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

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

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 $\{e_1, \dots, e_n\}$ 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. □

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) □

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) □

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) □

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, □

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

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

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

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

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, $\forall j \in \{1, \ldots, n\}$,

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

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

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

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

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

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

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 $f \colon U \to \mathbb{R}$.

($\implies$) 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

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

($\impliedby$) Suppose f is differentiable at $x \in U$. Then, ... $\square$

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)

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

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. $\square$

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*

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

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

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

(Area) Let $\Delta x_i = x_i - x_{i-1}$ *and* $\Delta y_j = y_j - y_{j-1}$*,* $\forall (i,j) \in \{1,\ldots,m\} \times \{1,\ldots,n\}$*. Then,* $\forall (i,j) \in \{1,\ldots,m\} \times \{1,\ldots,n\}$*, 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* $R_{i,j}$*,* $\forall (i,j) \in \{1,\ldots,m\} \times \{1,\ldots,n\}$ *and let* $S = \{(s_{i,j}, t_{i,j}) : (i,j) \in \{1,\ldots,m\} \times \{1,\ldots,n\}\}$*. 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

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

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

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

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 $S = \{x_1, \dots, x_n\}$ for some $n \in \mathbb{N}$. It follows that $S = \cup_{k=1}^{n} \{x_k\}$. We observe that $\forall k \in \{1, \dots, n\}$, $\{x_k\}$ 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. $\square$

Theorem 1.7.3 (One-Dimensional Riemann-Lebesgue Theorem). *Suppose that $[a, b] \subset \mathbb{R}$. Then, $f : [a, b] \to \mathbb{R}$ 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, $f : R \to \mathbb{R}$ 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 $f : R \to \mathbb{R}$ be a bounded function. Define, $\forall y \in [c, d]$, $f_y : [a, b] \to \mathbb{R}$ by*

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

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

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

(ii) The upper slice integral of f is the map $\overline{F} : [c, d] \to \mathbb{R}$ 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 $f : R \to \mathbb{R}$ 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 $f : R \to \mathbb{R}$ 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 $\chi_S : M \to \mathbb{R}$ 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 $\chi_S \colon R \to \mathbb{R}$ 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. $\square$

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 $f \colon U \to \mathbb{R}^m$ 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 $Jac(f) \colon U \to \mathbb{R}$ 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 $h \colon Z \to \mathbb{R}^2$ 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 $\varphi \colon U \to W$ is a C^1 diffeomorphism and $f \colon W \to \mathbb{R}$ 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. $\square$

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) $A_1, A_2, \cdots \in \mathcal{A} \implies \bigcup_{k\in\mathbb{N}} A_k \in \mathcal{A}$ ***(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) $A_1, A_2, \cdots \in \mathcal{A} \implies \bigcap_{k\in\mathbb{N}} A_k \in \mathcal{A}$.

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. □

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 $\mathcal{S} = \{S \in \mathcal{P}(X) : S$ is countable or X/S 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 $S_1, S_2, \dots \in \mathcal{S}$.

(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 $\exists N_1, N_2, \dots \in \mathbb{N}$ such that $S_{N_1}, S_{N_2}, \dots$ are uncountable. It follows that $X/S_{N_1}, X/S_{N_2}, \dots$ must be countable since $S_{N_1}, S_{N_2}, \dots \in \mathcal{S}$ 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. □

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 $A_1, A_2, \ldots, \subset \mathcal{A}$. Then, $A_1, A_2, \ldots$ 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. $\square$

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*

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

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 $\mathcal{A}$. $\square$

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 $f \colon \mathbb{R} \to \mathbb{R}$. Then,

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

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 $f : \mathbb{R} \to \mathbb{R}$. 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 $f : X \to Y$. 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 $f : X \to Y$ 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 $f : X \to Y$ 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 $2_\mathbb{R} : \mathbb{R} \to \mathbb{R}$ 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 $f : X \to Y$. 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 $f : X \to Y$, $g : Y \to W$. Then, the following identities hold:*

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

(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, $f^{-1}(\cup_{A\in\mathcal{A}} A) = \cup_{A\in\mathcal{A}} f^{-1}(A)$.

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

(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 $(g \circ f)^{-1}(A) = f^{-1}\left(g^{-1}(A)\right)$. □

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

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

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. □

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 (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 2.2.9 (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

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

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

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

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

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. □

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 $\forall p \in X$

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

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

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

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 $x < y$, $f(x) \le f(y)$;
(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]** $\square$

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 2.2.10 ($\mathcal{S}$-Measurable Function Arose from Function Composition)**.** *Suppose (X,S) is a measurable space, $k \in \mathbb{R}$, and $f\colon X \to \mathbb{R}$ is $\mathcal{S}$-Measurable. Then, kf, $|f|$, and f^2 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) Define $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) = x^2 \ \forall x \in \mathbb{R}$. Apply the argument in Part (i) to $f^2\colon X \to \mathbb{R}$, 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 2.2.11 (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.

($\Longleftarrow$) 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 2.2.12 (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. **[INCOMPLETE]** □

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)**.** *(a)* ***(Counting Measure)***

(b) ***Dirac Measure***

(c) ***Sum Measure***

(d)

(e)

(f)

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 $(X, \mathcal{S})$.

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) $\mu(D) \le \mu(E)$.
(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

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

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

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

□

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

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

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

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

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

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,

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

□

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]**

□

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 $|A \sqcup F| = |A| + |F|$. □

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)*

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

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

$$|D/F| = \frac{\epsilon}{2}. \qquad (*_1)$$

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 $|G| \leq \sum_{k\in\mathbb{N}} l(I_k)$ 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

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

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

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

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 $|D_k| \leq |D| < \infty$ 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 $E_1, E_2, \cdots \in \mathcal{L}$. 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)

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

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* $|A/B| = 0$.

Proposition 2.4.1 (Borel Measurability Implies Lebesgue Measurability)**.** *Suppose* $A \in \mathcal{P}(\mathbb{R})$ *is Borel measurable. Then, A is Lebseque measurable.*

Theorem 2.4.7 (Equivalence Conditions for 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. **[INCOMPLETE]**

□

Proposition 2.4.2 (Lebesgue Measurable Sets-Induced σ-algebra)**.** *The set*

$$\mathcal{L} = \{A \in \mathcal{P}(\mathbb{R}) : A \text{ is Lebesgue measurable}\}$$

is a σ*-algebra on* $\mathbb{R}$.

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.

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]**

$\square$

Definition 2.4.4 (Cantor set)**.**

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,

$$\begin{aligned} {f_k}^{-1}(\{k\}) &= \{x \in X : f_k(x) = k\} = \{x \in X : f(x) \in [k,\infty)\} = f^{-1}([k,\infty)) \in \mathcal{S}, \\ {f_k}^{-1}(\{-k\}) &= \{x \in X : f_k(x) = -k\} = \{x \in X : f(x) \in (-\infty,-k]\} = f^{-1}((-\infty,-k]) \in \mathcal{S}. \end{aligned}$$

Moreover, $\left\{\frac{m}{2^k}\right\} \in U$ and $\left\{\frac{m+1}{2^k}\right\} \in V$ imply that

$$\begin{aligned} {f_k}^{-1}\left(\left\{\frac{m}{2^k}\right\}\right) &= \left\{x \in X : f(x) \in [0,k) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right\} = f^{-1}\left(\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right) \in \mathcal{S}, \\ {f_k}^{-1}\left(\left\{\frac{m+1}{2^k}\right\}\right) &= \left\{x \in X : f(x) \in (-k,0) \cap \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right\} = f^{-1}\left(\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right)\right) \in \mathcal{S}, \end{aligned}$$

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

$$\begin{aligned} {f_k}^{-1}(B) &= {f_k}^{-1}(\tilde{B} \sqcup [B \cap f_k(X)]) \\ &= {f_k}^{-1}(\tilde{B}) \cup {f_k}^{-1}([B \cap f_k(X)]) \\ &= \emptyset \cup {f_k}^{-1}\left(\bigcup_{i=1}^{l} \{y_i\}\right) && \text{(For some } l \in \{1,\dots,card(f_k(X))\}\text{)} \\ &= \bigcup_{i=1}^{l} {f_k}^{-1}(\{y_i\}) \in \mathcal{S} && \text{(By the closure of } \mathcal{S} \text{ under countable union,)} \end{aligned}$$

where $y_i \in f_k(X)$ $\forall i \in \{1,\dots,l\}$ 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]**

□

Definition 2.5.3 (Lebesgue measurable function)**.**

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* $A = B \cup C$*.*

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]**

□

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. **[INCOMPLETE]** 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 $\int \chi_E d\mu \geq \mu(E)$.

($\leq$) 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, $\int \chi_E d\mu \leq \mathcal{L}(\chi_E, P) = \mu(E)$.

($\geq$) Consider arbitrary $\mathcal{S}$-partition $P = \{A_1, \ldots, A_m\}$ of X. Note that $\forall j \in \{1, \ldots, m\}$ $A_j \subset E \implies \inf_{x \in A_j} f(x) = 1$ and $A_j \not\subset E \implies \inf_{x \in A_j} f(x) = 0$. Then, ... ☐

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 $\forall x \in X\ \exists k \in \{1, \dots, n+1\}$ such that $x \in E_k$, since P' is a $\mathcal{S}$-partition of x. Fix $k \in \{1, \dots, n+1\}$. We observe that

$$\begin{aligned} x \in E_k \implies \chi_{E_k}(x) = 1 \text{ and } j \in \{1, \dots, n+1\} / \{k\} \implies \chi_{E_j}(x) = 0 \\ \implies f(x) = c_1(0) + c_2(0) + \cdots + c_{k-1}(0) + c_k(1) + c_{k+1}(0) + \cdots + c_{n+1}(0). \\ \therefore x \in E_k \implies f(x) = c_k. \qquad (*_0) \end{aligned}$$

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

$$\begin{aligned} \int f d\mu \geq \mathcal{L}(f, P') &= \sum_{k=1}^{n+1} \mu(E_k) \inf_{x \in E_k} f(x) \\ &= \sum_{k=1}^{n+1} \mu(E_k) c_k = \sum_{k=1}^{n} \mu(E_k) c_k. \qquad \text{(Since } c_{k+1} = 0\text{)} \end{aligned}$$

($\leq$) Consider arbitrary $\mathcal{S}$-partition $P = \{A_1, \dots, A_m\}$ of X. Fix $j \in \{1, \dots, m\}$. We claim that

$$\{f(x) : x \in A_j\} = \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\}. \qquad (*_1)$$

Consider arbitrary $p \in \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\}$. It follows that

$$\begin{aligned} \exists i \in \{1, \dots, n+1\} \text{ such that } p = c_i \text{ and } A_j \cap E_i \neq \emptyset \\ \implies \exists x \in A_j \text{ such that } x \in E_i \\ \implies f(x) = c_i. \qquad \text{(By } (*_0)\text{)} \\ \implies p = f(x) \in \{f(x) : x \in A_j\} \\ \therefore \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\} \subset \{f(x) : x \in A_j\}. \end{aligned}$$

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

$$\begin{aligned} \inf_{x \in A_j} f(x) = \inf \{f(x) : x \in A_j\} &= \inf \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\} \qquad \text{(By } (*_1)\text{)} \\ &= \min \{c_i : i \in \{1, \dots, n+1\} \text{ and } A_j \cap E_i \neq \emptyset\}, \qquad (*_2) \end{aligned}$$

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

$$\begin{aligned}
\sum_{j=1}^{m} \left[\sum_{k=1}^{n+1} \mu(A_j\cap E_k)c_k\right] &= \sum_{j=1}^{m} \left[\mu(A_j\cap E_1)c_1+\dots+\mu(A_j\cap E_{n+1})c_{n+1}\right]\\
&= (\mu(A_1\cap E_1)c_1+\dots+\mu(A_1\cap E_{n+1})c_{n+1})+\dots+(\mu(A_m\cap E_1)c_1+\dots+\mu(A_m\cap E_{n+1})c_{n+1})\\
&= c_1\left(\mu(A_1\cap E_1)+\dots+\mu(A_m\cap E_1)\right)+\dots+c_{n+1}\left(\mu(A_1\cap E_{n+1})+\dots+\mu(A_m\cap E_{n+1})\right)\\
&= \sum_{k=1}^{n+1} \left[c_k\left(\mu(A_1\cap E_k)+\dots+\mu(A_m\cap E_k)\right)\right] = \sum_{k=1}^{n+1}\left[c_k\sum_{j=1}^{m}\mu(A_j\cap E_k)\right]\\
&= \sum_{k=1}^{n+1}\left[c_k\mu(E_k)\right] = \sum_{k=1}^{n}\left[c_k\mu(E_k)\right],
\end{aligned}$$

where the second last equality is a result of $(*_3)$. Hence, we have that $\mathcal{L}(f,P)\leq\sum_{k=1}^{n}\left[c_k\mu(E_k)\right]$ for any arbitrary $\mathcal{S}$-partition P of X. It then follows, from definition, that

$$\int f d\mu \leq \sum_{k=1}^{n}\left[c_k\mu(E_k)\right].$$

Here, we complete the proof. □

Theorem 3.1.3 (Monotonicity of Integration)**.** *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)**.** *Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f \colon X \to [0, \infty]$ is $\mathcal{S}$-measurable. Then,*

$$\begin{aligned} \int f d\mu = \sup \Bigg\{ \sum_{j=1}^{m} c_j \mu(A_j) : &A_1, \ldots, A_m \in \mathcal{S} \textit{ are disjoint,} \\ &c_1, \ldots, c_m \in [0, \infty), \textit{ and} \\ &\forall x \in X,\ \sum_{j=1}^{m} c_j \chi_{A_j}(x) \leq f(x) \Bigg\}. \end{aligned} \tag{1}$$

Proof. **[INCOMPLETE]** 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

$$\begin{aligned} \mathcal{U} = \{ \sum_{j=1}^{m} c_j \mu(A_j) : &A_1, \ldots, A_m \in \mathcal{S} \text{ are disjoint,} \\ &c_1, \ldots, c_m \in [0, \infty), \text{ and} \\ &\forall x \in X,\ \sum_{j=1}^{m} c_j \chi_{A_j}(x) \leq f(x) \Bigg\} \end{aligned}$$

$\square$

Theorem 3.1.5 (Monotone Convergence Theorem)**.** *Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, and $\forall k \in \mathbb{N}$ $f_k \colon X \to [0, \infty]$ is $\mathcal{S}$-measurable and $0 \leq f_k(x) \leq f_{k+1}(x)$ $\forall x \in X$. Define $f \colon X \to [0, \infty]$ by*

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

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). *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. **[INCOMPLETE]**

□

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) $|f| = f^+ + f^-$.

Definition 3.1.5 (Integral of a real-valued function (with respect to a measure)). *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]**

$\square$

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]**

$\square$

Proposition 3.1.2 (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]**

$\square$

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]**

$\square$

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, $f\colon X \to$ is $\mathcal{S}$-measureable, and $E \in \mathcal{S}$. Then,*

$$\int_E f d\mu = \int \chi_E \cdot f d\mu.$$

Remark 3.2.1. *Note that it can be the case that $\chi_E f$ is not $\mathcal{S}$-measurable. So it could be the case that $\int \chi_E f d\mu$ and, hence, $\int_E f d\mu$ is undefined.*

Theorem 3.2.1 (Bounding an Integral on a Subset). *Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f\colon X \to$ is $\mathcal{S}$-measureable, and $E \in \mathcal{S}$.*

....

$$\left|\int_E f d\mu\right| \leq \mu(E) \cdot \sup_{x \in E} |f(x)|.$$

Proof. **[INCOMPLETE]**

□

Theorem 3.2.2 (Bounded Convergence Theorem).

Proof. **[INCOMPLETE]**

□

Definition 3.2.2 (Almost every).

Remark 3.2.2. *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.*

Corollary 3.2.1.

Theorem 3.2.3. *3.28*

$$\forall \epsilon > 0\ \exists \delta > 0 \text{ such that } B \in \mathcal{S} \text{ and } \mu(B) < \delta \implies \int_B g d\mu < \epsilon.$$

Proof. **[INCOMPLETE]**

□

Theorem 3.2.4.

Proof. **[INCOMPLETE]**

□

Theorem 3.2.5 (Dominated Convergence Theorem).

Proof. **[INCOMPLETE]**

□

Theorem 3.2.6.

Proof. **[INCOMPLETE]**

□

Definition 3.2.3 ($\int_a^b f$)**.**

Definition 3.2.4 ($||f||_1$; $\mathcal{L}^1(\mu)$)**.**

Theorem 3.2.7.

Proof. **[INCOMPLETE]**

□

Theorem 3.2.8.

Proof. **[INCOMPLETE]**

□

Definition 3.2.5 ($\mathcal{L}^1(\mathbb{R})$;$||f||_1$)**.**

Definition 3.2.6 (Step Function)**.**

Theorem 3.2.9.

Proof. **[INCOMPLETE]**

□

Theorem 3.2.10.

Proof. **[INCOMPLETE]**

□