Second Course in Analysis

Hsu-Hsiang Tsao

Spring of 2024

Contents

1 Multivariable Calculus

1.1 Linear Algebra

Definition 1.1.1 (Notations: $\mathcal{L}$, $\mathcal{M}$, n, m)**.** *Suppose $m, n \in \mathbb{N}$ and S, W are some sets. We denote by*

(i) $\mathcal{M}^{m,n}(S)$ the collection of all m-by-n matrix with entries in S;

(ii) $\mathcal{L}(S, W)$ the set of all linear transformations from S to W.

When no misunderstanding may arise and sufficient context is supplied, we denote, more briefly, $\mathcal{M}^{m,n}(\mathbb{R})$ and $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ by $\mathcal{M}$ and $\mathcal{L}$ respectively.

Unless otherwise stated, we assume that $m, n \in \mathbb{N}$.

Definition 1.1.2 (Transpose of a matrix)**.** *Suppose $A = (a_{k,l})_{k,l=1}^{m,n} \in \mathcal{M}^{m,n}(S)$ for some set S. Then, the transpose of A is the matrix*

$$A^T = (b_{i,j})_{i,j=1}^{n,m} \in \mathcal{M}^{n,m}(S)$$

where $b_{i,j} = a_{j,i}$ $\forall (i,j) \in \{1, \ldots n\} \times \{1, \ldots, m\}$. That is,

$$A^T = \begin{bmatrix} b_{1,1} & b_{1,2} & \ldots & b_{1,m} \\ b_{2,1} & b_{2,2} & \ldots & b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \ldots & b_{n,m} \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{2,1} & \ldots & a_{m,1} \\ a_{1,2} & a_{2,2} & \ldots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \ldots & a_{m,n} \end{bmatrix}.$$

Definition 1.1.3 (Matrix transformation)**.** *Suppose $m, n \in \mathbb{N}$ and $A \in \mathcal{M}^{m,n}(\mathbb{R})$. Then, the matrix transformation represented by A is the map $T_A \colon \mathbb{R}^n \to \mathbb{R}^m$ defined by*

$$T_A(v) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} v_j e_i \ \forall v \in \mathbb{R}^n,$$

where $v = \sum v_j e_j \in \mathbb{R}^n$ and $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{R}^n$.

Equivalently, T_A is defined by the matrix multiplication

$$T_A(v) = \left(A \cdot v^T\right)^T \ \ \forall v \in \mathbb{R}^n,$$

where $\forall v \in \mathbb{R}^n$ we treat v strictly as a 1-by-n matrix with entries in $\mathbb{R}$.

Proposition 1.1.1. *Matrix transformations are linear transformations.*

Proposition 1.1.2. *Suppose $m, n \in \mathbb{N}$. Then, $\mathcal{M}^{m,n}(\mathbb{R})$ is a vector space with* $\dim \mathcal{M}^{m,n}(\mathbb{R}) = mn$.

Proposition 1.1.3. *Suppose $m, n \in \mathbb{N}$. Then, $\mathcal{M}^{m,n}(\mathbb{R})$ and $\mathbb{R}^{mn}$ are isomorphic.*

Proposition 1.1.4. *Suppose $m, n \in \mathbb{N}$. Then, $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a vector space and* $\dim \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = nm$.

Proposition 1.1.5 (Canonical Isomorphism Induced by Matrix Transformation: $\mathcal{T}$)**.** *Suppose $m, n \in \mathbb{N}$. Then, $\mathcal{T} \colon \mathcal{M}^{m,n}(\mathbb{R}) \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ defined by $\mathcal{T}(A) = T_A \ \forall A \in \mathcal{M}^{m,n}(\mathbb{R})$ is an isomorphism*

Theorem 1.1.1 (Composition of Matrix Transformations)**.** *Suppose $m, k, n \in \mathbb{N}$, $A \in \mathcal{M}^{m,k}(\mathbb{R})$, and $B \in \mathcal{M}^{k,n}(\mathbb{R})$. Then, $T_A \circ T_B = T_{AB}$.*

Definition 1.1.4 (Norm on a vector space over $\mathbb{R}$). *Suppose V is a vector space over field $\mathbb{R}$. A norm on V is a map $||_V : V \to \mathbb{R}$ satisfying the following properties:*

(i) $|v|_V \geq 0 \ \forall v \in V$ with $|v|_V = 0 \iff v = 0_V$;

(ii) $|kv|_V = |k| \, |v|_V \ \ \forall k \in \mathbb{R}, v \in V$.

(iii) $|v + w|_V \leq |v|_V + |w|_V \ \ \forall v, w \in V$.

When no misunderstanding may arise and sufficient context is supplied, we may denote $||_V$, more briefly, by $||$.

Proposition 1.1.6 (Common Norms on $\mathbb{R}^n$). *Let $n \in \mathbb{N}$. Then, the following maps are norms on $\mathbb{R}^n$:*

(Euclidean Norm or l_2 norm) $||_2 : \mathbb{R}^n \to \mathbb{R}$ where

$$|x|_2 = \sqrt{\sum_{i=1}^{n} {x_i}^2} \ \forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

(Supremum norm or l_∞ norm) $||_\infty : \mathbb{R}^n \to \mathbb{R}$ where

$$|x|_\infty = \max \{|x_i| : i \in \{1, \ldots, n\}\} \ \ \forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

(l_1 norm) $||_1 : \mathbb{R}^n \to \mathbb{R}$ where

$$|x|_1 = \sum_{i=1}^{n} |x_i| \ \forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Definition 1.1.5 (Normed space). *A normed space is a vector space V along with a norm $||$ defined on V.*

Proposition 1.1.7 (Norm-Induced Metric; Normed Spaces are Metric Spaces). *Suppose $(V, ||)$ is a normed space. Then,*

(i) $d \colon V \times V \to \mathbb{R}$ defined by $d(v, w) = |v - w|$, $\forall (v, w) \in V \times V$, is a metric on V.

(ii) (v, d) is a metric space.

Definition 1.1.6 (Banach space). *A vector space is a Banach space if it is a complete normed space.*

Definition 1.1.7 (Operator norm and bounded operator). *Suppose V, W are normed spaces and $T \in \mathcal{L}(V, W)$. Then, the operator norm of T is*

$$||T|| = \sup \left\{ \frac{|T(v)|_W}{|v|_V} : v \neq 0_V \right\}.$$

An operator is bounded if its operator norm is finite.

Theorem 1.1.2 (Operator Norm Identity). *Suppose V, W are normed spaces where $T \in \mathcal{L}(V, W)$. Then,*

$$\begin{aligned} ||T|| &= \sup \{|T(v)| : |v| < 1\} \\ &= \sup \{|T(v)| : |v| \leq 1\} \\ &= \sup \{|T(v)| : |v| = 1\} \\ &= \inf \{M > 0 : v \in V \implies |T(v)| \leq M \, |v|\}. \end{aligned}$$

Proof. (INCOMPLETE) $\square$

Proposition 1.1.8 (Operator Norm Properties)**.** *Suppose V, W are normed spaces and $T \in \mathcal{L}(V, W)$. Then, the following statements hold:*

(i) $||T|| \geq 0$*;*

(ii) $||T|| = 0 \iff T = 0_{\mathcal{L}(V,W)}$*;*

(iii) Suppose U is a normed space and $S \in \mathcal{L}(U, V)$. Then, $||T \circ S|| \leq ||T|| \, ||S||$.

Proof. (ii) Suppose V, W are normed spaces and $T \in \mathcal{L}(V, W)$.

($\implies$) Suppose that $||T|| = 0$. It follows that $\frac{|T(v)|_W}{|v|_V} \leq 0$, $\forall v \in V/\{0_V\}$. We note that, by the definition of a norm, $|T(v)|_W \, , |v|_V \geq 0 \; \forall v \in V$. Thus, we obtain that, $\forall v \in V/\{0_V\}$,

$$\begin{aligned} \frac{|T(v)|_W}{|v|_V} \geq 0 \implies \frac{|T(v)|_W}{|v|_V} = 0 &\implies |T(v)|_W = 0 \\ &\implies T(v) = 0_W. \qquad \text{(By the definition of norm)} \end{aligned}$$

In addition, certainly $T(0_V) = 0_W$. Hence, we proved that $T(v) = 0_W$, $\forall v \in V$. That is, we show that $T = 0_{\mathcal{L}(V,W)}$, as desired.

($\impliedby$) Suppose $T = 0_{\mathcal{L}(V,W)}$. It follows that $T(v) = 0_W \; \forall v \in V$. Thus, we have that $|T(v)|_W = 0$, $\forall v \in V$, by the definition of a norm. As an immediate result, we obtain that

$$||T|| = \sup \left\{ \frac{|T(v)|_W}{|v|_V} : v \neq 0_V \right\} = \sup \{0 : v \neq 0_V\} = 0.$$

$\square$

Theorem 1.1.3 ($\mathcal{L}(V, W)$ is a Normed Space)**.** *Suppose V and W are normed spaces. Then, $\mathcal{L}(V, W)$ along with operator norm $|| \; || : \mathcal{L} \to \mathbb{R}$ is a normed space.*

Definition 1.1.8 (Comparability of norms)**.**

Proposition 1.1.9 (Comparability Induces an Equivalence Relation on the Set of Norms)**.**

Theorem 1.1.4 (All Norms on $\mathbb{R}^n$ are Comparable)**.**

Corollary 1.1.1 (Norms on Finite-Dimensional Normed Space are Comparable)**.**

Theorem 1.1.5 (Finite Operator Norm and Equivalent Conditions)**.** *Suppose V, W are normed spaces and $T \in \mathcal{L}(V, W)$. Then, the following conditions are equivalent:*

(i) $||T|| < \infty$*;*

(ii) T is uniformly continuous;

(iii) T is continuous;

(iv) T is continuous at the origin (that is, at 0_V).

Proof. (INCOMPLETE) Suppose V, W are normed spaces and $T \in \mathcal{L}(V, W)$.

($i \implies ii$): Suppose $||T|| < \infty$. By definition, we have that $\exists m > 0$ such that $m = ||T||$

$$\frac{|T(v)|_W}{|v|_V} \leq ||T|| \;\; \forall v \in V/\{0_V\} \implies |T(v)|_W \leq ||T|| \, |v|_V \;\; \forall v \in V/\{0_V\}$$

Consider arbitrary $v, w \in V$.

$(ii \implies iii)$ & $(iii \implies iv)$: Trivial.

$(iv \implies i)$: Suppose T is continuous at 0_V. $\square$

Theorem 1.1.6 (Characteristics of Linear Maps on Normed Spaces)**.** *Suppose $T \in \mathcal{L}(\mathbb{R}^n, W)$, where W is a normed space. Then,*

(i) T is continuous, and

(ii) T is an isomorphism implies T is a homeomorphism.

Proof. (INCOMPLETE) $\square$

Corollary 1.1.2. *Suppose V, W are finite-dimensional normed spaces. Then,*

(i) $T \in \mathcal{L}(V, W) \implies T$ is continuous, and

(ii) $\phi \in \mathcal{L}(V, W)$ is an isomorphism implies ϕ is a homeomorphism.

Proof. (INCOMPLETE) $\square$

Corollary 1.1.3. *(i) Suppose V is a finite-dimensional normed space with norms $||_a$ and $||_b$. Then, the identity map I on V is a homomorphism between the normed spaces $(V, ||_a)$ and $(V, ||_b)$.*

(ii) $\mathcal{T} \colon \mathcal{M} \to \mathcal{L}$ is a homeomorphism.

Proof. (INCOMPLETE) $\square$

Definition 1.1.9 (Conorm)**.**

Exercise 1.1.1 (Determine an Operator Norm)**.** *Consider the dilation map $T \colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(x, y) = (2x, y)$. Prove $||T|| = 2$.*

Proof. Outline: show $||T|| \leq 2$.

Find $(a, b) \in \mathbb{R}^2$ such that $\frac{|T(a,b)|}{|(a,b)|} = 2$.

Use the definition of sup to prove the conclusion, $\square$

1.2 Derivatives

Definition 1.2.1 ((Total) Derivative)**.** *Let $n, m \in \mathbb{N}$ and $U \subset \mathbb{R}^n$ be open. Suppose $f\colon U \to \mathbb{R}^m$.*

(i) The ***derivative (or total derivative)*** $(Df)_p$ ***of*** f ***at*** $p \in U$ *is a map, if it exists, $T\colon \mathbb{R}^n \to \mathbb{R}^m$ such that*

(a) T is a linear map, and

(b) T satisfies

$$f(p+v) = f(p) + T(v) + R(v) \text{ for all sufficiently small } v \in \mathbb{R}^n$$
$$\implies \lim_{|v|\to 0} \frac{R(v)}{|v|} = 0_{\mathbb{R}^m} \tag{Pugh}$$

or, equivalently,

$$f(p+v) = f(p) + T(v) + R(v) \text{ for all sufficiently small } v \in \mathbb{R}^n$$
$$\implies \lim_{v\to 0_{\mathbb{R}^n}} \frac{|R(v)|}{|v|} = 0, \tag{Rudin}$$

or, equivalently,

$$\lim_{v\to 0_{\mathbb{R}^n}} \frac{|f(p+v) - T(v) - f(p)|}{|v|} = 0, \tag{Rudin}$$

where $R(v) \in \mathbb{R}^m$ denotes the Taylor remainder for $f(p+v)$.

(ii) We say that f is ***differentiable at*** $p \in U$ *if $(Df)_p$ exists, and f is* ***differentiable*** *if f is differentiable at p, $\forall p \in U$.*

(iii) Let $E = \{p \in U : (Df)_p \text{ exists}\}$. We call the map $Df\colon E \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, defined by $[Df](p) = (Df)_p$ $\forall p \in E$, the ***derivative (or total derivative) of*** *f.*

Note that we may also denote Df by f'.

Remark 1.2.1. *Recall that $E \subset \mathbb{R}^n$ is open implies that, $\forall p \in E$, $\exists r_p > 0$ such that $q \in \mathbb{R}^n$ and $d(p,q) < r_p \implies q \in E$; that is, $N_{r_p}(p) \subset E$. In the above definition, by sufficiently small $v \in \mathbb{R}^n$, we mean that v is such that $d(p, p+v) < r_p$ so $p + v \in E$.*

Remark 1.2.2. *The choice of T is unique, since a limit is unique, provided it exists. See proof below.*

Definition 1.2.2 (Notations: e_i, u_j, f_j)**.** *Let $n, m \in \mathbb{N}$. Denote the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ by $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_m\}$, respectively.*

Suppose $U \subset \mathbb{R}^n$, $f\colon U \to \mathbb{R}^m$, and $\exists f_1, \ldots, f_m\colon U \to \mathbb{R}$ such that

$$f(p) = \sum_{j=1}^{m} f_j(p) e_j \ \forall p \in U.$$

Unless otherwise stated, we denote e_i the i-th standard basis vector of $\mathbb{R}^n$, $\forall i \in \{1, \ldots, n\}$, and u_j the j-th standard basis vector of $\mathbb{R}^m$, $\forall j \in \{1, \ldots, m\}$.

Similarly, we denote f_j the j-th component of f, $\forall j \in \{1, \ldots, m\}$.

Definition 1.2.3 (Partial derivative)**.** *Let $n, m \in \mathbb{N}$ and $U \subset \mathbb{R}^n$ be open, and denote $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_m\}$ the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Suppose $f\colon U \to \mathbb{R}^m$ and $f(x) = \sum_{i=1}^m f_i(x)u_i \ \forall x \in U$, where $f_j\colon U \to \mathbb{R}\ \forall j \in \{1, \ldots, m\}$.*

Suppose that $p \in U$, $i \in \{1, \ldots, m\}$, and $j \in \{1, \ldots, n\}$.

(i) The (i, j)-partial derivative or ij^{th} ***partial derivative of** f **at*** *$p \in U$ is*

$$\frac{\partial f_i(p)}{\partial x_j} = \lim_{t\to 0} \frac{f_i(p + te_j) - f_i(p)}{t} \in \mathbb{R},$$

provided the limit exists.

(ii) Let $E = \left\{p \in U : \frac{\partial f_i(p)}{\partial x_j} \text{ exists}\right\}$. We call the map $\frac{\partial f_i}{\partial x_j}\colon E \to \mathbb{R}$, defined by

$$\left[\frac{\partial f_i}{\partial x_j}\right](p) = \frac{\partial f_i(p)}{\partial x_j}\ \forall p \in E,$$

the ***(i,j)-partial derivative of*** *f.*

(iii) We call, $\forall j \in \{1, \ldots, n\}$,

$$\partial_{x_j} f(p) = (\partial_{x_j} f_1(p), \partial_{x_j} f_2(p), \ldots, \partial_{x_j} f_m(p)) \in \mathbb{R}^m$$

the partial derivative of f at p with respect to x_j, provided the individual (i, j)-partial derivatives of f at p exist.

(iv) We call, $\forall j \in \{1, \ldots, n\}$, the map $[\partial_{x_j} f]\colon E \to \mathbb{R}^m$ defined by

$$[\partial_{x_j} f](p) = \partial_{x_j} f(p)\ \forall p \in E$$

the partial derivative of f with respect to x_j.

We may also denote $\frac{\partial f_i}{\partial x_j}$, more briefly by $D_j f_i$ or $\partial_{x_j} f_i$. Similarly, we may denote $\frac{\partial f}{\partial x_j}$, more briefly by $D_j f$ or $\partial_{x_j} f$.

Definition 1.2.4 (Directional derivative)**.** *Let $n, m \in \mathbb{N}$ and $U \subset \mathbb{R}^n$ be open. Suppose that $f\colon U \to \mathbb{R}^m$, $p \in U$, and $v \in \mathbb{R}^n$. If the limit $\lim_{t\to 0} \frac{f(p+tv)-f(p)}{t}$ exists in $\mathbb{R}^m$, then*

(i) we say f is differentiable in the direction of v at p, and

(ii) we denote

$$D_v(p) = \lim_{t\to 0} \frac{f(p + tv) - f(p)}{t}$$

the directional derivative of f at p in the direction of v.

Theorem 1.2.1 (Differentiability Implies Continuity)**.** *Suppose $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}^m$ is differentiable at $p \in U$. Then, f is continuous at p.*

Proof. Suppose $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}^m$ is differentiable at $p \in U$. Suppose further that

$$f(p + v) = f(p) + (Df)_p(v) + R(v),\ \forall v \in \mathbb{R}^n \text{ such that } p + v \in U. \qquad (*)$$

By definition, we have that $(Df)_p \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $\lim_{|v|\to 0} \frac{R(v)}{|v|} = 0_{\mathbb{R}^m}$. We note that $(Df)_p$ is continuous since linear maps from one normed space to another are continuous. As an immediate result, we have that $||(Df)_p||$ is finite by a previous theorem. We observe that

$$\lim_{|v|\to 0} \left[||(Df)_p|| + \frac{|R(v)|}{|v|}\right] = ||(Df)_p|| \text{ and } \lim_{|v|\to 0} |v| = 0 \implies$$

$$\lim_{|v|\to 0} ||(Df)_p|| \cdot |v| + |R(v)| = \lim_{|v|\to 0} \left(||(Df)_p|| + \frac{|R(v)|}{|v|}\right) \cdot |v| = ||(Df)_p|| \cdot 0 = 0.$$

By definition, we obtain that $\forall \epsilon > 0 \; \exists \delta > 0$ such that $0 < |v| < \delta$ implies that

$$\begin{aligned}
&\big| ||(Df)_p|| \cdot |v| + |R(v)| \big| = ||(Df)_p|| \cdot |v| + |R(v)| < \epsilon \\
\implies &\epsilon > ||(Df)_p|| \cdot |v| + |R(v)| \geq |(Df)_p(v)| + |R(v)| && \text{(By the definition of } ||(Df)_p||\text{)} \\
&\geq |(Df)_p(v) + R(v)|\,. && \text{(By the Triangle Inequality)}
\end{aligned}$$

Thus, it holds that $\forall \epsilon > 0 \; \exists \delta > 0$ such that $p + v \in U$ and $0 < |(p+v) - p| = |v| < \delta$ implies that

$$|f(p+v) - f(p)| = |(Df)_p(v) + R(v)| < \epsilon. \qquad \text{(By (*))}$$

Hence, f is continuous at p by definition. (We note that if we replace $p + v$ with x, the above statement resembles precisely the familiar $\delta - \epsilon$ definition for the continuity of a function at a point) $\square$

Theorem 1.2.2 (Characterization of Derivative at a Point)**.** *Let $n, m \in \mathbb{N}$ and $U \subset \mathbb{R}^n$ be open. Suppose $f \colon U \to \mathbb{R}^m$ is differentiable at $p \in U$. Then, $(Df)_p \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is given by*

$$(Df)_p(u) = \lim_{t \to 0} \frac{f(p+tu) - f(p)}{t} \;\; \forall u \in \mathbb{R}^n.$$

Proof. Let $n, m \in \mathbb{N}$ and $U \subset \mathbb{R}^n$ be open. Suppose $f \colon U \to \mathbb{R}^m$ is differentiable at $p \in U$.

(Case 1): We observe that $(Df)_p(0_{\mathbb{R}^n}) = 0^{\mathbb{R}^m}$ since $(Df)_p$ is linear. It follows that, for $u = 0_{\mathbb{R}^n}$,

$$(Df)_p(u) = 0 = \lim_{t \to 0} \frac{f(p+tu) - f(p)}{t}$$

as desired.

(Case 2): Consider arbitrary $u \in \mathbb{R}^n / \{0_{\mathbb{R}^n}\}$ and suppose $\forall t \in \mathbb{R}$ such that $p + tu \in U$ we have that

$$f(p+tu) = f(p) + (Df)_p(tu) + R(tu).$$

By the differentiability of f at p, we obtain that

$$\begin{aligned}
&\lim_{|tu| \to 0} \frac{R(tu)}{|tu|} = 0_{\mathbb{R}^m} \implies \lim_{|t| \to 0} \frac{R(tu)}{|tu|} = 0_{\mathbb{R}^m} && \text{(Since } |u| \text{ is fixed)} \\
\implies &\forall \epsilon > 0 \; \exists \delta > 0 \text{ such that } 0 < |t| < \delta \implies \left| \frac{R(tu)}{|tu|} \right| = \frac{|R(tu)|}{|tu|} < \epsilon. && (*)
\end{aligned}$$

Consider arbitrary $\epsilon > 0$. Then, $\frac{\epsilon}{|u|} > 0$ since $u \neq 0_{\mathbb{R}^n}$ by assumption. By $(*)$, we obtain that

$$\begin{aligned}
\exists \delta > 0 \text{ such that } 0 < |t| < \delta \implies &\left| \frac{R(tu)}{t\,|u|} \right| = \frac{|R(tu)|}{|tu|} < \frac{\epsilon}{|u|} \\
\implies &\left| \frac{R(tu)}{t\,|u|} \cdot |u| \right| < \epsilon. \\
\therefore &\lim_{t \to 0} \frac{R(tu)}{t\,|u|} \, |u| = 0_{\mathbb{R}^m}. && (**)
\end{aligned}$$

Now, consider the limit $\lim_{t \to 0} \frac{f(p+tu)-f(p)}{t}$. We observe that

$$\begin{aligned}
\lim_{t \to 0} \frac{f(p+tu) - f(p)}{t} &= \lim_{t \to 0} \frac{(Df)_p(tu) + R(tu)}{t} && \text{(By assumption)} \\
&= \lim_{t \to 0} \frac{t(Df)_p(u) + R(tu)}{t} && \text{(Since } (Df)_p \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)\text{)} \\
= \lim_{t \to 0} \left[(Df)_p(u) + \frac{R(tu)}{t} \right] &= \lim_{t \to 0} \left[(Df)_p(u) + \frac{R(tu)}{t\,|u|} \, |u| \right] \\
&= (Df)_p(u) && \text{(By } (**)\text{)}
\end{aligned}$$

That is, we show that $\forall u \in \mathbb{R}^n$, $(Df)_p(u) = \lim_{t \to 0} \frac{f(p+tu)-f(p)}{t}$ and complete this proof. $\square$

Corollary 1.2.1 (Uniqueness of Total Derivative)**.** *Suppose $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}^m$ is differentiable at $p \in U$. Then, $T, \tilde{T} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ both satisfy the derivative properties implies that $T = \tilde{T}$.*

Proof. Suppose $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}^m$ is differentiable at $p \in U$. Suppose further that $T, \tilde{T} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ both satisfy the derivative properties. By the preceding theorem, we obtain that

$$T(u) = \lim_{t \to 0} \frac{f(p + tu) - f(p)}{t} = \tilde{T}(u) \ \forall u \in \mathbb{R}^n,$$

which implies, by definition, $T = \tilde{T}$. □

Theorem 1.2.3 (Existence of Total Derivative Implies the Existence of Partial Derivative)**.** *Let $n, m \in \mathbb{N}$ and $U \subset \mathbb{R}^n$ be open. Suppose $f\colon U \to \mathbb{R}^m$ is differentiable at $p \in U$. Then, all partial derivatives of f at p exist and they are the entries of the matrix that represents the total derivative $(Df)_p$ at p. That is,*

(i) $\forall (i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$, $\partial x_j f_i(p) \in \mathbb{R}$, and

(ii) $(Df)_p = T_{A_p}$, where $\forall (i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$ we have that $(A_p)_{i,j} = \partial_{x_j} f_i(p)$ and

$$A_p = \begin{bmatrix} \partial_{x_1} f_1(p) & \partial_{x_2} f_1(p) & \dots & \partial_{x_n} f_1(p) \\ \partial_{x_1} f_2(p) & \partial_{x_2} f_2(p) & \dots & \partial_{x_n} f_2(p) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(p) & \partial_{x_2} f_m(p) & \dots & \partial_{x_n} f_m(p) \end{bmatrix} \in \mathcal{M}^{m,n}(\mathbb{R}).$$

Proof. Outline: Apply the preceding theorem and let u be the standard basis vectors in $\mathbb{R}^n$. □

Proposition 1.2.1 (Derivative Identities)**.** *Suppose $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}^m$. Suppose further that $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ and f is differentiable at $p \in U$. Then,*

(i) $\forall j \in \{1, \dots, n\}$, $(Df)_p(e_j) = \partial_{x_j} f(p) \in \mathbb{R}^m$;

(ii) $(Df)_p(v) = D_v f(p) = \sum_{j=1}^{n} v_j \cdot \partial_{x_j} f(p) \in \mathbb{R}^m$.

(iii) $A \in \mathcal{M}^{m,n}(\mathbb{R})$ representing $(Df)_p$ has the identity

$$A = \begin{bmatrix} \partial_{x_1} f_1(p) & \partial_{x_2} f_1(p) & \dots & \partial_{x_n} f_1(p) \\ \partial_{x_1} f_2(p) & \partial_{x_2} f_2(p) & \dots & \partial_{x_n} f_2(p) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(p) & \partial_{x_2} f_m(p) & \dots & \partial_{x_n} f_m(p) \end{bmatrix} = \begin{bmatrix} \partial_{x_1} f(p)^T & \partial_{x_2} f(p)^T & \dots & \partial_{x_n} f(p)^T \end{bmatrix}.$$

Proof. Suppose $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}^m$. Suppose further that $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ and f is differentiable at $p \in U$.

(i) By Corollary 7 (Pugh 284), we obtain that $(Df)_p = T_A$, where T_A is the matrix transformation induced by

$$A = \begin{bmatrix} \partial_{x_1} f_1(p) & \partial_{x_2} f_1(p) & \dots & \partial_{x_n} f_1(p) \\ \partial_{x_1} f_2(p) & \partial_{x_2} f_2(p) & \dots & \partial_{x_n} f_2(p) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(p) & \partial_{x_2} f_m(p) & \dots & \partial_{x_n} f_m(p) \end{bmatrix} \in \mathcal{M}^{m,n}(\mathbb{R}).$$

Therefore, we then have, $\forall j \in \{1, \ldots, n\}$,

$$\begin{aligned}(Df)_p(e_j) &= T_A(e_j) = \left(A \cdot {e_j}^T\right)^T \\ &= \left(\partial_{x_j} f_1(p) \quad \partial_{x_j} f_2(p) \quad \ldots \quad \partial_{x_j} f_m(p)\right) = \partial_{x_j} f(p). \qquad \text{(By definition)}\end{aligned}$$

(ii) We observe that $\forall v \in \mathbb{R}^n$

$$\begin{aligned}(Df)_p(v) &= \lim_{t \to 0} \frac{f(p+tv) - f(p)}{t} && \text{(By Theorem 5 (Pugh 283))} \\ &= D_v f(p). && \text{(By definition of directional derivative)}\end{aligned}$$

In addition, $\forall v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, it follows from the linearity of $(Df)_p$ that

$$\begin{aligned}(Df)_p(v) = (Df)_p\left(\sum_{j=1}^n v_j e_j\right) &= \sum_{j=1}^n v_j [(Df)_p](e_j) \\ &= \sum_{j=1}^n v_j \cdot \partial_{x_j} f(p). && \text{(By Part (a))}\end{aligned}$$

(iii) The statement follows directly from the definition of the $\partial_{x_1} f(p), \partial_{x_2} f(p), \ldots, \partial_{x_n} f(p)$. ☐

Proposition 1.2.2. *Suppose $U \subset \mathbb{R}$ is an open interval and $f \colon U \to \mathbb{R}$. Then, $f'(x) \in \mathbb{R} \iff f$ is differentiable at $x \in U$.*

Proof. Suppose $U \subset \mathbb{R}$ is an open interval and $f \colon U \to \mathbb{R}$.

($\Longrightarrow$) Suppose that $f'(x) \in \mathbb{R}$, where $x \in U$. Suppose further that $f(x+v) = f(x) + f'(x)v + R(v)$ for $v \in \mathbb{R}$ such that $x + v \in U$. By definition, we have that

$$\lim_{v \to 0} \frac{f(x+v) - f(x)}{v} = f'(x) \in \mathbb{R},$$

which is equivalent to

$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ such that } |v| < \delta \implies \left|\frac{f(x+v) - f(x) - f'(x)v}{v}\right| < \epsilon,$$

which is also equivalent to

$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ such that } ||v| - 0| < \delta \implies \left|\frac{R(v)}{v} - 0\right| = \left|\frac{R(v)}{|v|} - 0\right| < \epsilon,$$

By definition, we have that $\lim_{|v| \to 0} \frac{R(v)}{|v|} = 0$ and, hence, f is differentiable at x with $(Df)_x \colon \mathbb{R} \to \mathbb{R}$ defined by

$$[(Df)_x](r) = f'(x) r \ \forall r \in \mathbb{R}.$$

($\Longleftarrow$) Suppose f is differentiable at $x \in U$. Then, ... ☐

Theorem 1.2.4 (Existence and continuity of Partial Derivatives Imply Existence of Total Derivative)**.** *Suppose $U \subset \mathbb{R}^n$ is open, and $f \colon U \to \mathbb{R}^m$. Suppose further that $\forall (i,j) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$ it holds that*

(i) $\forall p \in U$, $\partial_{x_j} f_i(p) \in \mathbb{R}$, and

(ii) $\partial_{x_j} f_i$ is continuous.

Then, f is differentiable on U.

Proof. (INCOMPLETE)

Outline 1 Let A be the matrix whose entries are the partial derivative.

2. Let T be the linear map represented by A

□

Definition 1.2.5 (Bilinear map)**.** *Suppose V, W, Z are vector spaces. Then, a map $B\colon V \times W \to Z$ is bilinear if*

(i) $\forall v \in V$, $B(v, \cdot)\colon W \to Z$ defined by $[B(v, \cdot)](w) = B(v, w)$ $\forall w \in W$ is linear, and

(ii) $\forall w \in W$, $B(\cdot, w)\colon V \to Z$ defined by $[B(\cdot, w)](v) = B(v, w)$ $\forall v \in V$ is linear.

Proposition 1.2.3 (Examples of Bilinear Maps)**.** *The usual multiplication on $\mathbb{R}$, dot product on $\mathbb{R}^n$, and matrix product are bilinear maps.*

Theorem 1.2.5 (Differentiation Rules)**.** ***(Linearity)*** *Suppose $c \in \mathbb{R}$, $U \subset \mathbb{R}^n$ is open, and $f\colon U \to \mathbb{R}^m$ and $g\colon U \to \mathbb{R}^m$ are differentiable at $p \in U$. Then, $f + cg$ is differentiable at p and*

$$(D(f + cg))_p = (Df)_p + c(Dg)_p.$$

(Chain Rule) *Suppose $f\colon U \to \mathbb{R}^m$ is differentiable at $p \in U$, and $W \subset \mathbb{R}^m$ is open with $f(U) \subset W$. Suppose further that $g\colon W \to \mathbb{R}^r$ is differentiable at $f(p)$. Then, $g \circ f\colon U \to \mathbb{R}^r$ is differentiable at p and*

$$(D[g \circ f])_p = (Dg)_{f(p)} \circ (Df)_p \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r).$$

(Leibniz Rule) *Suppose $\bullet\colon \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}^q$ is bilinear, where $f\colon U \to \mathbb{R}^m$ and $g\colon U \to \mathbb{R}^r$ are differentiable at $p \in U$. Then, $f \bullet g\colon U \to \mathbb{R}^q$ defined by*

$$[f \bullet g](u) = f(u) \bullet g(u) = \bullet(f(u), g(u)),\ \forall u \in U$$

is differentiable at p and

$$[(D(f \bullet g))_p](v) = [Df]_p(v) \bullet g(p) + f(p) \bullet [Dg]_p(v)\ \forall v \in U.$$

(Constant Map and Linear Map) *(i) Suppose $U \subset \mathbb{R}^n$ is open and $c \in \mathbb{R}^m$. Define $c_{\mathbb{R}}\colon U \to \mathbb{R}^m$ by $c_{\mathbb{R}}(u) = c$ $\forall u \in U$. Then, $c_{\mathbb{R}}$ is differentiable and $\forall p \in U$ $(Dc_{\mathbb{R}})_p = 0_{\mathbb{R}}$, where $0_{\mathbb{R}}\colon \mathbb{R}^n \to \mathbb{R}^m$ is defined by $0_{\mathbb{R}}(u) = 0_{\mathbb{R}^m}$ $\forall u \in U$.*

(ii) Suppose $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Then, T is differentiable and $\forall p \in \mathbb{R}^n$ $(DT)_p = T$.

Proof. (INCOMPLETE) □

Theorem 1.2.6 (Differentiability of Vector Function $\iff$ Component-wise Differentiability)**.** *Let $n, m \in \mathbb{N}$, $U \subset \mathbb{R}^n$ is open, and $f\colon U \to \mathbb{R}^m$. Then, f is differentiable at $p \in U \iff f_j$ is differentiable at p, $\forall j \in \{1, \ldots, m\}$.*

In addition, f is differentiable at $p \in U$ implies that

$$(Df_i)_p = \pi_i \circ (Df)_p\ \forall i \in \{1, \ldots, m\}$$

and

$$(Df)_p = \sum_{i=1}^{m} e_i (Df_i)_p = ((Df_1)_p, (Df_2)_p, \ldots, (Df_m)_p),$$

where $\forall i \in \{1, \ldots, m\}$ $\pi_i\colon \mathbb{R}^m \to \mathbb{R}$ is the projection map defined by $\pi_i(w_1, \ldots, w_m) = w_i$.

Proof. (INCOMPLETE) ☐

Definition 1.2.6 (Segment in $\mathbb{R}^n$)**.** *Let $p, q \in \mathbb{R}^n$. Then, the segment $[p, q]$ in $\mathbb{R}^n$ is*

$$[p, q] = \{(1 - \lambda)p + \lambda q : \lambda \in [0, 1]\}.$$

Theorem 1.2.7 (General Mean Value Theorem)**.** *Suppose $U \subset \mathbb{R}^n$ is open, $[p, q] \subset U$, and $f : U \to \mathbb{R}^m$ is differentiable. Then,*

$$|f(p) - f(q)| \leq M\,|p - q|,$$

where $M = sup\,\{||(Df)_q : q \in U||\}$.

Proof. (INCOMPLETE) ☐

Definition 1.2.7 (Integrating a Matrix and a derivative at a point)**.** *Suppose $[a, b] \subset \mathbb{R}$, $m, n \in \mathbb{N}$, $A \in \mathcal{M}^{m,n}(\mathbb{R})$. Then, we define*

(i)

$$\int_a^b Adt = \begin{bmatrix} \int_a^b a_{1,1}dt & \int_a^b a_{1,2}dt & \cdots \int_a^b a_{1,n}dt \\ \int_a^b a_{2,1}dt & \int_a^b a_{2,2}dt & \cdots \int_a^b a_{2,n}dt \\ \vdots & \vdots & \ddots & \vdots \\ \int_a^b a_{m,1}dt & \int_a^b a_{m,2}dt & \cdots \int_a^b a_{m,n}dt \end{bmatrix},$$

where $\forall (i, j) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$ $a_{i,j}$ is the (i,j)-entry of A.

(ii) Suppose $U \subset \mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is differentiable at $p \in U$ with derivative $(Df)_p$ at p. Then, we define

$$\int_b^a (Df)_p dt = \int_b^a B_p dt,$$

where $B_p \in \mathcal{M}^{m,n}(\mathbb{R})$ is the matrix representing $(Df)_p$.

Theorem 1.2.8 (C^1 Mean Value Theorem)**.** *Suppose $U \subset \mathbb{R}^n$ is open, $[p, q] \subset U$, and $f : U \to \mathbb{R}^m \in C^1$. Then,*

$$f(q) - f(p) = T \cdot (q - p)$$

where T is the average derivative of f on the segment $[p, q]$ with

$$T = \int_0^1 (Df)_{p+t(q-p)} dt \in \mathcal{M}^{m,n}(\mathbb{R}).$$

(Converse)

Proof. (INCOMPLETE) ☐

Corollary 1.2.2 (Connectedness, Differentiability, and Vanishing Derivative Implies Constantness)**.** *Suppose $U \subset \mathbb{R}^n$ is open and connected. Suppose further that $f : U \to \mathbb{R}^m$ is differentiable and $\forall p \in U$, $(Df)_p = \tilde{0}$, where $\tilde{0} : \mathbb{R}^n \to \mathbb{R}^m$ is defined by $\tilde{0}(v) = 0_{\mathbb{R}^m}$ $\forall v \in \mathbb{R}^n$. Then, f is contstant.*

Theorem 1.2.9 (Differentiation Past the Integral)**.** *Suppose $[a, b], (c, d) \subset \mathbb{R}$, $f : [a, b] \times (c, d) \to \mathbb{R}$ is continuous, $\frac{\partial f}{\partial y}(x, y) \in \mathbb{R}$ $\forall (x, y) \in [a, b] \times (c, d)$, and $\frac{\partial f}{\partial y} : [a, b] \times (c, d) \to \mathbb{R}$ is continuous. Then,* *(i)* $F\colon (c,d) \to \mathbb{R}$ *defined by*

$$F(y) = \int_a^b f(x,y)dx \ \forall y \in (c,d)$$

is of class C^1 *and*

(ii)

$$F'(y) = \int_a^b \frac{\partial f(x,y)}{\partial y} dx \ \forall y \in (c,d).$$

1.3 Higher Derivatives

Theorem 1.3.1 (Existence of Second Total Derivative Implies the Existence of Other Second Derivatives)**.** *(i) Suppose $U \subset \mathbb{R}^n$ is open, $f\colon U \to \mathbb{R}^m$, and $p \in U$. Suppose that $(D^2 f)_p$ exists. Then, $\forall k \in \{1, \ldots, m\}$*

(a) $(D^2 f_k)_p$ exists,

(b) $\forall i, j \in \{1, \ldots, n\}$ $\frac{\partial^2 f_k}{\partial x_i \partial x_j}(p) \in \mathbb{R}$, and

(c) $\forall i, j \in \{1, \ldots, n\}$

$$(D^2 f_k)_p(e_i, e_j) = \frac{\partial^2 f_k(p)}{\partial_{x_i} \partial_{x_j}}.$$

(ii)

Theorem 1.3.2 (Symmetry of Second Derivative)**.** *Suppose $U \subset \mathbb{R}$ is open, $f\colon U \to \mathbb{R}^m$, and $p \in U$. Suppose further that $(D^2 f)_p$ exists. Then, $(D^2 f)_p$ is symmetric; that is,*

$$\forall v, w \in U \;\; (D^2 f)_p(v, w) = (D^2 f)_p(w, v).$$

Corollary 1.3.1 (Existence of Second-Derivative Implies Equivalence of Second-Mixed Partials)**.**

Corollary 1.3.2 (Existence of r-th Derivative Implies Symetry and Equivalence of r-th Order Mixed Partials)**.**

Definition 1.3.1 (Class C^r)**.** *Suppose $U \subset \mathbb{R}^n$ and $r \in \mathbb{N}$. Then, $f\colon U \to \mathbb{R}^m$ is of class C^r and we write $f \in C^r$ if f is $r-$th order differentiable and $D^r f\colon U \to \mathcal{L}(U, Codomain(D^{r-1} f))$ is continuous.*

Definition 1.3.2 (Smoothness and class C^∞)**.** *Suppose $U \subset \mathbb{R}^n$. Then, $f\colon U \to \mathbb{R}^m$ is smooth or of class C^∞ and we write $f \in C^\infty$ if $f \in C^r$, $\forall r \in \mathbb{N}$.*

Example 1.3.1 (Examples and Non-examples of Smooth Functions)**.** $\cos, \sin, exp$, *and polynomials are smooth functions. Abs and the sign function are not smooth functions.*

Remark 1.3.1. *Let $r \in \mathbb{N}$. By the rules of differentiation, the functions in C^r are closed under the operations of linear combination, product, and composition, if defined.*

Proposition 1.3.1 (Containment Relationship Between the Function Spaces; Smoothness Hierarchy)**.** *Suppose $U \subset \mathbb{R}^n$ is open. Let $C_b = \{f : f \text{ is bounded}\}$, $\mathcal{R} = \{f : f \text{ is Riemann integrable}\}$, $C^0 = \{f : f \text{ is continuous}\}$. Then,*

$$C^\infty \subsetneq \cdots \subsetneq C^2 \subsetneq C^1 \subsetneq C^0 \subsetneq \mathcal{R} \subsetneq C_b.$$

Definition 1.3.3 (Pointwise convergence and uniform Convergence of functions on $\mathbb{R}^n$)**.** *Suppose $U \subset \mathbb{R}^n$, $f\colon U \to \mathbb{R}^m$, and $\forall k \in \mathbb{N}$ $f_k\colon U \to \mathbb{R}^m$. Then,*

(i) (f_k) converges to f pointwise and we write $f_k \to f$ if

$$\forall \epsilon > 0, \forall x \in U \; \exists N(\epsilon, x) \in \mathbb{N} \text{ such that } n > N(\epsilon, x) \implies |f_n(x) - f(x)| < \epsilon;$$

(ii) (f_k) converges to f uniformly on U and we write $f_k \rightrightarrows f$ if

$$\forall \epsilon > 0 \; \exists N(\epsilon) \in \mathbb{N} \text{ such that } x \in U \text{ and } n > N(\epsilon) \implies |f_n(x) - f(x)| < \epsilon.$$

Definition 1.3.4 (Pointwise convergence and uniform convergence of derivatives)**.** *Suppose $r \in \mathbb{N}$, $U \subset \mathbb{R}^n$, $f \colon U \to \mathbb{R}^m$, and $\forall k \in \mathbb{N}$ $f_k \colon U \to \mathbb{R}^m$. Then,*

(i) $(D^r f_k)$ converges to $D^r f$ pointwise and we write $D^r f_k \to D^r f$ if

$$\forall \epsilon > 0, \forall x \in U\ \exists N(\epsilon, x) \in \mathbb{N} \text{ such that } n > N(\epsilon, x) \implies ||(D^r f_k)_x - (D^r f)_x|| < \epsilon;$$

(ii) $(D^r f_k)$ converges to $D^r f$ uniformly on U and we write $D^r f_k \rightrightarrows D^r f$ if

$$\forall \epsilon > 0\ \exists N(\epsilon) \in \mathbb{N} \text{ such that } x \in U \text{ and } n > N(\epsilon) \implies ||(D^r f_k)_x - (D^r f)_x|| < \epsilon.$$

Definition 1.3.5 (Uniformly C^r convergent and Uniformly C^r Cauchy)**.** *Suppose $U \subset \mathbb{R}^n$ is open, $r \in \mathbb{N}$, and (f_k) is a sequence of functions in C^r where $f_k \colon U \to \mathbb{R}^m$ $\forall k \in \mathbb{N}$. Then, (f_k) is*

(i) uniformly C^r convergent if $\exists f \in C^r$ such that $f \colon U \to \mathbb{R}^m$ and

$$f_k \rightrightarrows f,\ Df_k \rightrightarrows Df, \text{ and } \ldots, D^r f_k \rightrightarrows D^r f;$$

(ii) uniformly C^r Cauchy if

$$\begin{aligned} &\forall \epsilon > 0\ \exists N \in \mathbb{N} \text{ such that } n, m \geq N \text{ and } x \in U \\ \implies &|f_n(x) - f_m(x)| < \epsilon,\ ||(Df_n)_x - (Df_m)_x|| < \epsilon,\ \ldots, ||(D^r f_n)_x - (D^r f_m)_x|| < \epsilon. \end{aligned}$$

Remark 1.3.2. *Covergence iff terms arbitrary close to limit*

Cauchyness iff terms arbitrary close to each other

Exercise 1.3.1. *Define, $\forall n \in \mathbb{N}$, $f_n \colon \mathbb{R}^2 \to \mathbb{R}$ by $f_n(x, y) = \sin\left(\frac{x+y}{n}\right)$ $\forall (x, y) \in \mathbb{R}^2$. Prove that (f_k) is not uniformly convergent on $\mathbb{R}^2$.*

Theorem 1.3.3 (Equivalence of Uniformly C^r Convergent and Uniformly C^r Cauchy)**.** *Suppose $U \subset \mathbb{R}^n$ is open, $r \in \mathbb{N}$, and (f_k) is a sequence of functions in C^r where $f_k \colon U \to \mathbb{R}^m$ $\forall k \in \mathbb{N}$. Then, (f_k) is uniformly C^r convergent iff (f_k) is uniformly C^r Cauchy.*

Proof. ($\implies$) We observe that $\forall n, m \in \mathbb{N}$,

$$\begin{aligned} |f_n - f_m| &= |f_n + f - f - f_m| \\ &\leq |f_n - f| + |f - f_m| && \text{(By triange inequality)} \end{aligned}$$

Then, $|f_n - f_m| \to 0$.

($\impliedby$) ☐

Definition 1.3.6 (C^r norm)**.** *Suppose $r \in \mathbb{N}$, $U \subset \mathbb{R}$ is open and $f \colon U \to \mathbb{R}^m \in C^1$. Then, the C^r norm of f is*

$$||f||_r = \max\left\{\sup_{x \in U}\{|f(x)|\}, \sup_{x \in U}\{||(Df)_x||\}, \ldots \sup_{x \in U}\{||D^r f_x||\}\right\}.$$

Theorem 1.3.4 (C^r Norm Induced Banach Space)**.** *Suppose $r \in \mathbb{N}$, $U \subset \mathbb{R}^n$. Then, $(C^r(U, \mathbb{R}^m), ||||_r)$ is a Banach space.*

Theorem 1.3.5 (C^r M-Test)**.**

Remark 1.3.3 (Methods of proving differentiability)**.** *The following is a list of methods through which one may prove the differentiability of a map in Euclidean spaces:*

1. *Via definition*
2. *Via differentiation rules*

3. Show that a map is component-wise differentiable (See Theorem 10 (Pugh 288))
4. Prove the existence and continuity of the partial derivatives of a map (See Theorem 8 (Pugh 284))
5. Via the converse of the C^1 Mean Value Theorem (See Theorem 12 (Pugh 289))
(??)6. Convergence of a sequence of uniformly C^r convergent functions implies the differentiability of the limit function

1.4 Implicit and Inverse Functions

Definition 1.4.1 (Contraction map)**.** *Suppose (M,d) is a metric space. Then, $f\colon M \to M$ is a contraction if*

$$\exists \theta \in [0,1) \text{ such that } p,q \in M \implies d(f(p),f(q)) \leq \theta \cdot d(p,q)$$

Definition 1.4.2 (Notation: f^n)**.** *] Suppose $n \in \mathbb{N}$, M is a metric space, and $f\colon M \to M$ is a contraction of M. Then, we denote*

$$f^n(x) = [f \circ f \circ \cdots \circ f](x) \ \forall x \in M,$$

where there are $n-1$ function compositions.

Remark 1.4.1. *Note that contraction depends on the metric of a given metric space. Geometrically, the repeated application of a contraction on a point will "make the image of the point and the point closer to each other".*

We remark that some texts refer to the above definition as a strict contraction since θ is strictly less than 1 and refer to a map with the above property a contraction for $\theta \leq 1$.

Theorem 1.4.1 (Banach Contraction Principal)**.** *Suppose M is a complete metric space and $f\colon M \to M$ is a contraction of M. Then,*

(i) f has a unique fix point $p \in M$. That is, $\exists! p$ such that $f(p) = p$;
(ii) $\forall x \in M$, $\lim_{n\to\infty} f^n(x) = p$;

Theorem 1.4.2 (Brouwer Fixed-Point Theorem)**.** *Suppose $B_m \subset \mathbb{R}^m$ is a closed unit ball and $f\colon B_m \to B_m$ is continuous. Then, f has a fixpoint $p \in B_m$.*

Remark 1.4.2. *We note that there exist several proofs for the Inverse Function Theorem and Implicit Function Theorem. One may prove either one of the theorems and apply it to prove the other. In fact, we prove the Implicit Function Theorem via the Inverse Function Theorem in lecture; where in Pugh, the text proved the Inverse Function Theorem via the Implicit Function Theorem.*

Definition 1.4.3 (C^r Diffeomorphism)**.** *Suppose $r \in \{1,2,\ldots,\infty\}$, $U \subset \mathbb{R}^n$ is open, and $f\colon U \to \mathbb{R}^m$. f is a C^r diffeomorphism if*

(i) f is a bijection, and
(ii) f and f^{-1} are C^r.

Proposition 1.4.1 (C^r Diffeomorphisms are Homeomorphisms)**.** *Suppose $U \subset \mathbb{R}^n$ is open, and $f\colon U \to \mathbb{R}^m$ is a C^r diffeomorphism for some $r \in \{1,2,\ldots,\infty\}$. Then, f is a homeomorphism.*

Proof. Suppose $U \subset \mathbb{R}^n$ is open, and $f\colon U \to \mathbb{R}^m$ is a C^r diffeomorphism for some $r \in \{1,2,\ldots,\infty\}$. By definition, we have that f^{-1} exists, where f and f^{-1} are C^r bijections. By definition, f and f^{-1} are continuously differentiable and, hence, continuous. Hence, f and f^{-1} are continuous bijections. By definition, f is a homeomorphism. □

Theorem 1.4.3 (Inverse Function Theorem)**.** *Suppose $U \subset \mathbb{R}^n$ is open, $f\colon U \to \mathbb{R}^n$ is C^1, and $(Df)_{x_0}$ is invertible where $x_0 \in U$. Then, there exist open sets $U_0 \subset U$ containing x_0 and $V \subset \mathbb{R}^n$ containing $f(x_0)$ such that*

(i) f restricted to U_0 onto V_0 is a bijection,
(ii) $f^{-1}\colon V_0 \to U_0$ is differentiable at $f(x_0)$, and
(iii) $(Df^{-1})_{f(x_0)} = [(Df)_{x_0}]^{-1}$.

Theorem 1.4.4 (Implicit Function Theorem). *Suppose $E \subset \mathbb{R}^n$ is open, and $f\colon E \to \mathbb{R}$ is C^1. Suppose further that $y \in E$ satisfies $f(y) = 0$ and $\partial_{x_n} f(y) \neq 0$. Denote $\tilde{x} = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$ the projection of $x = (x_1, \ldots, x_{n-1}, x_n) \in \mathbb{R}^n$ onto $\mathbb{R}^{n-1}$, $\forall x \in E$. Then, there exist*

(a) open sets $U \subset \mathbb{R}^{n-1}$ containing $\tilde{y}$ and $V \subset E$ containing y and
(b) a map $g\colon U \to \mathbb{R}$

satisfying the following properties:

(i) $g(\tilde{y}) = y_n$;
(ii) $\{x \in V : f(x) = 0\} = \{(\tilde{x}, g(\tilde{x})) : \tilde{x} \in U\}$;
(iii) $\forall j \in \{1, 2, \ldots, n-1\}$

$$\partial_{x_j} g(\tilde{y}) = -\frac{\partial_{x_j} f(y)}{\partial_{x_n} f(y)}.$$

1.5 *The Rank Theorem

Omitted

1.6 *Lagrange Multipliers

Omitted

1.7 Multiple Integrals

Definition 1.7.1 (Grid; area and mesh of a grid; Riemann sum and Riemann integral on a rectangle)**.** *Suppose* $\mathbf{R} = [a,b] \times [c,d] \subset \mathbb{R}^2$ *for some* $[a,b],[c,d] \subset \mathbb{R}$*. Let* P *and* Q *be families of subsets of* $[a,b]$ *and* $[c,d]$*, respectively, such that*

$$P = \{[x_{i-1}, x_i] \subset \mathbb{R} : a = x_0 < x_1 < \cdots < x_m = b \text{ with } i \in \{1,\dots,m\}\},$$
$$Q = \{[y_{j-1}, y_j] \subset \mathbb{R} : c = y_0 < y_1 < \cdots < y_n = d \text{ with } j \in \{1,\dots,n\}\}.$$

(Grid) A ***grid of*** $\mathbf{R}$ ***formed by*** P ***and*** Q *is the set of rectangles* $R_{i,j}$

$$G = P \times Q = \left\{R_{i,j} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \subset \mathbb{R}^2 : (i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}\right\}.$$

(Area) Let $\Delta x_i = x_i - x_{i-1}$ *and* $\Delta y_j = y_j - y_{j-1}$*,* $\forall (i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}$*. Then,* $\forall (i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}$*, the* ***area of*** $R_{i,j}$ *is*

$$|R_{i,j}| = \Delta x_i \cdot \Delta y_j.$$

(Mesh) The ***mesh of a grid*** G *is the diameter of the largest rectangle in* G*; that is,*

$$mesh(G) = diam(R_{i^*,j^*}) = \sqrt{(\Delta x_{i^*})^2 + (\Delta y_{j^*})^2}$$

where $R_{i^*,j^*} = [x_{i^*-1}, x_{i^*}] \times [y_{j^*-1}, y_{j^*}] \in G$ *is such that* $|R_{i^*,j^*}| \geq |R_{i,j}| \ \forall R_{i,j} \in G$*.*

(Riemann Sum) Select $(s_{i,j}, t_{i,j}) \in R_{i,j}$ *to be a sample point of* $R_{i,j}$*,* $\forall (i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}$ *and let* $S = \{(s_{i,j}, t_{i,j}) : (i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}\}$*. Then, the* ***Riemann sum of*** $f\colon \mathbf{R} \to \mathbb{R}$ ***with respect to grid*** G ***and sample points*** S *is the iterated sum*

$$R(f,G,S) = \sum_{i=1}^{m} \sum_{j=1}^{n} f(s_{i,j}, t_{i,j}) \cdot |R_{i,j}|.$$

(Riemann Integrable and Riemann Integral) $f\colon \mathbf{R} \to \mathbb{R}$ *is* ***Riemann integrable*** *if*

$$\exists r \in \mathbb{R} \text{ such that } \lim_{mesh(G) \to 0} R(f,G,S) = r.$$

Such $r \in \mathbb{R}$ *is the* ***Riemann integral of*** f ***on*** $\mathbf{R}$*, provided it exists, and we write*

$$\int_R f = \lim_{mesh(G) \to 0} R(f,G,S).$$

Definition 1.7.2 (Upper and lower Darboux sum and Darboux Integral on $\mathbb{R}^2$)**.** *Suppose* $\mathbf{R} = [a,b] \times [c,d] \subset \mathbb{R}^2$ *for some* $[a,b],[c,d] \subset \mathbb{R}$*. Let* G *be a grid of* $\mathbf{R}$ *and* $f\colon \mathbf{R} \to \mathbb{R}$ *be a bounded function. Then,*

(i) the lower sum of f *with respect to grid* G *is*

$$L(f,G) = \sum_{R_{i,j} \in G} m_{i,j} |R_{i,j}| \text{ where } m_{i,j} = \inf_{(s,t) \in R_{i,j}} f(s,t);$$

(ii) the upper sum f *with respect to grid* G *is*

$$U(f,G) = \sum_{R_{i,j} \in G} M_{i,j} |R_{i,j}| \text{ where } M_{i,j} = \sup_{(s,t) \in R_{i,j}} f(s,t);$$

(iii) the lower integral of f *on* $\mathbf{R}$ *is*

$$\underline{\int}_{\mathbf{R}} f = \sup\{L(f,G) : G \text{ is a grid of } \mathbf{R}\};$$

(iv) the upper integral of f *on* $\mathbf{R}$ *is*

$$\overline{\int}_{\mathbf{R}} f = \inf\{U(f,G) : G \text{ is a grid of } \mathbf{R}\}.$$

Theorem 1.7.1 (Properties of Integrals and Integrable Functions on $\mathbb{R}^2$)**.** *(The space of Riemann integrable functions is a vector space)*

(Monotonicity)

(Linearity)

(Riemann integrability $\iff$ Darboux integrability)

Definition 1.7.3 (Diameter of a set in a metric space)**.** *Suppose $S \subset M$ and $S \neq \emptyset$, where M is a metric space. Then, the diameter of S is*

$$diam(S) = \sup_{p,q \in S} d(p,q).$$

Definition 1.7.4 (Oscillation of a real-valued function, on a rectangle, at a point)**.** *Suppose $R = [a,b] \times [c,d] \subset \mathbb{R}^2$ for some $[a,b], [c,d] \subset \mathbb{R}$, and $f \colon R \to \mathbb{R}$. Then, the oscillation of f at $z \in R$ is*

$$osc_z(f) = \lim_{r \to 0} diam(f(R_r(z))),$$

where $R_r(z)$ is a neighborhood of z with radius $r > 0$ contained in R.

Definition 1.7.5 (Zero set in $\mathbb{R}$)**.** *A zero set in $\mathbb{R}$ is a set $Z \subset \mathbb{R}$ such that $\forall \epsilon > 0$, there exists a countable covering of Z, via open intervals (a_i, b_i), such that*

$$\sum_{i=1}^{\infty} [b_i - a_i] < \epsilon.$$

Proposition 1.7.1 (Empty Set is a Zero Set)**.** *$\emptyset$ is a zero set of $\mathbb{R}$.*

Proof. Consider arbitrary $\epsilon > 0$. Let $I_k = \left(\epsilon - \frac{\epsilon}{2^{k+3}}, \epsilon + \frac{\epsilon}{2^{k+3}}\right)$, $\forall k \in \mathbb{N}$. It follows that $\mathcal{U} = \{I_k : k \in \mathbb{N}\}$ is a covering of $\emptyset$ via open intervals. In addition, we also have that

$$\begin{aligned}\sum_{k=1}^{\infty} \left(\epsilon + \frac{\epsilon}{2^{k+3}}\right) - \left(\epsilon - \frac{\epsilon}{2^{k+3}}\right) &= \sum_{k=1}^{\infty} 2 \cdot \frac{\epsilon}{2^{k+3}} \\ &= \frac{\epsilon}{4} \sum_{k=1}^{\infty} \left[\frac{1}{2}\right]^k = \frac{\epsilon}{4} \cdot \frac{1}{1 - \frac{1}{2}} = \frac{\epsilon}{2} < \epsilon\end{aligned}$$

by the Geometric series identity. By definition, $\emptyset$ is a empty in $\mathbb{R}$. □

Proposition 1.7.2 (Singletons in $\mathbb{R}$ are Zero Sets)**.** *Suppose $r \in \mathbb{R}$. Then, $\{r\}$ is a zero set.*

Proof. Suppose $r \in \mathbb{R}$. $\forall \epsilon > 0$, define $I_k = \left(r - \frac{\epsilon}{2^{k+3}}, r + \frac{\epsilon}{2^{k+3}}\right)$, $\forall k \in \mathbb{N}$. It follows that $\mathcal{U} = \{I_k : k \in \mathbb{N}\}$ is a converging of $\{r\}$ via open intervals, since $\{r\} \subset \cup_{k \in \mathbb{N}} I_k$. Furthermore, we have that

$$\begin{aligned}\sum_{k \in \mathbb{N}} \left[r + \frac{\epsilon}{2^{k+3}}\right] - \left[r - \frac{\epsilon}{2^{k+3}}\right] &= \sum_{k \in \mathbb{N}} \frac{2\epsilon}{2^{k+3}} \\ &= \frac{\epsilon}{4} \sum_{k \in \mathbb{N}} \left[\frac{1}{2}\right]^k \\ &= \frac{\epsilon}{4} \cdot \frac{1}{1 - \frac{1}{2}} && \text{(By Geometric series identity, since } \tfrac{1}{2} < 1\text{)} \\ &= \frac{\epsilon}{2} < \epsilon.\end{aligned}$$

By definition, $\{r\}$ is a zero set as desired. □

Proposition 1.7.3 (Union of Zero Sets is a Zero Set)**.** *Suppose $Z, W \subset \mathbb{R}$ are zero sets. Then, $Z \cup W \subset \mathbb{R}$ is a zero set.*

Proof. Suppose $Z, W \subset \mathbb{R}$ are zero sets. By definition, we have that, $\forall \epsilon > 0$, there exists countable coverings of Z and W respectively, via open intervals (a_i, b_i) and (c_j, d_j), such that

$$\sum_{i=1}^{\infty}[b_i - a_i], \sum_{j=1}^{\infty}[c_j - d_j] < \epsilon.$$

Consider arbitrary $\epsilon > 0$. Then, $\frac{\epsilon}{2} > 0$ and we have that $\exists \mathcal{U}_0 = \{(a_i, b_i) \subset \mathbb{R} : i \in \mathbb{N}\}$ and $\mathcal{U}_1 = \{(c_j, d_j) \subset \mathbb{R} : j \in \mathbb{N}\}$ such that $\mathcal{U}_0$ and $\mathcal{U}_1$ are coverings of Z and W via open intervals, respectively, with the property that

$$\sum_{i=1}^{\infty}[b_i - a_i], \sum_{j=1}^{\infty}[c_j - d_j] < \frac{\epsilon}{2}. \qquad (*)$$

Let $\mathcal{U} = \mathcal{U}_0 \cup \mathcal{U}_1$. By definition, we have that

$$Z \subset \cup_{i\in\mathbb{N}}(a_i, b_i) \text{ and } W \subset \cup_{j\in\mathbb{N}}(c_j, d_j),$$
$$\therefore Z \cup W \subset [\cup_{i\in\mathbb{N}}(a_i, b_i)] \cup [\cup_{j\in\mathbb{N}}(c_j, d_j)] = \cup_{J\in\mathcal{U}} J.$$

Thus, $\mathcal{U}$ is a covering of $Z \cup W$ via open intervals. Denote $|J|$ the length of the open interval J, $\forall J \in \mathcal{U}$; that is, $J = (a, b) \in \mathcal{U} \implies |J| = b - a$. Furthermore, we have that

$$\sum_{J\in\mathcal{U}} |J| = \sum_{i=1}^{\infty}[b_i - a_i] + \sum_{j=1}^{\infty}[c_j - d_j] < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \qquad \text{(By } (*))$$

By definition, $Z \cup W$ is a zero set as desired. ☐

Corollary 1.7.1 (Countable Union of Zero Set is a Zero Set)**.** *Suppose $Z_1, Z_2, \cdots \subset \mathbb{R}$ are zero sets. Then, $\cup_{k\in\mathbb{N}} Z_k$ is a zero set.*

Proof. **(DO: DOUBLE CHECK)** Suppose $Z_1, Z_2, \cdots \subset \mathbb{R}$ are zero sets. Let $U_n = \cup_{i=1}^{n} Z_i$, $\forall n \in \mathbb{N}$. It suffices to show, via induction, that for $n \in \{2, 3, \ldots\}$ we have that $U_n = Z_1 \cup \cdots \cup Z_n$ is a zero set. We observe that the base case for $n = 2$ holds by the preceding theorem.

Let k be an arbitrary natural number greater than or equal to 2 and assume that U_k is a zero set. It suffices to show that

$$U_k \text{ is a zero set implies } U_{k+1} \text{ is a zero set.}$$

By assumption,

$$U_{k+1} = Z_1 \cup \cdots \cup Z_k \cup Z_{k+1} = U_k \cup Z_{k+1}.$$

By assumption and the inductive hypothesis, Z_{k+1} and U_k are zero sets in $\mathbb{R}$. Applying the preceding proposition, we then obtain that $U_k \cup Z_{k+1} = U_{k+1}$ is a zero set. By mathematical induction, we conclude that $\forall n \in \mathbb{N}$, $Z_1 \cup \cdots \cup Z_n$ is a zero set and, therefore, $\cup_{n\in\mathbb{N}} Z_n$ is a zero set. ☐

Theorem 1.7.2 (Countable Set in $\mathbb{R}$ is a Zero Set)**.** *Suppose $S \subset \mathbb{R}$ is a countable set. Then, S is a zero set.*

Proof. Suppose $S \subset \mathbb{R}$ is a countable set. We recall that, by a previous proposition, a singleton in $\mathbb{R}$ is a zero set.

(Case #1): Suppose S is finite. It follows that $S = \{x_1, \ldots, x_n\}$ for some $n \in \mathbb{N}$. It follows that $S = \cup_{k=1}^{n} \{x_k\}$. We observe that $\forall k \in \{1, \ldots, n\}$, $\{x_k\}$ is a zero set. By the preceding corollary, we obtain that $S = \cup_{k=1}^{n} \{x_k\}$ is a zero set as desired.

(Case #2): Suppose that S is countably infinite. By definition, $\forall k \in \mathbb{N}\ \exists x_k \in S$ since S is countably infinite. Thus, $S = \cup_{k \in \mathbb{N}} \{x_k\}$. We observe that $\forall k \in \mathbb{N}$, $\{x_k\}$ is a zero set. By the preceding corollary, we obtain that $S = \cup_{k \in \mathbb{N}} \{x_k\}$ is a zero set as desired. □

Theorem 1.7.3 (One-Dimensional Riemann-Lebesgue Theorem)**.** *Suppose that $[a, b] \subset \mathbb{R}$. Then, $f\colon [a, b] \to \mathbb{R}$ is Riemann integrable $\iff$ f is bounded, and its set of discontinuities is a zero set in $\mathbb{R}$.*

Definition 1.7.6 (Zero set in $\mathbb{R}^2$)**.** *A zero set in $\mathbb{R}^2$ is a set $Z \subset \mathbb{R}^2$ such that $\forall \epsilon > 0$, there exists a countable covering of Z, via open rectangles S_l, such that*

$$\sum_l |S_l| < \epsilon.$$

Theorem 1.7.4 (Two-Dimensional Riemann-Lebesgue Theorem)**.** *Suppose that $R = [a, b] \times [c, d] \subset \mathbb{R}^2$. Then, $f\colon R \to \mathbb{R}$ is Riemann integrable $\iff$ f is bounded, and its set of discontinuities is a zero set in $\mathbb{R}^2$.*

Definition 1.7.7 (Slice integrals)**.** *Suppose that $R = [a, b] \times [c, d] \subset \mathbb{R}^2$ for some $[a, b], [c, d] \subset \mathbb{R}$, and $f\colon R \to \mathbb{R}$ be a bounded function. Define, $\forall y \in [c, d]$, $f_y\colon [a, b] \to \mathbb{R}$ by*

$$f_y(x) = f(x, y)\ \forall x \in [a, b].$$

(i) The lower slice integral of f is the map $\underline{F}\colon [c, d] \to \mathbb{R}$ defined by

$$\underline{F}(y) = \underline{\int_a^b} f_y(x)dx.$$

(ii) The upper slice integral of f is the map $\overline{F}\colon [c, d] \to \mathbb{R}$ defined by

$$\overline{F}(y) = \overline{\int_a^b} f_y(x)dx.$$

Theorem 1.7.5 (Fubini's Theorem)**.** *Suppose that $R = [a, b] \times [c, d] \subset \mathbb{R}^2$ for some $[a, b], [c, d] \subset \mathbb{R}$ and $f\colon R \to \mathbb{R}$ is Riemann integrable. Then,*

(i) the lower and upper slice integral $\underline{F}, \overline{F}$ are integrable, and

(ii)

$$\int_R f = \int_c^d \underline{F}(y)dy = \int_c^d \left[\underline{\int_a^b} f(x, y)dx\right] dy = \int_c^d \overline{F}(y)dy = \int_c^d \left[\overline{\int_a^b} f(x, y)dx\right] dy.$$

Corollary 1.7.2 (Interchanging Order of Integration)**.** *Suppose that $R = [a, b] \times [c, d] \subset \mathbb{R}^2$ for some $[a, b], [c, d] \subset \mathbb{R}$ and $f\colon R \to \mathbb{R}$ is Riemann integrable. Then,*

$$\int_c^d \left[\int_a^b f(x, y)dx\right] dy = \int_a^b \left[\int_c^d f(x, y)dy\right] dx.$$

Definition 1.7.8 (Characteristic function of a subset in a metric space)**.** *Suppose (M, d) is a metric space. Then, the characteristic function of $S \subset M$ is the map $\chi_S\colon M \to \mathbb{R}$ defined by*

$$\chi_S(p) = \begin{cases} 1 & p \in S \\ 0 & p \in M/S \end{cases} \forall p \in M.$$

Remark 1.7.1. *Note that the characteristic function of $S \subset M$ may be defined more generally if (M, d) is not a metric space.*

Proposition 1.7.4 (Discontinuity of Characteristic Function at the Boundary)**.** *Suppose M is a metric space. Then, the characteristic function χ_S of $S \subset M$ is discontinuous at $p \in M \iff p \in \partial S$.*

Proof. Suppose (M, d) is a metric space, $S \subset M$, and $p \in M$. Denote $\overline{S}$, $int(S)$, and ∂S the closure, interior, and boundary of S, respectively. By definition, we have that $\partial S = \overline{S}/int(S)$.

($\implies$) It suffices to prove that $p \notin \partial S \implies \chi_S$ is continuous at p, since the condition is equivalent to the desired statement.

Suppose that $p \notin \partial S = \overline{S}/int(S)$. It follows that $p \in M/\overline{S}$ or $p \in int(S)$. We note that here $M/\overline{S}$ is open, since it is the complement of a closed set $\overline{S}$. In addition, $int(S)$ is also open, since the interior of a set is open. Consider aribtrary $\epsilon > 0$.

If $p \in M/\overline{S}$, then $\exists r > 0$ such that $B_r(p) \subset M/\overline{S}$ such that

$$\begin{aligned} q \in M \text{ and } d(q,p) < r &\implies q \in B_r(p) \subset M/\overline{S} \\ &\implies |\chi_S(p) - \chi_S(q)| = |0 - 0| && (\text{Since } q,p \notin \overline{S} \implies q,p \notin S) \\ &= 0 < \epsilon. \end{aligned}$$

If $p \in int(S) \subset S$, then $\exists R > 0$ such that $B_R(p) \subset int(S)$ such that

$$\begin{aligned} q \in M \text{ and } d(p,q) < R &\implies q \in B_R(p) \subset int(S) \subset S \\ &\implies |\chi_S(p) - \chi_S(q)| = |1 - 1| && (\text{Since } q,p \in S) \\ &= 0 < \epsilon. \end{aligned}$$

In both cases, $p \notin \partial S$ implies χ_S is continuous at p by definition.

($\impliedby$) It suffices to prove that χ_S is continuous at $p \implies p \notin \partial S$, since the condition is equivalent to the desired statement.

Suppose that χ_S is continuous at p. Suppose, to the contrary, that $p \in \partial S$. By the continuity of χ_S at p, we obtain that

$$\exists r > 0 \text{ such that } q \in M \text{ and } d(p,q) < r \implies |\chi_S(p) - \chi_S(q)| < \frac{1}{2}.$$

Take some $Q \in B_r(p)$ such that $Q \in M/\overline{S}$. It follows that

$$Q \in M \text{ and } d(p,Q) < r \implies |\chi_S(p) - \chi_S(Q)| = |1 - 0| = 1 < \frac{1}{2}, \quad (\text{Since } Q \notin \overline{S} \implies Q \notin S)$$

which is a contradiction. Thus, we conclude that $p \notin \partial S$ as desired.

We note that such Q does exist. Suppose not. Then, we have that $\forall q \in B_r(p)\ q \notin M/\overline{S}$, which implies $q \in \overline{S}$. That is, $B_r(p) \subset \overline{S}$. By definition, p is interior to $\overline{S}$ and, thus, $p \in int(S)$. However, by assumption $p \in \partial(S) = \overline{S}/int(S) \implies p \notin int(S)$, which is a contradiction. Thus, we conclude such Q must exist. Here, we complete the proof. ☐

Corollary 1.7.3 (Differentiability of Characteristic Function)**.** *Suppose (M, d) is a metric space and $S \subset M$. Then, χ_S is differentiable on $M/\partial S$ and not differentiable on ∂S.*

Definition 1.7.9 (Bounded set in a metric space)**.** *Suppose (M, d) is a metric space. Then,*

(i) $S \subset M$ is bounded if $\exists p \in M$, $r > 0$ such that $S \subset B_r(p)$;
(ii) $S \subset M$ is unbounded if it is not bounded.

Corollary 1.7.4 (Riemann Integrability of Characteristic Function of a Bounded Set)**.** *Suppose (M, d) is a metric space and $S \subset M$ is bounded. Then, $R = [a, b] \times [c, d] \subset \mathbb{R}^2$ contains S implies $\chi_S : R \to \mathbb{R}$ is Riemann integrable on R.*

Definition 1.7.10 (Riemann measurable set; area and length of sets)**.** *Suppose $I \subset \mathbb{R}$ and $S \subset \mathbb{R}^2$ are bounded. Then,*

(i) S is Riemann measurable if $\int \chi_S$ exists;
(ii) If S is Riemann measurable, then the area of S is

$$|S| = area(S) = \int \chi_S.$$

(iii) I is Riemann measurable if $\int \chi_I$ exists;
(iv) If I is Riemann measurable, then the lenght of I is

$$|I| = length(I) = \int \chi_I.$$

Theorem 1.7.6 (Riemann Measurable $\iff$ Boundary is a Zero Set)**.** *Suppose $S \subset \mathbb{R}^2$ is bounded. Then, S is Riemann measurable $\iff$ ∂S is a zero set.*

Proof. Suppose $S \subset \mathbb{R}^2$ is bounded. By definition, S is Riemann measurable $\iff$ $\int \chi_S$ exists $\iff$ χ_S is Riemann integrable by definition $\iff$ the set of discontinuities of χ_S is a zero set by the Riemann-Lebesgue Theorem $\iff$ the boundary of S is a zero set since $\{p \in \mathbb{R}^2 : \chi_S \text{ is discontinuous at } p\} = \partial S$ by a previous proposition. □

Theorem 1.7.7 (Cavalieri's Principal)**.** *Suppose $R = [a, b] \times [c, d] \subset \mathbb{R}^2$, $S \subset R$, and ∂S is a zero set. Then, the area of S is given by*

$$area(S) = \int_a^b length(S_x)dx,$$

where S_x is the vertical slices of S at x.

Definition 1.7.11 (Jacobian of a Differentiable function)**.** *Suppose $U \subset \mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is differentiable at $z \in U$. Then, the* ***Jacobian of** f **at** z is*

$$Jac_z(f) = Det(A)$$

where $A \in \mathcal{M}^{m,n}(\mathbb{R})$ represents $(Df)_z$.

Suppose f is differentiable on U. Then, we call the map $Jac(f) : U \to \mathbb{R}$ defined by

$$[Jac(f)](z) = Jac_z(f) \ \forall z \in U$$

the ***Jacbobian of** f.*

Definition 1.7.12 (Maximum coordinate norm)**.**

Proposition 1.7.5 (Neighborhoods Under $\mathbb{R}^2$ Maximum Coordinate Norm are Squares)**.**

Lemma 1.7.1 (Lemma 34 for the Change of Variables Theorem)**.**

Lemma 1.7.2 (Lemma 35 for the Change of Variables Theorem)**.** *The image of a zero set $Z \subset \mathbb{R}^2$ under a Liptshitz function $h : Z \to \mathbb{R}^2$ is a zero set.*

Theorem 1.7.8 (Change of Variables)**.** *Let $U, W \subset \mathbb{R}^2$ be open and $R = [a, b] \times [c, d] \subset U$. Suppose that $\varphi : U \to W$ is a C^1 diffeomorphism and $f : W \to \mathbb{R}$ is Riemann integrable. Then,*

$$\int_R [f \circ \varphi] \cdot |Jac(\varphi)| = \int_{\varphi(R)} f.$$

Remark 1.7.2. *We observe that the Change of Variables Theorem allows for the computation of Riemann integrals, of some function, on a more general region in $\mathbb{R}^2$.*

2 Lebesgue Theory

2.1 Outer Measure on $\mathbb{R}$

Definition 2.1.1 (Length of an interval in $\mathbb{R}$)**.** *Suppose $I \subset \mathbb{R}$ is a real interval. Then, the length of I is*

$$l(I) = \begin{cases} b-a & \text{if } I = (b,a) \\ 0 & \text{if } I = \emptyset \\ \infty & \text{if } I = (a,\infty) \text{ or } I = (-\infty, a) \text{ for some } a \in \mathbb{R} \text{ or } I = (-\infty,\infty) \end{cases}$$

Definition 2.1.2 (Covering of a set in $\mathbb{R}$)**.** *Suppose $S \subset \mathbb{R}$. Then, a covering of S is a family $\mathcal{U}$ of open sets $I_k \subset \mathbb{R}$ such that $S \subset \bigcup_k I_k$.*

Definition 2.1.3 (Outer measure of a set in $\mathbb{R}$)**.** *Denote $\mathcal{P}(\mathbb{R})$ the power set of $\mathbb{R}$. The outer measure on $\mathbb{R}$ is the map $|| : \mathcal{P}(\mathbb{R}) \to [0,\infty]$ defined by*

$$|A| = \inf \left\{ \sum_{k\in\mathbb{N}} l(I_k) : \forall k \in \mathbb{N},\ I_k \subset \mathbb{R} \text{ is an open interval and } A \subset \bigcup_{k\in\mathbb{N}} I_k \right\} \ \forall A \in \mathcal{P}(\mathbb{R})$$

or, equivalently,

$$|A| = \inf \left\{ \sum_{I\in\mathcal{U}} l(I) : \mathcal{U} \text{ is a countable covering of } A \text{ via open intervals} \right\} \ \forall A \in \mathcal{P}(\mathbb{R}).$$

We say that $|A|$ is the outer measure of A, $\forall A \in \mathcal{P}(\mathbb{R})$.

Theorem 2.1.1 (Outer Measure Preserves Order; Monotonicity of Outer Measure)**.** *Suppose $A \subset B \subset \mathbb{R}$. Then $|A| \leq |B|$.*

Proof. Suppose $A \subset B \subset \mathbb{R}$. Denote, respectively, S_A and S_B the sets

$$\left\{ \sum_{I\in\mathcal{U}} l(I) : \mathcal{U} \text{ is a countable covering of } A \text{ via open intervals} \right\},$$
$$\left\{ \sum_{I\in\mathcal{U}} l(I) : \mathcal{U} \text{ is a countable covering of } B \text{ via open intervals} \right\}.$$

By definition, we have that

$$|A| = \inf S_A \text{ and } |B| = \inf S_B.$$

For any countable covering $\mathcal{U}$ of B via open intervals, $\mathcal{U}$ is also a covering of A via open intervals since $A \subset B$, and we have that $\sum_{I\in\mathcal{U}} l(I) \in S_A$. Hence, we have that $S_B \subset S_A$. As an immediate result, we obtain that $\inf S_A \leq \inf S_B$. That is, $|A| \leq |B|$ as desired. ☐

Theorem 2.1.2 (Countable and Finite Subadditivity of Outer Measure)**.** *Suppose $A_1, A_2, \cdots \subset \mathbb{R}$. Then,*

$$\left| \bigcup_{k\in\mathbb{N}} A_k \right| \leq \sum_{k\in\mathbb{N}} |A_k|.$$

Theorem 2.1.3. *Countable and finite sets in $\mathbb{R}$ have outer measures of 0.*

Proof. Suppose $A \subset \mathbb{R}$ is countable. Then, we may write $A = \{a_k : k \in \mathbb{N}\}$. Consider arbitrary $\epsilon > 0$. Let $I_k = \left(a_k - \frac{\epsilon}{2^{k+3}}, a_k + \frac{\epsilon}{2^{k+3}}\right) \ \forall k \in \mathbb{N}$. We observe that $A \subset \bigcup_{k\in\mathbb{N}} I_k$ and

$$\begin{aligned}\sum_{k\in\mathbb{N}} l(I_k) &= \sum_{k\in\mathbb{N}} a_k + \frac{\epsilon}{2^{k+3}} - a_k + \frac{\epsilon}{2^{k+3}} = \sum_{k\in\mathbb{N}} \frac{2\epsilon}{2^{k+3}} \\ &= \frac{\epsilon}{4} \cdot \sum_{k\in\mathbb{N}} \left[\frac{1}{2}\right]^k = \frac{\epsilon}{4} \cdot \frac{1}{1-\frac{1}{2}} = \frac{\epsilon}{2} < \epsilon. \qquad \text{(By Geometric series identity)}\end{aligned}$$

That is, we have $\sum_{k\in\mathbb{N}} l(I_k) < \epsilon \ \forall \epsilon > 0$. It follows that $\sum_{k\in\mathbb{N}} l(I_k) = 0$, for otherwise we would obtain that $\sum_{k\in\mathbb{N}} l(I_k) > 0 \implies \sum_{k\in\mathbb{N}} l(I_k) > \sum_{k\in\mathbb{N}} l(I_k)$, which is a contradiction. □

Corollary 2.1.1 (Outer Measure of $\mathbb{Q}$ and $\mathbb{Z}$)**.** *$\mathbb{Q}$ and $\mathbb{Z}$ have outer measure 0.*

Proof. By Theorem 2.8, $|\mathbb{Q}|, |\mathbb{Z}| = 0$ since $\mathbb{Q}$ and $\mathbb{Z}$ are countable. □

Definition 2.1.4 (Translation of a set in $\mathbb{R}$)**.** *Suppose $t \in \mathbb{R}$ and $A \subset \mathbb{R}$. Then, the translation $t + A$ is defined by*

$$t + A = \{t + a : a \in A\}.$$

Theorem 2.1.4 (Translation Invariance of Outer Measure)**.** *Suppose $t \in \mathbb{R}$ and $A \subset \mathbb{R}$. Then, $|t + A| = |A|$.*

Proof. **[INCOMPLETE]**

□

Definition 2.1.5 (Open Cover; finite subcover)**.** *Suppose (M, d) is a metric space and $A \subset M$.*

(i) An open cover of A is a collection $\mathcal{U}$ of open sets of M such that

$$A \subset \bigcup_{I\in\mathcal{U}} I.$$

If $\mathcal{U}$ is an open cover of A, then We say that $\mathcal{U}$ covers A.

(ii) Suppose $\mathcal{U}$ is an open cover of A. Then, $\mathcal{V} \subset \mathcal{U}$ is a subcover of $\mathcal{U}$ if $\mathcal{V}$ is a cover of A.

(iii) A subcover of an open cover of A is finite if the subcover contains finitely many elements.

Theorem 2.1.5 (Heine-Borel Theorem)**.** *Suppose $A \subset \mathbb{R}$ is closed and bounded. Then, every open cover of A has a finite subcover.*

Proof. **[INCOMPLETE]**

□

Theorem 2.1.6 (Outer Measure of a Closed Interval)**.** *Suppose $a, b \in \mathbb{R}$ with $a < b$. Then, $|[a, b]| = b - a$.*

Proof. **[INCOMPLETE]**

□

Theorem 2.1.7 (Nondegenerate Intervals are Uncountable)**.** *Suppose $I \subset \mathbb{R}$ is an interval such that $card(I) > 1$. Then, I is uncountable.*

Proof. Suppose $I \subset \mathbb{R}$ is an interval such that $card(I) > 1$. Hence, I contains at least 2 distinct elements, say a and b. Without loss of generality, suppose that $a < b$.

It follows that $[a, b] \subset I$ implies that $|[a, b]| = b - a \leq |I|$ by Theorem 2.14 and the monotonicity of outer measure. It follows that $|I| > 0$. Recall that countable subsets of $\mathbb{R}$ have outer measure 0. By the contrapositive of the above statement, I is uncountable.

□

Theorem 2.1.8 (Nonadditivity of Outer Measure)**.** *There exist disjoint set $A, B \subset \mathbb{R}$ such that $|A \cup B| \neq |A| + |B|$.*

Proof. **[INCOMPLETE]**

□

2.2 Measurable Spaces and Functions

Remark 2.2.1. *The order defined by set containment* ($A \subset B$) *is a partial ordering, since there exist sets that are not subsets of each other.*

Definition 2.2.1 (Power set)**.** *Suppose X is a set. Then, the power set of X is*

$$\mathcal{P}(X) = \{S : S \subset X\}.$$

Remark 2.2.2. *By Theorem 2.22 (Axler 25), we note that there does not exist a map $\mu \colon \mathcal{P}(\mathbb{R}) \to [0, \infty]$ satisfying all the desired properties of a measure. Hence, we loosen the requirement for the domain of a measure, and it suffices to define a measure on a σ-algebra.*

Theorem 2.2.1 (Nonexistence of Extension of length to all subsets of $\mathbb{R}$)**.** *There does not exist a function μ that satisfies all the following properties:*

(i) $\mu \colon \mathcal{P}(\mathbb{R}) \to [0, \infty]$,
(ii) $I \subset \mathbb{R}$ *is an open interval implies that* $\mu(I) = l(I)$,
(iii) $A_1, A_2, \ldots$ *is a sequence of disjoint subsets of* $\mathbb{R}$ *implies that* $\mu\left(\cup_{k \in \mathbb{N}} A_k\right) = \sum_{k \in \mathbb{N}} \mu(A_k)$,
(iv) $A \subset \mathbb{R}, t > 0 \implies \mu(t + A) = \mu(A)$.

[Theorem 2.22 (Axler 25)]

Remark 2.2.3. *The counter-example provided in Theorem 28 may be used to prove the above theorem.*

Definition 2.2.2 (σ-algebra)**.** *Suppose X is a set. $\mathcal{A} \subset \mathcal{P}(X)$ is a σ-algebra of X if*

(i) $\emptyset \in \mathcal{A}$,
(ii) $A \in \mathcal{A} \implies X/A \in \mathcal{A}$ ***(Closure under complementation)***,
(iii) $A_1, A_2, \dots \in \mathcal{A} \implies \bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A}$ ***(Closure under countable union)***.

Theorem 2.2.2 (Properties of $\sigma-$Algebra)**.** *Suppose X is a set and $\mathcal{A}$ is a σ-algebra of X. Then,*

(i) $X \in \mathcal{A}$,
(ii) $A, B \in \mathcal{A} \implies A \cup B,\ A \cap B,\ A/B \in \mathcal{A}$,
(iii) $A_1, A_2, \dots \in \mathcal{A} \implies \bigcap_{k \in \mathbb{N}} A_k \in \mathcal{A}$.

Proof. Suppose X is a set and $\mathcal{A}$ is a σ-algebra of X.

(i) By the definition of σ-algebra, $\emptyset \in \mathcal{A} \implies X/\emptyset = X \in \mathcal{A}$.

(ii) Suppose $A, B \in \mathcal{A}$. (a) Let $A = I_1$, $B = I_2$, and $I_k = \emptyset$, $\forall k \in \mathbb{N}/\{1, 2\}$. It follows, from the closure of $\mathcal{A}$ under countable union, that $A \cup B = \cup_{k=1}^{\infty} I_k \in \mathcal{A}$.

(b) We note that $X/A, X/B \in \mathcal{A}$ by the closure of $\mathcal{A}$ under complementation. It follows that

$$\begin{aligned} & [X/A] \cup [X/B] \in \mathcal{A} && \text{(By Part (a))} \\ \implies & A \cap B = X/([X/A] \cup [X/B]) \in \mathcal{A}, \end{aligned}$$

by De Morgan's Law and the closure of $\mathcal{A}$ under complementation.

(c) This property follows immediately from De Morgan's Law and the closure of $\mathcal{A}$ under complementation. ☐

Proposition 2.2.1 (Smallest and Largest σ-algebra on a Set)**.** *Suppose X is a set. Then,*

(i) $\{\emptyset, X\}$ *is smallest σ-algebra on X, and*

(ii) $\mathcal{P}(X)$ *is the largest σ-algebra on X.*

Proof. Suppose X is a set.

(i) Let $\mathcal{S} = \{\emptyset, X\}$. We observe that $\emptyset \in \mathcal{S}$ by assumption. In addition, we also have that $X/\emptyset = X, X/X = \emptyset \in X$. Lastly, suppose $I_k \in \mathcal{S}\ \forall k \in \mathbb{N}$. We note that $\cup_{k\in\mathbb{N}} I_k = \emptyset$ or $\cup_{k\in\mathbb{N}} I_k = X$. In either case, $\cup_{k\in\mathbb{N}} I_k \in \mathcal{S}$. By definition, $\mathcal{S}$ is a σ-algebra on X.

Consider arbitrary σ-algebra $\mathcal{A}$ on X. We must have that $\emptyset, X \in \mathcal{A}$, by the definition and property of a σ-algebra. Thus, we have that $\mathcal{S} \subset \mathcal{A}$. Hence, we conclude $\mathcal{S}$ is indeed the smallest σ-algebra on X.

(ii) Let $\mathcal{S} = \mathcal{P}(X)$. We observe that $\emptyset \in \mathcal{S}$ since $\emptyset \subset X$. In addition, $S \in \mathcal{S}$ implies $X/S \subset X$, which implies $X/S \in \mathcal{S}$. Lastly, suppose $I_k \in \mathcal{S}\ \forall k \in \mathbb{N}$. We note that $\cup_{k\in\mathbb{N}} I_k \subset X \implies \cup_{k\in\mathbb{N}} I_k \in \mathcal{S}$. By definition, $\mathcal{S}$ is a σ-algebra on X.

Consider arbitrary σ-algebra $\mathcal{A}$ on X. By definition, $\mathcal{A} \subset \mathcal{P}(X) = \mathcal{S}$. Hence, we conclude $\mathcal{S}$ is indeed the largest σ-algebra on X. $\square$

Remark 2.2.4. *Suppose X is a set. We refer to the σ-algebra on X that is contained in every σ-algebra on X as the smallest σ-algebra on X. Similarly, we refer to the σ-algebra on X that contains every σ-algebra on X as the largest σ-algebra on X.*

Proposition 2.2.2. *Suppose X is a set. Then, $\{S \in \mathcal{P}(X) : S \text{ is countable or } X/S \text{ is countable}\}$ is a σ-algebra on X.*

Proof. Suppose X is a set. Let $\mathcal{S} = \{S \in \mathcal{P}(X) : S \text{ is countable or } X/S \text{ is countable}\}$.

(i) We observe that $\emptyset$ is finite and, therefore, countable. It follows that $\emptyset \in \mathcal{A}$.

(ii) Consider arbitrary $S \in \mathcal{S}$. By assumption, we have that $S = X/[X/S]$ is countable or X/S is countable. As an immediate result, $X/S \subset \mathcal{S}$.

(iii) Suppose that $S_1, S_2, \cdots \in \mathcal{S}$.

(Case #1) Suppose that $\forall k \in \mathbb{N}\ S_k$ is countable. It follows that $\cup_{k=1}^{\infty} S_k$ is countable since the countable union of countable sets is countable. Therefore, $\cup_{k=1}^{\infty} S_k \in \mathcal{S}$.

(Case #2) Suppose $\exists N_1, N_2, \cdots \in \mathbb{N}$ such that $S_{N_1}, S_{N_2}, \ldots$ are uncountable. It follows that $X/S_{N_1}, X/S_{N_2}, \ldots$ must be countable since $S_{N_1}, S_{N_2}, \cdots \in \mathcal{S}$ by assumption. We observe, since a subset of a countable set is countable, that

$$\begin{aligned}
\cap_{k=1}^{\infty}[X/S_k] \subset X/S_{N_1} &\implies \cap_{k=1}^{\infty}[X/S_k] \text{ is countable} \\
&\implies X/\left[\cup_{k=1}^{\infty} S_k\right] = \cap_{k=1}^{\infty}[X/S_k] \text{ is countable} && \text{(By De Morgan's Law)} \\
&\implies \cup_{k=1}^{\infty} S_k \in \mathcal{S}. && \text{(By assumption)}
\end{aligned}$$

By definition, $\mathcal{S}$ is indeed a σ-algebra on X. $\square$

Definition 2.2.3 (Measurable space; measurable set)**.** *(i) A measurable space is an order pair $(X, \mathcal{S})$, where X is a set and $\mathcal{S}$ is a σ-algebra of X.*

(ii) $S \in \mathcal{S}$ is a $\mathcal{S}$-measurable set, or a measurable set if sufficient context about $\mathcal{S}$ is supplied.

Theorem 2.2.3 (Existence of the Smallest σ-algebra)**.** *Suppose X is a set and $\mathcal{S} \subset \mathcal{P}(X)$. Then, the intersection of all σ-algebra on X containing $\mathcal{S}$ is the smallest σ-algebra on X containing $\mathcal{S}$.*

Proof. Suppose X is a set and $\mathcal{S} \subset \mathcal{P}(X)$. Let $\mathcal{A}$ be the intersection of all σ-algebra on X containing $\mathcal{S}$.

(i) We observe that all σ-algebra on X containing $\mathcal{S}$ contains $\emptyset$ implies that $\emptyset \in \mathcal{A}$.

(ii) Consider arbitrary $A \in \mathcal{A}$. It follows that A is also contained in all σ-algebra on X containing $\mathcal{S}$. As an immediate result of the closure of a σ-algebra under complementation, X/A is contained in all σ-algebra on X containing $\mathcal{S}$ and, hence, in $\mathcal{A}$.

(iii) Suppose $A_1, A_2, \dots, \subset \mathcal{A}$. Then, $A_1, A_2, \dots$ are certainly also contained in all σ-algebra on X containing $\mathcal{S}$. Therefore, $\cup_{k\in\mathbb{N}} A_k$ is also contained in all σ-algebra on X containing $\mathcal{S}$, by the closure of a σ-algebra under countable union. It follows that $\cup_{k\in\mathbb{N}} A_k \in \mathcal{A}$ as desired.

By definition, $\mathcal{A}$ is a σ-algebra on X and certainly $\mathcal{A}$ is contained in any σ-algebra on X by assumption. Hence, we conclude that $\mathcal{A}$ is the smallest σ-algebra on X. □

Example 2.2.1. *Suppose X is a set and $\mathcal{A} = \{\{x\} : x \in X\}$. Then, the smallest σ-algebra on X containing $\mathcal{A}$ is*

$$\mathcal{S} = \{S \subset X : S \text{ is countable or } X/S \text{ is countable}\}.$$

Proof. By a previous proposition, $\mathcal{S}$ is a σ-algebra on X. We observe that any singleton is countable. It follows that every singleton of X is in $\mathcal{S}$. That is, $\mathcal{S}$ contains $\mathcal{A}$. It suffices to show that every σ-algebra on X containing $\mathcal{A}$ contains $\mathcal{S}$.

Consider arbitrary σ-algebra $\mathcal{U}$ on X containing $\mathcal{A}$. We note that $\mathcal{A} \subset \mathcal{U}$ implies

$$x_1, x_2, \dots \in X \implies \{x_1\}, \{x_2\}, \dots \in \mathcal{U} \implies \bigcup_{k\in\mathbb{N}} \{x_k\} \in \mathcal{U}. \quad \text{(By the closure of a } \sigma\text{-algebra)}$$

That is, $S \subset X$ is countable implies $S \in \mathcal{U}$, since every countable $S \subset X$ is a countable union of some singletons in X. Furthermore, $X/S \subset X$ is countable for some $S \subset X$ implies $S \in X$, for we have that $X/S \in \mathcal{U}$ implies $X/[X/S] = S \in \mathcal{U}$ by the closure of $\mathcal{U}$ under complementation.

That is, we showed that for any $S \subset X$ such that S is countable or X/S is countable, it holds that $S \in \mathcal{U}$. By definition, $\mathcal{S} \subset \mathcal{U}$ for any σ-algebra $\mathcal{U}$ on X. Therefore, $\mathcal{S}$ is indeed the smallest σ-algebra on X containing $\mathcal{A}$. □

Example 2.2.2.

Definition 2.2.4 (Borel sets; Borel Algebra)**.** *Let $\mathcal{B}$ be the smallest σ-algebra on $\mathbb{R}$ containing all open sets in $\mathbb{R}$. Then, $B \in \mathcal{B}$ is a Borel set.*

We refer to $\mathcal{B}$ as the Borel Algebra.

Proposition 2.2.3 (Borel Sets Menu Theorem)**.** *(i) Suppose $S \subset \mathbb{R}$ is closed. Then, S is a Borel set.*

(ii) Suppose $S \subset \mathbb{R}$ is countable. Then, S is a Borel set.

(iii) Suppose $S \subset \mathbb{R}$ is an half-open interval. Then, S is a Borel set.

(iv) Suppose $f : \mathbb{R} \to \mathbb{R}$. Then,

$$S = \{x \in \mathbb{R} : f \text{ is continuous at } x\} \subset \mathbb{R}$$

is a Borel set.

Proof. Recall that, by definition, a Borel set is a set in the smallest σ-algebra $\mathcal{B}$ on $\mathbb{R}$ containing all open sets of $\mathbb{R}$.

(i) Suppose $S \subset \mathbb{R}$ is closed. It follows that $\mathbb{R}/S$ is open and, hence, contained in $\mathcal{B}$. By the closure of $\mathcal{B}$ under complementation $\mathbb{R}/[\mathbb{R}/S] = S \in \mathcal{B}$. Thus, S is a Borel set.

(ii) Suppose $S \subset \mathbb{R}$ is countable. It follows that $S = \cup_{k\in\mathbb{N}} \{x_k\}$, where $x_k \in \mathbb{R}\ \forall k \in \mathbb{N}$. Recall that singletons are closed. It follows that $\forall k \in \mathbb{N}\ \{x_k\} \in \mathcal{B}$ by Part (i). By the closure of $\mathcal{B}$ under countable union, $S = \cup_{k\in\mathbb{N}} \{x_k\} \in \mathcal{B}$. Thus, S is a Borel set.

(iii) Suppose $S \subset \mathbb{R}$ is an half-open interval. Without loss of generality, suppose that $S = [a, b)$ for some $a, b \in \mathbb{R}$ and $a < b$. We note that we may express $[a, b) = \cap_{k\in\mathbb{N}} \left(a - \frac{1}{k}, b\right)$. Furthermore, $\forall k \in \mathbb{N}$, $\left(a - \frac{1}{k}, b\right) \in \mathcal{B}$, by assumption, since it is open. It follows, from Theorem 2.25 (Axler 27), that $[a, b) \in \mathcal{B}$. Therefore, S is a Borel set.

(iv) **[INCOMPLETE]** Suppose $f \colon \mathbb{R} \to \mathbb{R}$. Let $S = \{x \in \mathbb{R} : f \text{ is continuous at } x\} \subset \mathbb{R}$. ☐

Example 2.2.3 (Measurable Space)**.** $(\mathbb{R}, \{\emptyset, \mathbb{R}\})$ *is a measurable Space.*

Definition 2.2.5 (Inverse image of a set under a function)**.** *Suppose X, Y are sets and $f \colon X \to Y$. Then, the inverse image of $A \subset Y$ under f is*

$$f^{-1}(A) = \{x \in X : f(x) \in A\}.$$

Proposition 2.2.4 (Inverse Image (Preimage) Property)**.** *Suppose X, Y are sets and $f \colon X \to Y$ and $x \in A \subset Y$. Then,*

(i) $x \in f^{-1}(A) \iff f(x) \in A$*,*

(ii) but it is not necessarily true that $f(x) = a \iff f^{-1}(\{a\}) = \{x\}$*, where* $a \in A$*.*

Proof. Suppose X, Y are sets and $f \colon X \to Y$ and $x \in A \subset Y$.

(i) ($\implies$) Suppose $x \in f^{-1}(A)$. By definition, we have that $f(x) \in A$. ($\impliedby$) Suppose $f(x) \in A$. Then, $x \in f^{-1}(A)$ by definition.

(ii) Consider the constant function $2_\mathbb{R} \colon \mathbb{R} \to \mathbb{R}$ defined by $2_\mathbb{R}(x) = 2\ \forall x \in \mathbb{R}$. We note that $f(1) = 2$ but $f^{-1}(\{2\}) = \mathbb{R} \neq \{1\}$. ☐

Remark 2.2.5. *Suppose X, Y are sets and $f \colon X \to Y$. We remark that the statement*

$$f^{-1}(a) = x \iff f(x) = a \qquad \text{(Where } x \in X \text{ and } a \in Y\text{)}$$

holds if and only if f is invertible. Hence, in the general case, the equivalence needs not to hold as f^{-1} would not necessarily be well-defined as a function. We remark that additional care is required to distinguish between an inverse image of a set under a function and an image of a set under an inverse function.

Theorem 2.2.4 (Inverse Image Identities)**.** *Suppose X, Y, W are sets, and $f \colon X \to Y$, $g \colon Y \to W$. Then, the following identities hold:*

(i) $A \subset Y \implies f^{-1}(Y/A) = X/f^{-1}(A)$*;*
(ii) $\mathcal{A} \subset \mathcal{P}(Y) \implies f^{-1}(\cup_{A\in\mathcal{A}} A) = \cup_{A\in\mathcal{A}} f^{-1}(A)$*;*
(iii) $\mathcal{A} \subset \mathcal{P}(Y) \implies f^{-1}(\cap_{A\in\mathcal{A}} A) = \cap_{A\in\mathcal{A}} f^{-1}(A)$*;*
(iv) $A \subset W \implies (g \circ f)^{-1}(A) = f^{-1}\left(g^{-1}(A)\right)$*.*

Proof. Suppose X, Y, W are sets, and $f \colon X \to Y$, $g \colon Y \to W$.

(i) Suppose that $A \subset Y$. Consider arbitrary $p \in f^{-1}(Y/A)$. It follows that

$$f(p) \in Y/A \iff f(p) \notin A \iff p \notin f^{-1}(p) \iff p \in X/f^{-1}(A).$$

(ii) Suppose that $\mathcal{A} \subset \mathcal{P}(Y)$. It follows that

$$
\begin{aligned}
p \in f^{-1}\left(\bigcup_{A\in\mathcal{A}} A\right) &\iff f(p) \in \bigcup_{A\in\mathcal{A}} A \iff f(p) \in \tilde{A} \\
&\iff p \in f^{-1}(\tilde{A}) \iff p \in \bigcup_{A\in\mathcal{A}} f^{-1}(A),
\end{aligned}
$$

for some $\tilde{A} \in \mathcal{A}$. Therefore, $f^{-1}\left(\cup_{A\in\mathcal{A}} A\right) = \cup_{A\in\mathcal{A}} f^{-1}(A)$.

(iii) Suppose that $\mathcal{A} \subset \mathcal{P}(Y)$. It follows that

$$
\begin{aligned}
p \in f^{-1}\left(\bigcap_{A\in\mathcal{A}} A\right) &\iff f(p) \in \bigcap_{A\in\mathcal{A}} A \iff f(p) \in A,\ \forall A \in \mathcal{A} \\
&\iff p \in f^{-1}(A),\ \forall A \in \mathcal{A} \iff p \in \bigcap_{A\in\mathcal{A}} f^{-1}(A).
\end{aligned}
$$

Therefore, $f^{-1}\left(\cap_{A\in\mathcal{A}} A\right) = \cap_{A\in\mathcal{A}} f^{-1}(A)$.

(iv) Suppose that $A \subset W$. We observe that

$$
\begin{aligned}
p \in (g\circ f)^{-1}(A) &\iff [g\circ f](p) \in A \iff g(f(x)) \in A \\
&\iff f(x) \in g^{-1}(A) \iff x \in f^{-1}(g^{-1}(A)).
\end{aligned}
$$

Hence, we conclude $(g\circ f)^{-1}(A) = f^{-1}\left(g^{-1}(A)\right)$. $\square$

Definition 2.2.6 (Measurable function)**.** *Suppose $(X, \mathcal{S})$ is a measurable space. Then, $f\colon X \to \mathbb{R}$ is a $\mathcal{S}$-measurable function if $\forall B \in \mathcal{B}$*

$$f^{-1}(B) \in \mathcal{S},$$

where $\mathcal{B}$ is the collection of all Borel sets.

Proposition 2.2.5 (Function Measurablity on Trivial σ-algebra Implies Function Constantness)**.** *Let X be a set and $\mathcal{S} = \{\emptyset, X\}$. Suppose that $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. Then, f is constant.*

Proof. Consider the measurable space $(X, \mathcal{S})$, where X is set and $\mathcal{S} = \{\emptyset, X\}$. Suppose that $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. By the $\mathcal{S}$-measurability of f, we have that for any Borel set $B \subset \mathbb{R}$

$$f^{-1}(B) \in \mathcal{S} \implies f^{-1}(B) = \emptyset \text{ or } f^{-1}(B) = X.$$

In particular, $\forall r \in \mathbb{R}$, we have that $f^{-1}(\{r\}) = \emptyset$ or $f^{-1}(\{r\}) = X$, since any closed subset of $\mathbb{R}$ is a Borel set by Example 2.30 (Axler 29). Suppose, to the contrary, that $\forall r \in \mathbb{R}$ we have that $f^{-1}(\{r\}) = \emptyset$. It follows that

$$
\begin{aligned}
&\forall r \in \mathbb{R},\ f^{-1}(\{r\}) = \{x \in X : f(x) = r\} = \emptyset && \text{(By definition)} \\
&\implies \forall r \in \mathbb{R}, \forall x \in X\ f(x) \neq r \\
&\implies f(X) \not\subset \mathbb{R} = Codomain(f),
\end{aligned}
$$

which is a contradiction since the range of a function is a subset of its codomain. It follows that $\exists R \in \mathbb{R}$ such that $f^{-1}(R) = X$. That is, we have that

$$f^{-1}(R) = \{x \in X : f(x) = R\} = X \implies \forall x \in X\ f(x) = R.$$

By definition, f is constant as desired. $\square$

Proposition 2.2.6 (Function Measurability on Power Set)**.** *Suppose X is a set. Then, $f\colon X \to \mathbb{R}$ is $\mathcal{P}(X)$-measurable.*

Proof. Suppose X is a set and $f\colon X \to \mathbb{R}$. Consider arbitrary Borel set $B \in \mathcal{B}$. It follows, from definition, that $f^{-1}(B) = \{x \in X : f(x) \in B\} \subset X \implies f^{-1}(B) \in \mathcal{P}(X)$. By definition, f is $\mathcal{P}(X)$-measurable. □

Proposition 2.2.7 (Identities for Intervals in $\mathbb{R}$)**.** *Suppose $a, b \in \mathbb{R}$ and $a < b$. Then,*

(i) $(a, b) = \bigcup_{k\in\mathbb{N}} \left(a, b - \frac{1}{k}\right] = \bigcup_{k\in\mathbb{N}} \left[a + \frac{1}{k}, b\right]$;

(ii) $(a, b] = \bigcap_{k\in\mathbb{N}} \left(a, b + \frac{1}{k}\right)$ *and* $[a, b) = \bigcap_{k\in\mathbb{N}} \left(a - \frac{1}{k}, b\right)$;

(iii) $(-\infty, a) = \bigcup_{k\in\mathbb{N}} \left(-k, a - \frac{1}{k}\right]$ *and* $(a, \infty) = \bigcup_{k\in\mathbb{N}} \left[a + \frac{1}{k}, k\right)$.

Proof. **[INCOMPLETE]** □

Example 2.2.4.

Definition 2.2.7 (Characteristic function)**.** *Suppose X is a set. Then, the characteristic function of $E \subset X$ is the map $\chi_E\colon X \to \mathbb{R}$ defined by*

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}.$$

We also refer to the Characteristic function as the indicator function.

Proposition 2.2.8 (Product of Characteristic Functions)**.** *Suppose X is a set and $E, A \subset X$. Then, $\chi_E \chi_A = \chi_{E\cap A}$.*

Proof. Suppose X is a set and $E, A \subset X$. Fix $x \in X$. We observe that

$$\begin{aligned} x \in E, A &\implies [\chi_E\chi_A](x) = 1 \cdot 1 = 1 = \chi_{E\cap A}(x), \\ x \notin E, A &\implies [\chi_E\chi_A](x) = 0 \cdot 0 = 0 = \chi_{E\cap A}(x), \\ x \in E \text{ and } x \notin A &\implies [\chi_E\chi_A](x) = 1 \cdot 0 = 0 = \chi_{E\cap A}(x), \\ x \notin E \text{ and } x \in A &\implies [\chi_E\chi_A](x) = 0 \cdot 1 = 0 = \chi_{E\cap A}(x). \end{aligned}$$

Thus, $\chi_E \chi_A = \chi_{E\cap A}$. □

Proposition 2.2.9 (Decomposition of Characteristic Function on a Disjoint Union)**.** *Suppose X is a set, I is a set of indexes (countable or uncountable), and $\{A_i : i \in I\}$ is a collection of disjoint subsets of X. Then,*

$$\chi_{\bigsqcup_{k\in I} A_k} = \sum_{k\in I} \chi_{A_k}.$$

Proof. Suppose X is a set, I is a set of indexes (countable or uncountable), and $\{A_i : i \in I\}$ is a collection of disjoint subsets of X. Fix $x \in X$.

Here, we have two cases to consider; namely when $x \notin \bigsqcup_{k\in I} A_k$ and when $x \in \bigsqcup_{k\in I} A_k$.

(Case #1): Suppose x does not live in the union. It follows that, $\forall k \in I$,

$$x \notin A_k \implies \chi_{A_k}(x) = 0,$$

which implies that $\sum_{k\in I} \chi_{A_k}(x) = 0 = \chi_{\bigsqcup_{k\in I} \chi_{A_k}}(x)$.

(Case #2): Suppose that x lives in the disjoint union. It follows that $\exists! N \in I$ such that $x \in A_N$. It follows that $\chi_{A_N}(x) = 1$ and $\chi_{A_m}(x) = 0$ $\forall m \in I/\{N\}$. As an immediate result, we yield $\sum_{k\in I} \chi_{A_k}(x) = \chi_{A_N}(x) = 1 = \chi_{\bigsqcup_{k\in I} \chi_{A_k}}(x)$.

Thus, we obtain the desired result and complete the proof. □

Proposition 2.2.10 (Image and Preimage of the Characteristic Function)**.** *Suppose $(X, \mathcal{S})$ is a measurable space and $E \subset X$. Let $\chi_E \colon X \to \mathbb{R}$ be the characteristic function of E on X. Then,*

(i) $\forall A \in \mathcal{P}(X),\ \chi_E(A) = \begin{cases} \{0\} & \text{if } A \subset X/E \\ \{1\} & \text{if } A \subset E \\ \{0,1\} & \text{if } A \cap E \neq \emptyset \text{ and } A \cap [X/E] \neq \emptyset \\ \emptyset & \text{if } A = \emptyset \end{cases}$

(ii) $\forall B \in \mathcal{P}(\mathbb{R})\ {\chi_E}^{-1}(B) = \begin{cases} E & \text{if } 0 \notin B \text{ and } 1 \in B \\ X/E & \text{if } 1 \notin B \text{ and } 0 \in B \\ X & \text{if } \{0,1\} \subset B \\ \emptyset & \text{if } \{0,1\} \cap B = \emptyset \end{cases}$

Proof. **[INCOMPLETE]** □

Proposition 2.2.11 (Equivalent Condition for $\mathcal{S}$-Measurablity of the Characteristic Function)**.** *Suppose $(X, \mathcal{S})$ is a measurable space and $E \subset X$. Then, the characteristic function $\chi_E \colon X \to \mathbb{R}$ of E on X is $\mathcal{S}$-measurable $\iff E \in \mathcal{S}$.*

Proof. **[INCOMPLETE]** □

Remark 2.2.6.

Theorem 2.2.5 (Criterion for $\mathcal{S}$-Measurability of a Function)**.** *Suppose $(X, \mathcal{S})$ is a measurable space and $f \colon X \to \mathbb{R}$ satisfies*

$$\forall a \in \mathbb{R},\ f^{-1}((a, \infty)) \in S.$$

Then, f is a $\mathcal{S}$-measurable function.

Proof. Suppose $(X, \mathcal{S})$ is a measurable space and $f \colon X \to \mathbb{R}$ satisfies

$$\forall a \in \mathbb{R},\ f^{-1}((a, \infty)) \in S.$$

Let $\mathcal{T} = \{A \in \mathcal{P}(X) : f^{-1}(A) \in \mathcal{S}\}$. It suffices to show that $\mathcal{T}$ is a σ-algebra, for we may show that $\mathcal{B} \subset \mathcal{T}$. That is, we may then prove that $\forall B \in \mathcal{B},\ f^{-1}(B) \in \mathcal{S}$.

(i) We observe that $f^{-1}(\emptyset) = \{x \in X : f(x) \in \emptyset\} = \emptyset \in \mathcal{S}$, since $\forall x \in X\ f(x) \notin \emptyset$. It follows that $\emptyset \in \mathcal{T}$ by the definition of $\mathcal{T}$.

(ii) Consider arbitrary $A \in \mathcal{T}$. It follows that $f^{-1}(A) \in \mathcal{S}$. In addition, we have that

$$\begin{aligned} f^{-1}(\mathbb{R}/A) &= X/f^{-1}(A) && \text{(By an identity of preimage)} \\ &\in \mathcal{S} && \text{(By the closure of } \mathcal{S} \text{ under complementation)} \\ \implies \mathbb{R}/A &\in \mathcal{T} && \text{(By the definition of } \mathcal{T}\text{)} \end{aligned}$$

Therefore, we conclude that $\mathcal{T}$ is closed under complementation.

(iii) Suppose $A_1, A_2, \cdots \in \mathcal{T}$. It follows that $f^{-1}(A_1), f^{-1}(A_2), \cdots \in \mathcal{S}$. Hence,

$$\begin{aligned} f^{-1}(\cup_{k \in \mathbb{N}} A_k) &= \cup_{k \in \mathbb{N}} f^{-1}(A_k) && \text{(By an identity of preimage)} \\ &\in \mathcal{S} && \text{(By the closure of } \mathcal{S} \text{ under countable union)} \\ \implies \cup_{k \in \mathbb{N}} A_k &\in \mathcal{T} && \text{(By the definition of } \mathcal{T}\text{)} \end{aligned}$$

Thus, $\mathcal{T}$ is closed under countable union and, therefore, is a σ-algebra on $\mathbb{R}$ by definition.

By assumption, we have that $(a, \infty) \in \mathcal{T}$ $\forall a \in \mathbb{R}$. It follows that $(-\infty, a] \in \mathcal{T}$ $\forall a \in \mathbb{R}$ by (ii). We observe that $\forall a, b \in \mathbb{R}$ with $a < b$, $(a, \infty) \cap (-\infty, b] = (a, b]$ implies $(a, \infty) \cap (-\infty, b] \in \mathcal{T}$, by the closure of $\mathcal{T}$ under countable intersection. By the identities of intervals in $\mathbb{R}$, $\forall a, b \in \mathbb{R}$ with $a < b$, it holds that

$$(-\infty, a) = \bigcup_{k \in \mathbb{N}} \left(-k, a - \frac{1}{k}\right] \text{ and } (a, b) = \bigcup_{k \in \mathbb{N}} \left(a, b - \frac{1}{k}\right],$$

which then implies (a, b), $(-\infty, a) \in \mathcal{T}$ by the closure of $\mathcal{T}$ under countable union. That is, we showed that $\mathcal{T}$ contains all open intervals and, therefore, open subsets of $\mathbb{R}$, since open subsets of $\mathbb{R}$ are countable unions of open intervals, which is contained in $\mathcal{T}$ by its closure under countable union.

Thus, $\mathcal{T}$ is a σ-algebra on $\mathbb{R}$ containing all open subsets of $\mathbb{R}$. It follows that $\mathcal{B} \subset \mathcal{T}$ since $\mathcal{B}$ is the smallest σ-algebra on $\mathbb{R}$ containing all open subsets of $\mathbb{R}$. Hence, $\mathcal{T}$ contains all Borel sets implying that $\forall B \in \mathcal{B}$ $f^{-1}(B) \in \mathcal{S}$. By definition, f is $\mathcal{S}$-measurable as desired. $\square$

Definition 2.2.8 (Borel Measurable function)**.** *Suppose $X \subset \mathbb{R}$. Then, a function $f \colon X \to \mathbb{R}$ is Borel measurable if*

$$f^{-1}(B) \in \mathcal{B}, \forall B \in \mathcal{B}.$$

Theorem 2.2.6 (Continuity Implies Borel Measurability)**.** *Suppose $B \in \mathcal{B}$ and $f \colon B \to \mathbb{R}$ is continuous. Then, f is Borel Measurable.*

Proof. Suppose $X \in \mathcal{B}$ and $f \colon X \to \mathbb{R}$ is continuous. It suffices to show that $f^{-1}((a, \infty)) \in \mathcal{B}$ $\forall a \in \mathbb{R}$ and apply a previous theorem.

Fix $a \in \mathbb{R}$. By the continuity of f, we have that $\forall p \in X$

$$\begin{aligned} \forall \epsilon > 0\ \exists \delta_p > 0 \text{ such that } x \in X \text{ and } |x - p| < \delta_p \\ \implies |f(x) - f(p)| = |f(p) - f(x)| < \epsilon. \end{aligned}$$

It follows that $p \in f^{-1}((a, \infty)) \iff f(p) \in (a, \infty)$ implies that

$$\begin{aligned} \exists \delta_p > 0 \text{ such that } x \in (p - \delta_p, p + \delta_p) \cap X \implies |f(p) - f(x)| < f(p) - a \\ \implies f(x) > a \iff f(x) \in (a, \infty) \iff x \in f^{-1}((a, \infty)) \\ \implies (p - \delta_p, p + \delta_p) \cap X \subset f^{-1}((a, \infty)) \end{aligned}$$

$$\implies \bigcup_{p \in f^{-1}((a,\infty))} [(p - \delta_p, p + \delta_p) \cap X] = \left[\bigcup_{p \in f^{-1}((a,\infty))} (p - \delta_p, p + \delta_p)\right] \cap X \subset f^{-1}((a, \infty))$$

by the distributivity of set operation. Denote $\mathcal{U}$ the union $\bigcup_{p \in f^{-1}((a,\infty))} (p - \delta_p, p + \delta_p)$. We observe that

$$f^{-1}((a, \infty)) \subset \mathcal{U} \cap X$$

since for any $p \in f^{-1}((a, \infty)) \subset X$, $p \in (p - \delta_p, p + \delta) \subset \mathcal{U}$ and certainly $p \in X$. Thus, we obtain that

$$f^{-1}((a, \infty)) = \mathcal{U} \cap X.$$

Recall, by definition, that $\mathcal{B}$ is the smallest σ-algebra on $\mathbb{R}$ containing all open sets of $\mathbb{R}$. It follows $\mathcal{U} \in \mathcal{B}$ since it is the union of open intervals, which is open. Hence, $\mathcal{U}, X \in \mathcal{B} \implies f^{-1}((a, \infty)) = \mathcal{U} \cap X \in \mathcal{B}$, by the closure of $\mathcal{B}$ under finite intersection. That is, we showed that $f^{-1}((a, \infty)) \in \mathcal{B}$, $\forall a \in \mathbb{R}$. By Theorem 2.39 (Axler 32), f is $\mathcal{B}$-measurable or, equivalently, Borel measurable. $\square$

Definition 2.2.9 (Increasing function)**.** *Suppose $X \subset \mathbb{R}$ and $f\colon X \to \mathbb{R}$ is a function. Then,*

(i) f is increasing if $\forall x, y \in X$ such that $x < y$, $f(x) \leq f(y)$;
(ii) f is strictly increasing if $\forall x, y \in X$ such that $x < y$, $f(x) < f(y)$.

Theorem 2.2.7 (Increasing Implies Borel Measurability)**.** *Suppose $X \in \mathcal{B}$ and $f\colon X \to \mathbb{R}$ is increasing. Then, f is Borel measurable.*

Proof. **[INCOMLETE]** □

Theorem 2.2.8 (Regularity Properties of $\mathcal{S}$-Measurable Functions)**.** *Suppose $(X, \mathcal{S})$ is a measurable space and $f\colon X \to \mathbb{R}$ is an $\mathcal{S}$-measurable function.*

(i) Suppose that $f(X) \subset Y \subset \mathbb{R}$ and $g\colon Y \to \mathbb{R}$ is Borel-measurable. Then, $g \circ f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable.

(ii) Suppose that $g\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. Then,
(a) $f + g$, $f - g$, fg are $\mathcal{S}$-measurable, and
(b) $g(x) \neq 0\ \forall x \in X \implies f/g$ is $\mathcal{S}$-measurable.

Proof. **[INCOMPLETE]** Suppose $(X, \mathcal{S})$ is a measurable space and $f\colon X \to \mathbb{R}$ is an $\mathcal{S}$-measurable function.

(i) Suppose that $f(X) \subset Y \subset \mathbb{R}$ and $g\colon Y \to \mathbb{R}$ is Borel-measurable. It follows that

$$\begin{aligned} \forall B \in \mathcal{B}\ [g \circ f]^{-1}(B) &= f^{-1}(g^{-1}(B)) && \text{(By a preimage identity)} \\ \implies [g \circ f]^{-1}(B) &\in \mathcal{S} \end{aligned}$$

since g is Borel measurable implies $g^{-1}(B) \in \mathcal{B}$ and f is $\mathcal{S}$-measurable implies that indeed $f^{-1}(g^{-1}(B)) \in \mathcal{S}$. Hence, $g \circ f$ is $\mathcal{S}$-measurable by definition.

(ii) Suppose that $g\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable.

(a) We claim that

$$(f + g)^{-1}((a, \infty)) = \bigcup_{r \in \mathbb{Q}} \left(f^{-1}((r, \infty)) \cap g^{-1}((a - r, \infty))\right).$$

By the below proposition, $-g\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. It follows, from the above result, that $f - g$ is $\mathcal{S}$-measurable since $f - g = f + (-g)$.

$fg = \frac{(f+g)^2 - f^2 - g^2}{2}$.

(b) □

Proposition 2.2.12 ($\mathcal{S}$-Measurable Function Arose from Function Composition)**.** *Suppose (X, S) is a measurable space, $k \in \mathbb{R}$, $p \in (0, \infty)$, and $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-Measurable. Then, kf, $|f|$, and f^p are $\mathcal{S}$-Measurable.*

Proof. Suppose (X, S) is a measurable space, $k \in \mathbb{R}$, and $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable.

(i) Define $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) = kx\ \forall x \in \mathbb{R}$. We observe that $\mathbb{R} \in \mathcal{B}$ and g is continuous implies that g is Borel measurable by Theorem Theorem 2.41 (Axler 33). It follows that $kf\colon X \to \mathbb{R}$, defined by $[kf](x) = [g \circ f](x)\ \forall x \in X$, is $\mathcal{S}$-measurable by the compositional regularity of $\mathcal{S}$-measurable functions.

(ii) Define $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) = |x|$ $\forall x \in \mathbb{R}$. Apply the argument in Part (i) to $|f| : X \to \mathbb{R}$, defined by $|f|(x) = g(f(x))$ $\forall x \in X$, and we obtain the desired result.

(iii) Suppose $p \in (0, \infty)$. Define $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) = x^p$ $\forall x \in \mathbb{R}$. Apply the argument in Part (i) to $f^p\colon X \to \mathbb{R}$, defined by $[f^2](x) = g(f(x))$ $\forall x \in X$, and we obtain the desired result. ☐

Theorem 2.2.9 ($\mathcal{S}$-Measurablity of Limit Function of $\mathcal{S}$-Measurable Functions)**.** *Suppose $(X, \mathcal{S})$ is a measurable space and $f_k\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable, $\forall k \in \mathbb{N}$. Suppose that $\lim_{k\to\infty} f_k(x)$ exists, $\forall x \in X$. Then, $f\colon X \to \mathbb{R}$, defined by*

$$f(x) = \lim_{k\to\infty} f_k(x) \ \forall x \in X$$

is $\mathcal{S}$-measurable.

Proof. **[INCOMLETE]** ☐

Definition 2.2.10 (Borel sets in $[-\infty, \infty]$)**.** *A set $B \subset [-\infty, \infty]$ is a Borel set if $B \cap \mathbb{R} \in \mathcal{B}$.*

Proposition 2.2.13 (Characteriztion of Borel Sets in $[-\infty, \infty]$)**.** *$C \subset [-\infty, \infty]$ is a Borel set $\iff$ $\exists B \in \mathcal{B}$ such that $C = B \cup S$ for some $S \in \{\emptyset, \{-\infty\}, \{\infty\}, \{-\infty, \infty\}\}$.*

Proof. Suppose that $C \subset [-\infty, \infty]$.

($\implies$) It suffices to prove the contrapositive of the statement. That is, it remains to show that $\forall B \in \mathcal{B}$, $\forall S \in \{\emptyset, \{-\infty\}, \{\infty\}, \{-\infty, \infty\}\}$ $C \neq B \cup S$ implies that $C \cap \mathbb{R} \notin \mathcal{B}$. Suppose, to the contrary, that $C \cap \mathbb{R} \in \mathcal{B}$. Then, we have that

$$\begin{aligned}
\mathbb{R}/[C \cap \mathbb{R}] &= \mathbb{R}/C \cup \mathbb{R}/\mathbb{R} && \text{(By De Morgan's Law)} \\
&= \mathbb{R}/C \in \mathcal{B} && \text{(By the closure of } \mathcal{B} \text{ under complementation)} \\
\implies \mathbb{R}/[\mathbb{R}/C] &= C \in \mathcal{B}, && \text{(By the closure of } \mathcal{B} \text{ under complementation)} \\
\implies C &\neq C \cup \emptyset && \text{(By assumption, } \forall B \in \mathcal{B}\ C \neq B \cup \emptyset)
\end{aligned}$$

which is a contradiction, since we would then obtain that $C \neq C$. Hence, we conclude that $C \cap \mathbb{R} \notin \mathcal{B}$ and prove the contrapositive as desired.

($\impliedby$) Suppose $\exists B \in \mathcal{B}$ such that $C = B \cup S$ for some $S \in \{\emptyset, \{-\infty\}, \{\infty\}, \{-\infty, \infty\}\}$. It follows that $C \cap \mathbb{R} = [B \cup S] \cap \mathbb{R} = [B \cap \mathbb{R}] \cup [\mathbb{R} \cap S] = [B \cap \mathbb{R}] \cup \emptyset = B \cap \mathbb{R} = B$, since $\forall B \in \mathcal{B}$ we have $B \subset \mathbb{R}$. Hence, indeed $C \cap \mathbb{R} \in \mathcal{B}$. By definition, $C \subset [-\infty, \infty]$ is a Borel set. ☐

Proposition 2.2.14 (Borel Algebra on $[-\infty, \infty]$)**.** *Let $\mathcal{B}_\infty$ be the collection of all Borel sets on $[-\infty, \infty]$. Then, $\mathcal{B}_\infty$ is a σ-algebra on $[-\infty, \infty]$.*

We refer to $\mathcal{B}_\infty$ as the Borel Algebra on $[-\infty, \infty]$ or the extended Borel Algebra.

Proof. **[INCOMPLETE]** ☐

Definition 2.2.11 (Measurable function on $[-\infty, \infty]$)**.** *Suppose $(X, \mathcal{S})$ is a measurable space. Then, $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable if*

$$\forall B \in \mathcal{B}_\infty \ f^{-1}(B) \in \mathcal{S}$$

Theorem 2.2.10 (Criterion for $\mathcal{S}$-Measurability of a Function on $[-\infty, \infty]$)**.** *Suppose $(X, \mathcal{S})$ is a measurable space and $f\colon X \to [-\infty, \infty]$ satisfies*

$$\forall a \in \mathbb{R},\ f^{-1}((a, \infty]) \in S.$$

Then, f is a $\mathcal{S}$-measurable function.

Proof. **[INCOMPLETE]** □

Theorem 2.2.11. *Suppose $(X, \mathcal{S})$ is a measurable space and $f_1, f_2, \ldots$ is a sequence of $\mathcal{S}$-measurable functions from X to $[-\infty, \infty]$. Then, $g, h \colon X \to [-\infty, \infty]$, defined by*

$$g(x) = \inf_{k \in \mathbb{N}} f_k(x) \text{ and } h(x) = \sup_{k \in \mathbb{N}} f_k(x) \; \forall x \in X,$$

are $\mathcal{S}$-measurable.

Proof. Suppose $(X, \mathcal{S})$ is a measurable space and $(f_n)_{n \in \mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions from X to $[-\infty, \infty]$. Define $h, g \colon X \to [-\infty, \infty]$ by

$$h(x) = \sup_{j \in \mathbb{N}} f_j(x) \text{ and } g(x) = \inf_{j \in \mathbb{N}} f_j(x), \; \forall x \in X$$

respectively. Fix $a \in \mathbb{R}$. We first claim that

$$h^{-1}((a, \infty]) = \bigcup_{j \in \mathbb{N}} {f_j}^{-1}((a, \infty]). \tag{$*_0$}$$

Fix $p \in h^{-1}((a, \infty])$. It follows that $\sup_{j \in \mathbb{N}} f_j(p) \in (a, \infty]$ by definition and, thus,

$$\exists N \in \mathbb{N} \text{ such that } \sup_{j \in \mathbb{N}} f_j(p) - \left(\sup_{j \in \mathbb{N}} f_j(p) - a \right) < f_N(p)$$

by an equivalent condition of supremum and since $\left(\sup_{j \in \mathbb{N}} f_j(p) - a\right) > 0$. Hence, we yield that $f_N(p) \in (a, \infty]$, which then implies that

$$p \in {f_N}^{-1}((a, \infty]) \subset \bigcup_{j \in \mathbb{N}} {f_j}^{-1}((a, \infty]). \tag{By definition}$$

$$\therefore h^{-1}((a, \infty]) \subset \bigcup_{j \in \mathbb{N}} {f_j}^{-1}((a, \infty]).$$

Now, fix $q \in \bigcup_{j \in \mathbb{N}} {f_j}^{-1}((a, \infty])$. Then, there exists $M \in \mathbb{N}$ such that $f_M(q) > a$. Certainly, we have

$$\begin{aligned} h(q) = \sup_{j \in \mathbb{N}} f_j(q) \geq f_M(q) > a &\implies q \in h^{-1}((a, \infty]) \\ &\implies \bigcup_{j \in \mathbb{N}} {f_j}^{-1}((a, \infty]) \subset h^{-1}((a, \infty]) \end{aligned}$$

and conclude $(*_0)$ holds. We note that, $\forall j \in \mathbb{N}$, ${f_j}^{-1}((a, \infty]) \in \mathcal{S}$ since f_j is $\mathcal{S}$-measurable by hypothesis. It follows that $h^{-1}((a, \infty]) \in \mathcal{S}$ as well, by $(*_0)$ and the closure of $\mathcal{S}$ under countable union. By Theorem 2.52 (Axler 37), h is $\mathcal{S}$-measurable.

That is, we showed that for any arbitrary sequence $(F_k)_{k \in \mathbb{N}}$ of $\mathcal{S}$-measurable functions defined from X to $[-\infty, \infty]$ satisfies the property that the function defined by $\sup_{j \in \mathbb{N}} F_j(x) \; \forall x \in X$ is $\mathcal{S}$-measureable.

By an identity of infimum, we have that

$$g(x) = \inf_{j \in \mathbb{N}} f_j(x) = -\sup_{j \in \mathbb{N}} \left[-f_j(x) \right], \forall x \in X.$$

By the regularity of $\mathcal{S}$-measurable functions, $(-f_j)_{j \in \mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions defined from X to $[-\infty, \infty]$. By the preceding result, $-g \colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable. It follows, from the regularity of $\mathcal{S}$-measurable functions, that $g = -(-g)$ is $\mathcal{S}$-measurable as desired. □

Proposition 2.2.15 (Restriction of a $\mathcal{S}$-Measurable Function is $\mathcal{S}$-measurable)**.** *Suppose $(X, \mathcal{S})$ is a measurable space, $E \in \mathcal{S}$, and $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. Then, $f|_E\colon E \to \mathbb{R}$ is $\mathcal{S}$-measurable.*

Proof. Suppose $(X, \mathcal{S})$ is a measurable space, $E \in \mathcal{S}$, and $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. Consider the restriction $f|_E\colon E \to \mathbb{R}$ of f to E. Fix $B \in \mathcal{B}$. It follows that

$$\begin{aligned} {f|_E}^{-1}(B) = \{x \in E : [f|_E](x) \in B\} = \{x \in E : f(x) \in B\} & & \text{(By definition)} \\ = \left\{x \in X : x \in E \text{ and } x \in f^{-1}(B)\right\} & \\ = E \cap f^{-1}(B), & \end{aligned}$$

which is contained in $\mathcal{S}$ by the closure of $\mathcal{S}$ under finite intersection, since f is $\mathcal{S}$-measurable and $E \in \mathcal{S}$ by hypothesis. By definition, $f|_E$ is $\mathcal{S}$-measurable. $\square$

Proposition 2.2.16. *Suppose $(X, \mathcal{S})$ is a measurable space, $E \in \mathcal{S}$, and $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. Suppose further that $\tilde{f}\colon X \to \mathbb{R}$ is a function such that*

(i) $\tilde{f}$ is identically f in X/E, and

(ii) $\forall B \in \mathcal{B}$, $\left\{x \in E : \tilde{f}(x) \in B\right\} \in \mathcal{S}$.

Then, $\tilde{f}$ is $\mathcal{S}$-measurable.

Proof. Suppose $(X, \mathcal{S})$ is a measurable space, $E \in \mathcal{S}$, and $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable. Suppose further that $\tilde{f}\colon X \to \mathbb{R}$ is a function such that $\forall x \in X/E$, $\tilde{f}(x) = f(x)$ and $\forall B \in \mathcal{B}$, $\left\{x \in E : \tilde{f}(x) \in B\right\} \in \mathcal{S}$.

Fix $B \in \mathcal{B}$. It follows, from our hypotheses, that

$$\begin{aligned} \tilde{f}^{-1}(B) = \left\{x \in X : \tilde{f}(x) \in B\right\} &= \left\{x \in X/E : \tilde{f}(x) \in B\right\} \cup \left\{x \in E : \tilde{f}(x) \in B\right\} \\ &= \{x \in X/E : f(x) \in B\} \cup \left\{x \in E : \tilde{f}(x) \in B\right\} \\ &= [f|_{X/E}]^{-1}(B) \cup \left\{x \in E : \tilde{f}(x) \in B\right\}. \end{aligned}$$

We note, by a proposition above, that $X/E \in \mathcal{S}$ and f is $\mathcal{S}$-measurable implies that $f|_{X/E}\colon X/E \to \mathbb{R}$ is $\mathcal{S}$-measurable. Thus, $[f|_{X/E}]^{-1}(B) \in \mathcal{S}$, implying that $\tilde{f}^{-1}(B) \in \mathcal{S}$, by the closure of $\mathcal{S}$ under finite union. By definition, $\tilde{f}$ is $\mathcal{S}$-measurable as desired. $\square$

Theorem 2.2.12 (Modifying a Borel Measurable Function on a Countable Set)**.** *Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is Borel measurable, $E \subset \mathbb{R}$ is countable, and $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$ is identically f in $\mathbb{R}/E$. Then, $\tilde{f}$ is Borel measurable.*

Proof. Consider the measurable space $(\mathbb{R}, \mathcal{B})$. Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is Borel measurable, $E \subset \mathbb{R}$ is countable, and $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$ satisfies $\forall x \in \mathbb{R}/E$, $\tilde{f}(x) = f(x)$.

By Example 2.30 (Axler 29), $E \in \mathcal{B}$ since it is countable. Fix $B \in \mathcal{B}$. We observe that $\left\{x \in E : \tilde{f}(x) \in B\right\} \subset E$, which implies that it is countable and, therefore, contained in $\mathcal{B}$. By a proposition above, $\tilde{f}$ is Borel measurable. $\square$

Theorem 2.2.13 (Modifying a Lebesgue Measurable Function on a Set of Lebesgue Measure Zero)**.** *Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable, $E \subset \mathbb{R}$ has Lebesgue measure zero, and $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$ identically f in $\mathbb{R}/E$. Then, $\tilde{f}$ is Lebesgue measurable.*

Proof. Consider the measurable space $(\mathbb{R}, \mathcal{L})$. Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable, $E \subset \mathbb{R}$ has Lebesgue measure zero, and $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$ satisfies $\forall x \in \mathbb{R}/E$, $\tilde{f}(x) = f(x)$.

Since E has Lebesgue measure zero, $E \in \mathcal{L}$. Fix $B \in \mathcal{B}$. We observe that $\left\{x \in E : \tilde{f}(x) \in B\right\} \subset E$, which implies that it also has Lebesgue measure zero and, therefore, is contained in $\mathcal{L}$. By a proposition above, $\tilde{f}$ is Lebesgue measurable. $\square$

2.3 Measures and Their Properties

Definition 2.3.1 (Measure on a measurable space)**.** *Suppose $(X, \mathcal{S})$ is a measurable space. A measure on $(X, \mathcal{S})$ is a map $\mu\colon \mathcal{S} \to [0, \infty]$ such that*

(a) $\mu(\emptyset) = 0$, and
(b) for every sequence of disjoint sets $E_1, E_2, \cdots \in \mathcal{S}$ it holds that

$$\mu\left(\bigsqcup_{k\in\mathbb{N}} E_k\right) = \sum_{k\in\mathbb{N}} \mu(E_k).$$

Remark 2.3.1. *Again, we require countability in the above definition to eliminate the case that $\mu(\mathbb{R}) = 0$ for some measure μ.*

Proposition 2.3.1 (Measure Menu Theorem)**.** *The following maps are measures.*

(a) ***(Counting Measure)*** *Suppose X is a set. Then, the counting measure $\mu\colon \mathcal{P}(X) \to [0, \infty]$ on $(X, \mathcal{P}(X))$ is defined by*

$$\mu(E) = \begin{cases} card(E) & \text{if } E \text{ is a finite set} \\ \infty & \text{if } E \text{ is not a finite set} \end{cases}, \ \forall E \in \mathcal{P}(X).$$

(b) ***(Dirac Measure)*** *Suppose $(X, \mathcal{S})$ is a measurable space and $c \in X$. Then, the Dirac Measure δ_c with respect to c is defined by*

$$\delta_c(E) = \begin{cases} 1 & c \in E \\ 0 & c \notin E \end{cases}, \ \forall E \in \mathcal{S}.$$

(c) ***(Sum Measure)*** *Suppose $(X, \mathcal{S})$ is a measurable space, and $w\colon X \to [0, \infty]$ is a function. Then, the sum measure $\mu\colon \mathcal{S} \to [0, \infty]$ is defined by*

$$\mu(E) = \sum_{x\in E} w(x), \ \forall E \in \mathcal{S}.$$

(d) Suppose $(X, \mathcal{S})$ is a measurable space, where $\mathcal{S}$ is the σ-algebra on X consisting of all subsets of X that are countable or have a countable complement in X. Define $\mu\colon \mathcal{S} \to [0, \infty]$ by

$$\mu(E) = \begin{cases} 0 & \text{if } E \text{ is countable} \\ 3 & \text{if } E \text{ is uncountable} \end{cases}, \ \forall E \in \mathcal{S}.$$

(e) ***(Trivial Measure)*** *Suppose $(X, \mathcal{S})$ is a measurable space. Then, the trivial measure $\mu\colon \mathcal{S} \to [0, \infty]$ on $(X, \mathcal{S})$ is defined by*

$$\mu(E) = 0, \ \forall E \in \mathcal{S}.$$

Definition 2.3.2 (Measure space)**.** *A measure space is a measurable space $(X, \mathcal{S})$ along with a measure μ on $(X, \mathcal{S})$.*

That is, a measure space is an order triple $(X, \mathcal{S}, \mu)$, where $(X, \mathcal{S})$ is a measurable space and μ is a measure on $(X, \mathcal{S})$.

Proposition 2.3.2 (Measure Space of Measure Zero $\iff$ Identically Trivial Measure)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Then, $\mu(X) = 0 \iff \mu$ is identically the trivial measure.*

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space.

$(\implies)$: Suppose that $\mu(X) = 0$. By the monotonicity of measure, $\mu(E) \leq \mu(X) = 0\ \forall E \in \mathcal{S}$. Thus, we have that $\mu(E) = 0\ \forall E \in \mathcal{S}$. By definition, μ is the trivial measure.

$(\impliedby)$: Suppose μ is the trivial measure. It follows that $\mu(E) = 0\ \forall E \in \mathcal{S}$. In particular, $X \in \mathcal{S} \implies \mu(X) = 0$. $\square$

Theorem 2.3.1 (Monotonicity of Measure; Measure of Set Difference)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, $D, E \in \mathcal{S}$, and $D \subset E$. Then, the following statements hold:*

(a) $\mu(D) \leq \mu(E)$.
(b) Suppose that $\mu(D) < \infty$. Then, $\mu(E/D) = \mu(E) - \mu(D)$.

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, $D, E \in \mathcal{S}$, and $D \subset E$.

(a) We observe that $E = E/D \sqcup D$ implies that

$$\begin{aligned} \mu(E) = \mu(E/D \sqcup D) &= \mu(E/D) + \mu(D) \\ &\geq \mu(D). && \text{(Since } \mu(E/D) \geq 0) \end{aligned}$$

(b) Suppose further that $\mu(D) < \infty$. From Part (a) we obtain that $\mu(E) = \mu(E/D) + \mu(D)$. Hence, it follows that

$$\mu(E/D) = \mu(E) - \mu(D).$$

$\square$

Theorem 2.3.2 (Countable Subadditivity of Measure)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $E_1, E_2, \dots \in \mathcal{S}$. Then,*

$$\mu\left(\bigcup_{k\in\mathbb{N}} E_k\right) \leq \sum_{k\in\mathbb{N}} \mu(E_k).$$

Proof. **[DO: PROVE $\bigcup_{k\in\mathbb{N}} E_k = \bigsqcup_{k\in\mathbb{N}} E_k/D_k$]** Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $E_1, E_2, \dots \in \mathcal{S}$. Let $D_1 = \emptyset$ and $D_k = E_1 \cup E_2 \cup \dots \cup E_{k-1}\ \forall k \in \mathbb{N}/\{1\}$. We then obtain that

$$\bigcup_{k\in\mathbb{N}} E_k = \bigsqcup_{k\in\mathbb{N}} E_k/D_k$$

which implies that

$$\begin{aligned} \mu\left(\bigcup_{k\in\mathbb{N}} E_k\right) = \mu\left(\bigsqcup_{k\in\mathbb{N}} E_k/D_k\right) &= \sum_{k\in\mathbb{N}} \mu(E_k/D_k) \\ &\leq \sum_{k\in\mathbb{N}} \mu(E_k) \end{aligned}$$

by the monotonicity of measure since $E_k/D_k \subset E_k\ \forall k \in \mathbb{N}$. $\square$

Theorem 2.3.3 (Measure of Increasing Union)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, and $E_1, E_2, \dots \in \mathcal{S}$ satisfies $E_1 \subset E_2 \subset \dots$. Then,*

$$\mu\left(\bigcup_{k\in\mathbb{N}} E_k\right) = \lim_{k\to\infty} \mu(E_k).$$

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, and $E_1, E_2, \dots \in \mathcal{S}$ satisfies $E_1 \subset E_2 \subset \dots$. Here, we have two cases to consider, namely $\mu(E_N) = \infty$ for some $N \in \mathbb{N}$ and $\mu(E_k) < \infty$ $\forall k \in \mathbb{N}$.

Suppose the first case holds. Then, $\mu\left(\cup_{k\in\mathbb{N}} E_k\right) = \infty$ by the monotonicity of measure since $E_N \subset \cup_{k\in\mathbb{N}} E_k$ with $\mu(E_N) = \infty$. Similarly, by the monotonicity of measure, we have that $\forall k \in \{n \in \mathbb{N} : n \geq N\}$ $\mu(E_k) = \infty$, which yields that $\lim_{k\to\infty} \mu(E_k) = \infty = \mu\left(\cup_{k\in\mathbb{N}} E_k\right)$ as desired.

Now, suppose the second case holds. Let $E_0 = \emptyset$. It follows that $\cup_{k\in\mathbb{N}} E_k = \sqcup_{k\in\mathbb{N}} E_k / E_{k-1}$. As an immediate result,

$$\begin{aligned}\mu \bigcup_{n\in\mathbb{N}} E_n = \mu\left(\bigsqcup_{n\in\mathbb{N}} E_n/E_{n-1}\right) &= \sum_{n=1}^{\infty} \mu(E_n/E_{n-1}) \\ = \lim_{k\to\infty} \sum_{n=1}^{k} \mu(E_n/E_{n-1}) = \lim_{k\to\infty} &\sum_{n=1}^{k} [\mu(E_n) - \mu(E_{n-1})] \\ = \lim_{k\to\infty} [\mu(E_1) - \mu(E_0)] + \dots + [\mu(E_{k-1}) - \mu(E_{k-2})] &+ [\mu(E_k) - \mu(E_{k-1})] \\ &= \lim_{k\to\infty} \mu(E_k).\end{aligned}$$

$\square$

Theorem 2.3.4 (Measure of Decreasing Intersection)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, and $E_1, E_2, \dots \in \mathcal{S}$ satisfies $E_1 \supset E_2 \supset \dots$ and $\mu(E_1) < \infty$. Then,*

$$\mu\left(\bigcap_{k\in\mathbb{N}} E_k\right) = \lim_{k\to\infty} \mu(E_k).$$

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, and $E_1, E_2, \dots \in \mathcal{S}$ satisfies $E_1 \supset E_2 \supset \dots$ and $\mu(E_1) < \infty$. We observe, by De Morgan's Law, that

$$\forall k \in \mathbb{N}\ E_1 / \bigcap_{k\in\mathbb{N}} E_k = \bigcup_{k\in\mathbb{N}} E_1/E_k.$$

In addition, we have that

$$E_1/E_1 \subset E_1/E_2 \subset E_1/E_3 \subset \dots,$$

where $E_1/E_k \in \mathcal{S}$ $\forall k \in \mathbb{N}$ by the closure of $\mathcal{S}$ under complementation. Hence, applying Theorem 2.59 (Axler 43), we yield that

$$\begin{aligned}\mu\left(E_1 / \bigcap_{k\in\mathbb{N}} E_k\right) = \mu\left(\bigcup_{k\in\mathbb{N}} E_1/E_k\right) &= \lim_{k\to\infty} \mu\left(E_1/E_k\right) \\ \implies \mu(E_1) - \mu\left(\bigcap_{k\in\mathbb{N}} E_k\right) &= \lim_{k\to\infty} [\mu(E_1) - \mu(E_k)] && \text{(By Theorem 2.57 (Axler 42))} \\ \implies \mu(E_1) - \mu\left(\bigcap_{k\in\mathbb{N}} E_k\right) &= \mu(E_1) - \lim_{k\to\infty} \mu(E_k) && \text{(Since } \lim_{k\to\infty}\mu(E_k) < \infty\text{)} \\ \implies \mu\left(\bigcap_{k\in\mathbb{N}} E_k\right) &= \lim_{k\to\infty} \mu(E_k).\end{aligned}$$

We note that $\lim_{k\to\infty} \mu(E_k) < \infty$ holds. Suppose otherwise. If $\mu(E_1) = 0$, then we obtain that $\mu(E_k) = 0$ $\forall k \in \mathbb{N}$ by the monotonicity of measure and since $(E_k)_{k\in\mathbb{N}}$ is decreasing. Thus, we obtain that $\lim_{k\to\infty} \mu(E_k) = 0$, which is a contradiction.

Hence, we suppose that $\mu(E_1) > 0$. By the definition of a limit, we have that

$$\forall \epsilon > 0 \; \exists N \in \mathbb{N} \text{ such that } n \in \mathbb{N} \text{ and } n > N \implies \mu(E_n) > \epsilon.$$

In particular,

$$\exists N \in \mathbb{N} \text{ such that } n \in \mathbb{N} \text{ and } n > N \implies \mu(E_n) > \mu(E_1)$$

which is a contradiction since $E_1 \supset E_k \; \forall k \in \mathbb{N}$ implies that $\mu(E_1) \geq \mu(E_k)$ by the monotonicity of measure. Hence, the limit is indeed finite. Here, we complete the proof. □

Theorem 2.3.5 (Measure of Union). *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $D, E \in \mathcal{S}$ satisfies $\mu(D \cap E) < \infty$. Then,*

$$\mu(D \cup E) = \mu(D) + \mu(E) - \mu(D \cap E).$$

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $D, E \in \mathcal{S}$ satisfies $\mu(D \cap E) < \infty$. Let $A = D \cap E$. We note that $D \cup E = [D/A] \sqcup [E/A] \sqcup A$ which yields that

$$\begin{aligned} \mu(D \cup E) &= \mu([D/A] \sqcup [E/A] \sqcup A) \\ &= \mu([D/A]) + \mu([E/A]) + \mu(A) \\ &= [\mu(D) - \mu(A)] + [\mu(E) - \mu(A)] + \mu(A) && \text{(By Theorem 2.57 (Axler 42))} \\ \implies \mu(D \cup E) &= \mu(D) + \mu(E) - \mu(D \cap E). \end{aligned}$$

□

2.4 Lebesgue Measure

Theorem 2.4.1 (Openness Implies Finite Additivity of Outer Measure Under Disjoint Union)**.** *Suppose $A, G \subset \mathbb{R}$ are disjoint and G is open. Then,*

$$|A \sqcup G| = |A| + |G|.$$

Proof. **[INCOMPLETE]**

$\square$

Theorem 2.4.2 (Closedness Implies Finite Additivity of Outer Measure Under Disjoint Union)**.** *Suppose $A, F \subset \mathbb{R}$ are disjoint and F is closed. Then,*

$$|A \sqcup F| = |A| + |F|.$$

Proof. Suppose $A, F \subset \mathbb{R}$ are disjoint and F is closed. Consider arbitrary sequence $(I_k)_{k\in\mathbb{N}}$ of open intervals of $\mathbb{R}$ such that $[A \sqcup F] \subset \bigcup_{k\in\mathbb{N}} I_k$. Here, we denote $\bigcup_{k\in\mathbb{N}} I_k$ by G.

Suppose that $|F| = \infty$. It follows that $F \subset A \sqcup F$ implies that $|A \sqcup F| = \infty$ since $|F| \leq |A \sqcup F|$ by the monotonicity of outer measure. In addition, we also have that $|A| + |F| = \infty$. Thus, we obtain that $|A \sqcup F| = |A| + |F|$.

Now, suppose that $|F| < \infty$. We observe that G is an open set since the union of open sets is open. In addition, $\forall a \in A$, $a \in G$ and $a \notin F$. Thus, it holds that $A \subset G/F$. Moreover, $G/F = G \cap [\mathbb{R}/F]$ is an open set since the intersection of finitely many open sets is open. It follows that

$$\begin{aligned}
& |A| \leq |G/F| && \text{(By the monotonicity of outer measure)} \\
\implies & |A| + |F| \leq |G/F| + |F| && \\
& = (|G| - |F|) + |F| = |G| && \text{(Theorem 2.62 (Axler 47))}
\end{aligned}$$

since $G = F \sqcup (G/F)$ and G/F is open. As an immediate result,

$$|A| + |F| \leq |G| \leq \sum_{k\in\mathbb{N}} l(I_k). \qquad \text{(By the definition of outer measure)}$$

It then holds that $|A| + |F|$ is a lower bound for the set

$$\left\{ \sum_{j\in\mathbb{N}} l(I_j) : \{I_j : j \in \mathbb{N}\} \text{ is a cover of } A \sqcup F \text{ by open intervals} \right\}.$$

Since $|A \sqcup F|$ is the least upper bound of the above set by definition, we then obtain that $|A \sqcup F| \geq |A| + |F|$. Recall that, by the subadditivity of outer measure, $|A \sqcup F| \leq |A| + |F|$. Combining the above results, we conclude that $|A \sqcup F| = |A| + |F|$. $\square$

Corollary 2.4.1 (Additivity of Outer Measure of Certain Sets)**.** *Suppose $A, G \subset \mathbb{R}$ are disjoint.* ***DOUBLE CHECK***

(i) Suppose that G or A/G is open. Then, $|A/G| = |A| - |G|$.

(ii) Suppose that $A \cup G$ or $\mathbb{R}/G$ is closed. Then, $|A \cup G| = |A| - |\mathbb{R}/G|$.

Theorem 2.4.3 (Approximating a Borel Set Via a Smaller Closed Set)**.** *(a)*

$$\mathcal{L} = \{D \in \mathcal{P}(\mathbb{R}) : \forall\epsilon > 0, \text{ there exists a closed set } F \subset D \text{ such that } |D/F| < \epsilon\}$$

is a σ-algebra on $\mathbb{R}$.

(b) Suppose $B \in \mathcal{B}$. Then, $\forall \epsilon > 0$ $\exists F \in \mathcal{P}(\mathbb{R})$ such that

(i) $F \subset B$ is closed, and

(ii) $|B/F| < \epsilon$.

Proof. **(a)** Consider the set

$$\mathcal{L} = \{D \in \mathcal{P}(\mathbb{R}) : \forall \epsilon > 0, \text{ there exists a closed set } F \subset D \text{ such that } |D/F| < \epsilon\}.$$

We observe that $\emptyset$ is closed, $\emptyset \subset \emptyset$, and $\emptyset/\emptyset = \{x \in \emptyset : x \notin \emptyset\} = \emptyset$. It follows that $\forall \epsilon > 0$ $|\emptyset/\emptyset| = |\emptyset| = 0 < \epsilon$, which implies that $\emptyset \in \mathcal{L}$.

Here, we claim that $\mathcal{L}$ is closed under countable intersections. Suppose that $D_1, D_2, \cdots \in \mathcal{L}$ and fix $\epsilon > 0$. Then, we have that $\exists F_1, F_2, \cdots \in \mathcal{P}(\mathbb{R})$ such that $\forall k \in \mathbb{N}$ it holds that $F_k \subset D_k$, F_k is closed, and $|D_k/F_k| < \frac{1}{2^k}$.

We observe that $\bigcap_{k\in\mathbb{N}} F_k$ is closed since the intersection of closed sets is closed. In addition, $\bigcap_{k\in\mathbb{N}} F_k \subset \bigcap_{k\in\mathbb{N}} D_k$. Fix $p \in \left[\bigcap_{k\in\mathbb{N}} D_k\right] / \left[\bigcap_{k\in\mathbb{N}} F_k\right]$. We observe that that $p \notin \bigcap_{k\in\mathbb{N}} F_k$ implies that $\exists m \in \mathbb{N}$ such that $p \notin F_m$, where $p \in \bigcap_{k\in\mathbb{N}} D_k$ implies that $p \in D_m$. It follows that $p \in D_m/F_m \subset \bigcup_{k\in\mathbb{N}}[D_k/F_k]$. Hence, we have that

$$\left[\bigcap_{k\in\mathbb{N}} D_k\right] / \left[\bigcap_{k\in\mathbb{N}} F_k\right] \subset \bigcup_{k\in\mathbb{N}} [D_k/F_k].$$

As an immediate result, we obtain that

$$\begin{aligned}
\left|\left[\bigcap_{k\in\mathbb{N}} D_k\right] / \left[\bigcap_{k\in\mathbb{N}} F_k\right]\right| &\leq \left|\bigcup_{k\in\mathbb{N}} [D_k/F_k]\right| && \text{(By the monotonicity of outer measure)} \\
&\leq \sum_{k\in\mathbb{N}} |D_k/F_k| && \text{(By the subadditivity of outer measure)} \\
< \sum_{k\in\mathbb{N}} \frac{\epsilon}{2^k} &= \epsilon \left[\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k - 1\right] \\
&= \epsilon \left[\frac{1}{1-\frac{1}{2}} - 1\right] = \epsilon.
\end{aligned}$$

Thus, $\bigcap_{k\in\mathbb{N}} D_k \in \mathcal{L}$, and we conclude that indeed $\mathcal{L}$ is closed under countable intersection.

It remains to show that $\mathcal{L}$ is closed under complementation. Consider arbitrary $D \in \mathcal{L}$. Here, we have two cases to consider; namely, $|D| < \infty$ and $|D| = \infty$.

Suppose the first case holds and fix $\emptyset > 0$. Since $D \in \mathcal{L}$, there exists a closed set $F \subset D$ such that

$$|D/F| = \frac{\epsilon}{2}. \tag{$*_1$}$$

Recall, from elementary real analysis, that if $S \subset \mathbb{R}$ is bounded below and w is a lower bound of S, then $w = \inf(S) \iff \forall \epsilon > 0$ $\exists x \in S$ such that $x < w + \epsilon$. By the definition,

$$|D| = \inf \left\{\sum_{k\in\mathbb{N}} l(J_k) : \{J_k \subset \mathbb{R} : k \in \mathbb{N}\} \text{ is a countable covering of } D \text{ by open intervals}\right\}.$$

It follows that there exists a covering $\{I_k \subset \mathbb{R} : k \in \mathbb{N}\}$ of D by open intervals such that

$$\sum_{k\in\mathbb{N}} l(I_k) < |D| + \frac{\epsilon}{2}.$$

Let $G = \bigcup_{k\in\mathbb{N}} I_k$. Then G is open, $G \supset D$, and $|G| \le \sum_{k\in\mathbb{N}} l(I_k)$ by definition since $\{I_k \subset \mathbb{R} : k \in \mathbb{N}\}$ covers G. Moreover, we have that

$$|G| < |D| + \frac{\epsilon}{2} \implies |G/D| = |G| - |D| < \frac{\epsilon}{2}. \qquad \text{(By Theorem 2.57 (Axler 42); } (*_2))$$

We observe that $\mathbb{R}/G$ is closed and $\mathbb{R}/G \subset \mathbb{R}/D$ since $D \subset G$. It follows, from Theorem 2.57 (Axler 42), that

$$\begin{aligned}
&[\mathbb{R}/D]/[\mathbb{R}/G] = [\mathbb{R}/D] \cap [(\mathbb{R})/(\mathbb{R}/G)] = [\mathbb{R}/D] \cap G = G/D \subset G/F && \text{(Since } F \subset D) \\
&\implies |[\mathbb{R}/D]/[\mathbb{R}/G]| \le |G/F| = |G| - |F| = |G| - |D| + |D| - |F| \\
&\qquad\qquad = |G/D| + |D/F| \\
&\qquad\qquad < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. && \text{(By } (*_1) \text{ and } (*_2))
\end{aligned}$$

Thus, $\mathbb{R}/D \in \mathcal{L}$, and we conclude that $\mathcal{L}$ is closed under complementation for the case when $|D| < \infty$.

Now, consider the case where $|D| = \infty$. Let $D_k = D \cap [-k, k]$ $\forall k \in \mathbb{N}$. We claim that $D_k \in \mathcal{L}$ $\forall k \in \mathbb{N}$. Fix $k \in \mathbb{N}$ and $\epsilon > 0$. Since $D \in \mathcal{L}$, we have that $\exists F \in \mathcal{P}(\mathbb{R})$ such that $F \subset D$, F is closed, and $|D/F| < \epsilon$.

Consider the set $F_k = F \cap [-k, k]$. We observe that $F_k \subset D_k$, and F_k is closed since the intersection of closed sets is closed. Fix $p \in D_k/F_k$. Then, we have that $p \in D, [-k, k]$ and $p \notin F$. Thus, $p \in (D/F) \cap [-k, k]$, and $D_k/F_k \subset (D/F) \cap [-k, k]$. As an immediate result, we obtain that

$$\begin{aligned}
|D_k/F_k| = |[D \cap [-k, k]]/[F \cap [-k, k]]| &\le |(D/F) \cap [-k, k]| \\
&\le |D/F| < \epsilon \qquad \text{(Since } (D/F) \cap [-k, k] \subset D/F)
\end{aligned}$$

by the monotonicity of outer measure. Thus, by the definition of $\mathcal{L}$, $D_k \in \mathcal{L}$ $\forall k \in \mathbb{N}$. We note that by the previous case, $\forall k \in \mathbb{N}$, $D_k \in \mathcal{L}$ and $|D_k| \le |D| < \infty$ implies that $\mathbb{R}/D_k \in \mathcal{L}$. It follows that

$$D = \bigcup_{k\in\mathbb{N}} D_k \implies \mathbb{R}/D = \bigcap_{k\in\mathbb{N}} [\mathbb{R}/D_k] \in \mathcal{L}$$

since $\mathcal{L}$ is closed under countable intersection. Hence, indeed $\mathcal{L}$ is closed under complementation.

To prove that $\mathcal{L}$ is a σ-algebra on $\mathbb{R}$, it remains to show that $\mathcal{L}$ is closed under countable union. Suppose that $E_1, E_2, \dots \in \mathcal{L}$. It follows that

$$\begin{aligned}
&\forall k \in \mathbb{N}, \mathbb{R}/D_k \in \mathcal{L} && \text{(By the closure of } \mathcal{L} \text{ under complementation)} \\
&\implies \bigcap_{k\in\mathbb{N}} [\mathbb{R}/E_k] \in \mathcal{L} && \text{(By the closure of } \mathcal{L} \text{ under countable intersection)} \\
&\implies \mathbb{R}/\left[\bigcap_{k\in\mathbb{N}} [\mathbb{R}/E_k]\right] \in \mathcal{L} && \text{(By the closure of } \mathcal{L} \text{ under complementation)} \\
&\implies \bigcup_{k\in\mathbb{N}} E_k \in \mathcal{L}. && \text{(Since } \textstyle\bigcup_{k\in\mathbb{N}} E_k = \mathbb{R}/\left[\bigcap_{k\in\mathbb{N}}[\mathbb{R}/E_k]\right] \text{ by De Morgan's Law)}
\end{aligned}$$

Thus, we conclude that $\mathcal{L}$ is closed under complementation, and $\mathcal{L}$ is indeed a σ-algebra on $\mathbb{R}$ by definition. Here, complete the proof for this part. $\square$

Proof. **(b)** Suppose $B \in \mathcal{B}$. By Part (a), the set

$$\mathcal{L} = \{D \in \mathcal{P}(\mathbb{R}) : \forall \epsilon > 0, \text{ there exists a closed set } F \subset D \text{ such that } |D/F| < \epsilon\}$$

is a σ-algebra on $\mathbb{R}$. To prove the desired claim, it suffices to show that all open sets of $\mathbb{R}$ are contained in $\mathcal{L}$, for this implies that $\mathcal{B} \subset \mathcal{L}$ and the desired statement follows immediately.

Consider arbitrary open set E of $\mathbb{R}$. Then, $\mathbb{R}/E$ is closed. Fix $\epsilon > 0$ and consider $F = \mathbb{R}/E$. We observe that $F \subset \mathbb{R}/E$, F is closed, and $|[\mathbb{R}/E]/F| = |\emptyset| = 0 < \epsilon$. Therefore, $\mathbb{R}/E \in \mathcal{L}$. By the closure of $\mathcal{L}$ under complementation, $E \in \mathcal{L}$ as desired.

That is, we proved that $\mathcal{L}$ is a σ-algebra on $\mathbb{R}$ containing all open subsets of $\mathbb{R}$. It follows that $\mathcal{B} \subset \mathcal{L}$, since $\mathcal{B}$ is the smallest σ-algebra on $\mathbb{R}$ containing all open subsets of $\mathbb{R}$ by definition.

As an immediate result, $\forall B \in \mathcal{B}$ it holds that $\forall \epsilon > 0$ there exist $F \in \mathcal{P}(\mathbb{R})$ such that $F \subset B$, F is closed, and $|B/F| < \epsilon$. Here, we complete the proof.

□

Theorem 2.4.4 (Borel Set Implies Additivity of Outer Measure)**.** *Suppose $A, B \subset \mathbb{R}$ are disjoint and $B \in \mathcal{B}$. Then,*

$$|A \cup B| = |A| + |B|.$$

Proof. Suppose $A, B \subset \mathbb{R}$ are disjoint and $B \in \mathcal{B}$. We observe, by the subadditivity of outer measure, that $|A \cup B| \leq |A| + |B|$. It remains to show that $|A \cup B| \geq |A| + |B|$.

Since $B \in \mathcal{B}$, there exists a closed set $F \subset \mathbb{R}$ such that $F \subset B$ and $|B/F| < \epsilon$, by Theorem 2.65 (Axler 48). It follows that $A \cup F \subset A \cup B$, which implies that

$$\begin{aligned}
|A \cup B| &\geq |A \cup F| && \text{(By the monotonicity of outer measure)} \\
&= |A| + |F| && \text{(By Theorem 2.63 (Axler 48) since } F \text{ is closed)} \\
&= |A| + [|B| - |B/F|] && \text{(By Theorem (Axler 48) 2.63 since } B = [B/F] \cup F) \\
&> |A| + |B| - \epsilon.
\end{aligned}$$

Suppose, to the contrary, that $|A \cup B| < |A| + |B|$. It must then hold that $0 < |A| + |B| - |A \cup B|$ and, hence,

$$|A \cup B| > |A| + |B| - [|A| + |B| - |A \cup B|] = |A \cup B|,$$

which is a contradiction. Therefore, we conclude that $|A \cup B| \geq |A| + |B|$ and, thus, $|A \cup B| = |A| + |B|$. □

Theorem 2.4.5 (Existsence of Non-Borel Set with Finite Outer Measure)**.** *There exists $B \in \mathcal{P}(\mathbb{R})$ such that $|B| < \infty$ and $B \notin \mathcal{B}$.*

Proof. By Theorem 2.18 (Axler 21), there exist disjoint subsets A, B of $\mathbb{R}$ such that $|A \cup B| \neq |A| + |B|$. We observe that A, B must have finite outer measures, for otherwise $|A \cup B| = \infty$ by the monotonicity of outer measure and $|A| + |B| = \infty = |A \cup B|$, which is a contradiction.

By the contrapositive of Theorem 2.66 (Axler 50), $|A \cup B| \neq |A| + |B|$ then implies that $B \notin \mathcal{B}$.

□

Theorem 2.4.6. *Outer measure is a measure on* $(\mathbb{R}, \mathcal{B})$.

Proof. Consider the measurable space $(\mathbb{R}, \mathcal{B})$. Denote $|| : \mathcal{B} \to [0, \infty]$ the outer measure. To prove the desired claim, it suffices to show that (i) $|\emptyset| = 0$ and (ii) $||$ is additive under the countable union of disjoint sets.

(i) We observe that $|\emptyset| = 0$ by Example 2.3 (Axler 15) since $\emptyset$ is finite.

(ii) Consider arbitrary sequence $(D_k)_{k\in\mathbb{N}}$ of disjoint sets in $\in \mathcal{B}$. We observe that by induction and Theorem 2.66 (Axler 50)

$$\forall n \in \mathbb{N} \ \left|\bigcup_{k=1}^{n} D_k\right| = \sum_{k=1}^{n} |D_k| .$$

In addition, $\forall n \in \mathbb{N}$, $\cup_{k=1}^{n} D_k \subset \cup_{k\in\mathbb{N}} D_k$ implies that $\sum_{k=1}^{n} |D_k| = |\cup_{k=1}^{n} D_k| \leq |\cup_{k\in\mathbb{N}} D_k|$ by the monotonicity of outer measure. As an immediate result,

$$\lim_{n\to\infty} \sum_{k=1}^{n} |D_k| = \sum_{k\in\mathbb{N}} |D_k| \leq |\cup_{k\in\mathbb{N}} D_k| .$$

Note that the above inequality is a result of elementary analysis. Lastly, by the subadditivity of outer measure, we yield that $|\cup_{k\in\mathbb{N}} D_k| \leq \sum_{k\in\mathbb{N}} |D_k|$. Therefore, we conclude that $|\cup_{k\in\mathbb{N}} D_k| = \sum_{k\in\mathbb{N}} |D_k|$. That is, we showed that $||$ is additive under the countable union of disjoint sets in $\mathcal{B}$.

By definition, $||$ is a measure on the measurable space $(\mathbb{R}, \mathcal{B})$ as desired. □

Definition 2.4.1 (Lebesgue measure)**.** *Lebesgue measure is the measure on* $(\mathbb{R}, \mathcal{B})$ *that assigns to each Borel set its outer measure.*

Definition 2.4.2 (Lebesgue measurable set)**.** $A \in \mathcal{P}(\mathbb{R})$ *is Lebesue measurable if* $\exists B \in \mathcal{B}$ *such that* $B \subset A$ *and* $|A/B| = 0$.

Theorem 2.4.7 (Characterization of Lebesgue measurability)**.** *Suppose* $A \in \mathcal{P}(\mathbb{R})$*. Then, the following statements are equivalent:*

(a) A *is Lebesgue measurable.*

(b) $\forall \epsilon > 0 \ \exists F \in \mathcal{P}(\mathbb{R})$ *such that* $F \subset A$ *is closed and* $|A/F| < \epsilon$.

(c) $\exists F_1, F_2, \cdots \in \mathcal{P}(\mathbb{R})$ *such that* $F_k \subset A$ *is closed* $\forall k \in \mathbb{N}$ *and* $|A/ \cup_{k\in\mathbb{N}} F_k| = 0$.

(d) $\exists B \in \mathcal{B}$ *such that* $B \subset A$ *and* $|A/B| = 0$.

(e) $\forall \epsilon > 0 \ \exists G \in \mathcal{P}(\mathbb{R})$ *such that* G *is open,* $G \supset A$*, and* $|G/A| < \epsilon$.

(f) $\exists G_1, G_2, \cdots \in \mathcal{P}(\mathbb{R})$ *such that* $G_k \supset A$ *is open* $\forall k \in \mathbb{N}$ *and* $\left|\left[\bigcap_{k\in\mathbb{N}} G_k\right]/A\right| = 0$

(g) $\exists B \in \mathcal{B}$ *such that* $B \supset A$ *and* $|B/A| = 0$.

Proof. Suppose $A \in \mathcal{P}(\mathbb{R})$. We observe that (a) is equivalent to (d) by definition. Here, we prove the desired equivalences by showing that $(b) \implies (c) \implies (d) \implies (b)$ and $(b) \implies (e) \implies (f) \implies (g) \implies (b)$.

Let $\tilde{\mathcal{L}}$ be the set such that

$$\tilde{\mathcal{L}} = \{A \in \mathcal{P}(\mathbb{R}) : \forall \epsilon > 0 \ \exists F \in \mathcal{P}(\mathbb{R}) \text{ such that } F \subset A,\ F \text{ is closed, and } |A/F| < \epsilon\} .$$

By Theorem 2.65, $\tilde{\mathcal{L}}$ is a σ-algebra on $\mathbb{R}$. Moreover, for any set $E \in \mathcal{P}(\mathbb{R})$ with outer measure zero, we have that $E \in \tilde{\mathcal{L}}$, since for all $\epsilon > 0$ we may take $F = \emptyset \subset E$, which is closed, to satisfy the condition $|E/F| = |E| - |F| = 0 < \epsilon$.

$(b \implies c)$: Suppose (b) holds. It follows that $\forall k \in \mathbb{N}$

$$\exists F_k \in \mathcal{P}(\mathbb{R}) \text{ such that } F_k \subset A,\ F_k \text{ is closed, and } |A/F_k| < \frac{1}{k}.$$

We observe that, $\forall k \in \mathbb{N}$,

$$\begin{aligned} F_k \subset \bigcup_{k\in\mathbb{N}} F_k &\implies A/F_k \supset A/\left[\bigcup_{k\in\mathbb{N}} F_k\right] \\ &\implies \frac{1}{k} > |A/F_k| \geq \left|A/\left[\bigcup_{k\in\mathbb{N}} F_k\right]\right|, \qquad \text{(By the monotoncity of outer measure)} \end{aligned}$$

which then implies that $\left|A/\left[\bigcup_{k\in\mathbb{N}} F_k\right]\right| = 0$. Suppose otherwise that $\left|A/\left[\bigcup_{k\in\mathbb{N}} F_k\right]\right| > 0$. As an immediate result $\frac{1}{\left|A/\left[\bigcup_{k\in\mathbb{N}} F_k\right]\right|} > k\ \forall k \in \mathbb{N}$.

Let $\tilde{k} = \left\lceil \frac{1}{\left|A/\left[\bigcup_{k\in\mathbb{N}} F_k\right]\right|} \right\rceil + 1$ so that $\tilde{k} > \frac{1}{\left|A/\left[\bigcup_{k\in\mathbb{N}} F_k\right]\right|}$ and $\tilde{k} \in \mathbb{N}$. It follows that $\frac{1}{\left|A/\left[\bigcup_{k\in\mathbb{N}} F_k\right]\right|} > \tilde{k}$, which is a contradiction. Here, $(F_k)_{k\in\mathbb{N}}$ satisfies the desired conditions for (c), and we conclude that (c) holds.

$(c \implies d)$: Suppose (c) holds. It follows that $\exists F_1, F_2, \cdots \subset A$ such that $\forall k \in \mathbb{N}$ $F_k \subset A$ is closed, and it holds that $\left|A/\bigcup_{k\in\mathbb{N}} F_k\right| = 0$. We observe that $\bigcup_{k\in\mathbb{N}} F_k$ is a Borel set, since $\mathcal{B}$ is closed under countable union and $\forall k \in \mathbb{N}$ $F_k \in \mathcal{B}$ as it is closed. Thus, $\bigcup_{k\in\mathbb{N}} F_k$ satisfies the condition for (d), and we conclude (d) holds.

$(d \implies b)$: Suppose (d) holds. It follows that there exists a Borel set B such that $B \subset A$ and $|A/B| = 0$. Since $\mathcal{B} \subset \tilde{\mathcal{L}}$, $B \in \tilde{\mathcal{L}}$. Recall from earlier that we showed that $\tilde{\mathcal{L}}$ contains all subsets of $\mathbb{R}$ with outer measure zero. Hence, A/B is in $\tilde{\mathcal{L}}$. We observe that $A = [A/B] \cup B$, which then implies that $A \in \tilde{\mathcal{L}}$, by the closure of $\tilde{\mathcal{L}}$ under finite union. Since $A \in \tilde{\mathcal{L}}$, (b) holds as desired.

Here, we showed that $(b) \iff (c) \iff (d)$.

$(b \implies e)$: Suppose b holds. It follows that $A \in \tilde{\mathcal{L}}$, which implies that $\mathbb{R}/A \in \tilde{\mathcal{L}}$ and therefore

$$\forall \epsilon > 0\ \exists F \in \mathcal{P}(\mathbb{R}) \text{ such that } F \subset \mathbb{R}/A \text{ is closed and } |[\mathbb{R}/A]/F| < \epsilon.$$

Fix $\epsilon > 0$. We observe that $\mathbb{R}/F$ is open and $A \subset \mathbb{R}/F$ since $F \subset \mathbb{R}/A$. Moreover, $(\mathbb{R}/F)/A = (\mathbb{R}/A)/F$ by elementary set theory. Hence,

$$|(\mathbb{R}/A)/F| = |(\mathbb{R}/F)/A| < \epsilon.$$

For all $\epsilon > 0$, let $G = \mathbb{R}/F$. We conclude that $\forall \epsilon > 0$, there exists an open set G such that $A \subset G$ and $|G/A| < \epsilon$. Thus, (e) holds.

$(e \implies f)$. Suppose e holds. Then, $\forall k \in \mathbb{N}$ $\exists G_k \supset A$ such that G_k is open and $|G_k/A| < \frac{1}{k}$. We observe that $\forall k \in \mathbb{N}$

$$\bigcap_{k\in\mathbb{N}} [G_k/A] \subset G_k/A \implies \left|\bigcap_{k\in\mathbb{N}} [G_k/A]\right| \leq |G_k/A| < \frac{1}{k}$$

implying that $\left|\bigcap_{k\in\mathbb{N}} [G_k/A]\right| = 0$. To see that the equality holds, we suppose, to the contrary, that $\left|\bigcap_{k\in\mathbb{N}} [G_k/A]\right| > 0$. It follows that $\forall k \in \mathbb{N}$ $\frac{1}{\left|\bigcap_{k\in\mathbb{N}} [G_k/A]\right|} > k$. Let $K = \left\lceil \frac{1}{\left|\bigcap_{k\in\mathbb{N}} [G_k/A]\right|} \right\rceil + 1$ so that $K > \frac{1}{\left|\bigcap_{k\in\mathbb{N}} [G_k/A]\right|}$. Since $K \in \mathbb{N}$, we observe that we also have that $\frac{1}{\left|\bigcap_{k\in\mathbb{N}} [G_k/A]\right|} > K$, which is a contradiction. Thus, (f) holds.

($f \implies g$): Suppose (f) holds. Then,

$$\exists G_1, G_2, \cdots \supset A \text{ such that } \forall k \in \mathbb{N}\ G_k \text{ is open, and } \left| \left[\bigcap_{k \in \mathbb{N}} G_k \right] / A \right| = 0.$$

Let $B = \bigcap_{k \in \mathbb{N}} G_k$. Here, B is a Borel set since it is a countable intersection of open sets, which are Borel sets, and is contained in $\mathcal{B}$ by the closure of $\mathcal{B}$ under countable intersection. Moreover, $B \supset A$ since $B \supset G_1 \supset A$. Here, (g) holds.

($g \implies b$): Suppose (G) holds. Then, $\exists B \in \mathcal{B}$ such that $B \supset A$ and $|B/A| = 0$. Note that $B/A \in \tilde{\mathcal{L}}$ since it has outer measure zero. Then, we yield that $\mathbb{R}/(B/A) \in \tilde{\mathcal{L}}$. Moreover, it holds that $\mathcal{B} \subset \tilde{\mathcal{L}} \implies B \in \tilde{\mathcal{L}}$. Thus, $A \in \tilde{\mathcal{L}}$ by the closure of $\tilde{\mathcal{L}}$ under finite intersection since $A = B \cap (\mathbb{R}/[B/A])$. ($b$) follows immediately since $A \in \tilde{\mathcal{L}}$.

Here, we showed all the desired equivalences and complete the proof. □

Proposition 2.4.1 (Lebesgue Measurable Sets-Induced σ-algebra)**.** *Let*

$$\mathcal{L} = \{A \in \mathcal{P}(\mathbb{R}) : A \textit{ is Lebesgue measurable}\}, \textit{ and}$$
$$\tilde{\mathcal{L}} = \{E \in \mathcal{P}(\mathbb{R}) : \forall \epsilon > 0\ \exists F \in \mathcal{P}(\mathbb{R}) \textit{ such that } F \subset E,\ F \textit{ is closed, and } |E/F| < \epsilon\}.$$

Then, $\mathcal{L} = \tilde{\mathcal{L}}$, and $\mathcal{L}$ is a σ-algebra on $\mathbb{R}$.

Proof. Let $\mathcal{L}$ be the collection of Lebesgue measurable sets of $\mathbb{R}$. Let $\tilde{\mathcal{L}}$ be the set

$$\tilde{\mathcal{L}} = \{E \in \mathcal{P}(\mathbb{R}) : \forall \epsilon > 0\ \exists F \in \mathcal{P}(\mathbb{R}) \text{ such that } F \subset E,\ F \text{ is closed, and } |E/F| < \epsilon\}.$$

Fix $A \in \mathcal{P}(\mathbb{R})$. By Theorem 2.71, we have that A is Lebesgue measurable $\iff$

$$\forall \epsilon > 0\ \exists F \in \mathcal{P}(\mathbb{R}) \text{ such that } F \subset A,\ F \text{ is closed, and } |A/F| < \epsilon.$$

Thus, $A \in \mathcal{L} \iff A \in \tilde{\mathcal{L}}$. That is, $\mathcal{L} = \tilde{\mathcal{L}}$. By Theorem 2.65, $\tilde{\mathcal{L}}$ is a σ-algebra on $\mathbb{R}$. Thus, $\mathcal{L}$ is a σ-algebra on $\mathbb{R}$ as desired. □

Definition 2.4.3 ($\mathcal{L}$; Lebesgue algebra)**.** *We call the set*

$$\mathcal{L} = \{A \in \mathcal{P}(\mathbb{R}) : A \textit{ is Lebesgue measurable}\}$$

the Lebesgue algebra.

Proposition 2.4.2 (Borel Measurability Implies Lebesgue Measurability)**.** *Suppose $A \in \mathcal{P}(\mathbb{R})$ is Borel measurable. Then, A is Lebsegue measurable.*

Proof. Suppose $A \in \mathcal{P}(\mathbb{R})$ is Borel measurable. Let $B = A$. Then, we have that $B \in \mathcal{B}$, $B \subset A$, and $|A/B| = |\emptyset| = 0$. By definition, A is Lebesgue measurable as desired. □

Proposition 2.4.3 (Zero Outer Measure Implies Lebesgue Measurability)**.** *Suppose $E \in \mathcal{P}(\mathbb{R})$ and $|E| = 0$. Then, E is Lebesgue measurable.*

Proof. Suppose $E \in \mathcal{P}(\mathbb{R})$ and $|E| = 0$. Let $B = \emptyset$. We observe that $B \in \mathcal{B}$ since it is closed, $B \subset E$, and $|E/B| = |E| = 0$. By definition, E is Lebesgue measurable as desired. □

Proposition 2.4.4. *$\mathcal{L}$ is the smallest σ-algebra on $\mathbb{R}$ containing $\mathcal{B}$ and all subsets of $\mathbb{R}$ with outer measure* 0.

Theorem 2.4.8 (Outer Measure is a Measure on Lebesgue Measurable Sets)**.** *Outer measure is a measure on* $(\mathbb{R}, \mathcal{L})$.

Proof. **[INCOMPLETE]**

□

Definition 2.4.4 (Cantor set)**.**

Proposition 2.4.5 (Translation and Dilation of a $\mathcal{B}$-Partition ($\mathcal{L}$-Partition))**.** *Let $\mathcal{S} \in \{\mathcal{B}, \mathcal{L}\}$ and consider the measurable space $(\mathbb{R}, \mathcal{S})$. Suppose that $P = \{A_1, \ldots, A_n\}$ is a $\mathcal{S}$-partition of $\mathbb{R}$.*

(i) If $t \in \mathbb{R}$, then $P - t = \{A_k - t : k \in \{1, \ldots, n\}\}$ is a $\mathcal{S}$-partition of $\mathbb{R}$.

(ii) If $t \in \mathbb{R}/\{0\}$, then $tP = \{tA_k : k \in \{1, \ldots, n\}\}$ is a $\mathcal{S}$-partition of $\mathbb{R}$.

Proof. Let $\mathcal{S} \in \{\mathcal{B}, \mathcal{L}\}$ and consider the measurable space $(\mathbb{R}, \mathcal{S})$. Fix $t \in \mathbb{R}$ and a $\mathcal{S}$-partition $P = \{A_1, \ldots, A_n\}$ of $\mathbb{R}$.

(i) By Exercise 2B.7 (Axler 38) and Exercise 2D.8 (Axler 60), Borel sets and Lebesgue measurable sets are translation invariant. Thus, $A_k - t \in \mathcal{S}$, $\forall k \in \{1, \ldots, n\}$. We also observe that $A_1 - t, \ldots, A_n - t$ remain disjoint, for otherwise we would obtain that $A_1, \ldots, A_n$ are not disjoint, which is a contradiction.

It remains to show that the disjoint union of $A_1 - t, \ldots, A_n - t$ equals $\mathbb{R}$.

First, certainly $\bigsqcup_{k=1}^{n}[A_k - t] \subset \mathbb{R}$, since $A_k \subset \mathbb{R}$ $\forall k \in \{1, \ldots, n\}$. Now, fix $x \in \mathbb{R}$. It follows that $x + t \in \mathbb{R}$ and, hence, $\exists N \in \{1, \ldots, n\}$ such that $x + t \in A_N$ since P is a $\mathcal{S}$-partition of $\mathbb{R}$. As an immediate result, $(x + t) - t = x \in A_N - t \subset \bigsqcup_{k=1}^{n}[A_k - t]$ and, hence, $\mathbb{R} \subset \bigsqcup_{k=1}^{n}[A_k - t]$.

Therefore, we conclude that $\bigsqcup_{k=1}^{n}[A_k - t] = \mathbb{R}$ and $P - t$ is a $\mathcal{S}$-partition of $\mathbb{R}$ by definition.

(ii) Suppose $t \neq 0$. By Exercise 2B.8 (Axler 38) and Exercise 2D.9 (Axler 60), Borel sets and Lebesgue measurable sets are dilation invariant. Thus, $tA_k \in \mathcal{S}$, $\forall k \in \{1, \ldots, n\}$. We also observe that $tA_1, \ldots, tA_n$ remain disjoint, for otherwise we would obtain that $A_1, \ldots, A_n$ are not disjoint, which is a contradiction.

It remains to show that the disjoint union of $tA_1, \ldots, tA_n$ equals $\mathbb{R}$.

First, certainly $\bigsqcup_{k=1}^{n}[tA_k] \subset \mathbb{R}$, since $A_k \subset \mathbb{R}$ $\forall k \in \{1, \ldots, n\}$. Now, fix $x \in \mathbb{R}$. It follows that $\frac{x}{t} \in \mathbb{R}$ and, hence, $\exists N \in \{1, \ldots, n\}$ such that $\frac{x}{t} \in A_N$ since P is a $\mathcal{S}$-partition of $\mathbb{R}$. As an immediate result, $\left[\frac{x}{t}\right] \cdot t = x \in tA_N \subset \bigsqcup_{k=1}^{n}[tA_k]$ and, hence, $\mathbb{R} \subset \bigsqcup_{k=1}^{n}[tA_k]$.

Therefore, we conclude that $\bigsqcup_{k=1}^{n}[tA_k] = \mathbb{R}$ and tP is a $\mathcal{S}$-partition of $\mathbb{R}$ by definition. □

Lemma 2.4.1. *Suppose $(X, \mathcal{S})$ is a measurable space and $(f_n)_{n\in\mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions from X to $[-\infty, \infty]$. Let $k \in \mathbb{N}$. Define $h_k, g_k \colon X \to [-\infty, \infty]$ by*

$$h_k(x) = \sup_{j\in\mathbb{N}, j\geq k} f_j(x) \text{ and } g_k(x) = \inf_{j\in\mathbb{N}, j\geq k} f_j(x), \ \forall x \in X$$

respectively. Then, h_k and g_k are $\mathcal{S}$-measurable.

Proof. Suppose $(X, \mathcal{S})$ is a measurable space and $(f_n)_{n\in\mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions from X to $[-\infty, \infty]$. Let $k \in \mathbb{N}$. Define $h_k, g_k \colon X \to [-\infty, \infty]$ by

$$h_k(x) = \sup_{j\in\mathbb{N}, j\geq k} f_j(x) \text{ and } g_k(x) = \inf_{j\in\mathbb{N}, j\geq k} f_j(x), \ \forall x \in X$$

respectively. Fix $a \in \mathbb{R}$. We first claim that

$$h_k{}^{-1}((a, \infty]) = \bigcup_{j=k}^{\infty} f_j{}^{-1}((a, \infty]). \tag{$*_0$}$$

Fix $p \in {h_k}^{-1}((a, \infty])$. It follows that $\sup_{j \in \mathbb{N}, j \geq k} f_j(p) \in (a, \infty]$ by definition and, thus,

$$\exists N \in \{k, k+1, \ldots\} \text{ such that } \sup_{j \in \mathbb{N}, j \geq k} f_j(p) - \left(\sup_{j \in \mathbb{N}, j \geq k} f_j(p) - a \right) < f_N(p)$$

by an equivalent condition of supremum and since $\left(\sup_{j \in \mathbb{N}, j \geq k} f_j(p) - a\right) > 0$. Hence, we yield that $f_N(p) \in (a, \infty]$, which then implies that

$$p \in f_N^{-1}((a, \infty]) \subset \bigcup_{j=k}^{\infty} {f_j}^{-1}((a, \infty]). \qquad \text{(By definition)}$$

$$\therefore {h_k}^{-1}((a, \infty]) \subset \bigcup_{j=k}^{\infty} {f_j}^{-1}((a, \infty]).$$

Now, fix $q \in \bigcup_{j=k}^{\infty} {f_j}^{-1}((a, \infty])$. Then, there exists $N \in \{k, k+1, \ldots\}$ such that $f_N(q) > a$. Certainly, we have

$$\begin{aligned} h_k(q) = \sup_{j \in \mathbb{N}, j \geq k} f_j(q) \geq f_N(q) > a &\implies q \in {h_k}^{-1}((a, \infty]) \\ &\implies \bigcup_{j=k}^{\infty} {f_j}^{-1}((a, \infty]) \subset {h_k}^{-1}((a, \infty]) \end{aligned}$$

and conclude $(*_0)$ holds. We note that, $\forall j \in \{k, k+1, \ldots\}$, ${f_j}^{-1}((a, \infty]) \in \mathcal{S}$ since f_j is $\mathcal{S}$-measurable by hypothesis. It follows that ${h_k}^{-1}((a, \infty]) \in \mathcal{S}$ as well, by $(*_0)$ and the closure of $\mathcal{S}$ under countable union. By Theorem 2.52 (Axler 37), h_k is $\mathcal{S}$-measurable.

That is, we showed that for any arbitrary sequence $(F_k)_{k \in \mathbb{N}}$ of $\mathcal{S}$-measurable functions defined from X to $[-\infty, \infty]$ satisfies the property that the function defined by $\sup_{j \in \mathbb{N}, j \geq k} F_j(x)\ \forall x \in X$ is $\mathcal{S}$-measureable.

By an identity of infimum, we have that

$$g_k(x) = \inf_{j \in \mathbb{N}, j \geq k} f_j(x) = - \sup_{j \in \mathbb{N}, j \geq k} \left[-f_j(x)\right], \forall x \in X.$$

By the regularity of $\mathcal{S}$-measurable functions, $(-f_k)_{j \in \mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions defined from X to $[-\infty, \infty]$. By the preceding result, $-g_k \colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable. It follows, from the regularity of $\mathcal{S}$-measurable functions, that $g_k = -(-g_k)$ is $\mathcal{S}$-measurable as desired. ☐

Theorem 2.4.9 (Extension for Theorem 2.48 (Axler 36))**.** *Suppose $(X, \mathcal{S})$ is a measurable space and $(f_k)_{k \in \mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions defined from X to $[-\infty, \infty]$. Suppose further that, $\forall x \in X$, $\lim_{k \to \infty} f_k(x)$ exists in $[-\infty, \infty]$. Then, $f \colon X \to [-\infty, \infty]$, defined by $f = \lim_{k \to \infty} f_k$, is $\mathcal{S}$-measurable.*

Proof. Suppose $(X, \mathcal{S})$ is a measurable space and $(f_k)_{k \in \mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions defined from X to $[-\infty, \infty]$. Suppose further that, $\forall x \in X$, $\lim_{k \to \infty} f_k(x)$ exists in $[-\infty, \infty]$. Define $f \colon X \to [-\infty, \infty]$ by $f = \lim_{k \to \infty} f_k$.

Since, $\forall x \in X$, the limit $\lim_{k \to \infty} f_k(x)$ exists in $[-\infty, \infty]$ by hypothesis, we have that $\lim_{k \to \infty} f_k(x) = \limsup_{k \to \infty} f_k(x)$ by an extended result of Theorem 10.7 (Ross 61). Therefore, $\lim_{k \to \infty} f_k = \limsup_{k \to \infty} f_k$. We note, by an equivalent definition of limit superior, that

$$\limsup_{k \to \infty} f_k(x) = \inf_{k \in \mathbb{N}} \left[\sup_{j \in \mathbb{N}, j \geq k} f_j(x) \right], \forall x \in X. \qquad (*)$$

Fix $k \in \mathbb{N}$. Define $h_k \colon X \to [-\infty, \infty]$ by $h_k(x) = \sup_{j \in \mathbb{N}, j \geq k} f_j(x)$ $\forall x \in X$. By a lemma above, h_k is $\mathcal{S}$-measurable. Thus, we have a sequence $(h_k)_{k \in \mathbb{N}}$ of $\mathcal{S}$-measurable functions defined from X to $[-\infty, \infty]$. Define $g \colon X \to [-\infty, \infty]$ by $g(x) = \inf_{k \in \mathbb{N}} h_k(x)$ $\forall x \in X$. By Theorem 2.53 (Axler 37), g is $\mathcal{S}$-measurable. Note that, by $(*)$, $f = g$. Thus, f is $\mathcal{S}$-measurable as desired. $\square$

Definition 2.4.5 (Limit superior; limit inferior)**.** *Suppose* $(x_k)_{k \in \mathbb{N}}$ *is a sequence in* $[-\infty, \infty]$*. Then,*

(i) the limit superior of $(x_k)_{k \in \mathbb{N}}$ *is defined by*

$$\limsup_{k \to \infty} x_k = \lim_{k \to \infty} \sup \{x_k, x_{k+1}, \ldots\},$$

(ii) the limit inferior of $(x_k)_{k \in \mathbb{N}}$ *is defined by*

$$\liminf_{k \to \infty} x_k = \lim_{k \to \infty} \inf \{x_k, x_{k+1}, \ldots\}.$$

Theorem 2.4.10. *Suppose that* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f_1, f_2, \ldots$ *is a sequence of* $\mathcal{S}$*-measurable functions on* X*. Define* $f \colon X \to [0, \infty]$ *by*

$$f(x) = \liminf_{k \to \infty} f_k(x) \ \forall x \in X.$$

Then, f *is* $\mathcal{S}$*-measurable.*

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f_1, f_2, \ldots$ is a sequence of $\mathcal{S}$-measurable functions on X. Define $f \colon X \to [0, \infty]$ by $f = \liminf_{k \to \infty} f_k$.

Define, $\forall k \in \mathbb{N}$, $g_k \colon X \to [-\infty, \infty]$ by $g_k = \inf_{j \in \mathbb{N}, j \geq k} f_k$. Thus, we have that $\lim_{k \to \infty} g_k = f$. Note that by a previous lemma, $\forall k \in \mathbb{N}$, g_k is $\mathcal{S}$-measurable. It follows, from another previous lemma, that $\lim_{k \to \infty} g_k = f$ is $\mathcal{S}$-measurable as desired. $\square$

Theorem 2.4.11 (Fatou's Lemma)**.** *Suppose that* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f_1, f_2, \ldots$ *is a sequence of nonnegative* $\mathcal{S}$*-measurable functions on* X*. Define* $f \colon X \to [0, \infty]$ *by* $f = \liminf_{k \to \infty} f_k$*. Then,*

$$\int f d\mu \leq \liminf_{k \to \infty} \int f_k d\mu.$$

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f_1, f_2, \ldots$ is a sequence of nonnegative $\mathcal{S}$-measurable functions on X. Define $f \colon X \to [0, \infty]$ by $f = \liminf_{k \to \infty} f_k$.

We first note that by the preceding lemma, f is $\mathcal{S}$-measurable. Define, $\forall k \in \mathbb{N}$, $g_k \colon X \to [-\infty, \infty]$ by $g_k = \inf_{j \in \mathbb{N}, j \geq k} f_j$. By a previous lemma, g_k is $\mathcal{S}$-measurable, $\forall k \in \mathbb{N}$. Hence, the expressions in the claim are well-defined.

To prove the desired claim, we first show that

$$\int f d\mu = \lim_{k \to \infty} \int \inf_{j \in \mathbb{N}, j \geq k} f_j d\mu. \tag{$*$}$$

Fix $l, h \in \mathbb{N}$ such that $l < h$. It follows that, $\forall x \in X$,

$$\begin{aligned} \{f_h(x), f_{h+1}(x), \ldots\} &\subset \{f_l(x), f_{l+1}(x), \ldots, f_h(x), f_{h+1}(x), \ldots\} \\ &\implies \inf_{j \in \mathbb{N}, j \geq h} f_j(x) \geq \inf_{j \in \mathbb{N}, j \geq l} f_j(x) \qquad \text{(By a property of infimum)} \\ &\implies g_h(x) \geq g_l(x). \qquad \text{(By the definition of } g_k\text{)} \end{aligned}$$

In addition, $g_1(x) \geq 0 \ \forall x \in X$, since $(f_k)_{k\in\mathbb{N}}$ is nonnegative by hypothesis. Hence, we have that $(g_k)_{k\in\mathbb{N}}$ is an increasing sequence of nonnegative $\mathcal{S}$-measurable functions. By the Monotone Convergence Theorem (Axler 78), we yield that

$$\begin{aligned} \lim_{k\to\infty} \int g_k d\mu &= \int \lim_{k\to\infty} g_k d\mu \\ \implies \lim_{k\to\infty} \int \inf_{j\in\mathbb{N}, j\geq k} f_j d\mu &= \int \lim_{k\to\infty} \inf_{j\in\mathbb{N}, j\geq k} f_j d\mu \\ &= \int \liminf_{k\to\infty} f_k d\mu = \int f d\mu. \qquad \text{(By definition and hypothesis)} \end{aligned}$$

Fix $k \in \mathbb{N}$. We observe that, $\forall n \in \mathbb{N}$ such that $n \geq k$

$$\int \inf_{j\in\mathbb{N}, j\geq k} f_j d\mu \leq \int f_n d\mu, \qquad \text{(By the monotonicity of integration for nonnegative functions)}$$

which implies that

$$\inf_{j\in\mathbb{N}, j\geq k} \int f_j d\mu \geq \int \inf_{j\in\mathbb{N}, j\geq k} f_j d\mu, \tag{$*_1$}$$

by the definition of infimum. As an immediate result, it holds that

$$\begin{aligned} \liminf_{k\to\infty} \int f_k d\mu &= \lim_{k\to\infty} \inf_{j\in\mathbb{N}, j\geq k} \int f_j d\mu && \text{(By definition)} \\ &\geq \lim_{k\to\infty} \int \inf_{j\in\mathbb{N}, j\geq k} f_j d\mu && \text{(By } (*_1)\text{)} \\ &= \int f d\mu. && \text{(By } *_0\text{)} \end{aligned}$$

$\square$

2.5 Convergence of Measurable Functions

Definition 2.5.1 (Pointwise convergence; uniform convergence)**.** *Suppose X is a set, $U \subset X$, $f_k \colon X \to \mathbb{R}$ $\forall k \in \mathbb{N}$, and $f \colon X \to \mathbb{R}$.*

(i) $(f_k)_{k\in\mathbb{N}}$ converges pointwise to f on U if

$$\forall x \in U \ \lim_{k\to\infty} f_k(x) = f(x);$$

(ii) $(f_k)_{k\in\mathbb{N}}$ converges uniformly to f on U if

$$\forall \epsilon > 0 \ \exists N \in \mathbb{N} \text{ such that } k \geq N \text{ and } x \in U \implies |f_k(x) - f(x)| < \epsilon.$$

If $(f_k)_{k\in\mathbb{N}}$ converges to f pointwise on U, then we write $f_k \to f$ pointwise on U. Similarly, if $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly on U, then we write $f_k \to f$ uniformly on U.

Theorem 2.5.1 (Uniform Convergence Preserves Continuity)**.** *Suppose $B \subset \mathbb{R}$, $f_k \colon B \to \mathbb{R}$ $\forall k \in \mathbb{N}$, and $f_k \to f$ on B, where $f \colon B \to \mathbb{R}$. Suppose $\forall k \in \mathbb{N}$ f_k is continuous at $b \in B$. Then, f is continuous at b.*

Theorem 2.5.2 (Egorov's Theorem; Pointwise Convergence of Measurable Functions Implies Almost Uniform Convergence)**.** *Let $(X, \mathcal{S}, \mu)$ be a measurable space with $\mu(X) < \infty$ and $f \colon X \to \mathbb{R}$. Suppose that, $\forall k \in \mathbb{N}$, $f_k \colon X \to \mathbb{R}$ is $\mathcal{S}$-measurable, and $f_k \to f$ pointwise on X. Then, $\forall \epsilon > 0$ $\exists E \in \mathcal{S}$ such that*

(i) $\mu(X/E) < \epsilon$, and

(ii) $f_k \to f$ uniformly on E.

Definition 2.5.2 (Simple function)**.** *Suppose X, Y are a set and $f \colon X \to Y$. Then, f is a simple function if $f(X)$ is finite.*

Theorem 2.5.3 (Approximation of $\mathcal{S}$-Measurable Function via Simple Functions)**.** *Suppose that $(X, \mathcal{S})$ is a measurable space and $f \colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable. Then, there exists a sequence $(f_k)_{k\in\mathbb{N}}$ of functions such that*

(i) $\forall k \in \mathbb{N}$ $f_k \colon X \to \mathbb{R}$ is simple and $\mathcal{S}$-measurable,

(ii) $k \in \mathbb{N}$ and $x \in X$ implies that $|f_k(x)| \leq |f_{k+1}(x)| \leq |f(x)|$,

(iii) $f_k \to f$ pointwise on X, and

(iv) if f is bounded, then $f_k \to f$ uniformly on X.

Proof. Suppose that $(X, \mathcal{S})$ is a measurable space and $f \colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable.

Define, $\forall k \in \mathbb{N}$, $f_k \colon X \to \mathbb{R}$ by

$$f_k(x) = \begin{cases} k & \text{if } f(x) \in [k, \infty) \\ -k & \text{if } f(x) \in (-\infty, -k] \\ \frac{m}{2^k} & \text{if } f(x) \in [0, k) \text{ and } m \in \mathbb{Z} \text{ is such that} f(x) \in \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right) \\ \frac{m+1}{2^k} & \text{if } f(x) \in (-k, 0) \text{ and } m \in \mathbb{Z} \text{ is such that} f(x) \in \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right) \end{cases}, \forall x \in X.$$

(i) Fix $k \in \mathbb{N}$ and Consider arbitrary $x \in X$ such that $f(x) \in [0, k)$. Hence, we then have that $0 \leq 2^k f(x) < 2^k k$. Note that $\exists! m \in \mathbb{Z}$ such that $m \leq 2^k f(x) < m + 1$, since every real number is bounded by an integer and its immediate successor.

We claim that $0 \leq m < 2^k k$. Suppose otherwise. Then, we yield $0 > m$ or $m \geq 2^k k$. Suppose $0 > m$. Then, it must hold that $m \leq -1$. It follows that $2^k f(x) < m + 1 \leq 0$, which is a contradiction since we have that $0 \leq 2^k f(x)$ by assumption. Now, suppose that $m \geq 2^k k$. Then, we obtain that $2^k f(x) > m \geq 2^k k$, which again is a contradiction as $2^k f(x) < 2^k k$ by assumption.

Hence, we conclude that there are at most $2^k k + 1$ $m \in \mathbb{Z}$ such that $f(x) \in \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)$. It follows that $card\left(f_k([0,k))\right) \leq 2^k k + 1$. Similarly, $card\left(f_k((-k,0))\right) \leq 2^k k + 1$. As an immediate result,

$$\begin{aligned} card(f_k(X)) &\leq card(f_k([0,k))) + card(f_k((-k,0))) + card(f_k([k,\infty))) + card(f_k((-\infty,-k]) \\ &\leq (1 + 2^k k) + (1 + 2^k k) + 1 + 1 = 4 + 2^{k+1} k. \end{aligned}$$

By definition, f_k is a simple function. Let

$$U = \left\{ \left\{ \frac{m}{2^k} \right\} : m \in \mathbb{Z} \text{ and } \exists x \in X \text{ such that } f(x) \in [0,k) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right) \right\},$$
$$V = \left\{ \left\{ \frac{m+1}{2^k} \right\} : m \in \mathbb{Z} \text{ and } \exists x \in X \text{ such that } f(x) \in (-k,0) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right) \right\}.$$

We observe that, by definition and the $\mathcal{S}$-measurablility of f,

$$f_k{}^{-1}(\{k\}) = \{x \in X : f_k(x) = k\} = \{x \in X : f(x) \in [k,\infty)\} = f^{-1}([k,\infty)) \in \mathcal{S},$$
$$f_k{}^{-1}(\{-k\}) = \{x \in X : f_k(x) = -k\} = \{x \in X : f(x) \in (-\infty,-k]\} = f^{-1}((-\infty,-k]) \in \mathcal{S}.$$

Moreover, $\left\{\frac{m}{2^k}\right\} \in U$ and $\left\{\frac{m+1}{2^k}\right\} \in V$ imply that

$$f_k{}^{-1}\left(\left\{\frac{m}{2^k}\right\}\right) = \left\{x \in X : f(x) \in [0,k) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right\} = f^{-1}\left(\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right) \in \mathcal{S},$$
$$f_k{}^{-1}\left(\left\{\frac{m+1}{2^k}\right\}\right) = \left\{x \in X : f(x) \in (-k,0) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right\} = f^{-1}\left(\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right) \in \mathcal{S},$$

since $[0,k) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right), (-k,0) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right) = \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)$. **[DO: REWORD TO AVOID ABUSE OF NOTATION AND ELABORATE]**

That is, we showed that $\forall y \in f_k(X)$ $f^{-1}(\{y\}) \in \mathcal{S}$. Consider arbitrary $B \in \mathcal{B}$. If $B \cap f_k(X) = \emptyset$. Then, $f_k{}^{-1}(B) = \{x \in X : f_k(x) \in B\} = \emptyset$. Now, suppose that $B \cap f_k(X) \neq \emptyset$. Let $\tilde{B} = B/[B \cap f_k(X)]$ so $B = \tilde{B} \sqcup [B \cap f_k(X)]$ and $\tilde{B} \cap f_k(X) = \emptyset$. It follows that

$$\begin{aligned} f_k{}^{-1}(B) &= f_k{}^{-1}(\tilde{B} \sqcup [B \cap f_k(X)]) \\ &= f_k{}^{-1}(\tilde{B}) \cup f_k{}^{-1}([B \cap f_k(X)]) \\ &= \emptyset \cup f_k{}^{-1}\left(\bigcup_{i=1}^{l} \{y_i\}\right) && \text{(For some } l \in \{1,\ldots,card(f_k(X))\}) \\ &= \bigcup_{i=1}^{l} f_k{}^{-1}(\{y_i\}) \in \mathcal{S} && \text{(By the closure of } \mathcal{S} \text{ under countable union,)} \end{aligned}$$

where $y_i \in f_k(X)$ $\forall i \in \{1,\ldots,l\}$ since $B \cap f_k(X) \subset f_k(X)$ and $f_K(X)$ is finite. By definition, f_k is $\mathcal{S}$-measurable.

(ii) This statement follows trivially from the definition of $(f_k)_{k\in\mathbb{N}}$.

(iii) Fix $k \in \mathbb{N}$. We observe that $f(x) = k \implies f_k(k) = k$, $f(x) = -k \implies f_k(-k) = -k$, and

$$\begin{aligned} f(x) \in [0,k) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right), \text{ where } m \in \mathbb{Z} &\implies f_k(x) = \frac{m}{2^k} \\ &\implies |f_k(x) - f(x)| < \frac{m+1}{2^k} - \frac{m}{2^k} = \frac{1}{2^k}, \\ f(x) \in (-k,0) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right), \text{ where } m \in \mathbb{Z} &\implies f_k(x) = \frac{m+1}{2^k} \\ &\implies |f_k(x) - f(x)| < \frac{m+1}{2^k} - \frac{m}{2^k} = \frac{1}{2^k}. \end{aligned}$$

That is, we obtain that

$$\forall x \in X \ f(x) \in [-k,k] \implies |f_k(x) - f(x)| < \frac{1}{2^k}.$$

Fix $x \in X$ and then fix $\epsilon > 0$. Let $N' = \left\lceil \log\left(\frac{1}{\epsilon}\right) / \log(2) \right\rceil + 1$ so that $\frac{1}{2^{N'}} < \epsilon$. Moreover, let $\tilde{N} = \lceil |f(x)| \rceil$. Take $N = \max\left\{N', \tilde{N}\right\}$. It follows that

$$k > N \text{ implies that } f(x) \in [-k,k] \implies |f_k(x) - f(x)| < \frac{1}{2^k} < \frac{1}{2^N} < \frac{1}{2^{N'}} < \epsilon.$$

Thus, $f_k \to f$ pointwise on X as desired.

(iv) Suppose that f is bounded. By definition, $\exists M > 0$ such that $|f(x)| \leq M \ \forall x \in X$. Let $\tilde{N} = \lceil M \rceil + 1$. Fix $\epsilon > 0$ and let $N' = \left\lceil \log\left(\frac{1}{\epsilon}\right) / \log(2) \right\rceil + 1$ so that $\frac{1}{2^{N'}} < \epsilon$. Choose $N = \max\left\{\tilde{N}, N'\right\}$ so $\forall x \in X \ x \in [-N, N]$. It follows that

$$x \in X \text{ and } k > N \implies f(x) \in [-k,k] \implies |f_k(x) - f(x)| < \frac{1}{2^k} < \frac{1}{2^N} < \frac{1}{2^{N'}} < \epsilon.$$

That is, we showed that

$$\forall \epsilon > 0 \ \exists N > 0 \text{ such that } x \in X \text{ and } k > N \implies |f_k(x) - f(x)| < \epsilon.$$

By definition, $f_k \to f$ uniformly on X as desired. Here, we complete the proof. □

Theorem 2.5.4 (Luzin's Theorem; Continuity of some Restriction of a Borel Measurable Function)**.** *Suppose $g\colon \mathbb{R} \to \mathbb{R}$ is Borel-measurable. Then, $\forall \epsilon > 0$ there exists a closed set $F \subset \mathbb{R}$ such that*

(i) $|\mathbb{R}/F| < \epsilon$, and

(ii) $g|_F$ is continuous on F.

Proof. **[INCOMPLETE]**

□

Theorem 2.5.5 (Continuous Extension of Continuous Function)**.** *Suppose $F \subset \mathbb{R}$ is closed and $g\colon F \to \mathbb{R}$ is continuous. Then, there exists a map $h\colon \mathbb{R} \to \mathbb{R}$ such that h is continuous and $h|_F = g$.*

Proof. **[INCOMPLETE]**

□

Theorem 2.5.6 (Second Version of Luzin's Theorem)**.** *Suppose $E \subset \mathbb{R}$ and $g\colon E \to \mathbb{R}$ is Borel-measurable. Then, $\forall \epsilon > 0$ there exists a closed set $F \subset E$ and a continuous map $h\colon \mathbb{R} \to \mathbb{R}$ such that*

(i) $|E/F| < \epsilon$, and

(ii) $h|_F = g|_F$.

Proof. **[INCOMPLETE]**

□

Theorem 2.5.7 (Lebesgue Measurability Implies Almost Borel Measurability)**.** *Suppose* $f\colon \mathbb{R} \to \mathbb{R}$ *is Lebesgue measurable. Then, there exists a Borel measurable function* $g\colon \mathbb{R} \to \mathbb{R}$ *such that*

$$|\{x \in \mathbb{R} : g(x) \neq f(x)\}| = 0.$$

Proof. **[INCOMPLETE]**

$\square$

Proposition 2.5.1 (A Lebesgue Measurable Set Differ from a Borel Set by a Set of Measure Zero)**.** *Suppose* $A \in \mathcal{P}(\mathbb{R})$*. Then,* A *is Lebesgue measurable if and only if* $\exists B \in \mathcal{B}$ *and* $C \in \mathcal{P}(\mathbb{R})$ *such that* $|C| = 0$ *and* $A = B \cup C$*.*

3 Integration

3.1 Integration with Respect to a Measure

Definition 3.1.1 ($\mathcal{S}$-partition). *Suppose $(X, \mathcal{S})$ is a measurable space. An $\mathcal{S}$-partition of X is a finite collection P of disjoint sets in $\mathcal{S}$ such that the union of the sets in P equals X.*

That is, $P = \{A_1, \ldots, A_m\}$ is a $\mathcal{S}$-partition of X if

(i) $A_k \in \mathcal{S}$ $\forall k \in \{1, \ldots, m\}$,

(ii) $i, j \in \{1, \ldots, m\}$ and $i \neq j$ implies that $A_i \cap A_j = \emptyset$,

(iii) $X = \bigsqcup_{k=1}^{m} A_k$.

Definition 3.1.2 (Lower Lebesgue sum of a non-negative function). *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f\colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $P = \{A_1, \ldots, A_m\}$ is a $\mathcal{S}$-partition of X. Then, the lower Lebesgue sum of f with respect to μ is*

$$\mathcal{L}(f, P) = \sum_{j=1}^{m} \left[\mu(A_j) \inf_{x \in A_j} f(x)\right].$$

Definition 3.1.3 (Integral of a nonnegative function). *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, and $f\colon X \to [0, \infty]$ is $\mathcal{S}$-measurable. Then, the integral of f with respect to μ is*

$$\int f d\mu = \sup \left\{\mathcal{L}(f, P) : P \text{ is a } \mathcal{S}\text{-partition of } X\right\}.$$

Theorem 3.1.1 (Integral of Characteristic Function). *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, and $E \in \mathcal{S}$. Then,*

$$\int \chi_E d\mu = \mu(E).$$

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, and $E \in \mathcal{S}$. It suffices to show that $\int \chi_E d\mu \leq \mu(E)$ and $\int \chi_E d\mu \geq \mu(E)$.

($\geq$) Consider the $\mathcal{S}$-partition of X $P = \{E, X/E\}$. Then,

$$\mathcal{L}(\chi_E, P) = \mu(E) \inf_{x \in E} \chi_E(x) + \mu(X/E) \inf_{x \in X/E} \chi_E(x) = \mu(E) \cdot 1 + \mu(X/E) \cdot 0 = \mu(E).$$

By definition, $\int \chi_E d\mu \geq \mathcal{L}(\chi_E, P) = \mu(E)$.

($\leq$) Consider arbitrary $\mathcal{S}$-partition $P = \{A_1, \ldots, A_m\}$ of X. Note that $\forall j \in \{1, \ldots, m\}$ $A_j \subset E \implies \inf_{x \in A_j} \chi_E(x) = 1$ and $A_j \not\subset E \implies \inf_{x \in A_j} \chi_E(x) = 0$. Then,

$$\mathcal{L}(\chi_E, P) = \sum_{j=1}^{m} \mu(A_j) \cdot \inf_{x \in A_j} \chi_E(x) \qquad \text{(By definition)}$$

$$= \sum_{j \in \{1, \ldots, m\}, A_j \subset E} \mu(A_j) \cdot \inf_{x \in A_j} \chi_E(x) + \sum_{j \in \{1, \ldots, m\}, A_j \not\subset E} \mu(A_j) \cdot \inf_{x \in A_j} \chi_E(x)$$

$$= \sum_{j \in \{1, \ldots, m\}, A_j \subset E} \mu(A_j) \cdot (1) + \sum_{j \in \{1, \ldots, m\}, A_j \not\subset E} \mu(A_j) \cdot (0)$$

$$= \sum_{j \in \{1, \ldots, m\}, A_j \subset E} \mu(A_j) = \mu\left(\bigcup_{j \in \{1, \ldots, m\}, A_j \subset E} A_j\right)$$

since $A_1, \ldots, A_m$ are disjoint. Note that $\bigcup_{j\in\{1,\ldots,m\}, A_j \subset E} A_j \subset E$ implies that the measure of the union is at most $\mu(E)$ by the monotonicity of μ. That is, we showed that $\mu(E)$ is an upper bound for the set

$$\{\mathcal{L}(\chi_E, P) : P \text{ is a } \mathcal{S}\text{-parition of } X\}.$$

By definition, $\int \chi_E d\mu \leq \mu(E)$ since $\int \chi_E d\mu$ is the least upper bound of the set above.

Here, we complete the proof. $\square$

Theorem 3.1.2 (Integral of Simple Function)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, $E_1, \ldots, E_n \in \mathcal{S}$ are disjoint, and $c_1, \ldots, c_n \in [0, \infty]$. Then,*

$$\int \left[\sum_{k=1}^{n} c_k \chi_{E_k}\right] d\mu = \sum_{k=1}^{n} c_k \mu(E_k).$$

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, $E_1, \ldots, E_n \in \mathcal{S}$ are disjoint, and $c_1, \ldots, c_n \in [0, \infty]$. Let $f = \sum_{k=1}^{n} c_k \chi_{E_k}$ where, $f\colon X \to [0, \infty]$. Moreover, let $E_{n+1} = X / [\cup_{k=1}^{n} E_k]$ and $c_{k+1} = 0$. It follows that $P' = \{E_1, \ldots, E_{n+1}\}$ is a $\mathcal{S}$-partition of X and $f = \sum_{k=1}^{n+1} c_k \chi_{E_k}$.

We observe that $\forall x \in X\ \exists k \in \{1, \ldots, n+1\}$ such that $x \in E_k$, since P' is a $\mathcal{S}$-partition of x. Fix $k \in \{1, \ldots, n+1\}$. We observe that

$$\begin{aligned} x \in E_k \implies \chi_{E_k}(x) = 1 \text{ and } j \in \{1, \ldots, n+1\} / \{k\} &\implies \chi_{E_j}(x) = 0 \\ \implies f(x) = c_1(0) + c_2(0) + \cdots + c_{k-1}(0) + c_k(1) &+ c_{k+1}(0) + \cdots + c_{n+1}(0). \\ \therefore x \in E_k &\implies f(x) = c_k. \qquad (*_0) \end{aligned}$$

To prove the desired statement, it suffices to show that

$$\int f d\mu \geq \sum_{k=1}^{n} c_k \mu(E_k) \text{ and } \int f d\mu \leq \sum_{k=1}^{n} c_k \mu(E_k).$$

($\geq$) By definition,

$$\int f d\mu = \sup \{\mathcal{L}(f, P) : P \text{ is a } \mathcal{S}\text{-partition of X}\}.$$

By $(*_0)$, we obtain that $\inf_{x \in E_k} f(x) = c_k$. As an immediate result, we obtain that

$$\begin{aligned} \int f d\mu \geq \mathcal{L}(f, P') &= \sum_{k=1}^{n+1} \mu(E_k) \inf_{x\in E_k} f(x) \\ &= \sum_{k=1}^{n+1} \mu(E_k) c_k = \sum_{k=1}^{n} \mu(E_k) c_k. \qquad \text{(Since } c_{k+1} = 0) \end{aligned}$$

($\leq$) Consider arbitrary $\mathcal{S}$-partition $P = \{A_1, \ldots, A_m\}$ of X. Fix $j \in \{1, \ldots, m\}$. We claim that

$$\{f(x) : x \in A_j\} = \{c_i : i \in \{1, \ldots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\}. \qquad (*_1)$$

Consider arbitrary $p \in \{c_i : i \in \{1, \ldots, n+1\}$ and $A_j \cap E_i \neq \emptyset\}$. It follows that

$$\begin{aligned} \exists i \in \{1, \ldots, n+1\} \text{ such that } p = c_i \text{ and } A_j \cap E_i \neq \emptyset& \\ \implies \exists x \in A_j \text{ such that } x \in E_i& \\ \implies f(x) = c_i.& \qquad \text{(By } (*_0)) \\ \implies p = f(x) \in \{f(x) : x \in A_j\}& \\ \therefore \{c_i : i \in \{1, \ldots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\} \subset \{f(x) : x \in A_j\}.& \end{aligned}$$

Now, consider arbitrary $q \in \{f(x) : x \in A_j\}$. It follows that $\exists x \in A_j$ such that $q = f(x)$. Since P' is a $\mathcal{S}$-partition of X, $\exists i \in \{1, \dots, n+1\}$ such that $x \in E_i$. Hence, we have that $q = f(x) = c_i$ by $(*_0)$ and $A_j \cap E_i \neq \emptyset$. Thus, $q \in \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\}$ and, hence,

$$\{f(x) : x \in A_j\} \subset \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\}.$$

Therefore, we obtained the equality as desired. It then follows that

$$\begin{aligned}\inf_{x \in A_j} f(x) = \inf\{f(x) : x \in A_j\} &= \inf\{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\} && \text{(By } (*_1)) \\ &= \min\{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\}, && (*_2)\end{aligned}$$

since the set is finite. Moreover, since P' and P are $\mathcal{S}$-partitions of X, we have that

$$\forall j \in \{1, \dots, m\},\ A_j = \bigsqcup_{i=1}^{n+1} (A_j \cap E_i), \text{ and } \forall k \in \{1, \dots, n+1\},\ E_k = \bigsqcup_{l=1}^{m} (A_l \cap E_k). \qquad (*_3)$$

It follows that

$$\begin{aligned}\mathcal{L}(f, P) &= \sum_{j=1}^{m} \mu(A_j) \inf_{x \in A_j} f(x) && \text{(By definition)} \\ &= \sum_{j=1}^{m} \left[\sum_{k=1}^{n+1} \mu(A_j \cap E_k)\right] \inf_{x \in A_j} f(x) = \sum_{j=1}^{m} \left[\sum_{k=1}^{n+1} \left[\mu(A_j \cap E_k) \inf_{x \in A_j} f(x)\right]\right] && \text{(By } (*_3)) \\ &= \sum_{j=1}^{m} \left[\sum_{k=1}^{n+1} \left[\mu(A_j \cap E_k) \min_{\{i \in \{1, \dots, n+1\} : A_j \cap E_i \neq \emptyset\}} c_i\right]\right] && \text{(By } (*_2)) \\ &\leq \sum_{j=1}^{m} \left[\sum_{k=1}^{n+1} \mu(A_j \cap E_k) c_k\right].\end{aligned}$$

To see why the last inequality holds, we first fix $k \in \{1, \dots, n+1\}$. We note that $\forall j \in \{1, \dots, m\}$

$$\begin{aligned}A_j \cap E_k = \emptyset &\implies \mu(A_j \cap E_k) = 0 \\ \implies \mu(A_j \cap E_k) &\min_{\{i \in \{1, \dots, n+1\} : A_j \cap E_i \neq \emptyset\}} c_i = 0 = \mu(A_j \cap E_k) c_k,\end{aligned}$$

where

$$\begin{aligned}A_j \cap E_k \neq \emptyset &\implies \exists x \in A_j \text{ such that } x \in E_k \implies f(x) = c_k && \text{(By } (*_0)) \\ &\implies c_k \in \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\} \\ &\implies \min_{\{i \in \{1, \dots, n+1\} : A_j \cap E_i \neq \emptyset\}} c_i \leq c_k,\end{aligned}$$

which then implies the desired inequality. In addition, we may write

$$\begin{aligned}\sum_{j=1}^{m} \left[\sum_{k=1}^{n+1} \mu(A_j \cap E_k) c_k\right] &= \sum_{k=1}^{n+1} \left[c_k \sum_{j=1}^{m} \mu(A_j \cap E_k)\right] && \text{(By a property of a nested finite sum)} \\ &= \sum_{k=1}^{n+1} [c_k \mu(E_k)] = \sum_{k=1}^{n} [c_k \mu(E_k)],\end{aligned}$$

where the second last equality is a result of $(*_3)$. Hence, we have that $\mathcal{L}(f, P) \leq \sum_{k=1}^{n} [c_k \mu(E_k)]$ for any arbitrary $\mathcal{S}$-partition P of X. It then follows, from definition, that

$$\int f d\mu \leq \sum_{k=1}^{n} [c_k \mu(E_k)].$$

Here, we complete the proof. $\square$

Theorem 3.1.3 (Monotonicity of Integration (for Nonnegative Functions))**.** *Suppose that* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f, g\colon X \to [0, \infty]$ *are* $\mathcal{S}$*-measurable and* $f(x) \leq g(x)$ $\forall x \in X$*. Then,*

$$\int f d\mu \leq \int g d\mu.$$

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f, g\colon X \to [0, \infty]$ are $\mathcal{S}$-measurable and $f(x) \leq g(x)$ $\forall x \in X$.

Consider arbitrary $\mathcal{S}$-partition $P = \{A_1, \ldots, A_m\}$ of X. We observe that $\forall j \in \{1, \ldots, m\}$ it holds that

$$\begin{aligned} &\forall x \in A_j, \ \inf_{x \in A_j} f(x) \leq f(x) \leq g(x) \\ \implies &\inf_{x \in A_j} f(x) \text{ is a lower bound of } \{g(x) : x \in A_j\} \\ \implies &\inf_{x \in A_j} f(x) \leq \inf_{x \in A_j} g(x), \qquad \text{(By definition)} \end{aligned}$$

which then yield that

$$\begin{aligned} \mathcal{L}(f, P) = \sum_{k=1}^{m} \mu(A_k) \inf_{x \in A_k} f(x) \leq \sum_{k=1}^{m} \mu(A_k) \inf_{x \in A_k} g(x) &= \mathcal{L}(g, P) \\ &\leq \int g d\mu \qquad \text{(By definition)} \end{aligned}$$

implying that $\int g d\mu$ is a upper bound for the set

$$\{\mathcal{L}(f, P) : P \text{ is a } \mathcal{S}\text{-partition of } X\}. \tag{$*$}$$

As an immediate result,

$$\int f d\mu \leq \int g d\mu$$

since $\int f d\mu$ is the least upper bound of the set in $(*)$. $\square$

Theorem 3.1.4 (Integral via Simple Functions (for Nonnegative Functions))**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f\colon X \to [0, \infty]$ *is* $\mathcal{S}$*-measurable. Then,*

$$\begin{aligned} \int f d\mu = \sup \Bigg\{ \sum_{j=1}^{m} c_j \mu(A_j) : &A_1, \ldots, A_m \in \mathcal{S} \text{ are disjoint,} \\ &c_1, \ldots, c_m \in [0, \infty), \text{ and} \\ &\forall x \in X, \ \sum_{j=1}^{m} c_j \chi_{A_j}(x) \leq f(x) \Bigg\}. \end{aligned}$$

Equivalently,

$$\int f d\mu = \sup \left\{ \int s d\mu : s\colon X \to [0, \infty) \text{ is a } \mathcal{S}\text{-measurable simple function, } 0 \leq s \leq f \right\}.$$

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to [0, \infty]$ is $\mathcal{S}$-measurable. Let $\mathcal{U}$ be the set

$$\begin{aligned} \mathcal{U} = \Bigg\{ \sum_{j=1}^{m} c_j \mu(A_j) : &A_1, \ldots, A_m \in \mathcal{S} \text{ are disjoint,} \\ &c_1, \ldots, c_m \in [0, \infty), \text{ and} \\ &\forall x \in X, \ \sum_{j=1}^{m} c_j \chi_{A_j}(x) \leq f(x) \Bigg\}. \end{aligned}$$

It suffice to show that $\int f d\mu \leq \sup \mathcal{U}$ and $\int f d\mu \geq \sup \mathcal{U}$.

($\geq$) Consider aribtrary disjoint sets $A_1, \dots, A_m \in \mathcal{S}$ and constants $c_1, \dots, c_m \in [0, \infty)$ such that

$$f(x) \geq \sum_{k=1}^{m} c_k \chi_{A_k}(x) \ \forall x \in X.$$

It follows that $\sum_{k=1}^{m} c_k \mu(A_k) \in \mathcal{U}$. We observe that, by the monotonicity of integration,

$$\int f d\mu \geq \int \sum_{k=1}^{m} c_k \chi_{A_k} d\mu = \sum_{k=1}^{m} c_k \mu(A_k) \qquad \text{(By Theorem 3.7 (Axler 76))}$$

Thus, we showed that $\int f d\mu$ is an upper bound for $\mathcal{U}$. It follows immediately that $\int f d\mu \geq \sup \mathcal{U}$.

($\leq$) Here, we have two cases to consider. Namely, the cases where $\forall A \in \mathcal{S}\ \mu(A) > 0 \implies \inf_{x \in A} f(x) < \infty$ and $\exists A \in \mathcal{S}$ such that $\mu(A) = 0$ and $\inf_{x \in A} f(x) = \infty$.

(Case #1): Suppose that $\forall A \in \mathcal{S}\ \mu(A) > 0 \implies \inf_{x \in A} f(x) < \infty$. By the contraposition of the hypothesis, we have that $\inf_{x \in A} f(x) = \infty \implies \mu(A) = 0$. Here, we adopt the convention $0 \cdot \infty = 0$.

Consider arbitrary $\mathcal{S}$-partition $P = \{A_1, \dots, A_m\}$ of X. We observe that

$$\sum_{k=1}^{m} \mu(A_k) \inf_{x \in A_k} f(x) = \sum_{k \in \{1, \dots, m\}, \mu(A_k) > 0} \mu(A_k) \inf_{x \in A_k} f(x).$$

Here, $\forall k \in \{1, \dots, m\}$ such that $\mu(A_k) > 0$ it holds that $\inf_{x \in A_k} f(x) < \infty$, since, $\forall k \in \{1, \dots, m\}$, the hypothesis implies that if A_k has a positive measure, then it corresponds to a finite infimum $\inf_{x \in A_k} f(x)$; where the contraposition of the hypothesis implies that if A_k corresponds to an infinite infimum $\inf_{x \in A_k} f(x)$, then it has measure 0.

By omitting indexes corresponding to nonpositive measure $\mu(A_j)$ and reindexing, we obtain the set of indexes $\{1, \dots, M\}$ such that

$$\sum_{j=1}^{M} \mu(A_j) \inf_{x \in A_j} f(x) = \sum_{k=1}^{m} \mu(A_k) \inf_{x \in A_k} f(x), \qquad (*)$$

where $M \in \{1, \dots, m\}$, $\forall j \in \{1, \dots, M\}\ \mu(A_j) > 0$ and $\inf_{x \in A_j} f(x) < \infty$. Let $c_j = \inf_{x \in A_j} f(x)$ $\forall j \in \{1, \dots, M\}$, so we have that $c_1, \dots, c_M \in [0, \infty)$.

Fix $x \in X$. We note that $\exists! n \in \{1, \dots, M\}$ such that $x \in A_n$, by the definition of a $\mathcal{S}$-partition. It follows that

$$\sum_{j=1}^{M} c_j \chi_{A_j}(x) = c_n = \inf_{x \in A_n} f(x) \leq f(x).$$

We then yield that

$$\mathcal{L}(f, P) = \sum_{j=1}^{m} \mu(A_j) \cdot \inf_{x \in A_j} f(x) \qquad \text{(By definition)}$$

$$= \sum_{j=1}^{M} \mu(A_j) \cdot \inf_{x \in A_j} f(x) = \sum_{j=1}^{M} \mu(A_j) c_j \in \mathcal{U}. \qquad \text{(By (*))}$$

Thus, $\sup \mathcal{U}$ is an upper bound for $\{\mathcal{L}(f, P) : P \text{ is a } \mathcal{S}\text{-partition of } X\}$. By definition, $\int f d\mu \leq \sup \mathcal{U}$ since $\int f d\mu$ is the least upper bound of the set above.

(Case #2): Suppose that $\exists A \in \mathcal{S}$ such that $\mu(A) = 0$ and $\inf_{x \in A} f(x) = \infty$. It follows that

$$\infty = \inf_{x \in A} f(x) \le f(x) \ \forall x \in A \implies f(x) = \infty \ \forall x \in A.$$

Fix $t \in (0, \infty)$. We observe that

$$\begin{aligned} & t\mu(A) \in \mathcal{U} && \text{(Since } f(x) = \infty > t\chi_A(x),\ \forall x \in X\text{)} \\ \implies & \sup \mathcal{U} \ge t\mu(A) && (**) \end{aligned}$$

implying that $\sup \mathcal{U} = \infty$, for otherwise we would yield that

$$\begin{aligned} \sup \mathcal{U} < \infty \implies \sup \mathcal{U} \in [0, \infty) \implies & \left[\frac{\sup \mathcal{U} + 1}{\mu(A)}\right] \mu(A) \in \mathcal{U} && \left(\text{Since } \tfrac{\sup \mathcal{U}+1}{\mu(A)} \in (0, \infty)\right) \\ \implies \sup \mathcal{U} \ge & \left[\frac{\sup \mathcal{U} + 1}{\mu(A)}\right] \mu(A) = \sup \mathcal{U} + 1, && \text{(By } (**)\text{)} \end{aligned}$$

which is a contradiction. Hence, $\int f d\mu \le \sup \mathcal{U} = \infty$.

Therefore, we conclude that $\int f d\mu = \sup \mathcal{U}$ as desired. Here, we complete this part of the proof.

To prove the last equality, it suffices to show that

$$\mathcal{U} = \left\{ \int s d\mu : s \colon X \to [0, \infty) \text{ is a simple } \mathcal{S}\text{-measurable function, } 0 \le s \le f \right\}.$$

□

Theorem 3.1.5 (Monotone Convergence Theorem)**.** *Suppose that* $(X, \mathcal{S}, \mu)$ *is a measure space, and* $(f_k)_{k \in \mathbb{N}}$ *is an increasing sequence of nonnegative* $\mathcal{S}$*-measurable functions, defined from* X *to* $[0, \infty]$*, converging pointwise to* $f \colon X \to [0, \infty]$ *on* X*. Then,*

$$\lim_{k \to \infty} \int f_k d\mu = \int f d\mu.$$

That is, suppose that, $\forall k \in \mathbb{N}$*,* $f_k \colon X \to [0, \infty]$ *is* $\mathcal{S}$*-measurable and* $0 \le f_k(x) \le f_{k+1}(x)$ $\forall x \in X$*. Then,*

$$\lim_{k \to \infty} \int f_k d\mu = \int f d\mu.$$

Proof. **[INCOMPLETE]**

□

Theorem 3.1.6 (Integral-Type Sum for Simple Functions)**.** *Suppose that* $(X, \mathcal{S}, \mu)$ *is a measure space. Suppose* $a_1, \ldots, a_m, b_1, \ldots, b_n \in [0, \infty]$ *and* $A_1, \ldots, A_m, B_1, \ldots, B_n \in \mathcal{S}$ *satisfies*

$$\sum_{j=1}^{m} a_j \chi_{A_j} = \sum_{k=1}^{n} b_k \chi_{B_k}.$$

Then,

$$\sum_{j=1}^{m} a_j \mu(A_j) = \sum_{k=1}^{n} b_k \mu(B_k).$$

Proof. **[INCOMPLETE]**

□

Theorem 3.1.7 (Integral of Linear Combination of Characteristic Functions)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $E_1, \ldots, E_n \in \mathcal{S}$, and $c_1, \ldots, c_n \in [0, \infty]$. Then,*

$$\int \left(\sum_{k=1}^{n} c_k \chi_{E_k} \right) d\mu = \sum_{k=1}^{n} c_k \mu(E_k).$$

Proof. **[INCOMPLETE]**

□

Theorem 3.1.8 (Additivity of Integration (of Nonnegative Functions))**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \colon X \to [0, \infty]$ are $\mathcal{S}$-measurable. Then,*

$$\int (f + g)\, d\mu = \int f d\mu + \int g d\mu.$$

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \colon X \to [0, \infty]$ are $\mathcal{S}$-measurable.

We first observe that the equality holds for simple nonnegative $\mathcal{S}$-measurable functions by Theorem 3.15 (Axler 80).

To see this, fix a nonnegative, $\mathcal{S}$-measurable, simple functions $S, T \colon X \to \mathbb{R}$. By definition, $S(X)$ and $T(X)$ are finite. Therefore, we may write, for some $n, m \in \mathbb{N}$, $S(X) = \{c_1, \ldots, c_n\}$ and $T(X) = \{r_1, \ldots, r_m\}$. Moreover, we also have that

$$S(x) = \sum_{k=1}^{n} c_k \cdot \chi_{E_k}(x) \ \forall x \in X,$$

where, $\forall k \in \{1, \ldots, n\}$, $E_k = S^{-1}(\{c_k\})$, $c_k \in [0, \infty) \subset [0, \infty]$, and $E_k \in \mathcal{S}$ since S is $\mathcal{S}$-measurable and $\{c_k\}$ is a Borel set. Similarly,

$$T(x) = \sum_{j=1}^{m} r_j \cdot \chi_{F_j}(x) \ \forall x \in X,$$

where, $\forall j \in \{1, \ldots, m\}$, $F_j = T^{-1}(\{r_j\})$, $r_j \in [0, \infty) \subset [0, \infty]$, and $F_j \in \mathcal{S}$.

Let $a_k = c_k$ and $A_k = E_k$, $\forall k \in \{1, \ldots, n\}$. In addition, let $a_{n+j} = r_j$ and $A_{n+j} = F_j$, $\forall j \in \{1, \ldots, m\}$. Thus, we have that

$$[S + T](x) = \sum_{k=1}^{n+m} a_k \chi_{A_k}(x) \ \forall x \in X,$$

where, $\forall k \in \{1, \ldots, n + m\}$,$c_k \in [0, \infty) \subset [0, \infty]$ and $A_k \in \mathcal{S}$.

By Theorem 3.15 (Axler 80), it holds that

$$\int S d\mu = \sum_{k=1}^{n} c_k \mu(E_k), \ \int T d\mu = \sum_{j=1}^{m} r_j \mu(F_j), \text{ and } \int [S + T] d\mu = \sum_{k=1}^{n+m} a_k \mu(A_k). \tag{$*$}$$

As an immediate result,

$$\begin{aligned} \int [S + T] d\mu = \sum_{k=1}^{n+m} a_k \mu(A_k) &= \sum_{k=1}^{n} a_k \mu(A_k) + \sum_{k=n+1}^{n+m} a_k \mu(A_k) \\ &= \sum_{k=1}^{n} c_k \mu(E_k) + \sum_{j=1}^{m} r_j \mu(F_j) && \text{(By assumption)} \\ &= \int S d\mu + \int T d\mu. && \text{(By ($*$))} \end{aligned}$$

Thus, additivity indeed holds for simple nonnegative $\mathcal{S}$-measurable functions.

We observe that the $\mathcal{S}$-measurability of f and g enables us, by Theorem 2.89 (Axler 65), to approximate the functions respectively via simple functions, $(f_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$, defined from X to $\mathbb{R}$, where $(f_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ are increasing sequences of $\mathcal{S}$-measurable functions. Moreover, $(f_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ converges pointwise to f and g, respectively, on X and consist of nonnegative functions by construction.

It follows that, $\forall k \in \mathbb{N}$, $f_k + g_k$ is $\mathcal{S}$-measurable by the regularity of $\mathcal{S}$-measurable functions (Theorem 2.46, Axler 35). Furthermore, $(f_k + g_k)_{k\in\mathbb{N}}$ is an increasing sequence of nonnegative functions.

As an immediate result of the Montone Convergence Theorem, we yield that

$$\lim_{k\to\infty} f_k d\mu = \int f d\mu, \ \lim_{k\to\infty} g_k d\mu = \int g d\mu, \text{ and } \lim_{k\to\infty} \int [f_k + g_k] d\mu = \int [f+g] d\mu. \qquad (*_1)$$

Here, we have two cases to consider; namely, when both $\int f d\mu$ and $\int g d\mu$ are finite, and when either $\int f d\mu$ or $\int g d\mu$ (or both) is infinite.

(Case #1): Suppose that $\int f d\mu, \int g d\mu < \infty$. It follows that

$$\begin{aligned}
\int f d\mu + \int g d\mu &= \lim_{k\to\infty} \int f_k d\mu + \lim_{k\to\infty} \int g_k d\mu && \text{(By } (*_1)\text{)} \\
&= \lim_{k\to\infty} \left[\int f_k d\mu + \int g_k d\mu \right] && \text{(Since the limits exist)} \\
&= \lim_{k\to\infty} \left[\int [f_k + g_k] d\mu \right] \\
&= \int [f+g] d\mu, && \text{(By } (*_1)\text{)}
\end{aligned}$$

where the third equality holds since, $\forall k \in \mathbb{N}$, f_k and g_k are simple nonnegative $\mathcal{S}$-measurable functions.

(Case #2): Without loss of generality, we assume that $\int f d\mu = \infty$. It follows that $\int f d\mu + \int g d\mu = \infty$. Since f and g are nonnegative, we have that $[f+g](x) \geq f(x)$ $\forall x \in X$. By the monotonicity of integration of nonnegative functions (Theorem 3.8, Axler 77), we yield that $\int [f+g] d\mu \geq \int f d\mu = \infty$. Therefore, $\int [f+g] d\mu = \infty = \int f d\mu + \int g d\mu$ as desired.

Here, we complete the proof. $\square$

Definition 3.1.4 (f^+; f^-)**.** *Suppose X is a set and $f\colon X \to [-\infty, \infty]$ is a function. Then, we define $f^+, f^-\colon X \to [-\infty, \infty]$ by*

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) \geq 0 \\ 0 & \text{if } f(x) < 0 \end{cases} \text{ and } f^-(x) = \begin{cases} 0 & \text{if } f(x) \geq 0 \\ -f(x) & \text{if } f(x) < 0 \end{cases}, \ \forall x \in X.$$

Proposition 3.1.1 (Function Decomposition)**.** *Suppose X is a set and $f\colon X \to [-\infty, \infty]$ is a function. Then,*

(i) $f = f^+ - f^-$, and

(ii) $|f| = f^+ + f^-$.

Proof. **[INCOMPLETE]** Suppose X is a set and $f\colon X \to [-\infty, \infty]$ is a function.

(i)

(ii) $\square$

Proposition 3.1.2 ($\mathcal{S}$-Measurability of f^+ and f^-)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measurable space and $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable. Then, f^+ and f^- are $\mathcal{S}$-measurable.*

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measurable space and $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable. To prove the desired claim, it suffices to show that $\forall a \in \mathbb{R}$ $f^+((a, \infty]),\ f^-((a, \infty]) \in \mathcal{S}$ and invoke Theorem 2.52 (Axler 37).

Fix $a \in \mathbb{R}$. Suppose that $a \geq 0$. We first observe that $\forall x \in X$

$$f^+(x) > a \geq 0 \iff f^+(x) = f(x) \iff f(x) > a \geq 0.$$

It follows that

$$\begin{aligned} {f^+}^{-1}((a, \infty]) = \big\{x \in X : f^+(x) \in (a, \infty]\big\} = \{x \in X : f(x) \in (a, \infty]\} \\ = f((a, \infty]) \in \mathcal{S}, \end{aligned}$$

since f is $\mathcal{S}$-measurable and $(a, \infty] = (a, \infty) \cup \{\infty\}$ implies $(a, \infty] \in \mathcal{B}_\infty$.

Similarly, we observe that, $\forall x \in X$,

$$f^-(x) > a \geq 0 \iff f^-(x) = -f(x) \iff -f(x) > a \geq 0.$$

Therefore,

$$\begin{aligned} {f^-}^{-1}((a, \infty]) = \big\{x \in X : f^-(x) \in (a, \infty]\big\} = \{x \in X : -f(x) \in (a, \infty]\} \\ = [-f]((a, \infty]) \in \mathcal{S}, \end{aligned}$$

since $-f$ is $\mathcal{S}$-measurable by regularity and $(a, \infty] \in \mathcal{B}_\infty$.

Now, suppose that $a < 0$. By definition, we have that

$$\begin{aligned} {f^+}^{-1}((a, \infty]) = \big\{x \in X : f^+(x) \in (a, \infty]\big\} \\ = \big\{x \in X : f^+(x) \in (a, 0]\big\} \sqcup \big\{x \in X : f^+(x) \in (0, \infty]\big\} \\ = \big\{x \in X : f^+(x) \in (a, 0]\big\} \sqcup {f^+}^{-1}((0, \infty]). \end{aligned}$$

We observe that by the construction of f^+, $\forall x \in X$,

$$f^+(x) \in (a, 0] \iff f^+(x) = 0 \iff f(x) < 0.$$

Therefore, we may write

$$\big\{x \in X : f^+(x) \in (a, 0]\big\} = \{x \in X : f(x) \in [-\infty, 0)\} = f^{-1}([-\infty, 0)),$$

where the above set is contained in $\mathcal{S}$ since $[-\infty, 0) \in \mathcal{B}_\infty$ and f is $\mathcal{S}$-measurable. By combining the above results, ${f^+}^{-1}((a, \infty]) \in \mathcal{S}$ by the closure of $\mathcal{S}$ under finite union, where ${f^+}^{-1}((0, \infty]) \in \mathcal{S}$ by the case where $a \geq 0$.

Similarly, by definition, we have that

$$\begin{aligned} {f^-}^{-1}((a, \infty]) = \big\{x \in X : f^-(x) \in (a, \infty]\big\} \\ = \big\{x \in X : f^-(x) \in (a, 0]\big\} \sqcup \big\{x \in X : f^-(x) \in (0, \infty]\big\} \\ = \big\{x \in X : f^-(x) \in (a, 0]\big\} \sqcup {f^-}^{-1}((0, \infty]). \end{aligned}$$

We observe that by the construction of f^-, $\forall x \in X$,

$$f^-(x) \in (a, 0] \iff f^-(x) = 0 \iff f(x) \geq 0.$$

Therefore, we may write

$$\{x \in X : f^-(x) \in (a, 0]\} = \{x \in X : f(x) \in [0, \infty]\} = f^{-1}([0, \infty]),$$

where the above set is contained in $\mathcal{S}$ since $[0, \infty] = [0, \infty) \cup \{\infty\} \in \mathcal{B}_\infty$ and f is $\mathcal{S}$-measurable. By combining the above results, $f^{-\,-1}((a, \infty]) \in \mathcal{S}$ by the closure of $\mathcal{S}$ under finite union, where $f^{-\,-1}((0, \infty]) \in \mathcal{S}$ by the case where $a \geq 0$.

Therefore, we conclude that f^+ and f^- are $\mathcal{S}$-measurable by definition as desired. □

Proposition 3.1.3 (Properties of f^+ and f^-). *Suppose X is a set and $f \colon X \to [-\infty, \infty]$. Then,*

(i) $\forall x \in X,\ f(x) \geq 0 \implies f^+ = f$ and f^- is identically 0;

(ii) $\forall x \in X,\ f(x) < 0 \implies f^+$ is identically 0 and $f = -f^-$;

(iii) $c \geq 0 \implies (cf)^+ = cf^+$ and $(cf)^- = cf^-$;

(iv) $c < 0 \implies (cf)^+ = -cf^-$ and $(cf)^- = -cf^+$;

(v) If $E \subset X$, then $(\chi_E f)^+ = \chi_E f^+$ and $(\chi_E f)^- = \chi_E f^-$;

(vi) If $g \colon X \to [-\infty, \infty]$, then it needs not to hold that $(f+g)^+ = f^+ + g^+$ and $(f-g)^- = f^- + g^-$.

Proof. Suppose X is a set and $f \colon X \to [-\infty, \infty]$.

(v) Suppose $E \subset X$. Fix $x \in X$.

(case #1): Suppose that $x \notin E$. It follows that $\chi_E(x) = 0$ and, hence,

$$\chi_E(x)f(x), \chi_E(x)f^+(x), \chi_E(x)f^-(x) = 0.$$

Moreover, $(\chi_E f)(x) \geq 0 \implies (\chi_E f)^+(x) = \chi_E(x)f(x)$ and $(\chi_E f)^-(x) = 0$ by definition. Thus, $(\chi_E f)^+(x) = (\chi_E f^+)(x)$ and $(\chi_E f)^-(x) = (\chi_E f^-)(x)$.

(Case #2): Suppose that $X \in E$. Then, $\chi_E(x) = 1$.

Suppose that $(\chi_E f)(x) \geq 0$. By definition, $(\chi_E f)^+(x) = (\chi_E f)(x)$ and $(\chi_E f)^-(x) = 0$. Furthermore, we must have that $f(x) \geq 0$, implying $f^+(x) = f(x)$ and $f^-(x) = 0$. By combining the above results, it holds that

$$(\chi_E f)^+(x) = (\chi_E f)(x) = (\chi_E f^+)(x) \text{ and } (\chi_E f)^-(x) = (\chi_E f^-)(x).$$

Suppose that $(\chi_E f)(x) < 0$. It follows that $f(x) < 0$, implying that $f^+(x) = 0$ and $f^-(x) = -f(x)$ by definition. In addition, $(\chi_E f)(x) < 0 \implies (\chi_E f)^+(x) = 0$ and $(\chi_E f)^-(x) = -(\chi_E f)(x)$ by definition. As an immediate result,

$$(\chi_E f^+)(x) = 0 = (\chi_E f)^+(x) \text{ and } (\chi_E f^-)(x) = (\chi_E f)^-(x)$$

Indeed, $(\chi_E f)^+ = \chi_E f^+$ and $(\chi_E f)^- = \chi_E f^-$.

□

Definition 3.1.5 (Integral of a real-valued function). *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f \colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable and such that $\int f^+ d\mu < \infty$ or $\int f^- d\mu < \infty$ (or both integrals are finite). Then, the integral of f with respect to μ is defined by*

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu.$$

Theorem 3.1.9 (Homogeneity of Integration)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f\colon X \to [-\infty, \infty]$ *is such that* $\int f d\mu$ *is defined. If* $c \in \mathbb{R}$, *then*

$$\int cf d\mu = c \int f d\mu.$$

Proof. **[INCOMPLETE]**

□

Theorem 3.1.10 (Additivity of Integration)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f, g\colon X \to \mathbb{R}$ *are* $\mathcal{S}$*-measurable such that* $\int |f| \, d\mu, \int |g| \, d\mu < \infty$. *Then,*

$$\int (f + g) \, d\mu = \int f d\mu + \int g d\mu.$$

Proof. **[INCOMPLETE]**

□

Proposition 3.1.4 (Linearity of Integration)**.**

Theorem 3.1.11 (Monotonicity of Integration)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f, g\colon X \to \mathbb{R}$ *are* $\mathcal{S}$*-measurable such that* $\int f d\mu, \int g d\mu$ *are defined. Suppose further that* $\forall x \in X$ $f(x) \leq g(x)$. *Then,*

$$\int f \mu \leq \int g d\mu.$$

Proof. **[INCOMPLETE]**

□

Theorem 3.1.12 (Absoulte Value of Integral and Integral of Absolute Value)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space and* $f\colon X \to [-\infty, \infty]$ *is such that* $\int f d\mu$ *is defined. Then,*

$$\left| \int f d\mu \right| \leq \int |f| \, d\mu.$$

Proof. **[INCOMPLETE]**

□

3.2 Limits of Integrals and Integrals of Limits

Definition 3.2.1 (Integration on a subset)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measureable, and $E \in \mathcal{S}$. Then,*

$$\int_E f d\mu = \int \chi_E \cdot f d\mu,$$

provided $\int \chi_E \cdot f d\mu$ is defined; otherwise $\int_E f d\mu$ is not defined.

Theorem 3.2.1 (Integral on Subset via Simple Functions (for Nonnegative Functions))**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $E \in \mathcal{S}$, and $f\colon X \to [0, \infty]$ is $\mathcal{S}$-measurable. Then,*

$$\int_E f d\mu = \sup \left\{ \int_E s d\mu : s\colon X \to [0, \infty) \text{ is a } \mathcal{S}\text{-measurable simple function, } 0 \leq s \leq f \right\}.$$

Proof. This is an equivalent definition of $\int_E f d\mu$. (See Royden's Real Analysis (79)) □

Proposition 3.2.1 (Decomposition of Integral (on Subset) of Function on $[-\infty, \infty]$)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, and $f\colon X \to [-\infty, \infty]$. Suppose further that $U, E \in \mathcal{S}$ with $U \subset E$, where $\int (\chi_E f)^+ d\mu < \infty$ or $\int (\chi_E f)^- d\mu < \infty$. Then,*

$$\int_E f d\mu = \int_{E/U} f d\mu + \int_U f d\mu.$$

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, and $f\colon X \to [-\infty, \infty]$. Suppose further that $U, E \in \mathcal{S}$ with $U \subset E$, where $\int (\chi_E f)^+ d\mu < \infty$ or $\int (\chi_E f)^- d\mu < \infty$. It follows that $\int (\chi_E f) d\mu$ is defined. By definition, $\int_E f d\mu$ is defined as well.

We observe that E/U and U are disjoint with $E = [E/U] \sqcup U$ implies that $\chi_E = \chi_{E/U} + \chi_U$ by a previous proposition. It then follows that

$$\begin{aligned}
\int_E f d\mu = \int \chi_E f d\mu &= \int (\chi_E f)^+ d\mu - \int (\chi_E f)^- d\mu && \text{(By definition)} \\
&= \int \chi_E f^+ d\mu - \int \chi_E f^- d\mu && \text{(By a previous proposition)} \\
&= \int (\chi_{E/U} + \chi_U) f^+ d\mu - \int (\chi_{E/U} + \chi_U) f^- d\mu && \text{(Since } \chi_E = \chi_{E/U} + \chi_U\text{)} \\
&= \int \chi_{E/U} f^+ d\mu + \int \chi_U f^+ d\mu - \int \chi_{E/U} f^- d\mu - \int \chi_U f^- d\mu, && \text{(By Theorem 3.16)}
\end{aligned}$$

since $\chi_{E/U} f^+$, $\chi_U f^+$, $\chi_{E/U} f^-$, and $\chi_U f^-$ are nonnegative $\mathcal{S}$-measurable functions. By the above result and a previous proposition, we then yield

$$\begin{aligned}
\int_E f d\mu &= \left(\int \chi_{E/U} f^+ d\mu - \int \chi_{E/U} f^- d\mu \right) + \left(\int \chi_U f^+ d\mu - \int \chi_U f^- d\mu \right) \\
&= \left(\int [\chi_{E/U} f]^+ d\mu - \int [\chi_{E/U} f]^- d\mu \right) + \left(\int [\chi_U f]^+ d\mu - \int [\chi_U f]^- d\mu \right) \\
&= \int \chi_{E/U} f d\mu + \int \chi_U f d\mu && \text{(By definition)} \\
&= \int_{E/U} f d\mu + \int_U f d\mu. && \text{(By definition)}
\end{aligned}$$

□

Lemma 3.2.1 (Identities Relating to Characteristic Function)**.** *Suppose X is a set, $E \subset X$, and $f\colon X \to Y$ where $Y \subset [-\infty, \infty]$. Then,*

(i) $\chi_X f = \chi_E f + \chi_{X/E} f$*, and*

(ii) $f = \chi_X f$*.*

Proof. Suppose X is a set, $E \subset X$, and $f\colon X \to Y$ where $Y \subset [-\infty, \infty]$.

(i) Fix $x \in X$. If $x \in E$, then

$$[\chi_X f](x) = (1) \cdot f(x) = (0) \cdot f(x) + (1) \cdot f(x) = [\chi_{X/E} f + \chi_E f](x).$$

If $x \notin E$, then

$$[\chi_X f](x) = (1) \cdot f(x) = (1) \cdot f(x) + (0) \cdot f(x) = [\chi_{X/E} f + \chi_E f](x).$$

Therefore, the equality holds.

(ii) Fix $x \in X$. It follows that $[\chi_X f](x) = (1) \cdot f(x) = f(x)$. Therefore, $\chi_X f = f$ as desired.

$\square$

Theorem 3.2.2 (Trivial Integral (of Nonnegative Function) on a Set of Measure Zero)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space $f\colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $E \in \mathcal{S}$ with $\mu(E) = 0$. Then,*

$$\int_E f d\mu = 0.$$

Theorem 3.2.3 (Bounding an Integral on a Subset)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measureable, $E \in \mathcal{S}$, and $\int_E f d\mu$ is defined. Then,*

$$\left| \int_E f d\mu \right| \le \mu(E) \sup_{x \in E} |f(x)| .$$

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measureable, $E \in \mathcal{S}$, and $\int_E f d\mu$ is defined. Let $c = \sup_{x \in E} |f(x)|$.

We observe that, $\forall x \in X$, $|[\chi_E \cdot f]|\,(x) = [\chi_E \cdot |f|](x) \le [c \cdot \chi_E](x)$. In addition, $|f|$ is $\mathcal{S}$-measurable by Example 2.45 since f is $\mathcal{S}$-meaurable.

We also have that the constant function $c_{\mathbb{R}}\colon X \to [0, \infty]$ defined by $c_{\mathbb{R}}(x) = c$ $\forall x \in X$ is $\mathcal{S}$-measurable. To see this, we first fix $B \in \mathcal{B}_\infty$, where $\mathcal{B}_\infty$ denotes the collection of Borel subsets of $[-\infty, \infty]$. It follows that

$$c_{\mathbb{R}}{}^{-1}(B) = \{x \in X : c_{\mathbb{R}}(x) \in B\} = \begin{cases} X & \text{if } c \in B \\ \emptyset & \text{if } c \notin B \end{cases} \in \mathcal{S}.$$

By definition, $c_{\mathbb{R}}$ is indeed $\mathcal{S}$-measurable. Furthermore, χ_E is $\mathcal{S}$-measurable since $E \in \mathcal{S}$. As an immediate result, $\chi_E \cdot |f| : X \to [0, \infty]$ and $\chi_E \cdot c_{\mathbb{R}}\colon X \to [0, \infty]$ are $\mathcal{S}$-measurable functions by the regularity of $\mathcal{S}$-measurable functions. Hence, we may apply Theorem 3.8 (Axler 77).

It then follows that

$$\begin{aligned} \left| \int_E f d\mu \right| &= \left| \int \chi_E \cdot f d\mu \right| && \text{(By definition)} \\ \le \int |\chi_E \cdot f|\, d\mu &= \int [\chi_E \cdot |f|] d\mu && \text{(By Theorem 3.23 (Axler 84))} \\ \le \int [\chi_E \cdot c_{\mathbb{R}}] d\mu &= \int [c \cdot \chi_E] d\mu && \text{(By Theorem 3.8 (Axler 77))} \\ &= c\mu(E). && \text{(By Theorem 3.15)} \end{aligned}$$

$\square$

Theorem 3.2.4 (Bounded Convergence Theorem)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space with* $\mu(X) < \infty$*. Suppose further that* $(f_k)_{k\in\mathbb{N}}$ *is a sequence of* $\mathcal{S}$*-measurable functions from* X *to* $\mathbb{R}$ *converging pointwise on* X *to* $f\colon X \to \mathbb{R}$*. If there exists* $c \in (0, \infty)$ *such that*

$$k \in \mathbb{N} \text{ and } x \in X \implies |f_k(x)| \leq c,$$

Then,

$$\lim_{k\to\infty} \int f_k d\mu = \int f d\mu.$$

Proof. **[INCOMPLETE]**

$\square$

Definition 3.2.2 (Almost every)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space. A set* $E \in \mathcal{S}$ *contains* μ*-almost every element of* X *if* $\mu(X/E) = 0$.

A statement is true for almost every element in X *if the statement is true for elements in a set* $E \in \mathcal{S}$ *containing* μ*-almost every element of* X.

We say, more succinctly, almost every, provided sufficient context regarding the measure μ *is supplied.*

Example 3.2.1. *Consider the measure space* $(\mathbb{R}, \mathcal{B}, \lambda)$*. Then, almost every real number is irrational since* $\forall x \in \mathbb{R}/\mathbb{Q}$ x *is irrational and* $\lambda(\mathbb{R}/[\mathbb{R}/\mathbb{Q}]) = \lambda(\mathbb{Q}) = 0$ *as countable subsets of* $\mathbb{R}$ *has Lebesgue measure* 0.

Remark 3.2.1. *Note that the assumption that* $f_k \to f$ *pointwise on* X *can be replaced by* $f_k \to f$ *almost everywhere as the measure of the almost everywhere set is the same as the measure of the set excluding the zero sets.*

Theorem 3.2.5 (Trivial Integral (of Nonnegative Function) on Subset $\iff$ Vanishes Almost Everywhere)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space,* $E \in \mathcal{S}$*, and* $f\colon X \to [0, \infty]$ *is* $\mathcal{S}$*-measurable. Then,*

$$\int_E f d\mu = 0 \iff f(x) = 0 \text{ almost everywhere in } E.$$

Theorem 3.2.6 (Equality Almost Everywhere Implies Equality of Integrals)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space,* $f, g\colon X \to [-\infty, \infty]$ *are* $\mathcal{S}$*-measurable, and* $g(x) = f(x)$ *for almost every* $x \in X$*. Then,* $\int f d\mu = \int g d\mu$ *or both integrals are undefined.*

Corollary 3.2.1. *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space, and* $f\colon X \to [-\infty, \infty]$ *is* $\mathcal{S}$*-measurable. Suppose further that* $E \in \mathcal{S}$ *contains almost every element of* X*, and* $\int f d\mu$ *is defined. Then,*

$$\int f d\mu = \int_E f d\mu.$$

Theorem 3.2.7 (Properties of Nonnegative $\mathcal{S}$-Measurable Functions with Finite Integral)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space,* $g\colon X \to [0, \infty]$ *is* $\mathcal{S}$*-measurable, and* $\int g d\mu < \infty$*. Then,*

(i) ***(Small Measure Implies Small Integral)***

$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ such that } B \in \mathcal{S} \text{ and } \mu(B) < \delta \implies \int_B g d\mu < \epsilon;$$

(ii) ***(Integrable Functions Lives Mostly on Sets with Finite Measure)***

$$\forall \epsilon > 0, \exists E \in \mathcal{S} \text{ such that } \mu(E) < \infty \text{ and } \int_{X/E} g d\mu < \epsilon.$$

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, $g\colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $\int g d\mu < \infty$.

(i) Let $\mathcal{U}$ be the set

$$\mathcal{U} = \left\{ \sum_{j=1}^{m} c_j \mu(A_j) : A_1, \ldots, A_m \in \mathcal{S} \text{ are disjoint}, \right.$$
$$c_1, \ldots, c_m \in [0, \infty), \text{ and}$$
$$\left. \forall x \in X,\ \sum_{j=1}^{m} c_j \chi_{A_j}(x) \leq g(x) \right\}.$$

By Theorem 3.9 (Axler 77), $\int g d\mu = \sup \mathcal{U}$, and we yield that $\sup \mathcal{U} \in \mathbb{R}$ since the integral of g is finite. In addition, we observe that

$$1 \cdot \chi_{\emptyset}(x) = 0 \leq g(x),\ \forall x \in X$$

which implies that $1 \cdot \mu(\emptyset) = 0 \in \mathcal{U}$. Thus, $\mathcal{U}$ is non-empty. Fix $\epsilon > 0$. By an equivalent definition of a supremum, we obtain that

$$\exists u \in \mathcal{U} \text{ such that } \int g d\mu - \frac{\epsilon}{2} < u, \tag{$*_0$}$$

where $u = \sum_{j=1}^{m} c_j \mu(A_j)$ for some disjoint sets $A_1, \ldots, A_m \in \mathcal{S}$ and real constants $c_1, \ldots, c_m \in [0, \infty)$ satisfying

$$\forall x \in X,\ \sum_{j=1}^{m} c_j \chi_{A_j}(x) \leq g(x).$$

Let $h = \sum_{j=1}^{m} c_j \chi_{A_j}$ with $h\colon X \to [0, \infty]$. We observe that here h is simple and $\mathcal{S}$-measurable since $A_1, \ldots, A_m \in \mathcal{S}$. Moreover, $\forall x \in X$ $0 \leq h(x) \leq g(x)$ and $\int h d\mu = u < \infty$ by Theorem 3.7 (Axler 76) and Theorem 3.8 (Axler 77), which implies that

$$\int g d\mu - \int h d\mu < \frac{\epsilon}{2}. \tag{By $(*_0)$; $(*_1)$}$$

Let $H = \max\{h(x) : x \in X\}$, which exists since h is simple implies its range is finite. Here, $H \geq 0$. If $H = 0$, then consider arbitrary $\delta > 0$; if $H > 0$, then consider $\delta = \frac{1}{2} \cdot \frac{\epsilon}{2H}$. Thus, we have $\delta > 0$ satisfying $H\delta < \frac{\epsilon}{2}$.

Fix $B \in \mathcal{S}$ and suppose that $\mu(B) < \delta$. We observe that $\chi_B[g-h], \chi_B h\colon X \to [0, \infty]$ are $\mathcal{S}$-measurable by the regularity of $\mathcal{S}$-measurable functions since g, h, χ_B are $\mathcal{S}$-meausrable. In addition, $\chi_B[g-h]$ and $\chi_B h$ are nonnegative. Thus, we may apply Theorem 3.16 (Axler 80) to decompose $\int_B g d\mu$. It follows that

$$\begin{aligned}
\int_B g d\mu &= \int \chi_B \left([g-h] + h\right) d\mu && \text{(By definition)} \\
&= \int \chi_B [g-h] d\mu + \int \chi_B h d\mu && \text{(By Theorem 3.16 (Axler 80))} \\
&\leq \int [g-h] d\mu + \int H \cdot \chi_B d\mu && \text{(By Theorem 3.8 (Axler 77))} \\
&= \int [g-h] d\mu + H\mu(B) && \text{(By Theorem 3.15 (Axler 80))} \\
&< \frac{\epsilon}{2} + H\delta && \text{(Since } \textstyle\int [g-h] d\mu = \int g d\mu - \int h d\mu \text{ and } \mu(B) < \delta) \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon. && \text{(Since } H\delta < \tfrac{\epsilon}{2})
\end{aligned}$$

To see why $\int [g-h]d\mu = \int g d\mu - \int h d\mu$, we first show that $\mathcal{L}(g-h,P) = \mathcal{L}(g,P) - \mathcal{L}(h,P)$ for any $\mathcal{S}$-partition P of X. Fix a $\mathcal{S}$-parition $P = \{A_1, \ldots, A_m\}$ of X. Then, fix $j \in \{1, \ldots, m\}$. We observe that,

$$\begin{aligned} &\forall x \in A_j,\ [g(x)-h(x)] + h(x) = g(x) \\ \implies \inf_{x\in A_j}([g(x)-h(x)] + h(x)) = \inf_{x\in A_j}[g(x)-h(x)] + \inf_{x\in A_j} h(x) &= \inf_{x\in A_j} g(x). \end{aligned}$$

As an immediate result,

$$\begin{aligned} \sum_{j=1}^m \mu(A_j) \inf_{x\in A_j} g(x) &= \sum_{j=1}^m \mu(A_j)\left(\inf_{x\in A_j}[g(x)-h(x)] + \inf_{x\in A_j} h(x)\right) \\ &= \sum_{j=1}^m \mu(A_j)\inf_{x\in A_j}[g(x)-h(x)] + \sum_{j=1}^m \mu(A_j)\inf_{x\in A_j} h(x) \\ &\implies \mathcal{L}(g,P) = \mathcal{L}(g-h,P) + \mathcal{L}(h,P). \end{aligned}$$

Let $\mathcal{P}$ denote the set of all $\mathcal{S}$-partition of X. It follows that,

$$\begin{aligned} \sup_{P\in\mathcal{P}} \mathcal{L}(g,P) = \sup_{P\in\mathcal{P}}[\mathcal{L}(g-h,P) + \mathcal{L}(h,P)] &= \sup_{P\in\mathcal{P}} \mathcal{L}(g-h,P) + \sup_{P\in\mathcal{P}} \mathcal{L}(h,P) \\ \implies \int g d\mu &= \int [g-h]d\mu + \int h d\mu. \quad \text{(By definition)} \\ \therefore \int [g-h]d\mu &= \int g d\mu - \int h d\mu. \end{aligned}$$

Here, we complete the proof.

(ii) [INCOMPLETE] □

Theorem 3.2.8 (Dominated Convergence Theorem)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f\colon X \to [-\infty,\infty]$ is $\mathcal{S}$-measurable, and $(f_k)_{k\in\mathbb{N}}$ is a sequence of $\mathcal{S}$-measurable functions from X to $[-\infty,\infty]$ converging pointwise to f on almost every $x \in X$. If there exists an $\mathcal{S}$-measurable function $g\colon X \to [0,\infty]$ such that*

$$\int g d\mu < \infty \text{ and } |f_k(x)|$$

for every $k \in \mathbb{N}$ and for almost every $x \in X$. Then,

$$\lim_{k\to\infty} \int f_k d\mu = \int f d\mu.$$

Proof. **[INCOMPLETE]**

□

Theorem 3.2.9 (Riemann Integrability $\iff$ Continuity at Almost Everywhere)**.** *Suppose $a < b$ and $f\colon [a,b] \to \mathbb{R}$ is bounded. Then,*

(i) f is Riemann-integrable $\iff$

$$|\{x \in [a,b] : f \text{ is not continuous at } x\}| = 0.$$

(ii) f is Riemann integrable implies that f is Lebesgue measurable and

$$\int_a^b f = \int_{[a,b]} f d\lambda.$$

Proof. **[INCOMPLETE]**

□

Corollary 3.2.2 (Riemann Integrability of Characteristic Function of Bounded Open Subset of $\mathbb{R}$)**.** *Suppose $a < b$ with $[a, b] \subset \mathbb{R}$ and $G \subset (a, b)$ is open. Then, χ_G is Riemann integrable on $[a, b]$.* ***[DO: Verify validity of the statement]***

Definition 3.2.3 ($\int_a^b f$)**.**

Definition 3.2.4 ($\mathcal{L}^1$-norm; Lebesgue space)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space.*

(i) Suppose $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable. The $\mathcal{L}^1$ norm of f (with respect to μ) is defined by

$$||f||_{1,\mu} = \int |f| \, d\mu.$$

(ii) The Lebesgue space (with respect to μ) is defined by

$$\mathcal{L}^1(\mu) = \left\{ f\colon X \to \mathbb{R} : f \text{ is } \mathcal{S}\text{-measurable and } ||f||_1 < \infty \right\}.$$

We denote $||f||_{1,\mu}$ and $\mathcal{L}^1(\mu)$ more succinctly, by $||f||_1$ and $\mathcal{L}^1$ respectively if sufficient context regarding the measure μ is supplied.

Definition 3.2.5 ($\mathcal{L}^1(\mathbb{R})$;$||f||_1$)**.** *Denote λ the Lebesgue measure on either $\mathcal{B}$ or $\mathcal{L}$, where $\mathcal{B}$ is the collection of Borel subsets of $\mathbb{R}$ and $\mathcal{L}$ is the collection of Lebesgue measurable subsets of $\mathbb{R}$. Then, we denote*

(i) $\mathcal{L}^1(\lambda)$ by $\mathcal{L}^1(\mathbb{R})$, and

(ii) $\int |f| \, d\lambda$ by $||f||_1$.

Proposition 3.2.2 (ℓ^1)**.** *Let $\mu\colon \mathcal{P}(\mathbb{N}) \to [0, \infty]$ be the counting measure on $\mathbb{N}$.*

(i) Suppose $x\colon \mathbb{N} \to \mathbb{R}$ is a map defined by a sequence $(x_k)_{k\in\mathbb{N}}$ of real numbers. Then, $||x||_1$ is given by

$$||x||_1 = \sum_{k\in\mathbb{N}} |x_k|.$$

(ii) $\mathcal{L}^1(\mu)$, which we denote by ℓ^1, is given by

$$\ell^1 = \left\{ x\colon \mathbb{N} \to \mathbb{R} : x \text{ is defined by a sequence } (x_k)_{k\in\mathbb{N}} \text{ of real numbers and } \sum_{k\in\mathbb{N}} |x_k| < \infty \right\}.$$

Theorem 3.2.10 (Properties of the $\mathcal{L}^1$ Norm)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \in \mathcal{L}^1(\mu)$. Then, the following statements hold.*

(Nonnegativity) $||f||_1 \geq 0$.

(Almost Positive Definiteness) $||f||_1 = 0 \iff f(x) = 0$ *for almost every* $x \in X$.

(Homogeneity) $\forall c \in \mathbb{R}$, $||cf||_1 = |c| \, ||f||_1$.

(Triangle Inequality) $||f + g||_1 \leq ||f||_1 + ||g||_1$.

Proof. **[INCOMPLETE]**

□

Remark 3.2.2. *We note that the $\mathcal{L}^1$ norm is almost a norm as it possesses all the properties of a norm except for positive definiteness.*

Theorem 3.2.11 (Approximating $\mathcal{L}^1(\mu)$ Function By Simple Function)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f \in \mathcal{L}^1(\mu)$. Then, $\forall \epsilon > 0$, there exists a simple function $g \in \mathcal{L}^1(\mu)$ such that*

$$||f - g||_1 < \epsilon.$$

Proof. **[INCOMPLETE]**

□

Definition 3.2.6 (Step Function)**.** *A step function is a map $g \colon \mathbb{R} \to \mathbb{R}$ of the form*

$$g = a_1 \chi_{I_1} + \cdots + a_n \chi_{I_n},$$

where, $\forall k \in \{1, \ldots, n\}$, I_k is an interval of $\mathbb{R}$ and $a_k \in \mathbb{R} / \{0\}$.

Theorem 3.2.12 (Approximating $\mathcal{L}^1(\mathbb{R})$ Function By Step Function and Continuous Function)**.** *Suppose $f \in \mathcal{L}^1(\mathbb{R})$. Then,*

(i) $\forall \epsilon > 0$, there exists a step function $g \in \mathcal{L}^1(\mathbb{R})$ such that

$$||f - g||_1 < \epsilon.$$

(ii) $\forall \epsilon > 0$ there exists a continuous map $g \colon \mathbb{R} \to \mathbb{R}$ such that

$$\{x \in \mathbb{R} : g(x) \neq 0\} \text{ is bounded and } ||f - g||_1 < \epsilon.$$

Proof. **[INCOMPLETE]**

□

Theorem 3.2.13 (Approximating $\mathcal{L}^1$ Function By Continuous Function)**.** *Suppose $f \in \mathcal{L}^1(\mathbb{R})$. Then, $\forall \epsilon > 0$ there exists a continuous map $g \colon \mathbb{R} \to \mathbb{R}$ such that*

(i) $||f - g||_1 < \epsilon$, and

(ii) $||x \in \mathbb{R} : g(x) \neq 0||$ is bounded.

Proof. **[INCOMPLETE]**

□

4 Differentiation

5 Product Measures

6 Banach Spaces

7 $\mathcal{L}^p$ Spaces

7.1 $\mathcal{L}^p(\mu)$

Definition 7.1.1 (Notation: $\mathbb{F}$)**.** *Unless otherwise specified, we denote $\mathbb{F}$ the real field $\mathbb{R}$ or the complex field $\mathbb{C}$.*

Definition 7.1.2 (p-norm, essential supremum, $\mathcal{L}^p$ space)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to \mathbb{F}$ is $\mathcal{S}$-measurable.*

(i) Suppose that $p \in (0, \infty)$. Then, the p-norm of f (with respect to μ) is defined by

$$||f||_{p,\mu} = \left(\int |f| \, d\mu \right)^{\frac{1}{p}}.$$

(ii) The essential supremum of f (with respect to μ) is defined by

$$||f||_{\infty,\mu} = \inf \{t > 0 : \{x \in X : |f(x)| > t\} \text{ has measure zero}\}.$$

(iii) Suppose that $p \in (0, \infty]$. Then, the $\mathcal{L}^p$ space (with respect to μ) is defined by

$$\mathcal{L}^p(\mu) = \Big\{ f\colon X \to \mathbb{F} : f \text{ is } \mathcal{S}\text{-measurable and } ||f||_{p,\mu} < \infty \Big\}.$$

We denote $||f||_{p,\mu}$, $||f||_{\infty,\mu}$, and $\mathcal{L}^p(\mu)$ more succinctly, by $||f||_p$, $||f||_\infty$ and $\mathcal{L}^p$ respectively if sufficient context regarding the measure μ is supplied.

Remark 7.1.1. *Note that the essential supremum is well-defined as $\{x \in X : |f(x)| > t\} = \{x \in X : |f|(x) \in (t, \infty)\} = |f|^{-1}((t,\infty)) \in \mathcal{S}$, since $|f|$ is $\mathcal{S}$-measurable by regularity.*

Theorem 7.1.1 (Properties of $||f||_\infty$)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to \mathbb{F}$ is $\mathcal{S}$-measurable. Then,*

(i) $||f||_\infty = \inf \{t > 0 : |f(x)| \le t$ for almost every $x \in X\}$;

(ii) $|f(x)| \le ||f||_\infty$ for almost every $x \in X$; particularly,

(a) $\mu(X) > 0 \implies ||f||_\infty > 0$ and $|f(x)| < ||f||_\infty$ for almost every $x \in X$, and

(b) $\mu(X) = 0 \implies ||f||_\infty = 0$ and $|f(x)| = ||f||_\infty$ for almost every $x \in X$;

(iii) $t \ge ||f||_\infty \implies t \ge |f(x)|$ for almost every $x \in X$.

Proof. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to \mathbb{F}$ is $\mathcal{S}$-measurable. Fix $t > 0$.

(i) We observe that, by definition, $\{x \in X : |f(x)| > t\}$ has measure zero $\iff$

$$X / \{x \in X : |f(x)| > t\} = \{x \in X : |f(x)| \le t\}$$

contains almost every $x \in X \iff |f(x)| \le t$ holds for almost every $x \in X$. The equality in (i) follows.

(ii) (Case #1): Suppose that $\mu(X) > 0$. By Part (i),

$$||f||_\infty = \inf \{t > 0 : |f(x)| \le t \text{ for almost every } x \in X\}.$$

Fix $k \in \mathbb{N}$. By an equivalent definition of infimum, $\exists t_k \in \{t > 0 : |f(x)| \leq t$ for almost every $x \in X\}$ such that $||f||_\infty + \frac{1}{k} > t_k > 0$. Hence,

$$\lim_{k\to\infty}\left(||f||_\infty + \frac{1}{k}\right) = ||f||_\infty > 0.$$

In addition, there exists $A_k \in \mathcal{S}$ such that, $\forall x \in A_k$, $t_k \geq |f(x)|$ and $\mu(X/A_k) = 0$, by definition. Let $Z = \bigcup_{k\in\mathbb{N}}[X/A_k]$. We observe that, by the countable subadditivity of μ,

$$\mu(Z) \leq \sum_{k\in\mathbb{N}} \mu(X/A_k) = \sum_{k\in\mathbb{N}} 0 = 0.$$

As an immediate result, X/Z is nonempty, for otherwise $\mu(X) = \mu([X/Z] \sqcup Z) = \mu(X/Z) + \mu(Z) = 0$, which is a contradiction. Fix $x \in X/Z = \bigcap_{k\in\mathbb{N}} A_k$. It follows that, $\forall k \in \mathbb{N}$, $||f||_\infty + \frac{1}{k} > t_k \geq |f(x)|$. Thus,

$$\lim_{k\to\infty}\left[||f||_\infty + \frac{1}{k}\right] = ||f||_\infty > |f(x)| = \lim_{k\to\infty} |f(x)|.$$

Thus, we showed that $||f||_\infty > |f(x)|$ $\forall x \in X/Z$ where $\mu(Z) = 0$. By definition, $||f||_\infty > |f(x)|$ for almost every $x \in X$.

(Case #2): Suppose that $\mu(X) = 0$. Then, we must have that μ is identically the trivial measure; that is, $\mu(E) = 0$ $\forall E \in \mathcal{S}$. By regularity, $|f|$ is also $\mathcal{S}$-measurable. By definition,

$$||f||_\infty = \inf\{t > 0 : \{x \in X : |f(x)| > t\} \text{ has measure zero }\} = \inf\{t > 0\} = 0,$$

since $\forall t > 0$ $\{x \in X : |f(x)| > t\} = |f|^{-1}((t,\infty)) \in \mathcal{S}$ has measure zero. We observe that $|f|^{-1}(\{0\}) \in \mathcal{S}$, where $\mu(X/|f|^{-1}(\{0\})) = 0$. Moreover, $\forall x \in |f|^{-1}(\{0\})$, $|f(x)| = 0 = ||f||_\infty$. Thus, $|f(x)| = ||f||_\infty$ for almost every $x \in X$ by definition.

(iii) Suppose $t \geq ||f||_\infty$. By Part (ii), $||f||_\infty \geq |f(x)|$ for almost every $x \in X$. Since $t \geq ||f||_\infty$, it holds that $t \geq |f(x)|$ for almost every $x \in X$. $\square$

Proposition 7.1.1 (Essential Supremum of Trivial Measure is Identially Zero). *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, where μ is the trivial measure. Then, for every $\mathcal{S}$-measurable function $f\colon X \to \mathbb{F}$, $||f||_\infty = 0$.*

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space, where μ is the trivial measure. Fix a $\mathcal{S}$-measurable function $f\colon X \to \mathbb{F}$.

By definition, $||f||_\infty = \inf\{t > 0 : \{x \in X : |f(x)| > t\}$ has measure zero$\}$. Since f is $\mathcal{S}$-measurable by hypothesis, it holds that, $\forall t > 0$, $\{x \in X : |f(x)| > t\} = |f|^{-1}((t,\infty)) \in \mathcal{S}$ and has measure zero. Thus, $||f||_\infty$ reduces to $\inf\{t > 0\}$, which equals 0. $\square$

Lemma 7.1.1 (Finite p-Norm Implies Finite Integral). *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $p \in (0,\infty)$, $f\colon X \to \mathbb{F}$ is $\mathcal{S}$-measurable. Then, $||f||_p < \infty \implies \int |f|^p \, d\mu < \infty$*

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to \mathbb{F}$ is $\mathcal{S}$-measurable. By definition

$$\begin{aligned} ||f||_p = \left(\int |f|^p \, d\mu\right)^{\frac{1}{p}} < \infty \implies \exists r \in \mathbb{R} \text{ such that } \left(\int |f|^p \, d\mu\right)^{\frac{1}{p}} &= r < \infty \\ \implies \int |f|^p \, d\mu = r^p &< \infty. \end{aligned}$$

$\square$

Proposition 7.1.2 (p-Norm for Counting Measure; ℓ^p)**.** *Suppose X is a set and $p \in (0, \infty)$.Let $\mu \colon \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ be the counting measure on X.*

(i) $\mathcal{L}^p(\mu)$, which we denote more succinctly by ℓ^p, is given by

$$\ell^p = \left\{ (a_1, a_2, \dots) : a_1, a_2, \dots \in \mathbb{F} \text{ and } \sum_{k \in \mathbb{N}} |a_k|^p < \infty \right\}.$$

(ii) $\mathcal{L}^\infty(\mu)$, which we denote more succinctly by ℓ^∞, is given by

$$\ell^\infty = \left\{ (a_1, a_2, \dots) : a_1, a_2, \dots \in \mathbb{F} \text{ and } \sup_{k \in \mathbb{N}} |a_k| < \infty \right\}.$$

Proof. **[INCOMPLETE]**

$\square$

Lemma 7.1.2 (Bounding some p-Norm; Homogeneity of p-Norm)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. Then, $f, g \in \mathcal{L}^p(\mu)$ and $a \in \mathbb{F}$ implies*

(i) $||f + g||_p{}^p \leq 2^p \left[||f||_p{}^p + ||g||_p{}^p \right]$, and

(ii) $||af||_p = |a| \cdot ||f||_p$.

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. Fix $f, g \in \mathcal{L}^p(\mu)$ and $a \in \mathbb{F}$. By definition, $f, g \colon X \to \mathbb{F}$, implying that $|f|, |g| : X \to \mathbb{R}$.

(i) We first observe that the map $\phi \colon \mathbb{R}_{\geq 0} \to \mathbb{R}$, defined by $\phi(x) = x^p \ \forall x \in \mathbb{R}_{\geq 0}$, is increasing. To see this, we note that $\phi'(x) = px^{p-1} > 0 \ \forall x \in \mathbb{R}_{\geq 0}$. It follows, from the Triangle inequality, that

$$\begin{aligned} |f(x) + g(x)| \leq |f(x)| + |g(x)| \implies |f(x) + g(x)|^p &\leq (|f(x)| + |g(x)|)^p \\ &\leq (2 \max \{|f(x)|, |g(x)|\})^p = 2^p \cdot \max \{|f(x)|, |g(x)|\}^p \\ &\leq 2^p \cdot (|f(x)|^p + |g(x)|^p). \end{aligned}$$

We note that $|f + g|$ and the function given by $2^p \cdot (|f(x)|^p + |g(x)|^p)$ are both $\mathcal{S}$-measurable by regularity, with nonnegative codomains. By Theorem 3.8 (Axler 77), we yield that

$$\begin{aligned} \int |f + g| \, d\mu &\leq \int 2^p \, |f|^p + |g|^p \, d\mu \\ &= 2^p \int |f|^p + |g|^p \, d\mu && \text{(By Theorem 3.20 (Axler 80))} \\ &= 2^p \cdot \left(\int |f|^p \, d\mu + \int |g|^p \, d\mu \right), && \text{(By Theorem 3.16 (Axler 80))} \end{aligned}$$

where we may apply Theorem 3.16 since $|f|^p$ and $|g|^p$ are $\mathcal{S}$-measurable, by regularity, with nonnegative codomain. As an immediate result,

$$\begin{aligned} \left(\left[\int |f + g| \, d\mu \right]^{\frac{1}{p}} \right)^p &\leq 2^p \cdot \left(\left[\left(\int |f|^p \, d\mu \right)^{\frac{1}{p}} \right]^p + \left[\left(\int |g|^p \, d\mu \right)^{\frac{1}{p}} \right]^p \right), \\ \therefore ||f + g||_p^p &\leq 2^p \cdot \left(||f||_p^p + ||g||_p^p \right). \end{aligned}$$

(ii) We observe that $|a| \in \mathbb{R}$, $|a|^p |f|^p$ is $\mathcal{S}$ measurable by regularity, and $\int |f|^p \, d\mu$ is defined since $|f|^p$ is nonnegative. Thus, we may apply Theorem 3.20 (Axler 80). By definition, we obtain that

$$\begin{aligned} ||af||_p = \left[\int |af|^p \, d\mu\right]^{\frac{1}{p}} &= \left[\int |a|^p |f|^p \, d\mu\right]^{\frac{1}{p}} \\ &= \left[|a|^p \int |f|^p \, d\mu\right]^{\frac{1}{p}} && \text{(By Theorem 3.20)} \\ &= |a| \, ||f||_p \, . && \text{(By definition)} \end{aligned}$$

□

Theorem 7.1.2 ($\mathcal{L}^p(\mu)$ is a vector space)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. Then, $(\mathcal{L}^p(\mu), +, \cdot)$ is a vector space over field $\mathbb{F}$, where $+$ and $\cdot$ denote the usual pointwise function addition and pointwise scalar-function multiplication, respectively.*

Proof. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. Consider $(\mathcal{L}^p(\mu), +, \cdot)$ over field $\mathbb{F}$, where $+$ and $\cdot$ denote the usual pointwise function addition and pointwise scalar-function multiplication, respectively.

Fix $f, g \in \mathcal{L}^p(\mu)$ and $a \in \mathbb{F}$. We first note that $|a| \in \mathbb{F}$. To prove that $\mathcal{L}^p(\mu)$ with the stated operations is a vector space, it suffices to show that $\mathcal{L}^p(\mu)$ is a subspace of the vector space $\mathbb{F}^X = \{f : f\colon X \to \mathbb{F}\}$. Hence, it remains to show that $\mathcal{L}^p(\mu)$ is closed under the two operations.

We observe, by the above proposition, that $||f + g||_p \leq 2^p(||f||^p + ||g||^p) < \infty$ since $f, g \in \mathcal{L}^p(\mu) \iff ||f||_p \, , ||g||_p < \infty$ by definition.

Moreover, by the above proposition, $||af||_p = |a| \cdot ||f||_p < \infty$, since $||f||_p < \infty$ and $|a| \in \mathbb{R}$. Thus, $f + g, af \in \mathcal{L}^p(\mu)$ by definition.

Indeed, $\mathcal{L}^p(\mu)$ is closed under the stated operations. Thus, we conclude $\mathcal{L}^p(\mu)$ is a subspace and, therefore, a vector space by definition. □

Definition 7.1.3 (Dual exponent)**.** *Suppose $p \in [1, \infty]$. Then, the dual exponent of p is $p' \in [1, \infty]$ such that*

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Theorem 7.1.3 (Young's Inequality)**.** *Suppose $p \in (1, \infty)$. Then, $a, b \geq 0$ implies that*

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

Proof. **[INCOMPLETE]**

□

Theorem 7.1.4 (Holder's Inequality)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $p \in [1, \infty]$, and $f, h\colon X \to \mathbb{F}$ are $\mathcal{S}$-measurable. Then,*

$$||fh||_1 \leq ||f||_p \, ||h||_p \, .$$

Proof. **[INCOMPLETE]**

□

Theorem 7.1.5 (Inequality Involving two p-Norms; $\mathcal{L}^p$ Spaces are Decreasing)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space with* $\mu(X) < \infty$ *and* $0 < p < q < \infty$*.*

(i) $f \in \mathcal{L}^q(\mu)$ *implies that*

$$||f||_p \leq \mu(X)^{\frac{q-p}{pq}} \cdot ||f||_q \,.$$

(ii) $\mathcal{L}^q(\mu) \subset \mathcal{L}^p(\mu)$*.*

Proof. **[INCOMPLETE]**

□

Theorem 7.1.6 (p-Norm as a Supremum of the Absolute Values of Integrals)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space,* $p \in [1, \infty)$*, and* $f \in \mathcal{L}^p(\mu)$*. Then,*

$$||f||_p = \sup\left\{\left|\int fhd\mu\right| : h \in \mathcal{L}^{p'}(\mu) \text{ and } ||h||_{p'} \leq 1\right\}.$$

Proof. **[INCOMPLETE]**

□

Theorem 7.1.7 (Minkowski's Inequality)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space,* $p \in [1, \infty]$*, and* $f, g \in \mathcal{L}^p(\mu)$*. Then,*

$$||f + g||_p \leq ||f||_p + ||g||_p \,.$$

Proof. **[INCOMPLETE]**

□

7.2 $L^p(\mu)$

Definition 7.2.1 ($\mathcal{Z}(\mu); \tilde{f}$)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $p \in (0, \infty]$. Then,*

(i) $\mathcal{Z}(\mu)$ is defined as

$$\mathcal{Z}(\mu) = \left\{ f\colon X \to \mathbb{F} : f \text{ is } \mathcal{S}\text{-measurable and identically } 0 \text{ almost everywhere} \right\}.$$

(ii) Suppose $f \in \mathcal{L}^p(\mu)$. Then, $\tilde{f} \subset \mathcal{L}^p(\mu)$ is defined as

$$\tilde{f} = \{f + z : z \in \mathcal{Z}(\mu)\}.$$

Proposition 7.2.1. *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $p \in (0, \infty]$. Then, $\mathcal{Z}(\mu)$ is a subspace of $\mathcal{L}^p(\mu)$.*

Definition 7.2.2 ($L^p(\mu)$)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $p \in (0, \infty]$.*

(i) $L^p(\mu)$ is defined by

$$L^p(\mu) = \left\{ \tilde{f} : f \in \mathcal{L}^p(\mu) \right\}.$$

(ii) Suppose $\tilde{f}, \tilde{g} \in L^p(\mu)$ and $a \in \mathbb{F}$. Then, $\tilde{f} + \tilde{g}$ and $a\tilde{f}$ are defined by

$$\tilde{f} + \tilde{g} = (f+g)^{\sim} and a\tilde{f} = (af)^{\sim}.$$

Definition 7.2.3 ($||\cdot||_p$ on $L^p(\mu)$)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $p \in (0, \infty]$. Then, $||\cdot||_p : L^p(\mu) \to \mathbb{R}$ on $L^p(\mu)$ is defined as*

$$\left|\left|\tilde{f}\right|\right|_p = ||f||_p, \; \forall f \in \mathcal{L}^p(\mu).$$

Theorem 7.2.1 ($L^p(\mu)$ is a Normed Vector Space)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $p \in [1, \infty]$. Then,*

(i) $||\cdot||_p$ is a norm on $L^p(\mu)$, and

(ii) $(L^p(\mu), +, \cdot)$ is a vector space, where $+$ and $\cdot$ denote the addition and scalar-function multiplication on $L^p(\mu)$, respectively.

Definition 7.2.4 ($L^p(E)$ for $E \subset \mathbb{R}$)**.** *Suppose $E \in \mathcal{B}$ (or $E \in \mathcal{L}$) and $p \in (0, \infty]$. Then, we denote $L^p(\lambda_E)$ by $L^p(E)$, where λ_E is the Lebesgue measure restricted to $\{B \in \mathcal{B} : B \in E\}$ (or $\{B \in \mathcal{L} : B \in E\}$).*

Theorem 7.2.2 ($\mathcal{L}^p(\mu)$ is Complete)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $p \in [1, \infty]$. Suppose $(f_k)_{k \in \mathbb{N}}$ is a sequence of functions in $\mathcal{L}^p(\mu)$ such that*

$$\forall \epsilon > 0, \; \exists n \in \mathbb{N} \text{ such that } j, k \in \mathbb{N} \text{ and } j, k \geq n \implies ||f_j - f_k||_p < \epsilon.$$

Then, there exists $f \in \mathcal{L}^p(\mu)$ such that

$$\lim_{k \to \infty} ||f_k - f||_p = 0.$$

To put simply and informally, Cauchy sequences in $\mathcal{L}^p(\mu)$ converges to an element in $\mathcal{L}^p(\mu)$.

Theorem 7.2.3 (Convergent Sequence in $\mathcal{L}^p(\mu)$ have Pointwise Convergent Subsequences)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $p \in [1, \infty]$. Suppose $f \in \mathcal{L}^p(\mu)$ and $(f_k)_{k \in \mathbb{N}}$ is a sequence of functions in $\mathcal{L}^p(\mu)$ such that $\lim_{k \to \infty} ||f_k - f||_p = 0$. Then, there exists a subsequence $(f_{(k_m)})_{m \in \mathbb{N}}$ such that*

$$\lim_{m \to \infty} f_{(k_m)}(x) = f(x)$$

for almost every $x \in X$.

Theorem 7.2.4 (L^P Space is a Banach Space)**.** *Suppose* $(X, \mathcal{S}, \mu)$ *is a measure space and* $p \in [1, \infty]$*. Then,* $L^p(\mu)$ *is a Banach space.*

Proof. $\square$

8 Hilbert Spaces

8.1 Inner Product Spaces

Definition 8.1.1 (Inner product; inner product space)**.** *Suppose that $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and V is a vector space over field $\mathbb{F}$.*

(a) The map $\langle\rangle : V \times V \to \mathbb{F}$, defined by $(u, v) \mapsto \langle u, v\rangle \in \mathbb{F}\ \forall (u, v) \in V \times V$, is an inner product on V if $\langle\rangle$ satisfies the following properties:

(Nonnegativity) $\langle v, v\rangle \geq 0\ \forall v \in V$;

(Positive-Definiteness) $\langle v, v\rangle = 0 \iff v = 0_V$;

(Linearity in the First Slot) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle\ \ \forall u, v, w \in V$

(Homogeneity in the First Slot) $\langle \lambda u, v\rangle = \lambda \langle u, v\rangle\ \ \forall \lambda \in F, u, v \in V$;

(Conjugate Symmetry) $\langle u, v\rangle = \overline{\langle u, v\rangle}\ \forall u, v \in V$.

(b) $(V, \langle\rangle)$ is an inner product space if $\langle\rangle$ is an inner product on V.

Proposition 8.1.1 (Inner Product Space Menu Theorem)**.**

Theorem 8.1.1 (Properties of Inner Product)**.** *Suppose V is an inner product space over field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ with an inner product $\langle\rangle : V \times V \to \mathbb{F}$. Then, the following statements hold:*

(1) $\forall u \in V$, $\phi_u \colon V \to \mathbb{F}$ defined by $\phi_u(v) = \langle v, u\rangle\ \ \forall v \in V$ is a linear map from V to $\mathbb{F}$.

(2) $\langle 0_V, u\rangle = 0$ and $\langle u, 0_V\rangle = 0\ \forall u \in V$.

(3) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle\ \ \forall u, v, w \in V$.

(4) $\langle u, kv\rangle = \overline{k}\langle u, v\rangle\ \ \forall k \in F$ and $u, v \in V$.

Proof. Suppose V is an inner product space over field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ with an inner product $\langle\rangle : V \times V \to \mathbb{F}$.

(1) Fix $u \in V$. Define $\phi_u : V \times V \to \mathbb{F}$ by $\phi_u(v) = \langle v, u\rangle\ \ \forall v \in V$. Fix $v, w \in V$ and $k \in \mathbb{F}$. It follows that

$$\begin{aligned}
\phi_u(v + w) &= \langle v + w, u\rangle = \langle v, u\rangle + \langle w, u\rangle && \text{(By the additivity of } \langle\rangle\text{)}\\
&= \phi_u(v) + \phi_u(w).\\
\phi_u(kv) &= \langle kv, u\rangle = k\langle v, u\rangle && \text{(By the homogeneity of } \langle\rangle\text{)}\\
&= k\phi_u(v).
\end{aligned}$$

By definition, ϕ_u is a linear map.

(2) Fix $u \in V$ and define ϕ_u as in part (1). It follows that

$$\begin{aligned}
\phi_u(0_V) &= \langle 0_V, u\rangle = 0 && \text{(Since } \phi_u \text{ is a linear map)}\\
&\implies \overline{\langle 0_V, u\rangle} = \overline{0}\\
&\implies \langle u, 0_V\rangle = 0. && \text{(By the conjugate symmetry of } \langle\rangle\text{)}
\end{aligned}$$

Thus, we obtain $\langle 0_V, u\rangle = 0 = \langle u, 0_V\rangle\ \ \forall u \in V$.

(3) Fix $u, v, w \in V$. Then, we have that

$$\begin{aligned}
\langle u, v+w\rangle &= \overline{\langle v+w, u\rangle} && \text{(By the conjugate symmetry of } \langle\rangle\text{)} \\
&= \overline{\langle v, u\rangle + \langle w, u\rangle} && \text{(By the addivity of } \langle\rangle\text{)} \\
&= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} && \text{(By the property of conjugate)} \\
&= \langle u, v\rangle + \langle u, w\rangle\,. && \text{(By the conjugate symmetry of } \langle\rangle\text{)}
\end{aligned}$$

(4) Fix $u, v \in V$ and $k \in \mathbb{F}$. It follows that

$$\begin{aligned}
\langle u, kv\rangle &= \overline{\langle kv, u\rangle} && \text{(By the conjugate symmetry of } \langle\rangle\text{)} \\
&= \overline{k\,\langle v, u\rangle} && \text{(By the homogeneity of } \langle\rangle\text{)} \\
&= \overline{k} \cdot \overline{\langle v, u\rangle} && \text{(By the property of conjugate)} \\
&= \overline{k}\,\langle u, v\rangle\,. && \text{(By the conjugate symmetry of } \langle\rangle\text{)}
\end{aligned}$$

$\square$

Proposition 8.1.2 (Inner Product Menu Theorem)**.**

Proposition 8.1.3 (Norm associated with an inner product)**.** *Suppose* $(V, \langle\rangle)$ *is an inner product space. Then,* $||\cdot|| : V \to \mathbb{R}$ *define by* $||f|| = \sqrt{\langle f, f\rangle}\ \forall f \in V$ *is a norm on* V *associated with* $\langle\rangle$*.*

Theorem 8.1.2 (Homogeneity of a Norm)**.**

Theorem 8.1.3 (Pythagorean Theorem)**.**

Theorem 8.1.4 (Cauchy-Schwartz Inequality)**.**

8.2 Orthogonality

Definition 8.2.1 (Hilbert space)**.** *An inner product space $(V, \langle\rangle)$ is a Hilbert space if V equipped with the $\langle\rangle$-induced norm is a Banach space.*

That is, an inner product space is a Hilbert space if it is complete with respect to the norm induced by its inner product.

Proposition 8.2.1 (Hilbert Space Menu Theorem)**.** *(i) Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Then, $L^2(\mu)$ with its usual inner product $\langle\rangle$, defined by $\langle f, g\rangle = \int f\overline{g}d\mu\ \forall (f,g) \in X \times X$, is a Hilbert space.*

(ii) Counting measure and $\mathbb{F}^n$

(iii) ℓ^2

(iv) Every closed subspace of a Hilbert space is a Hilbert space.

Remark 8.2.1 (Metric Topology in a Hilbert Space)**.** *Suppose V is a Hilbert space. To discuss the notions of limits, convergence of sequences, and metric topology in V, we consider, unless otherwise specified, the metric on V induced by the norm on V induced by the inner product of V.*

Definition 8.2.2 (Distance from a point to a set)**.** *Suppose $(V, ||\cdot||)$ is a normed space, $\mathcal{U} \subset V$ is nonempty, and $f \in V$. Then, the distance from f to $\mathcal{U}$ is defined by*

$$distance(f,\mathcal{U}) = \inf\{||f-g|| : g \in \mathcal{U}\}.$$

Definition 8.2.3 (Convex Set (in a vector space))**.** *Suppose V is a vector space and $\mathcal{U} \subset V$. Then, $\mathcal{U}$ is convex if*

$$f, g \in \mathcal{U} \implies \forall t \in [0,1],\ (1-t)f + tg \in \mathcal{U}.$$

Proposition 8.2.2 (Convexity of Vector Spaces and Their Subspaces)**.** *Suppose V is a vector space over field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $\mathcal{U} \subset V$ is a subspace of V. Then, V and $\mathcal{U}$ are convex.*

Proof. Suppose V is a vector space over field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Fix a subspace $\mathcal{U} \subset V$ of V. By definition, $\mathcal{U}$ is also a vector space. Therefore, $\mathcal{U}$ is closed under vector addition and scalar-vector multiplication.

Fix $f, g \in V$ and then fix $t \in [0,1]$. It follows that $(1-t)f$ and tg live in $\mathcal{U}$. As an immediate result, $(1-t)f + tg \in \mathcal{U}$. By definition, $\mathcal{U}$ is indeed convex.

Thus, we proved that any arbitrary subspace of V is convex. We note that V is also a subspace of V. Thus, V is also convex. ☐

Proposition 8.2.3 (Convexity of Open Ball in a Normed Vector Space)**.**

Theorem 8.2.1 (Distance to a Closed Convex Set is Attained in a Hilbert Space)**.** *Suppose V is a Hilbert space, $f \in V$, and $\mathcal{U} \subset V$ is nonempty, closed, and convex. Then, $\exists! g \in \mathcal{U}$ such that $||f-g|| = distance(f,\mathcal{U})$.*

Proof. Suppose V is a Hilbert space, $f \in V$, and $\mathcal{U} \subset V$ is nonempty, closed, and convex.

By definition, $distance(f,\mathcal{U}) = \inf\{||f-G|| : G \in \mathcal{U}\}$. By an equivalent definition of infimum, $\forall \epsilon > 0\ \exists g_\epsilon \in \mathcal{U}$ such that $distance(f,\mathcal{U}) + \epsilon > ||f - g_\epsilon||$. In particular,

$$\forall k \in \mathbb{N},\ \exists g_k \in \mathcal{U} \text{ such that } 0 \leq ||f-g_k|| - distance(f,\mathcal{U}) < \frac{1}{k}.$$

Thus, $k \to \infty$ implies that $||f - g_k|| - distance(f, \mathcal{U}) \to 0$. That is, we yield

$$\lim_{k\to\infty} ||f - g_k|| = distance(f, \mathcal{U}). \tag{$*_0$}$$

Here, we claim that $(g_k)_{k\in\mathbb{N}}$ is a Cauchy sequence of elements in $\mathcal{U}$. Fix $i, j \in \mathbb{N}$. We observe that

$$\begin{aligned}
0 \le ||g_i - g_j||^2 &= ||(f - g_j) - (f - g_i)||^2 \\
&= 2\,||(f - g_j)||^2 + 2\,||(f - g_i)||^2 - ||(f - g_j) + (f - g_i)||^2 && \text{(By Theorem 8.20 (Axler 220))} \\
&= 2\,||(f - g_j)||^2 + 2\,||(f - g_i)||^2 - 4\left|\left|f - \frac{g_i + g_j}{2}\right|\right|^2 && \text{(By homogeneity of inner product)} \\
&\le 2\,||(f - g_j)||^2 + 2\,||(f - g_i)||^2 - 4distance(f, \mathcal{U}), && (*_1)
\end{aligned}$$

where the last inequality holds by definition as $\mathcal{U}$ is convex implies that $\frac{g_i + g_j}{2} \in \mathcal{U}$. By $(*_0)$, $i \to \infty$ and $j \to \infty$ implies that $||g_i - g_j||^2 \to 0$. Thus, we have that

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } i, j > N \implies ||g_i - g_j||^2 < \epsilon.$$

Fix $\epsilon > 0$. Then, $\epsilon^2 > 0$ and we have that

$$\exists N \in \mathbb{N} \text{ such that } i, j > N \implies ||g_i - g_j||^2 < \epsilon^2 \implies ||g_i - g_j|| < \epsilon,$$

since $||g_i - g_j|| \ge 0$. Hence, we conclude that $(g_k)_{k\in\mathbb{N}}$ is indeed a Cauchy sequence by definition.

Since V is a Hilbert space by hypothesis, V is a Banach space and, therefore, complete with respect to the metric induced by its norm. Thus, $(g_k)_{k\in\mathbb{N}}$ converges. Moreover, $\exists g \in \mathcal{U}$ such that $\lim_{k\to\infty} g_k = g$ since $\mathcal{U}$ is closed by hypothesis.

Fix $\epsilon > 0$. By definition, $\exists N \in \mathbb{N}$ such that $k > N \implies ||g_k - g|| < \epsilon$. Thus, $k > N$ implies that

$$\begin{aligned}
&||(f - g) - (f - g_k)|| = ||g_k - g|| < \epsilon \\
\implies &\; |||f - g|| - ||f - g_k||| \le ||(f - g) - (f - g_k)|| < \epsilon && \text{(By Triangle Inequality)} \\
\implies &\; ||f - g_k|| - ||f - g|| < \epsilon.
\end{aligned}$$

That is, we showed that $\forall \epsilon > 0$, $\exists G \in \mathcal{U}$ such that $||f - g|| + \epsilon > ||f - G||$. By an equivalent definition of infimum, $||f - g|| = \inf \{||f - G|| : G \in \mathcal{U}\} = distance(f, \mathcal{U})$ as desired.

To prove the uniqueness of g, we first suppose that $\exists h \in \mathcal{U}$ possessing the same properties as g. Thus, we have that $h = distance(f, \mathcal{U})$. We note that the inequality $(*_1)$ holds for arbitrary $G \in \mathcal{U}$. It follows that

$$\begin{aligned}
0 =\le ||g - h|| &\le 2\,||f - g||^2 + 2\,||f - h||^2 - 4distance(f, \mathcal{U})^2 \\
&= 2distance(f, \mathcal{U})^2 + 2distance(f, \mathcal{U})^2 - 4distance(f, \mathcal{U})^2 = 0.
\end{aligned}$$

Thus, $||g - h|| = 0$ and, therefore, $||g - h||^2 = \langle g - h, g - h\rangle = 0$. By the positive definiteness of an inner product, $g - h = 0_V$, the zero vector of V. We conclude that $g = h$ as desired, proving the uniqueness of g. □

Remark 8.2.2. *The above theorem need not hold for a non-Hilbert Banach space.*

Remark 8.2.3. *Not all subspaces are closed.*

Definition 8.2.4 (Orthonormal projection)**.** *Suppose V is a Hilbert space and $\mathcal{U} \subset V$ is nonempty, closed, and convex. Then, the orthogonal projection of V onto $\mathcal{U}$ is the map $P_{\mathcal{U}} : V \to V$ defined by $P_{\mathcal{U}}(f) = g$, where $g \in \mathcal{U}$ is the unique element such that $||f - g|| = distance(f, \mathcal{U})$.*

We may denote $P_{\mathcal{U}}(f)$, more succinctly, by $P_{\mathcal{U}} f$.

Proposition 8.2.4 (Properties of Orthogonal Projection)**.** *Suppose V is a Hilbert space and $\mathcal{U} \subset V$ is nonempty, closed, and convex. Then,*

(i) $\forall f \in V$, $P_{\mathcal{U}} f = f \iff f \in \mathcal{U}$, *and*

(ii) $P_{\mathcal{U}} \circ P_{\mathcal{U}} = P_{\mathcal{U}}$.

Proof. Suppose V is a Hilbert space and $\mathcal{U} \subset V$ is nonempty, closed, and convex.

(i) Fix $f \in V$.

($\Longrightarrow$): Suppose that $P_{\mathcal{U}} f = f$. By definition, $P_{\mathcal{U}} f$ is the unique element in $\mathcal{U}$ such that $P_{\mathcal{U}} f = distance(f,\mathcal{U})$. It follows trivially that $f \in \mathcal{U}$.

($\Longleftarrow$): Suppose that $f \in \mathcal{U}$. We observe that

$$distance(f,\mathcal{U}) = \inf \{||f - G|| : G \in \mathcal{U}\} \leq ||f - f|| = 0.$$

In addition, the values of a norm are nonnegative implies that $distance(f,\mathcal{U})$ is also nonnegative. Thus, we must have that $distance(f,\mathcal{U}) = 0 = ||f - f||$. By Theorem 8.28, f is the unique element such that the above equality holds. By definition, $P_{\mathcal{U}} f = f$.

(ii) Fix $f \in V$. It follows that $P_{\mathcal{U}}(f) \in \mathcal{U}$ by definition. Thus, $[P_{\mathcal{U}} \circ P\mathcal{U}](f) = P_{\mathcal{U}}(P_{\mathcal{U}}(f)) = P_{\mathcal{U}}(f)$ by Part (i). Thus, $P_{\mathcal{U}} \circ P_{\mathcal{U}} = P_{\mathcal{U}}$. □

Theorem 8.2.2 (Orthogonal Projection Onto Closed Subspace)**.** *Suppose V is a Hilbert space, $\mathcal{U} \subset V$ is a closed subspace of V, and $f \in V$. Then, the following statements hold.*

(i) $\forall g \in \mathcal{U}$, $f - P_{\mathcal{U}} f$ *is orthogonal to* g.

(ii) Suppose $h \in \mathcal{U}$ *and* $f - h$ *is orthogonal to* g, $\forall g \in \mathcal{U}$. *Then,* $h = P_{\mathcal{U}} f$.

(iii) $P_{\mathcal{U}} : V \to V$ *is a linear map.*

(iv) $||P_{\mathcal{U}} f|| \leq ||f||$, *where* $||P_{\mathcal{U}} f|| = ||f|| \iff f \in \mathcal{U}$.

Definition 8.2.5 (Orthogonal complement)**.**

Theorem 8.2.3 (Properties of Orthogonal complement)**.**

Theorem 8.2.4 (Orthogonal complement of the Orthogonal complement)**.**

Theorem 8.2.5 (Necessary and Sufficient Considtion for Denseness of Subspace)**.**

Theorem 8.2.6 (Orthogonal Decomposition in Closed Hilbert Subspace)**.**

Theorem 8.2.7.

Theorem 8.2.8 (Riesz Representation Theorem)**.**

8.3 Orthonormal Basis

Definition 8.3.1 (Orthonormal family)**.**

Theorem 8.3.1 (Finite Orthonormal Families)**.**

Definition 8.3.2 (Unordered sum)**.**

9 Real and Complex Measure

10 Linear Maps on Hilbert Spaces

11 Fourier Analysis

12 Probability Measures

Definition 12.0.1 (Probability space)**.**

Theorem 12.0.1.

Definition 12.0.2 (Random variable)**.**

Useful Propositions From Elementary Analysis

Proposition 12.0.1 (Open Subset of $\mathbb{R}$ is a Countable Union of Open Intervals)**.** *Suppose $U \subset \mathbb{R}$ is open. Then, $\exists I_1, I_2, \dots \subset \mathbb{R}$ such that $U = \bigcup_{k\in\mathbb{N}} I_k$, where $I_k \subset \mathbb{R}$ is an open interval $\forall k \in \mathbb{N}$.*

Proposition 12.0.2 (Equivalent Condition for Uniform Convergence)**.** *Suppose X is a set , $f\colon X \to \mathbb{C}$, and $(f_n)_{n\in\mathbb{N}}$ is a sequence of functions defined from X to $\mathbb{C}$. Then,*

$$f_n \to f \text{ uniformly on } S \subset X \iff \lim_{n\to\infty} \sup_{x\in S} |f_n(x) - f(x)| = 0.$$

Proof. Suppose X is a set, $f\colon X \to \mathbb{C}$, and $(f_n)_{n\in\mathbb{N}}$ is a sequence of functions defined from X to $\mathbb{C}$. Let $S \subset X$

($\Longrightarrow$) Suppose that $f_n \to f$ uniformly on S. By definition, we have that

$$\forall \epsilon > 0 \; \exists N > 0 \text{ such that } n > N \text{ and } x \in S \implies |f_n(x) - f(x)| < \epsilon$$

which yields that, $\forall \epsilon > 0$, $n > N$ implies ϵ is an upper bound for $\{|f_n(x) - f(x)| : x \in S\}$ and, hence,

$$\forall \epsilon > 0 \; \exists N > 0 \text{ such that } n > N \implies \sup_{x\in S} |f_n(x) - f(x)| \le \epsilon.$$

Fix $\epsilon > 0$. Then, $\frac{\epsilon}{2} > 0$ and, by the above statement, we have that

$$\exists N > 0 \text{ such that } n > N \implies \sup_{z\in S} |f_n(x) - f(x)| = \left| \sup_{x\in S} |f_n(x) - f(x)| - 0 \right| \le \frac{\epsilon}{2} < \epsilon.$$

By definition, $\lim_{n\to\infty} \sup_{x\in S} |f_n(x) - f(x)| = 0$.

($\Longleftarrow$) Suppose that $\lim_{n\to\infty} \sup_{x\in S} |f_n(x) - f(x)| = 0$. By definition, we have that

$$\begin{aligned}\forall \epsilon > 0 \; \exists N > 0 \text{ such that } n > N &\implies \sup_{x\in S} |f_n(x) - f(x)| < \epsilon \\ &\implies n > N \text{ and } x \in S \implies |f_n(x) - f(x)| < \epsilon. \qquad \text{(By the definition of sup)}\end{aligned}$$

By definition, $f_n \to f$ uniformly on S as desired. □

Proposition 12.0.3 (Equivalent Condition for Supremum and Infimum)**.** *Suppose $S \subset \mathbb{R}$ is nonempty.*

(i) Suppose further that S is bounded above. Then, an upper bound U of S equals $\sup S$ $\iff$

$$\forall \epsilon > 0 \; \exists x \in S \text{ such that } U - \epsilon < x.$$

(ii) Suppose further that S is bounded below. Then, a lower bound L of S equals $\inf S$ $\iff$

$$\forall \epsilon > 0 \; \exists x \in S \text{ such that } x < L + \epsilon.$$

Proposition 12.0.4. *Suppose $(s_n)_{n\in\mathbb{N}}$ is a sequence of real numbers.*

(i) Suppose $(s_n)_{n\in\mathbb{N}}$ converges. Then,

$$\inf_{k\in\mathbb{N}} s_k = \lim_{k\to\infty} s_k = \sup_{k\in\mathbb{N}} s_k.$$

(ii) Suppose $\sup_{k\in\mathbb{N}} s_k$ *and* $\inf_{k\in\mathbb{N}} s_k$ *exist in $\mathbb{R}$ and* $\sup_{k\in\mathbb{N}} s_k = \inf_{k\in\mathbb{N}} s_k$. *Then, $\forall k \in \mathbb{N}$* $s_k = \sup_{k\in\mathbb{N}} s_k = \inf_{k\in\mathbb{N}} s_k$.