \frac{1}{h} \int_{[x,x+h]} f(t)dt = f(x) \tag{1.4}$$

for almost every $x \in \mathbf{R}$, and

$$\lim_{h \to 0^+} \frac{1}{h} \int_{[x-h,x]} f(t)dt = f(x) \tag{1.5}$$

for almost every $x \in \mathbf{R}$.

We can see that Theorem 1.6.9 follows from Theorem 1.6.11:

Proof of Theorem 1.6.9. By Theorem 1.6.10, F is continuous. Now we show that F is almost everywhere differentiable. For almost every $x \in \mathbf{R}$, we have the right limit

$$\begin{aligned} F'(x+) &= \lim_{y \to x; y \in \mathbf{R} \cap (x,+\infty)} \frac{F(y) - F(x)}{y - x} \\ &= \lim_{y \to x; y \in \mathbf{R} \cap (x,+\infty)} \frac{1}{y-x} \Big(\int_{[-\infty,y]} f(t)dt - \int_{[-\infty,x]} f(t)dt \Big) \\ &= \lim_{y \to x; y \in \mathbf{R} \setminus \{x\}} \frac{1}{y-x} \int_{[x,y]} f(t)dt. \end{aligned}$$

After a change of variables (replace y with $x + h$), we can write

$$F'(x+) = \lim_{y \to x; y \in \mathbf{R} \cap (x,+\infty)} \frac{1}{y-x} \int_{[x,y]} f(t)dt = \lim_{h \to 0^+} \frac{1}{h} \int_{[x,x+h]} f(t)dt = f(x).$$

A similar argument shows that

$$F'(x-) = \lim_{y \to x; y \in \mathbf{R} \cap (-\infty,x)} \frac{1}{x-y} \int_{[y,x]} f(t)dt = \lim_{h \to 0^+} \frac{1}{h} \int_{[x-h,x]} f(t)dt = f(x)$$

for almost every $x \in \mathbf{R}$. Thus the limit exists, F is almost everywhere differentiable, and $F'(x) = f(x)$ for almost every $x \in \mathbf{R}$. $\square$

We will just prove the first fact (1.4); the second fact (1.5) is similar.The conclusion (1.4) we want to prove is a *convergence theorem*: an assertion that for all functions f in a given *class*, a certain *sequence of linear expressions* $T_h f$ converge in some *sense* to specified *limit*. In this case, we assert that absolutely integrable functions f (class), and right averages $T_h f(x) := \frac{1}{h}\int_{[x,x+h]} f(t)dt$ (sequence of linear expressions) converges pointwise almost everywhere (sense) to f.

There is a general and very useful argument to prove such convergence theorems, known as the *density argument*. This argument requires two ingredients, which we state informally as follows:

(i) A verification of the convergence result for some "dense subclass" of "nice" functions f, such as continuous functions, smooth functions, simple functions, etc. By "dense", we mean that a general function f in the original class can be approximated to arbitrary accuracy in a suitable sense by a function in the nice subclass.

(ii) A quantitative estimate that upper bounds the maximal fluctuation of the linear expressions $T_h f$ in terms of the "size" of the function f (where the precise definition of "size" depends on the nature of the approximation in the first ingredient).

Once one has these two ingredients, it is usually not too hard to put them together to obtain the desired convergence theorem for general functions f (not just those in the dense subclass).

We illustrate this with a simple example:

Proposition 1.6.12 (Translation is continuous in L^1). *Let $f : \mathbf{R}^d \to \mathbf{C}$ be an absolutely integrable function, and for each $h \in \mathbf{R}^d$, let $f_h : \mathbf{R}^d \to \mathbf{C}$ be the shifted function*

$$f_h(x) := f(x-h).$$

Then f_h converges in L^1 norm to f as $h \to 0$, thus

$$\lim_{h\to 0}\int_{\mathbf{R}^d} |f_h(x) - f(x)| dx = 0.$$

Proof. Proof omitted. □

Proposition 1.6.13. *Let $f : \mathbf{R}^d \to \mathbf{C}$, $g : \mathbf{R}^d \to \mathbf{C}$ be Lebesgue measurable functions such that f is absolutely integrable and g is essentially bounded (i.e. bounded outside of a null set). Then the* convolution *$f * g : \mathbf{R}^d \to \mathbf{C}$ defined by the formula*

$$f * g(x) = \int_{\mathbf{R}^d} f(y)g(x-y)dy$$

is well-defined (in the sense that the integrand on the right-hand side is absolutely integrable) and the $f * g$ is a bounded, continuous function.

Proof. We first show that the convolution $f * g$ is well-defined. Since g is essentially bounded, we can find an $M > 0$ such that $|g(x)| \leq M$ for a.e. $x \in \mathbf{R}^d$. Then

$$f * g(x) = \int_{\mathbf{R}^d} |f(y)g(x-y)|dy \leq \int_{\mathbf{R}^d} M|f(y)|dy = M\int_{\mathbf{R}^d} |f(y)|dy < \infty.$$

It follows that $f * g$ is bounded.

To show that $f * g$ is continuous, we first show that $f * g = g * f$. Let $z = x - y$, then

$$f * g(x) = \int_{\mathbf{R}^d} f(y)g(x-y)dy = \int_{\mathbf{R}^d} f(x-z)g(z)dz = g * f(x).$$

Since there exists $M > 0$ such that $|g| \leq M$ for a.e. x, we have

$$\begin{aligned}
|f * g(x) - f * g(y)| &= \Big|\int_{\mathbf{R}^d} f(x-z)g(z)dz - \int_{\mathbf{R}^d} f(y-z)g(z)dz\Big| \\
&\leq \int_{\mathbf{R}^d} |f(x-z)g(z) - f(y-z)g(z)|dz \\
&= \int_{\mathbf{R}^d} |g(z)||f(x-z) - f(y-z)|dz \\
&\leq M\int_{\mathbf{R}^d} |f(x-z) - f(y-z)|dz.
\end{aligned}$$

Use translation invariance (Proposition 1.3.26), we have

$$M\int_{\mathbf{R}^d} |f(x-z) - f(y-z)|dz = M\int_{\mathbf{R}^d} |f(z) - f(z+(y-x))|dz.$$

By Proposition 1.6.12, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|f(z) - f(z + (y - x))\|_{L^1(\mathbf{R}^d)} \leq \varepsilon/M$, so that $|f * g(x) - f * g(y)| \leq \varepsilon$ whenever $|x - y| \leq \delta$. Thus $f * g$ is continuous. Furthermore, since $f * g$ is bounded, the continuity implies uniformly continuity. $\square$

The above proposition is illustrative of a more general intuition, which is that convolutions tend to be *smoothing* in nature; the convolution $f * g$ of two functions is usually at least as *regular* as, and often more regular than, either of the two factors f, g.

This smoothing phenomenon gives rise to an important fact, namely the *Steinhaus theorem*:

Theorem 1.6.14 (Steinhaus theorem). *Let $E \subset \mathbf{R}^d$ be a Lebesgue measurable set of positive measure. Then the set $E - E := \{x - y : x, y \in E\}$ contains an open neighbourhood of the origin.*

Proof. Suppose that E is bounded, let $-E := \{-y : y \in E\}$. We also have the convolution $1_E * 1_{-E}$ that

$$1_E * 1_{-E}(x) = \int_{\mathbf{R}^d} 1_E(y) 1_{-E}(x - y) dy,$$

which is continuous.

We first show that for every $x \in E$ such that $1_E * 1_{-E}(x) > 0$, we have $x \in E - E$. $1_E * 1_{-E}(x) > 0$ implies $1_E(y)1_{-E}(x - y) = 1$ for every $y \in E$, then we have $y - x \in E$, so that $x = y - (y - x) \in E - E$. Use this conclusion, we see that $0 \in E - E$: for every $y \in E$, we have

$$1_E * 1_{-E}(0) = \int_{\mathbf{R}^d} 1_E(y) 1_{-E}(-y) dy = \int_E 1_{-E}(-y) dy = m(E) > 0.$$

Then by the continuity of $1_E * 1_{-E}$, for every open set $V \subset (0, +\infty)$ that contains $1_E * 1_{-E}(0)$, there exists a neighbourhood U of $0 \in E - E$ such that $1_E * 1_{-E}(U) \subset V$. This means that $U \subset E - E$, as desired.

Now we suppose that E is unbounded. Consider the open ball $B(0, n)$ centered at the origin with radius n. Then $E \cap B(0, n)$ is bounded. Then from above, the set $E - E \cap B(0,n) = \{x - y : x, y \in E \cap B(0,n)\}$ contains an open neighbourhood of the origin for all $n \geq 1$. As $n \to \infty$, we have $E - E$ contains an open neighbourhood of the origin. $\square$

Proposition 1.6.15. *A* homomorphism $f : \mathbf{R}^d \to \mathbf{C}$ *is a map with the property that* $f(x+y) = f(x) + f(y)$ *for all* $x, y \in \mathbf{R}^d$*. Then*

(i) *All measurable homomorphisms are continuous.*

(ii) f *is a measurable homomorphism if and only if it takes the form* $f(x_1, \cdots, x_d) = x_1 z_1 + \cdots + x_d z_d$ *for all* $x_1, \cdots, x_d \in \mathbf{R}$ *and some complex coefficients* $z_1, \cdots, z_d$*.*

Now we return to the Lebesgue differentiation theorem, and apply the density argument. The dense subclass result is already contained in Corollary 1.6.8, which asserts that (1.4) holds for all continuous functions f. The quantitative estimate we will need is the following special case of the *Hardy-Littlewood maximal inequality*:

Lemma 1.6.16 (One-sided Hardy-Littlewood maximal inequality)**.** *Let* $f : \mathbf{R} \to \mathbf{C}$ *be an absolutely integrable function, and let* $\lambda > 0$*. Then*

$$m(\{x \in \mathbf{R} : \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} |f(t)| dt \geq \lambda\}) \leq \frac{1}{\lambda} \int_{\mathbf{R}} |f(t)| dt.$$

We can see that Theorem 1.6.11 follows from Lemma 1.6.16:

Proof of Theorem 1.6.11. Let $f : \mathbf{R} \to \mathbf{C}$ be absolutely integrable, and let $\varepsilon, \lambda > 0$ be arbitrary. Then by Littlewood's second principle, we can find a function $g : \mathbf{R} \to \mathbf{C}$ which is continuous and compactly supported, with

$$\int_{\mathbf{R}} |f(x) - g(x)| dx \leq \varepsilon.$$

Applying the one-sided Hardy-Littlewood maximal inequality, we conclude that

$$m(\{x \in \mathbf{R} : \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} |f(x) - g(x)| dt \geq \lambda\}) \leq \frac{\varepsilon}{\lambda}.$$

In a similar spirit, from Markov's inequality (Lemma 1.3.28) we have

$$m(\{x \in \mathbf{R} : |f(x) - g(x)| \geq \lambda\}) \leq \frac{\varepsilon}{\lambda}.$$

By subadditivity, we conclude that for all $x \in \mathbf{R}$ outside of a set E of measure at most $2\varepsilon/\lambda$, one has both

$$\frac{1}{h} \int_{[x,x+h]} |f(t) - g(t)| dt < \lambda \tag{1.6}$$

and

$$|f(x) - g(x)| < \lambda \tag{1.7}$$

for all $h > 0$.

Now let $x \in \mathbf{R} \setminus E$. From the dense subclass result (Corollary 1.6.8) applied to the continuous function g, we have

$$\Big| \frac{1}{h} \int_{[x,x+h]} g(t) dt - g(x) \Big| < \lambda$$

whenever h is sufficiently close to 0. Combining this with (1.6), (1.7), and the triangle inequality, we conclude that

$$\Big| \frac{1}{h} \int_{[x,x+h]} f(t) dt - f(x) \Big| < 3\lambda$$

for all x outside of a set of measure $2\varepsilon/\lambda$. Keeping λ fixed and sending ε to zero, we conclude that

$$\limsup_{h \to 0} \Big| \frac{1}{h} \int_{[x,x+h]} f(t) dt - f(x) \Big| < 3\lambda$$

for almost every $x \in \mathbf{R}$. If we then let λ go to zero along a countable sequence (e.g. $\lambda := 1/n$ for $n = 1, 2, \cdots$), we conclude that

$$\limsup_{h \to 0} \Big| \frac{1}{h} \int_{[x,x+h]} f(t) dt - f(x) \Big| = 0$$

for almost every $x \in \mathbf{R}$, and the claim follows. □

The only remaining task is to establish the one-sided Hardy-Littlewood maximal inequality. We will do so by using the *rising sun lemma*:

Lemma 1.6.17 (Rising sun lemma)**.** *Let $[a, b]$ be a compact interval, and let $F : [a, b] \to \mathbf{R}$ be a continuous function. Then one can find an at most countable family of disjoint non-empty open intervals $I_n = (a_n, b_n)$ in $[a, b]$ with the following properties:*

(i) *For each n, either $F(a_n) = F(b_n)$, or else $a_n = a$ and $F(b_n) \geq F(a_n)$.*

(ii) *If $x \in (a, b]$ does not lie in any of the intervals I_n, then one must have $F(y) \leq F(x)$ for all $x \leq y \leq b$.*

This lemma is proven using the following basic fact:

Lemma 1.6.18. *Any open subset U of $\mathbf{R}$ can be written as the union of at most countably many disjoint non-empty open intervals, whose endpoints lie outside of U.*

Proof. Since U is open, for every $x \in U$ we can find points

$$a_x := \inf\{a \leq x : (a, x) \subset U\}, \quad \text{and} \quad b_x := \sup\{b \geq x : (x, b) \subset U\}$$

such that $I_x := (a_x, b_x)$ be a maximal open subinterval of U and containing x. To show that I_x are disjoint, we suppose for sake of contradiction that I_x and I_y intersect. Then $I_x \cup I_y$ is also a subinterval of U and containing x, y. Since I_x and I_y are maximal subinterval of U, we have $I_x \cup I_y \subset I_x$ and $I_x \cup I_y \subset I_y$. This implies that $I_x = I_y$. Thus for every two maximal open subinterval of U, they are either same or disjoint. Since we have $U = \bigcup_{x \in I_x} I_x$, the remaining work is to show that there are at most countably many I_x. Since every I_x must containing at least one rational, and I_x are either disjoint or same, thus there are countably many disjoint I_x. □

Then we can use above lemma to prove the rising sun lemma:

Proof of Lemma 1.6.17. Proof omitted. □

Lemma 1.6.19 (Two-sided Hardy-Littlewood maximal inequality)**.** *Let $f : \mathbf{R} \to \mathbf{C}$ be an absolutely integrable function, and let $\lambda > 0$. Then*

$$m(\{x \in \mathbf{R} : \sup_{x \in I} \frac{1}{|I|} \int_I |f(t)|dt \geq \lambda\}) \leq \frac{2}{\lambda} \int_{\mathbf{R}} |f(t)|dt,$$

where the supremum ranges over all intervals I of positive length that contain x.

Proof. By upwards monotonicity, it will suffice to show that

$$m(\{x \in [-n,n] : \sup_{x \in I \subset [-n,n]} \frac{1}{|I|} \int_I |f(t)|dt \geq \lambda\}) \leq \frac{2}{\lambda} \int_{\mathbf{R}} |f(t)|dt$$

for every $n > 0$. By modifying λ by an epsilon, we may replace the non-strict inequality here with strict inequality:

$$m(\{x \in [-n,n] : \sup_{x \in I \subset [-n,n]} \frac{1}{|I|} \int_I |f(t)|dt > \lambda\}) \leq \frac{2}{\lambda} \int_{\mathbf{R}} |f(t)|dt. \tag{1.8}$$

Fix $[-n,n]$. We apply the rising sun lemma to the function $F : [-n,n] \to \mathbf{R}$ defined as

$$F(x) := \int_{[-n,x]} |f(t)|dt - (x+n)\lambda.$$

By Theorem 1.6.10, F is continuous, and so we can find an at most countable sequence of intervals $I_k = (a_k, b_k)$ with the properties given by the rising sun lemma. Let a, b be endpoints of I and $a < b$. From the second property of that lemma, we observe that

$$\{x \in (-n,n] : \sup_{x \in I \subset (-n,n]} \frac{1}{|I|} \int_I |f(t)|dt > \lambda\} \subset \bigcup_k I_k,$$

since the property $\frac{1}{|I|} \int_I |f(t)|dt > \lambda$ can be rearranged as $F(b) > F(a)$. By countable additivity, we may thus upper bound the left-hand side of by

(1.8) by $\sum_k (b_k - a_k)$. On the other hand, since $F(b_k) - F(a_k) \geq 0$, we have

$$\int_{I_k} |f(t)| dt \geq \lambda(b_k - a_k)$$

and thus

$$\sum_k (b_k - a_k) \leq \frac{1}{\lambda} \sum_k \int_{I_k} |f(t)| dt.$$

As the I_k are disjoint intervals in I, we may apply monotone convergence and monotonicity to conclude that

$$\sum_k \int_{I_k} |f(t)| dt \leq \int_I |f(t)| dt,$$

and the claim follows. □

Lemma 1.6.20. *(Rising sun inequality) Let $f : \mathbf{R} \to \mathbf{R}$ be an absolutely integrable function, and let $f^* : \mathbf{R} \to \mathbf{R}$ be the one-sided signed Hardy-Littlewood maximal function*

$$f^*(x) := \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} f(t) dt.$$

Then one has

$$\lambda m(\{f^*(x) > \lambda\}) \leq \int_{x:f^*(x)>\lambda} f(x) dx$$

for all real λ (note here that we permit λ to be zero or negative), and this inequality implies Lemma 1.6.16.

Proof. We first suppose that $\lambda = 0$. Then the assertion degenerate into to prove that

$$\int_{\{x \in \mathbf{R}: f^*(x) > 0\}} f(x) dx \geq 0.$$

Since $h > 0$, $f^*(x) > 0$ implies that

$$\sup_{h>0} \int_{[x,x+h]} f(t) dt > 0$$

This is sufficient to show that

$$\int_{\{x\in[a,b]:f^*(x)>0\}} f(x)dx \geq 0$$

for any compact interval $[a,b]$.

Fix $[a,b]$. We apply the rising sun lemma to the function $F:[a,b]\to\mathbf{R}$ defined as

$$F(x) := \int_{[a,x]} f(t)dt.$$

By Theorem 1.6.10, F is continuous, and so we can find an at most countable sequence of intervals $I_n = (a_n, b_n)$ with the properties given by the rising sun lemma. From the second property of that lemma, we observe that x must be contained in some of I_n, since $F(x) \geq F(x+h)$ contradict $f^*(x) > 0$. Then for every $x \in I_n$, we can find an $h > 0$ such that $x + h \leq b_n$ such that $F(x+h) - F(x) = \int_{[x,x+h]} f(t)dt > 0$. Since for every $x \in [a,b]$ either lies in I_n or not, we can conclude that

$$\int_{\{x\in[a,b]:f^*(x)>0\}} f(x)dx = \sum_n \int_{I_n} f(t)dt = \sum_n F(b_n) - F(a_n) \geq 0$$

and the claim with $\lambda = 0$ follows.

Now we consider arbitrary λ. Notice that

$$\begin{aligned} f^*(x) - \lambda &= \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} f(t)dt - \lambda \\ &= \sup_{h>0} \frac{1}{h} \int_{[x,x+h]} f(t) - \lambda dt \\ &= (f-\lambda)^*(x). \end{aligned}$$

Let $g := f - \lambda$. Then the claim is equivalent to show that

$$\int_{x:g^*(x)>0} g(x)dx = \int_{x:g^*(x)>0} f(x)dx - \lambda m(\{g^*(x) > 0\}) \geq 0,$$

which is our degeneration case. $\square$

1.6.2 The Lebesgue differentiation theorem in higher dimensions

Now we extend the Lebesgue differentiation theorem to higher dimensions. Theorem 1.6.9 does not have an obvious high-dimensional analogue, but Theorem 1.6.11 does:

Theorem 1.6.21 (Lebesgue differentiation theorem in general dimension)**.** *Let $f : \mathbf{R}^d \to \mathbf{C}$ be an absolutely integrable function. Then for almost every $x \in \mathbf{R}^d$, one has*

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)| dy = 0 \tag{1.9}$$

and

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y) = f(x)$$

where $B(x,r) := \{y \in \mathbf{R}^d : |x-y| < r\}$ is the open ball of radius r centered at x.

A point x for which (1.9) holds is called a *Lebesgue point* of f; thus, for an absolutely integrable function f, almost every point in $\mathbf{R}^d$ will be a Lebesgue point for f.

Lemma 1.6.22. *Call a function $f : \mathbf{R}^d \to \mathbf{C}$* locally integrable *if, for every $x \in \mathbf{R}^d$, there exists an open neighbourhood of x on which f is absolutely integrable. Then*

(i) *f is locally integrable if and only if $\int_{B(0,r)} |f(x)|dx < \infty$ for all $r > 0$.*

(ii) *Theorem 1.6.21 implies a generalisation of itself in which the condition of absolute integrability of f is weakened to local integrability.*

Proof. (i) Suppose that f is locally integrable. Fix r. For every $x \in \overline{B(0,r)}$, where $\overline{B(0,r)}$ is closed ball, there is an n such that

$$\int_{B(x,n)} |f(t)|dt < \infty,$$

and

$$B(0,r) \subset \overline{B(0,r)} \subset \bigcup_{x\in\overline{B(0,r)}} B(x,n)$$

By Heine-Borel theorem, there is finitely many $B_k(x,n)$ such that

$$B(0,r) \subset \bigcup_{k=1}^{N} B_k(x,n).$$

Then we have

$$\int_{B(0,n)} |f(t)|dt \le \int_{\bigcup_{k=1}^N B_k(x,n)} |f(t)|dt \le \sum_{k=1}^{N} \int_{B_k(x,n)} |f(t)|dt < \infty.$$

Since this is hold for every $r > 0$, the claim follows.

Conversely, suppose $\int_{B(0,r)} |f(x)|dx < \infty$ for all $r > 0$. Then for every $x \in \mathbf{R}^d$ we can find an enough large r such that $x \in B(0,r)$ and there is neighbourhood $B(x,n)$ of x contained in $B(0,r)$. Thus we have

$$\int_{B(x,n)} |f(t)|dt \le \int_{B(0,r)} |f(t)|dt < \infty.$$

□

Lemma 1.6.23. *For each $h > 0$, let E_h be a subset of $B(0,h)$ with the property that $m(E_h) \ge cm(B(0,h))$ for some $c > 0$ independent of h. If $f : \mathbf{R}^d \to \mathbf{C}$ is locally integrable, and x is a Lebesgue point of f, then*

$$\lim_{h\to 0} \frac{1}{m(E_h)} \int_{x+E_h} f(y)dy = f(x).$$

Proof. From Lebesgue differentiation theorem, for a.e. $x \in \mathbf{R}^d$, we have

$$\lim_{h\to 0} \frac{1}{m(B(x,h))} \int_{B(x,h)} |f(y) - f(x)|dy = 0.$$

Since $E_h \subset B(0,h)$ follows that $E_h + x \subset B(x,h)$, by the monotonicity, we have

$$\lim_{h\to 0} \frac{1}{m(B(x,h))} \int_{B(x,h)} |f(y) - f(x)| dy \geq \lim_{h\to 0} \frac{c}{m(E_h)} \int_{E_h+x} |f(y) - f(x)| dy.$$

This implies that

$$\lim_{h\to 0} \frac{1}{m(E_h)} \int_{E_h+x} |f(y) - f(x)| dy = 0.$$

By the triangle inequality, we have

$$\Big| \frac{1}{m(E_h)} \int_{E_h+x} f(y) - f(x) dy \Big| \leq \frac{1}{m(E_h)} \int_{E_h+x} |f(y) - f(x)| dy.$$

Thus we have

$$\lim_{h\to 0} \frac{1}{m(E_h)} \int_{x+E_h} f(y) dy = f(x).$$

□

To prove Theorem 1.6.21, we use the density argument. The dense subclass case is easy:

Theorem 1.6.24. *Theorem 1.6.21 holds whenever f is continuous.*

Proof. Since f is continuous, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(y) - f(x)| \leq \varepsilon$ whenever $|y - x| < \delta$. Then $m(B(x,\delta))$ is finite, and we have

$$\frac{1}{m(B(x,\delta))} \int_{B(x,\delta)} |f(y) - f(x)| dy \leq \frac{1}{m(B(x,\delta))} \int_{B(x,\delta)} \varepsilon dy \leq \varepsilon$$

whenever $|y - x| < \delta$. Thus the claim follows. □

The quantitative estimate needed is the following:

Theorem 1.6.25 (Hardy-Littlewood maximal inequality)**.** *Let $f : \mathbf{R}^d \to \mathbf{C}$ be an absolutely integrable function, and let $\lambda > 0$. Then*

$$m(\{x \in \mathbf{R}^d : \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)| dy \geq \lambda\}) \leq \frac{C_d}{\lambda} \int_{\mathbf{R}^d} |f(t)| dt$$

for some constant $C_d > 0$ depending only on d.

Use the density argument, we show that Theorem 1.6.25 implies Theorem 1.6.21:

Proof of Theorem 1.6.21. Let $f : \mathbf{R}^d \to \mathbf{C}$ be absolutely integrable, and let $\varepsilon, \lambda > 0$ be arbitrary. Then by Littlewood's second principle, we can find a function $g : \mathbf{R}^d \to \mathbf{C}$ which is continuous, with

$$\int_{\mathbf{R}^d} |f(x) - g(x)| dx \leq \varepsilon.$$

Applying the Hardy-Littlewood maximal inequality, we conclude that

$$m(\{x \in \mathbf{R}^d : \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - g(y)| dy \geq \lambda\}) \leq \frac{C_d \varepsilon}{\lambda}.$$

In a similar spirit, from Markov's inequality we have

$$m(\{x \in \mathbf{R}^d : |f(x) - g(x)| \geq \lambda\}) \leq \frac{\varepsilon}{\lambda}.$$

By subadditivity, we conclude that for all $x \in \mathbf{R}^d$ outside of a set E of measure at most $(C_d + 1)\varepsilon/\lambda$, one has both

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - g(y)| dy < \lambda \tag{1.10}$$

and

$$|f(x) - g(x)| < \lambda \tag{1.11}$$

for all $r > 0$.

Now let $x \in \mathbf{R}^d \setminus E$. From the dense subclass result (Theorem 1.6.24) applied to the continuous function g, we have

$$\Big| \frac{1}{m(B(x,r))} \int_{B(x,r)} g(y) dy - g(x) \Big| < \lambda$$

as $r \to 0$. Combining this with (1.10), (1.11), and the triangle inequality, we conclude that

$$\left|\frac{1}{m(B(x,r))}\int_{B(x,r)} f(y)dy - f(x)\right| < 3\lambda.$$

for all r sufficiently close to zero. In particular, we have

$$\limsup_{r\to 0}\left|\frac{1}{m(B(x,r))}\int_{B(x,r)} f(y)dy - f(x)\right| < 3\lambda$$

for all x outside of a set of measure $(C_d+1)\varepsilon/\lambda$. Keeping λ fixed and sending ε to zero, we conclude that

$$\limsup_{r\to 0}\left|\frac{1}{m(B(x,r))}\int_{B(x,r)} f(y)dy - f(x)\right| < 3\lambda.$$

for almost every $x \in \mathbf{R}^d$. If we then let λ go to zero along a countable sequence, we conclude that

$$\limsup_{r\to 0}\left|\frac{1}{m(B(x,r))}\int_{B(x,r)} f(y)dy - f(x)\right| = 0$$

for almost every $x \in \mathbf{R}^d$, and the claim follows. □

In the one-dimensional case, this estimate was established via the rising sun lemma. Unfortunately, that lemma relied heavily on the ordered nature of $\mathbf{R}$, and does not have an obvious analogue in higher dimensions. Instead, we will use the following covering lemma.

Definition 1.6.26 (c times the ball)**.** Given an open ball $B = B(x,r)$ in $\mathbf{R}^d$ and a real number $c > 0$, we write $cB := B(x, cr)$ for the ball with the same center as B, but c times the radius. Note that $|cB| = c^d|B|$ for any open ball $B \subset \mathbf{R}^d$ and any $c > 0$.

Lemma 1.6.27 (Vitali-type convering lemma)**.** *Let $B_1, \cdots, B_n$ be a finite collection of open balls in $\mathbf{R}^d$ (not necessarily disjoint). Then there exists a subcollection $B_1', \cdots, B_m'$ of disjoint balls in this collec-*

tion, such that

$$\bigcup_{i=1}^{n} B_i \subset \bigcup_{j=1}^{m} 3B'_j.$$

In particular, by finite subadditivity,

$$m\Big(\bigcup_{i=1}^{n} B_i\Big) \leq 3^d \sum_{j=1}^{m} m(B'_j).$$

Proof. Proof omitted. □

Now we can prove the Hardy-Littlewood inequality, which we do with the constant $C_d := 2^d$:

Proof of Theorem 1.6.25. Let K be arbitrary compact set, let $\varepsilon, r > 0$. For every $x \in K$, we have $K \subset \bigcup_{x\in K} \varepsilon B(x, r)$. By completeness of K, there exists finitely many balls $(\varepsilon B(x_i, r_i))_{i=1}^n$ covering K, i.e., $K \subset \bigcup_{i=1}^n \varepsilon B(x_i, r_i)$.

Denote $B(x_i, r_i)$ as B_i. Let B'_1 be such that

$$m(B'_1) := \max(m(B_1), \cdots, m(B_n)).$$

§1.7 Outer measures, pre-measures, and product measures

In this text so far, we have focused primarily on one specific example of a countably additive measure, namely Lebesgue measure. This measure was constructed from a more primitive concept of Lebesgue outer measure, which in turn was constructed from the even more primitive concept of elementary measure.

It turns out that both of these constructions can be abstracted. In this section, we will give the Carathéodory extension theorem, which constructs a countably additive measure from any abstract outer measure; this generalises the construction of Lebesgue measure from Lebesgue outer measure. One can in turn construct outer measures from another concept known as a pre-measure, of which elementary measure is a typical example.

With these tools, one can start constructing many more measures, such as Lebesgue-Stieltjes measures, product measures, and Hausdorff measures. With a little more effort, one can also establish the Kolmogorov extension theorem, which allows one to construct a variety of measures on infinite-dimensional spaces, and is of particular importance in the foundations of probability theory, as it allows one to set up probability spaces associated to both discrete and continuous random processes, even if they have infinite length.

The most important result about product measure, beyond the fact that it exists, is that one can use it to evaluate iterated integrals, and to interchange their order, provided that the integrand is either unsigned or absolutely integrable. This fact is known as the Fubini-Tonelli theorem, and is an absolutely indispensable tool for computing integrals, and for deducing higher-dimensional results from lower-dimensional ones.

1.7.1 Outer measures and the Carathéodory extension theorem

We begin with the abstract concept of an outer measure.

Definition 1.7.1 (Abstract outer measure)**.** Let X be a set. An *abstract outer measure* (or *outer measure* for short) is a map $\mu^* : 2^X \to [0, +\infty]$ that assigns an unsigned extended real number $\mu^*(E) \in [0, +\infty]$ to every set $E \subset X$ which obeys the following axioms:

(i) (Empty set) $\mu^*(\emptyset) = 0$.

(ii) (Monotonicity) If $E \subset F$, then $\mu^*(E) \leq \mu^*(F)$.

(iii) (Countable subadditivity) If $E_1, E_2, \dots \subset X$ is a countable sequence of subsets of X, then $\mu^*(\bigcup_{n=1}^\infty E_n) \leq \sum_{n=1}^\infty \mu^*(E_n)$.

Thus, for instance, Lebesgue outer measure m^* is an outer measure. On the other hand, Jordan outer measure $m^{*,(J)}$ is only finitely subadditive rather than countably subadditive and thus is not, strictly speaking, an outer measure; for this reason this concept is often referred to as *Jordan outer content* rather than *Jordan outer measure*.

In Definition 1.2.2, we used Lebesgue outer measure together with the notion of an open set to define the concept of Lebesgue measurability. This definition is not available in our more abstract setting, as we do not necessarily have the notion of an open set. An alternative definition of measurability was put forth in Proposition 1.2.24, but this still required the notion of a box or an elementary set, which is still not available in the setting. Nevertheless, we can modify that definition to give an abstract definition of measurability:

Definition 1.7.2 (Carathéodory measurability)**.** Let μ^* be an outer measure on a set X. A set $E \subset X$ is said to be *Carathéodory measurable* with respect to μ^* if one has

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E)$$

for every set $A \subset X$.

Example 1.7.3 (Null sets are Carathéodory measurable)**.** Suppose that E is a null set for an outer measure μ^* (i.e. $\mu^*(E) = 0$). Then E is Carathéodory measurable with respect to μ^*.

Proof. From monotonicity, for every $A \subset E$, we have $\mu^*(A) = 0$. Then from empty set,

$$\mu^*(A \cap E) + \mu^*(A \setminus E) = \mu^*(A) + \mu^*(\emptyset) = 0.$$

Thus E is Carathéodory measurable. □

Proposition 1.7.4 (Compatibility with Lebesgue measurability)**.** *A set $E \subset \mathbf{R}^d$ is Carathéodory measurable with respect to Lebesgue outer measurable if and only if it is Lebesgue measurable.*

Proof. The "only if" part is immediately comes from Proposition 1.2.24.

We show the "if" part. By subadditivity, it is sufficient to show that

$$m^*(A) \geq m^*(A \cap E) + m^*(A \setminus E).$$

By outer regularity (Lemma 1.2.11), we can find a sequence of open sets $O_1, O_2, \cdots$ containing E such that

$$m^*(A) \geq m(O_n) - \frac{1}{n}.$$

Define the G to be a G_δ set as $G := \bigcap_{n=1}^\infty O_n$, and we have $A \subset G \subset O_n$. This implies that $m^*(A) = m(G)$. Then we have

$$m^*(A) = m(G) = m(G \cap E) + m(G \setminus E).$$

Since G containing A, we have $m(G \cap E) \geq m^*(A \cap E)$ and $m(G \setminus E) \geq m^*(A \setminus E)$. Thus

$$m^*(A) \geq m^*(A \cap E) + m^*(A \setminus E),$$

as desired. □

Lemma 1.7.5. *Let $\mathcal{B}$ be a Boolean algebra on a set X. Then $\mathcal{B}$ is a σ-algebra if and only if it is closed under countable disjoint unions, which means that $\bigcup_{n=1}^\infty E_n \in \mathcal{B}$ whenever $E_1, E_2, E_3, \cdots \in \mathcal{B}$ are a countable sequence of disjoint sets in $\mathcal{B}$.*

Proof. We prove the "if" part. Let $F_1, F_2, F_3, \cdots \in \mathcal{B}$ be arbitrary countable sequence. Let $E_n := F_n \setminus \bigcup_{k=1}^{n-1} F_k$. Then we obtain a sequence of countable disjoint sets $E_1, E_2, E_3 \cdots \in \mathcal{B}$. From the hypothesis, we have

$$\bigcup_{n=1}^\infty F_n = \bigcup_{n=1}^\infty \left(F_n \setminus \bigcup_{k=1}^{n-1} F_k \right) = \bigcup_{n=1}^\infty E_n \in \mathcal{B}.$$

This shows that $\mathcal{B}$ is a σ-algebra. The "only if" part is trivial. □

The construction of Lebesgue measure can then be abstracted as follows:

Theorem 1.7.6 (Carathéodory extension theorem)**.** *Let $\mu^* : 2^X \to [0, +\infty]$ be an outer measure on a set X, let $\mathcal{B}$ be the collection of all subsets of X that are Carathéodory measurable with respect to μ^*, and let $\mu : \mathcal{B} \to [0, +\infty]$ be the restriction of μ^* to $\mathcal{B}$ (thus $\mu(E) := \mu^*(E)$ whenever $E \in \mathcal{B}$). Then $\mathcal{B}$ is a σ-algebra, and μ is a measure.*

Proof. Proof omitted. □

1.7.2 Pre-measures

In previous notes, we saw that finitely additive measures, such as elementary measure or Jordan measure, could be extended to a countably additive measure, namely Lebesgue measure. It is natural to ask whether this property is true in general. In other words, given a finitely additive measure $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ on a Boolean algebra $\mathcal{B}_0$, is it possible to find a σ-algebra $\mathcal{B}$ refining $\mathcal{B}_0$, and a countably additive measure $\mu : \mathcal{B} \to [0, +\infty]$ that extends μ_0?

There is an obvious necessary condition in order for μ_0 to have a countably additive extension, namely that μ_0 already has to be countably additive within $\mathcal{B}_0$. More precisely, suppose that $E_1, E_2, \dots \in \mathcal{B}_0$ were disjoint sets such that their union $\bigcup_{n=1}^\infty E_n$ was also in $\mathcal{B}_0$. Then, in order for μ_0 to be extendible to a countably additive measure, it is clearly necessary that

$$\mu_0\Big(\bigcup_{n=1}^\infty E_n\Big) = \sum_{n=1}^\infty \mu_0(E_n).$$

Using the Carathéodory extension theorem, we can show that this necessary condition is also sufficient. More precisely, we have

Definition 1.7.7 (Pre-measure)**.** A *pre-measure* on a Boolean algebra $\mathcal{B}_0$ is a finitely additive measure $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ with the property that $\mu_0(\bigcup_{n=1}^\infty E_n) = \sum_{n=1}^\infty \mu_0(E_n)$ whenever $E_1, E_2, \dots \in \mathcal{B}_0$ are disjoint sets such that $\bigcup_{n=1}^\infty E_n$ is in $\mathcal{B}_0$.

We show that this definition can be relaxed as following:

Proposition 1.7.8 (Relaxed definitions of pre-measure)**.**

(i) *The requirement that μ_0 is finitely additive can be relaxed to the condition that $\mu_0(\emptyset) = 0$ without affecting the definition of a pre-measure.*

(ii) *The condition $\mu_0(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu_0(E_n)$ can be relaxed to $\mu_0(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu_0(E_n)$ without affecting the definition of a pre-measure.*

Proof. (i) Let μ_0 be as part (i), then for $E, F \in \mathcal{B}_0$, we have $E \cup F \cup \emptyset \cup \emptyset \cup \cdots$ still lies in $\mathcal{B}_0$. Then by the property, we have

$$\begin{aligned} \mu_0(E \cup F \cup \emptyset \cup \emptyset \cup \cdots) &= \mu_0(E) + \mu_0(F) + \mu_0(\emptyset) + \mu_0(\emptyset) + \cdots \\ &= \mu_0(E) + \mu_0(F). \end{aligned}$$

Thus μ_0 is finitely additive measure.

(i) Let μ_0 be as part (ii). Since μ_0 is finitely additive, it follows the monotonicity. This implies that

$$\mu_0\Big(\bigcup_{n=1}^{\infty} E_n\Big) \geq \mu_0\Big(\bigcup_{n=1}^{k} E_n\Big) = \sum_{n=1}^{k} \mu_0(E_n)$$

for every $k \geq 1$. Taking the k to infinity, we have

$$\mu_0\Big(\bigcup_{n=1}^{\infty} E_n\Big) \geq \sum_{n=1}^{\infty} \mu_0(E_n).$$

Thus μ_0 is pre-measure. □

Example 1.7.9. The elementary measure (on the elementary Boolean algebra) is a pre-measure.

Proof. Let $m : \overline{\mathcal{E}[\mathbf{R}^d]} \to [0, +\infty]$ be elementary measure as Example 1.4.26. Let $E_1, E_2, E_3 \cdots \in \overline{\mathcal{E}[\mathbf{R}^d]}$ be all elementary such that $\bigcup_{n=1}^{\infty} E_n \in \overline{\mathcal{E}[\mathbf{R}^d]}$. By Lemma 1.2.6, we have $m(E) = m^*(E)$ for every elementary set $E \in$

$\overline{\mathcal{E}[\mathbf{R}^d]}$, where m^* is Lebesgue outer measure. Then we have $m(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} m(E_n)$. If there is some co-elementary sets in $E_1, E_2, \cdots$; without loss of generality, let E_k be co-elementary set. Then we have

$$m\Big(\bigcup_{n=1}^{\infty} E_n\Big) \leq m(E_k) = \infty,$$

this follows that $m(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} m(E_n)$. □

Example 1.7.10 (A finitely additive measure is not a pre-measure)**.** Let $X = \mathbf{N}$, let $\mathcal{B}_0 = 2^{\mathbf{N}}$. Define $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ as $\mu_0(E) = 0$ if $E \subset X$ is finite, and $\mu_0(E) = 1$ if $E \subset X$ is infinite. This is easy to see that μ_0 is finitely additive measure but is not a pre-measure for that $\mu_0(\bigcup_{n=1}^{\infty}\{n\}) \geq \sum_{n=1}^{\infty} \mu_0(\{n\}) = 0$.

Theorem 1.7.11 (Hahn-Kolmogorov theorem)**.** *Every pre-measure $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ on a Boolean algebra $\mathcal{B}_0$ in X can be extended to a countably additive measure $\mu : \mathcal{B} \to [0, +\infty]$.*

Proof. Proof omitted. □

Let us call the measure μ constructed in the above proof the *Hahn-Kolmogorov extension* of the pre-measure μ_0. Thus, for instance, from Proposition 1.7.4, the Hahn-Kolmogorov extension of elementary measure (with the convention that co-elementary sets have infinite elementary measure) is Lebesgue measure. This is not quite the unique extension of μ_0 to a countably additive measure, though. For instance, one could restrict Lebesgue measure to the Borel σ-algebra, and this would still be a countably additive extension of elementary measure. However, the extension is unique within its own σ-algebra:

Proposition 1.7.12. *Let $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ be a pre-measure, let $\mu : \mathcal{B} \to [0, +\infty]$ be the Hahn-Kolmogorov extension of μ_0, and let $\mu' : \mathcal{B}' \to [0, +\infty]$ be another countably additive extension of μ_0. Suppose also that μ_0 is σ-finite, which means that one can express the whole space X as the countable union of sets $E_1, E_2, \dots \in \mathcal{B}_0$ for which $\mu_0(E_n) < \infty$ for all n. Then μ and μ' agree on their common domain of definition. In other words, $\mu(E) = \mu'(E)$ for all $E \in \mathcal{B} \cap \mathcal{B}'$.*

Proof. For any set $E \subset X$, define the outer measure $\mu^*(E)$ of E to be the quantity

$$\mu^*(E) := \inf \Big\{ \sum_{n=1}^{\infty} \mu_0(E_n) : E \subset \bigcup_{n=1}^{\infty} E_n; E_n \in \mathcal{B}_0 \text{ for all } n \Big\}.$$

We first show that $\mu'(E) \leq \mu^*(E)$ for all $E \in \mathcal{B}'$. Let $\varepsilon > 0$. From definition, we have

$$\mu^*(E) \geq \sum_{n=1}^{\infty} \mu_0(E_n) - \varepsilon \geq \mu_0\Big(\bigcup_{n=1}^{\infty} E_n \Big) - \varepsilon.$$

Since μ' is extended from μ_0, we have $\mu_0 = \mu'$ for all $E \in \mathcal{B}'$, thus $\mu^*(E) \geq \mu'(\bigcup_{n=1}^{\infty} E_n) - \varepsilon$. By monotonicity and ε was arbitrary, we have $\mu^*(E) \geq \mu'(E)$.

Then we show that $\mu'(E) \geq \mu^*(E)$ for every $E \in \mathcal{B} \cap \mathcal{B}'$.

1.7.3 Lebesgue-Stieltjes measure

Now we use the Hahn-Kolmogorov extension theorem to construct a variety of measures. We begin with Lebesgue-Stieltjes measure.

Theorem 1.7.13 (Existence of Lebesgue-Stieltjes measure)**.** *Let $F : \mathbf{R} \to \mathbf{R}$ be a monotone non-decreasing function, and define the left and right limits*

$$F_-(x) := \sup_{y<x} F(y), \qquad F_+(x) := \inf_{y>x} F(y);$$

thus one has $F_-(x) \le F(x) \le F_+(x)$ *for all* x. *Let* $\mathcal{B}[\mathbf{R}]$ *be the Borel* σ*-algebra on* $\mathbf{R}$. *Then there exists a unique Borel measure* $\mu_F : \mathcal{B}[\mathbf{R}] \to [0, +\infty]$ *such that*

$$\begin{aligned} \mu_F([a,b]) &= F_+(b) - F_-(a), \quad \mu_F([a,b)) = F_-(b) - F_-(a), \\ \mu_F((a,b]) &= F_+(b) - F_+(a), \quad \mu_F((a,b)) = F_+(b) - F_+(a), \end{aligned} \tag{1.12}$$

for all $-\infty < a < b < \infty$, *and*

$$\mu_F(\{a\}) = F_+(a) - F_-(a) \tag{1.13}$$

for all $a \in \mathbf{R}$.

Proof. In our proof, we allow intervals to be unbounded, which means we including the half-infinite intervals $[a, +\infty], (a, +\infty), (-\infty, a], (\infty, a)$ and the doubly infinite interval $(-\infty, +\infty)$ as intervals.

Define the F-volume $|I|_F \in [0, +\infty]$ of any interval I to be the required value of $\mu_F(I)$ given by (1.12) and let $|\emptyset|_F = 0$. In particular, let $F_-(+\infty) = \sup_{y\in\mathbf{R}} F(y)$ and $F_+(-\infty) = \inf_{y\in\mathbf{R}} F(y)$.

i) Additivity property of F*-volume:*

We show that F-volume obeys the additivity property for any disjoint intervals I, J that share a common endpoint. We only need to focus on the common endpoint, and have

$$\begin{aligned} |[a,b]|_F + |(b,c]|_F &= F_+(b) - F_-(a) + F_+(c) - F_+(b) \\ &= F_+(c) - F_-(a) \\ &= |[a,c]|_F, \\ |[a,b)|_F + |[b,c]|_F &= F_-(b) - F_-(a) + F_+(c) - F_-(b) \\ &= F_+(c) - F_-(a) \\ &= |[a,c]|_F. \end{aligned}$$

Then we have $|I \cup J|_F = |I|_F + |J|_F$, and it follows that $|I|_F = |I_1|_F + \cdots + |I_k|_F$, where $I_1, \cdots, I_k$ is a partition of I.

Let $\mathcal{B}_0$ be the Boolean algebra generated by the (possibly infinite) intervals, then $\mathcal{B}_0$ consists of those sets that can be expressed as a finite union of intervals. We can define a measure μ_0 on this algebra by declaring

$$\mu_0(E) = |I_1|_F + \cdots + |I_k|_F$$

whenever $E = I_1 \cup \cdots \cup I_k$ is the disjoint union of finitely many intervals. We show that this measure is well-defined and obeys the additivity property.

ii) Well-definedness of μ_0:

To show that μ_0 is well-defined, we see that the intersection of any two intervals is still an interval. If E be partitioned into different ways, i.e., $E = I_1 \cup \cdots \cup I_n = I'_1 \cup \cdots \cup I'_{n'}$, we have $I_i = \bigcup_{j=1}^{n'} I_i \cap I'_j$ for every i where $I_i \cap I'_1, \cdots, I_k \cap I'_{n'}$ is a partition of I_i. Then by the additivity of F-volume, we have

$$\mu_0(E) = \sum_{i=1}^{n} |I_i|_F = \sum_{i=1}^{n} \sum_{j=1}^{n'} |I_i \cap I'_j|_F = \sum_{j=1}^{n'} |I_j|_F.$$

This shows that μ_0 is well-defined.

iii) Finitely additivity of μ_0:

To show that μ_0 is finitely additive: Let $E, F \in \mathcal{B}_0$ be disjoint, where $E = I_1 \cup \cdots \cup I_m$ and $F = J_1 \cup \cdots \cup J_{m'}$, then $E \cup F \in \mathcal{B}_0$ and there are finitely many disjoint intervals such that $E \cup F = K_1 \cup \cdots \cup K_n$. We see that $I_1 \cap K_k, \cdots, I_m \cap K_k, J_1 \cap K_k, \cdots, J_{m'} \cap K_k$ is a partition of K_k for every $1 \le k \le n$. Thus we have

$$\begin{aligned}
\mu_0(E \cup F) &= \sum_{k=1}^{n} |K_k|_F \\
&= \sum_{k=1}^{n} \Big(\sum_{i=1}^{m} |I_i \cap K_k|_F + \sum_{j=1}^{m'} |J_j \cap K_k|_F \Big) \\
&= \sum_{i=1}^{m} \sum_{k=1}^{n} |I_i \cap K_k|_F + \sum_{j=1}^{m'} \sum_{k=1}^{n} |J_j \cap K_k|_F \\
&= \sum_{i=1}^{m} |I_i|_F + \sum_{j=1}^{m'} |J_j|_F \\
&= \mu_0(E) + \mu_0(F).
\end{aligned}$$

This shows that μ_0 is finitely additive.

iv) μ_0 is a pre-measure:

We now claim that μ_0 is a pre-measure; thus we suppose that $E \in \mathcal{B}_0$ is the disjoint union of countably many sets $E_1, E_2, \cdots \in \mathcal{B}_0$, and wish to show that

$$\mu_0(E) = \sum_{n=1}^{\infty} \mu_0(E_n).$$

By splitting up E into intervals and then intersecting each of the E_n with these intervals and using finite additivity, we may assume that E is a single interval. By splitting up the E_n into their component intervals and using finite additivity, we may assume that E_n are also individual intervals. By finite additivity, we have

$$\mu_0(E) \geq \sum_{n=1}^{N} \mu_0(E_n)$$

for any N, so it suffices to show that

$$\mu_0(E) \leq \sum_{n=1}^{\infty} \mu_0(E_n).$$

We check the inner regularity that

$$\mu_0(E) = \sup_{K \subset E} \mu_0(K) \tag{1.14}$$

where K ranges over all compact intervals contained in E. From monotonicity, which follows from finite additivity, it is suffices to show that

$$\mu_0(E) \leq \sup_{K \subset E} \mu_0(K).$$

Suppose that E is bounded. Let $\overline{E}$ be the closure of E, and consider $\overline{E} \setminus E$.