# 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*

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

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

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

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

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

**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) □

**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, $\forall v \in V/\{0_V\}$,

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

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

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

□

**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 $m = ||T||$

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

Consider arbitrary $v, w \in V$.

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

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

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

*(i) $T$ is continuous, and*

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

*Proof.* (INCOMPLETE) $\square$

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

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

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

*Proof.* (INCOMPLETE) $\square$

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

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

*Proof.* (INCOMPLETE) $\square$

**Definition 1.1.9** (Conorm)**.**

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

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

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

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

## 1.2 Derivatives

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

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

*(a) $T$ is a linear map, and*

*(b) $T$ satisfies*

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

*or, equivalently,*

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

*or, equivalently,*

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

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

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

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

*Note that we may also denote $Df$ by $f'$.*

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

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

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

*Suppose $U \subset \mathbb{R}^n$, $f\colon U \to \mathbb{R}^m$, and $\exists f_1, \dots, f_m\colon U \to \mathbb{R}$ 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 $\mathbb{R}^n$, $\forall i \in \{1, \dots, n\}$, and $u_j$ the $j$-th standard basis vector of $\mathbb{R}^m$, $\forall j \in \{1, \dots, m\}$.*

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

**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

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

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

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

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

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

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

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

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

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

as desired.

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

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

By the differentiability of $f$ at $p$, we obtain that

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

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

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

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

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

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

**Corollary 1.2.1** (Uniqueness of Total Derivative). *Suppose $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^m$ 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 $f : U \to \mathbb{R}^m$ 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 $f : U \to \mathbb{R}^m$ 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 $f : U \to \mathbb{R}^m$. 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 $f : U \to \mathbb{R}^m$. 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 $f\colon U \to \mathbb{R}$.

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

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

which is equivalent to

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

which is also equivalent to

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

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

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

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

**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 $p$, $\forall j \in \{1, \dots, m\}$.*

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

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

*and*

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

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

*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)*

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

*where $\forall (i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$ $a_{i,j}$ 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 $C^\infty$)**.** *Suppose* $U \subset \mathbb{R}^n$*. Then,* $f\colon U \to \mathbb{R}^m$ *is smooth or of class* $C^\infty$ *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*

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

*(ii) uniformly $C^r$ Cauchy if*

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

**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 $\forall n, m \in \mathbb{N}$,

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

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*

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

**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, $f: M \to M$ 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 $f: M \to M$ 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 $\theta$ 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 $f: M \to M$ 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 $f: B_m \to B_m$ 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 $f: U \to \mathbb{R}^m$. $f$ 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 $f: U \to \mathbb{R}^m$ 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 $f: U \to \mathbb{R}^m$ is a $C^r$ diffeomorphism for some $r \in \{1, 2, \ldots, \infty\}$. By definition, we have that $f^{-1}$ exists, where $f$ and $f^{-1}$ are $C^r$ bijections. By definition, $f$ and $f^{-1}$ are continuously differentiable and, hence, continuous. Hence, $f$ and $f^{-1}$ are continuous bijections. By definition, $f$ is a homeomorphism. □

**Theorem 1.4.3** (Inverse Function Theorem)**.** *Suppose $U \subset \mathbb{R}^n$ is open, $f: U \to \mathbb{R}^n$ 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) $f^{-1}: V_0 \to U_0$ 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*

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

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

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

**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

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

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

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

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 $Z_1, Z_2, \dots \subset \mathbb{R}$ are zero sets. Then, $\cup_{k\in\mathbb{N}} Z_k$ is a zero set.*

*Proof.* **(DO: DOUBLE CHECK)** Suppose $Z_1, Z_2, \dots \subset \mathbb{R}$ 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 $n \in \{2, 3, \dots\}$ we have that $U_n = Z_1 \cup \dots \cup Z_n$ 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,

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

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 $\forall n \in \mathbb{N}$, $Z_1 \cup \dots \cup Z_n$ is a zero set and, therefore, $\cup_{n\in\mathbb{N}} Z_n$ is a zero set. $\square$

**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 $\overline{F}\colon [c,d] \to \mathbb{R}$ defined by*

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

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

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

*(ii)*

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

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

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

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

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

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

**Proposition 1.7.4** (Discontinuity of Characteristic Function at the Boundary)**.** *Suppose $M$ is a metric space. Then, the characteristic function $\chi_S$ 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 $\partial S$ 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 $\chi_S$ is continuous at $p$ by definition.

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

Suppose that $\chi_S$ is continuous at $p$. Suppose, to the contrary, that $p \in \partial S$. By the continuity of $\chi_S$ 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, $\chi_S$ is differentiable on $M/\partial S$ and not differentiable on $\partial S$.*

**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 $\iff$ $\partial S$ is a zero set.*

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

**Theorem 1.7.7** (Cavalieri's Principal)**.** *Suppose $R = [a,b] \times [c,d] \subset \mathbb{R}^2$, $S \subset R$, and $\partial S$ 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 $A_1, A_2, \dots \subset \mathbb{R}$. 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. $\square$

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

**Definition 2.1.4** (Translation of a set in $\mathbb{R}$).

**Theorem 2.1.4** (Translation Invariance of Outer Measure).

**Definition 2.1.5** (Open Cover).

**Theorem 2.1.5** (Heine-Borel Theorem).

**Theorem 2.1.6** (Outer Measure of a Closed Interval).

**Theorem 2.1.7** (Nondegenerate Intervals are Uncountable).

**Theorem 2.1.8** (Nonadditivity of Outer Measure).

## 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 $\sigma$-algebra.*

**Theorem 2.2.1** (Nonexistence of Extension of length to all subsets of $\mathbb{R}$)**.** *There does not exist a function $\mu$ 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) $A_1, A_2, \dots$ 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** ($\sigma$-algebra)**.** *Suppose $X$ is a set. $\mathcal{A} \subset \mathcal{P}(X)$ is a $\sigma$-algebra of $X$ if*

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

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

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

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

(i) By the definition of $\sigma$-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 complement. 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 complement.

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

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

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

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

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

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

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

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

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

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

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

(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 $\sigma$-algebra on $X$. $\square$

**Definition 2.2.3** (Measurable space; measurable set)**.** *(i) A measurable space is an order pair $(X, \mathcal{S})$, where $X$ is a set and $\mathcal{S}$ is a $\sigma$-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 $\sigma$-algebra)**.** *Suppose $X$ is a set. Then, the intersection of all $\sigma$-algebra on $X$ is the smallest $\sigma$-algebra on $X$.*

*Proof.* Suppose $X$ is a set. Let $\mathcal{A}$ be the intersection of all $\sigma$-algebra on $X$.

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

(ii) Consider arbitrary $A \in \mathcal{A}$. It follows that $A$ is also contained in all $\sigma$-algebra on $X$. As an immediate result of the closure of a $\sigma$-algebra under complement, $X/A$ is contained in all $\sigma$-algebra on $X$ 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 $\sigma$-algebra on $X$. Therefore, $\cup_{k\in\mathbb{N}} A_k$ is also contained in all $\sigma$-algebra on $X$, by the closure of a $\sigma$-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 $\sigma$-algebra on $X$ and certainly $\mathcal{A}$ is contained in any $\sigma$-algebra on $X$ by assumption. Hence, we conclude that $\mathcal{A}$ is the smallest $\sigma$-algebra on $X$. □

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

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

*Proof.* By a previous proposition, $\mathcal{S}$ is a $\sigma$-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 $\sigma$-algebra on $X$ containing $\mathcal{A}$ contains $\mathcal{S}$.

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

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

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 complement.

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 $\sigma$-algebra $\mathcal{U}$ on $X$. Therefore, $\mathcal{S}$ is indeed the smallest $\sigma$-algebra on $X$ containing $\mathcal{A}$. □

**Example 2.2.2.**

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

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

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

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

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

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

*is a Borel set.*

*Proof.* Recall that, by definition, a Borel set is a set in the smallest $\sigma$-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 complement $\mathbb{R}/[\mathbb{R}/S] = S \in \mathcal{B}$. Thus, $S$ is a Borel set.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

*(i) $A \subset Y \implies f^{-1}(X/A) = Y/f(A)$;*
*(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 $f\colon X \to Y$, $g\colon Y \to W$.

(i) Suppose that $A \subset Y$.

$f^{-1}(X/A) = Y/f(A)$;
(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) $A \subset W \implies (g \circ f)^{-1}(A) = f^{-1}\left(g^{-1}(A)\right)$. $\square$

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

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

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

**Proposition 2.2.5** (Function Measurablity on Trivial $\sigma$-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. $\square$

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

**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*

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

**Theorem 2.2.5** ($\mathcal{S}$-Measurability and Containment of Inverse Image of Unbounded Above Open Intervals in $\mathcal{S}$)**.**

**Definition 2.2.8** (Borel Measurable function)**.**

**Theorem 2.2.6** (Continuity Implies Borel Measurability)**.**

**Definition 2.2.9** (Increasing function)**.**

**Theorem 2.2.7** (Increasing Implies Borel Measurability)**.**

**Theorem 2.2.8** ($\mathcal{S}$-Measurable Functions Arose from $\mathcal{S}$-Measurable Functions )**.**

**Theorem 2.2.9** (Limit of $\mathcal{S}$-Measurable Functions)**.**

**Definition 2.2.10** (Borel sets of $[-\infty, \infty]$ )**.**

**Definition 2.2.11** (Measurable function)**.**

**Theorem 2.2.10** (Condition for Function Measurablity)**.**

**Theorem 2.2.11.**

## 2.3 Measures and Their Properties

**Definition 2.3.1** (Measure on a measurable space).

**Proposition 2.3.1** (Measure Menu Theorem).

**Definition 2.3.2** (Measure space).

**Theorem 2.3.1** (Monotonicity of Measure; Measure of Set Difference).

**Theorem 2.3.2** (Countable Subadditivity of Measure).

**Theorem 2.3.3** (Measure of Increasing Union).

**Theorem 2.3.4** (Measure of Decreasing Intersection).

**Theorem 2.3.5** (Measure of Union).

## 2.4 Lebesgue Measure

## 2.5 Convergence of Measurable Functions