# MAT237: Multivariable Calculus

Tianyu Du

Friday 12th April, 2019

## Contents

# 1 Limits, continuity, and related topics

## 1.1 Open and Closed Sets, Boundary, Interior and Closure

**Definition 1.1.1.** Let $\mathbf{a} \in \mathbb{R}^n$, and $r > 0$. The **open ball with centre a and radius** $r$ is defined as

$$\mathcal{B}(r, \mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x} - \mathbf{a}|| < r\}. \tag{1.1.1}$$

**Definition 1.1.2.** The **sphere with centre a and radius** $r$ is defined as

$$\partial\mathcal{B}(r, \mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x} - \mathbf{a}|| = r\} \tag{1.1.2}$$

**Definition 1.1.3.** Let $S \subset \mathbb{R}^n$, $S$ is **bounded** if

$$\exists r > 0 \ s.t. \ S \subset \mathcal{B}(r, \mathbf{0}) \tag{1.1.3}$$

**Definition 1.1.4.** Let $S \subset \mathbb{R}^n$, then the **complement** of $S$ is defined as

$$S^c := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \notin S\} \tag{1.1.4}$$

**Definition 1.1.5.** Let $S \subset \mathbb{R}^n$, the **interior** of $S$ is defined as

$$S^{int} := \{\mathbf{x} \in \mathbb{R}^n : \exists \varepsilon > 0 \ s.t. \ \mathcal{B}(\varepsilon, \mathbf{x}) \subset S\} \tag{1.1.5}$$

**Definition 1.1.6.** The **boundary** of $S$ is defined as

$$\partial S := \{\mathbf{x} \in \mathbb{R}^n : \forall \varepsilon > 0 \ \mathcal{B}(\varepsilon, \mathbf{x}) \cap S \neq \varnothing \wedge \mathcal{B}(\varepsilon, \mathbf{x}) \cap S^c \neq \varnothing\} \tag{1.1.6}$$

**Theorem 1.1.1.** *A point $\mathbf{x} \in S$ is either a boundary point or a interior point.*

**Definition 1.1.7.** The **closure** of $S$ is defined as

$$\overline{S} := S^{int} \cup \partial S \tag{1.1.7}$$

**Theorem 1.1.2.** *For any $S \subset \mathbb{R}^n$*

$$S^{int} \subset S \subset \overline{S} \tag{1.1.8}$$

**Theorem 1.1.3.** *For any $S \subset \mathbb{R}^n$*

$$\partial S = \partial(S^c) \tag{1.1.9}$$

**Definition 1.1.8.** A set $S \subset \mathbb{R}^n$ is **open** if $S = S^{int}$. $S$ is **closed** if $S = \overline{S}$.

**Theorem 1.1.4.**

$$S \text{ is closed} \iff S^c \text{ is open} \tag{1.1.10}$$

*Proof.*

$$S \text{ is closed} \iff \partial S \subset S \iff \partial(S^c) \subset S \tag{1.1.11}$$

$$\iff \text{no point of } S^c \text{ is a boundary point} \iff S^c \text{ is open} \tag{1.1.12}$$

■

## 1.2 Limits and Continuity

### 1.2.1 Limits of Multivariable Functions

**Definition 1.2.1.** Let $S \subset \mathbb{R}^n$, $\mathbf{f} : S \to \mathbb{R}^k$, and $\mathbf{a} \in S$, then

$$\lim_{\mathbf{x} \to \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{L} \in \mathbb{R}^k \tag{1.2.1}$$

is defined as

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ s.t. \ \forall \mathbf{x} \in S, \ 0 < ||\mathbf{x} - \mathbf{a}|| < \delta \implies ||\mathbf{f}(\mathbf{x}) - \mathbf{L}|| < \varepsilon \tag{1.2.2}$$

For this definition to be non-trivial, we need **a** not be an isolated point,

$$\forall \delta > 0, \ \exists \mathbf{x} \in S \ s.t. \ ||\mathbf{x} - \mathbf{a}|| \in (0, \delta) \tag{1.2.3}$$

**Theorem 1.2.1** (Limit Laws)**.** Let $S \subset \mathbb{R}^n$ and $\mathbf{a} \in \mathbb{R}^n$ satisfying (1.3.3) And $f, g : S \to \mathbb{R}$, $L, M \in \mathbb{R}$ such that

$$\lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = L \tag{1.2.4}$$

$$\lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = M \tag{1.2.5}$$

then

$$\lim_{\mathbf{x} \to \mathbf{a}} [f(\mathbf{x}) + g(\mathbf{x})] = L + M \tag{1.2.6}$$

$$\lim_{\mathbf{x} \to \mathbf{a}} [f(\mathbf{x}) \cdot g(\mathbf{x})] = LM \tag{1.2.7}$$

**Theorem 1.2.2** (Squeeze Theorem)**.** Let $S \subset \mathbb{R}^n$, $\mathbf{a} \in \mathbb{R}^n$ satisfies (1.3.3). Suppose that $f, g, h : S \to \mathbb{R}$ and there exists $p > 0$ and $L \in \mathbb{R}$ such that

$$\forall \mathbf{x} \in S \cap \mathcal{B}(p, \mathbf{a}) \ f(\mathbf{x}) \leq g(\mathbf{x}) \leq h(\mathbf{x}) \tag{1.2.8}$$

and

$$\lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = \lim_{\mathbf{x} \to \mathbf{a}} h(\mathbf{x}) = L \tag{1.2.9}$$

then

$$\lim_{\mathbf{x} \to \mathbf{a}} g(\mathbf{x}) = L \tag{1.2.10}$$

**Corollary 1.2.1.** Let $g, h : S \to \mathbb{R}$ and

$$|g(\mathbf{x})| \leq h(\mathbf{x}) \ \forall \mathbf{x} \in S \tag{1.2.11}$$

$$\text{and } \lim_{\mathbf{x} \to \mathbf{a}} h(\mathbf{x}) = 0 \tag{1.2.12}$$

then

$$\lim_{\mathbf{x} \to \mathbf{a}} g(\mathbf{x}) = 0 \tag{1.2.13}$$

**Theorem 1.2.3.** Assume that $S \subset \mathbb{R}^n$ and let $\mathbf{a} \in \mathbb{R}^n$ satisfying (1.3.3). Let $\mathbf{f} : S \to \mathbb{R}^k$, then

$$\lim_{\mathbf{x} \to \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{L} \iff \lim_{\mathbf{x} \to \mathbf{a}} f_j(\mathbf{x}) = L_j \ \forall j \tag{1.2.14}$$

### 1.2.2 Continuity

**Definition 1.2.2.** Let $S \subset \mathbb{R}^n$ and $\mathbf{f} : S \to \mathbb{R}^k$. $\mathbf{f}$ is **continuous at $\mathbf{a} \in S$** if

$$\lim_{\mathbf{x} \to \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) \tag{1.2.15}$$

and $\mathbf{f}$ is **continuous** if $\mathbf{f}$ is continuous at every point in $S$.

**Theorem 1.2.4** (Basic Properties of Continuity)**.** Assume that $S \subset \mathbb{R}^n$ and $\mathbf{a} \in S$,

(i) If $\mathbf{f} : S \to \mathbb{R}^k$ is continuous at $\mathbf{a}$, then every component of $\mathbf{f}$, $f_j : S \to \mathbb{R}$, is continuous at $\mathbf{a}$.

(ii) If $\mathbf{f}, \mathbf{g} : S \to \mathbb{R}^k$ are continuous at $\mathbf{a}$, then $\mathbf{f} + \mathbf{g}$ is continuous at $\mathbf{a}$.

(iii) If $f, g : S \to \mathbb{R}$ continuous, then $fg$ is continuous and $\frac{f}{g}$ is continuous given $g(\mathbf{a}) \neq 0$.

(iv) A composition of continuous functions is continuous.

(v) The elementary functions of a single variable (trigonometric functions and their inverses, polynomials, exponential and log) are continuous on their domains.

### 1.2.3 Continuous Functions and Open Sets

**Theorem 1.2.5.** Assume that $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^k$, then the following are equivalent

(i) $\mathbf{f}$ is continuous;

(ii) For every open set $\mathcal{O} \subset \mathbb{R}^k$, $\mathbf{f}^{-1}(\mathcal{O})$ is also open;

(iii) For every closed set $\mathcal{C} \subset \mathbb{R}^k$, $\mathbf{f}^{-1}(\mathcal{C})$ is also closed.

## 1.3 Sequences and Completeness

**Definition 1.3.1.** A sequence $\{\mathbf{a}_j\}_j$ in $\mathbb{R}^n$ **converges to the limit $\mathbf{L} \in \mathbb{R}^n$** if

$$\forall \varepsilon > 0 \; \exists J \in \mathbb{N}, \; s.t. \; \forall j \geq J \implies ||\mathbf{a}_j - \mathbf{L}|| < \varepsilon \tag{1.3.1}$$

**Theorem 1.3.1.**

$$\lim_{j \to \infty} \mathbf{a}_j = \mathbf{L} \iff \lim_{j \to \infty} ||\mathbf{a}_j - \mathbf{L}|| = 0 \tag{1.3.2}$$

**Theorem 1.3.2.** Let $\{a_{jk}\}_j$ be a sequence in $\mathbb{R}^n$ where $k \in [n]$, and let $\mathbf{L} = (L_1, \ldots, L_n) \in \mathbb{R}^n$, then

$$\lim_{j \to \infty} \mathbf{a}_j = \mathbf{L} \iff \lim_{j \to \infty} a_{jk} = L_k \; \forall k \in [n] \tag{1.3.3}$$

*Proof Idea.*

$$\forall j \in [n], \; |a_j - L_j| \leq ||\mathbf{a} - \mathbf{L}|| \leq n \max_{k \in [n]} |a_k - L_k| \tag{1.3.4}$$

■

**Axiom 1.1** (the Completeness Axiom)**.** Every *bounded* and *nonempty* set of *real numbers* has a *least upper bound* (**supremum**) and a *greatest lower bound* (**infimum**).

**Theorem 1.3.3** (Monotone Sequence Theorem)**.** Every bounded nondecreasing sequence of real numbers converges to a limit.

*Proof Idea.* Note that such sequence converges to its supremum $S$.
Let $\varepsilon > 0$, there exists $j^*$ such that

$$S - \varepsilon < a_{j^*} \leq S \tag{1.3.5}$$

take such $j^*$ and by the nondecreasing property,

$$\forall j \geq j^* \; a_j > S - \varepsilon \tag{1.3.6}$$

which implies $|S - a_j| < \varepsilon$. ■

**Definition 1.3.2.** A **subsequence** of a sequence $\{\mathbf{a}_j\}_{j\geq j_0}$ in $\mathbb{R}^n$ is a sequence constructed as $\{a_{k_j}\}_j$, such that $\{k_j\}_j$ is a *strictly increasing* sequence bounded below by $j_0$.

**Remark 1.3.1.** Subsequences can be constructed using strictly increasing transformations.

**Proposition 1.3.1.** If $\{\mathbf{a}_j\}_j$ is a sequence in $\mathbb{R}^n$ converges to $\mathbf{L}$, then (i) any subsequence of it converges to the (ii) same limit.

*Proof Idea.* Suppose not and reach a contradiction. ■

**Theorem 1.3.4** (Bounded Sequence Theorem in $\mathbb{R}$)**.** Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.

*Proof.* Let $\{a_j\}_j$ be a bounded sequence.
For each $j \in \mathbb{N}$, define $b_{k_j} := \inf_{k>k_j} a_k$.
Note that $\{b_j\}$ is non-decreasing and bounded, so it converges to some limit $\ell$.
Let $\{a_{k_j}\}_j$ denote a subsequence of the original sequence, define $k_0 = j_0$, and indices are constructed in a recurrent way.
Suppose every index before $k_j$ has been chosen, we choose $k_{j+1}$ to be the index such that

$$b_{k_j} \leq a_{k_{j+1}} < b_{k_j} + \frac{1}{j} \tag{1.3.7}$$

by construction, $\{a_{k_j}\}_j$ is bounded by both $\{b_{k_j}\}_j$ and $\{b_{k_j} + \frac{1}{j}\}_j$, and both bounding sequences converge to $\ell$. So $\{a_{k_j}\}_j$ converges to $\ell$ by *squeeze theorem.* ■

**Theorem 1.3.5** (Bounded Sequence Theorem)**.** Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

*Proof.* Let $\{\mathbf{a}_j\}_j$ be a bounded sequence.
Applying the previous theorem iteratively, we can construct a subsequence of $\{\mathbf{a}_{k_j}\}_j$ such that $\{\mathbf{a}_{k_j} \cdot \mathbf{e}_1\}_j$ is bounded and convergent.
Then we apply the previous theorem iteratively on the constructed convergent subsequences to construct new subsequences with more convergent components. ■

## 1.4 Compactness

### 1.4.1 Compactness

**Definition 1.4.1** (Heine-Borel Property)**.** A set $S$ is **compact** if every *open* covering of $S$ has a *finite* sub-covering.

**Definition 1.4.2** (Sequentially Compact)**.** A set $S \subset \mathbb{R}^n$ is **compact** if every sequence in $S$ has a subsequence that converges to a limit in $S$.

**Proposition 1.4.1.** If $\{\mathbf{x}_j\}_j$ is a *convergent* sequence in a *closed* set $S \subset \mathbb{R}^n$, the then limit of this sequence is in $S$.

*Proof Idea.* Let $\mathbf{x} := \lim_{j\to\infty} \mathbf{x}_j$, and we wish to show $\mathbf{x} \in S$.
Equivalently, we can show $\mathbf{x} \in \overline{S}$, and that's

$$\forall \varepsilon > 0 \ \mathcal{B}(\varepsilon, \mathbf{x}) \cap S \neq \varnothing \tag{1.4.1}$$

this is immediately true by the definition of sequence convergence. There must be some points in the sequence, thus in $S$, belongs to $\mathcal{B}(\varepsilon, \mathbf{x})$. ■

**Theorem 1.4.1** (Bolzano-Weierstrass)**.** Let $S \subset \mathbb{R}^n$,

$$S \text{ is compact } \iff S \text{ is closed and bounded} \tag{1.4.2}$$

*Proof Idea.*
( $\Longleftarrow$ ) Suppose $S$ is closed and bounded, boundedness ensures such sequence converges, and closeness ensures the limit point of sequence is in $S$.
( $\Longrightarrow$ ) Prove by *modus tollens*.
**Case (i):** $S$ is not bounded, then

$$\forall R > 0 \ \exists \mathbf{x} \in S \backslash \mathcal{B}(R, \mathbf{0}) \tag{1.4.3}$$

and above $\mathbf{x}(R)$ depends on $R$, we can construct a sequence using $\mathbf{x}(j)$ such that the $||\mathbf{x}||$ is ever increasing and it does not have a limit.
**Case (ii):** $S$ is not closed, we can construct a sequence with subsequence converges to $\mathbf{x} \in \partial S \backslash S$, which is nonempty because $S$ is not closed. ■

### 1.4.2 the Extreme Value Theorem

**Theorem 1.4.2** (the Extreme Value Theorem)**.** Assume $K$ is a compact subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ is continuous.
Then (i)

$$f(K) \text{ is compact} \tag{1.4.4}$$

and (ii) the infimum and supremum of $f(\mathbf{x})$ on $K$ are attainable.

$$\exists \ \overline{\mathbf{x}}, \underline{\mathbf{x}} \in K \ s.t. \begin{cases} f(\overline{\mathbf{x}}) = \sup_{\mathbf{x}\in K} f(\mathbf{x}) \\ f(\underline{\mathbf{x}}) = \inf_{\mathbf{x}\in K} f(\mathbf{x}) \end{cases} \tag{1.4.5}$$

*Proof.* Let $\{y_j\}_j$ be a sequence in $f(K)$, and we can find a sequence $\{\mathbf{z}_j\}_j$ in $K$ such that $y_j = f(\mathbf{z}_j)$ (by definition of image). Because $K$ is compact, there exists a subsequence of $\{\mathbf{z}_j\}_j$ converges to $\mathbf{z}^* \in K$. Since $f$ is continuous, we can conclude there a subsequence, sharing the same indices, such that $f(\mathbf{z}_j) \to f(\mathbf{z}^*)$ (*Proposition 1.5.2*). Obviously $f(\mathbf{z}^*) \in f(K)$, *so* $f(K)$ *is compact.*
Since $f(K)$ is compact, by *Proposition 1.5.3*, $\sup_{\mathbf{x}\in K} f(\mathbf{x}) \in f(K)$. By definition of image, $\exists \mathbf{x} \in K$ such that $f(\mathbf{x}) = \sup_{\mathbf{x}\in K} f(\mathbf{x})$, *supremum attainability shown.*
Proof for infimum attainability is the same. ■

**Proposition 1.4.2.** Assume that $\{\mathbf{z}_j\}_j$ is a sequence in a set $S \subset \mathbb{R}^k$, and $f$ is a continuous real-valued function defined on $S$, then

$$\mathbf{z}_j \to \mathbf{z} \implies f(\mathbf{z}_j) \to f(\mathbf{z}) \tag{1.4.6}$$

**Proposition 1.4.3.** If $S$ is a compact set in $\mathbb{R}$, then $\sup S$ and $\inf S$ both in $S$.

*Proof Idea.* Suppose $\sup S \notin S$, by definition of supremum,

$$\forall \varepsilon \ \exists x \in S \ s.t \ \sup S - \varepsilon < x \leq \sup S \tag{1.4.7}$$

note that such $x \in \mathcal{B}(\varepsilon, \sup S)$. Also, similarly,

$$\forall \varepsilon > 0 \ \exists x \notin S \ s.t. \ \sup S < x < \sup S + \varepsilon \tag{1.4.8}$$

so such $x \in \mathcal{B}(\varepsilon, \sup S)$. We conclude

$$\forall \varepsilon > 0 \ \mathcal{B}(\varepsilon, \sup S) \cap S \neq \varnothing \wedge \mathcal{B}(\varepsilon, \sup S) \cap S^c \neq \varnothing \tag{1.4.9}$$

which means $\sup S \in \partial S$. Thus if $\sup S \notin S$, $S$ cannot be closed and this contradicts our assumption that $S$ is compact.
The proof for $\inf S \in S$ is similar. ■

### 1.4.3 Uniform Continuity

**Definition 1.4.3.** Let $S \subset \mathbb{R}^n$, a function $\mathbf{f} : S \to \mathbb{R}^k$ is **uniformly continuous** if

$$\forall \varepsilon \ \exists \delta > 0 \ s.t. \ \forall \mathbf{x}, \ \mathbf{y} \in S, \ ||\mathbf{x} - \mathbf{y}|| < \delta \implies ||\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})|| < \varepsilon \tag{1.4.10}$$

**Remark 1.4.1.** In the definition of *continuity*, value of $\delta$ can depend on $\mathbf{x}$. But in the definition of *uniform continuity*, one $\delta$ has to work for every $\mathbf{x}$.

**Theorem 1.4.3.** If $K$ is a compact subset of $\mathbb{R}^n$, and $\mathbf{f} : K \to R^k$ is continuous, then $\mathbf{f}$ is uniformly continuous.

## 1.5 the Intermediate Value Theorem

**Definition 1.5.1.** A set $S \subset \mathbb{R}^n$ is **path-connected** (**arcwise connected/ pathwise connected**) if for every $\mathbf{x}, \mathbf{y} \in S$, there exists a continuous function $\gamma : [0,1] \to S$ such that

$$\gamma(0) = \mathbf{x}, \ \gamma(1) = \mathbf{y} \tag{1.5.1}$$

**Example 1.5.1.** Convex sets are path-connected, a path can be constructed using the convex combination,

$$\gamma(t) := (1-t)\mathbf{x} + t\mathbf{y} \tag{1.5.2}$$

**Proposition 1.5.1.** Let $S_1, S_2 \subset \mathbb{R}^n$ be two path-connected sets, and $S_1 \cap S_2 \neq \varnothing$. Then $S_1 \cup S_2$ is path-connected.

*Proof.* Take $\mathbf{z} \in S_1 \cap S_2$, and let $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ be two connecting paths between $\mathbf{x}, \mathbf{z}$ and $\mathbf{z}, \mathbf{y}$ respectively. Then define $\gamma : [0,1] \to S_1 \cup S_2$ as

$$\gamma(t) := \mathbb{1}\{t \in [0, \frac{1}{2})\} \times \tilde{\gamma}_1(2t) + \mathbb{1}\{t \in [\frac{1}{2}, 1]\} \times \tilde{\gamma}_2(2(t - \frac{1}{2})) \tag{1.5.3}$$

■

**Theorem 1.5.1** (the intermediate Value Theorem)**.** Assume that $S$ is a path-connected subset of $\mathbb{R}^n$ and that $f : S \to \mathbb{R}$ is continuous. Let $\mathbf{a}, \mathbf{b} \in S$.
Then for every $t \in (\min\{f(\mathbf{a}), f(\mathbf{b})\}, \max\{f(\mathbf{a}), f(\mathbf{b})\})$, there exists $\mathbf{c} \in S$ such that $f(\mathbf{c}) = t$.

*Proof.* Let $\mathbf{a}, \mathbf{b} \in S$. WLOG, assume $f(\mathbf{a}) < f(\mathbf{b})$. Let $t$ be an arbitrary value in $(f(\mathbf{a}), f(\mathbf{b}))$.
Since $S$ is path-connected, let $\vec{\varphi} : [0,1] \to S$ be a continuous function such that $\vec{\varphi}(0) = \mathbf{a}$ and $\vec{\varphi}(1) = \mathbf{b}$.
Then we can construct composite $f \circ \vec{\varphi} : [0,1] \to \mathbb{R}$, then apply the Intermediate Value Theorem in $\mathbb{R}$. We can conclude that $\exists \eta \in (0,1)$ *s.t.* $f \circ \vec{\varphi}(\eta) = t$. And $\vec{\varphi}(\eta) \in S$ is the point desired. ∎

# 2 Differentiation and related topics

## 2.1 Differentiation of Real-Valued Functions

### 2.1.1 Single Variable Case

**Definition 2.1.1** (Equivalent Definitions of Differentiability)**.** Let $S \subset \mathbb{R}$ open, and $f : S \to \mathbb{R}$ is said to be **differentiable at** $x \in S$ if

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \text{ exists} \tag{2.1.1}$$

or there exists $m \in \mathbb{R}$ such that

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - mh}{h} = 0 \tag{2.1.2}$$

or there exists $m \in \mathbb{R}$ and $E(h) : \mathbb{R} \to \mathbb{R}$ such that

$$f(x+h) = f(x) + mh + E(h), \ \lim_{h \to 0} \frac{E(h)}{h} = 0 \tag{2.1.3}$$

If $f$ is differentiable at $x$, we define the **derivative** $f'(x) := m$.

### 2.1.2 Differentiability of Real-valued Functions Defined on $\mathbb{R}^n$

**Definition 2.1.2.** Let $S$ be an open subset of $\mathbb{R}^n$, and $f : S \to \mathbb{R}$ is **differentiable at** $\mathbf{x} \in S$ if

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})}{||\mathbf{h}||} \text{ exists} \tag{2.1.4}$$

or there exists $\mathbf{m} \in M_{1 \times n}$ such that

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - \mathbf{m} \cdot \mathbf{h}}{||\mathbf{h}||} = 0 \tag{2.1.5}$$

or there exists $\mathbf{m} \in M_{1 \times n}$ and $E(\mathbf{h})$ such that

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \mathbf{m} \cdot \mathbf{h} + E(\mathbf{h}), \ \lim_{\mathbf{h} \to \mathbf{0}} \frac{E(\mathbf{h})}{||\mathbf{h}||} = 0 \tag{2.1.6}$$

If $f$ is differentiable at $\mathbf{x}$, we define its gradient as $\nabla f(\mathbf{a}) := \mathbf{m}$.

**Theorem 2.1.1.** Assume that $f : S \to \mathbb{R}$, where $S$ is an open subset of $\mathbb{R}^n$, and that $\mathbf{x} \in S$. If $f$ is *differentiable* at $\mathbf{x}$, then $f$ is continuous at $\mathbf{x}$.

*Proof.*

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \mathbf{m}\cdot\mathbf{h} + E(\mathbf{h}) \tag{2.1.7}$$

Note that when $||\mathbf{h}|| \leq 1$,

$$E(\mathbf{h}) \leq \frac{|E(\mathbf{h})|}{||\mathbf{h}||} \tag{2.1.8}$$

By the *Squeeze Theorem*, $\lim_{\mathbf{h}\to\mathbf{0}} E(\mathbf{h}) = 0$. Also, $\lim_{\mathbf{h}\to\mathbf{0}} \mathbf{m}\cdot\mathbf{h} = 0$. Thus

$$\lim_{\mathbf{h}\to\mathbf{0}} f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = 0 \tag{2.1.9}$$

so $f$ is continuous at $\mathbf{x}$. ■

### 2.1.3 Partial Differentiability

**Definition 2.1.3.** Let $S$ be an open subset of $\mathbb{R}^n$, and $f : S \to \mathbb{R}$. The **$j$-th partial derivative of $f$ at $\mathbf{x}$** is defined as

$$\frac{\partial f(\mathbf{x})}{\partial x_j} := \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{e}_j) - f(\mathbf{x})}{h} \tag{2.1.10}$$

**Theorem 2.1.2.** Let $f$ be a function $S \to \mathbb{R}$, where $S$ is an open subset of $\mathbb{R}^n$. If $f$ is differentiable at a point $\mathbf{x} \in S$, then (i) $\frac{\partial f}{\partial x_j}$ exists at $\mathbf{x}$ for every $j \in [n]$ and (ii)

$$\nabla f(\mathbf{x}) = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})(\mathbf{x}) \tag{2.1.11}$$

**Theorem 2.1.3.** Assume $f$ is a function $S \to \mathbb{R}$ for some open $S \subset \mathbb{R}^n$. If all partial derivatives of $f$ exist and are continuous at every point of $S$, then $f$ is differentiable in $S$.

**Definition 2.1.4.** A function $f : S \to R$ is said to be **of class $C^1$** if all partial derivatives of $f$ exist and continuous at every point of $S$.

### 2.1.4 Directional Derivatives

**Definition 2.1.5.** A **direction** in $\mathbb{R}^n$ is represented by a unit vector $\mathbf{u}$. And given such a unit vector, the **directional derivative of $f$ at $\mathbf{x}$ in the direction of $\mathbf{u}$** is defined as

$$\partial_{\mathbf{u}} f(\mathbf{x}) := \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{u}) - f(\mathbf{x})}{h} \tag{2.1.12}$$

**Theorem 2.1.4.** If $f$ is differentiable at a point $\mathbf{x}$, then $\partial_{\mathbf{u}} f(\mathbf{x})$ exists for every unit vector $\mathbf{u}$, and moreover

$$\partial_{\mathbf{u}} f(\mathbf{x}) = \mathbf{u}\cdot\nabla f(\mathbf{x}) \tag{2.1.13}$$

## 2.2 Differentiation

**Definition 2.2.1.** Assume $S$ is an open subset of $\mathbb{R}^n$. Given function $\mathbf{f} : S \to \mathbb{R}^m$, we say that $\mathbf{f}$ is differentiable at a point $\mathbf{a} \in S$ if there exists $M \in M_{m\times n}$ such that

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) = M\mathbf{h} + \mathbf{E}(\mathbf{h}),\ \lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{E}(\mathbf{h})}{||\mathbf{h}||} = \mathbf{0} \in \mathbb{R}^m \tag{2.2.1}$$

If such $M$ exists, we define the **Jacobian matrix** of $\mathbf{f}$ at $\mathbf{a}$ as

$$D\mathbf{f}(\mathbf{a}) := M \tag{2.2.2}$$

**Definition 2.2.2.** Given a differentiable function $f : S \to \mathbb{R}$, where $S$ is an open subset of $\mathbb{R}^n$, at a point $\mathbf{a}$ we define the **differential of $f$ at a** as

$$df|_{\mathbf{a}}(\mathbf{h}) := \nabla f(\mathbf{a}) \cdot \mathbf{h} \tag{2.2.3}$$

**Remark 2.2.1.** The differential is discussed only for real-valued functions here.

**Remark 2.2.2.** The differential can be used for linear approximations for small $\mathbf{h}$.

$$f(\mathbf{x} + \mathbf{h}) \approx f(\mathbf{x}) + df|_{\mathbf{a}}(\mathbf{h}) \tag{2.2.4}$$

## 2.3 the Chain Rule

**Theorem 2.3.1** (the Chain Rule)**.** Let $S_n \subset \mathbb{R}^n$ and $T_m \subset \mathbb{R}^m$, given functions $\mathbf{g} : S_n \to \mathbb{R}^m$ and $\mathbf{f} : T_m \to \mathbb{R}^\ell$. Also let $\mathbf{a} \in S_n$ such that $\mathbf{g}$ is differentiable at $\mathbf{a}$ and $\mathbf{f}$ is differentiable at $\mathbf{g}(\mathbf{a})$[1]. Then

$$\underbrace{D(\mathbf{f} \circ \mathbf{g})(\mathbf{a})}_{\ell \times n} = \underbrace{D(\mathbf{f})(\mathbf{g}(\mathbf{a}))}_{\ell \times m} \underbrace{D\mathbf{g}(\mathbf{a})}_{m \times n} \tag{2.3.1}$$

**Example 2.3.1.**

$$\frac{d}{d\mathbf{x}} ||\mathbf{x}|| = \frac{\mathbf{x}}{||\mathbf{x}||} \tag{2.3.2}$$

**Definition 2.3.1.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is called **homogeneous of degree** $\alpha$ if

$$f(\lambda \mathbf{x}) = \lambda^\alpha f(\mathbf{x}) \ \forall \mathbf{x} \neq \mathbf{0}, \ \lambda \in \mathbb{R}_{++} \tag{2.3.3}$$

**Theorem 2.3.2** (the Euler's Theorem of Homogeneous Functions)**.** If $f : \mathbb{R}^n \to \mathbb{R}$ is a homogeneous equation of degree $\alpha$, then

$$\nabla f(\mathbf{x}) \cdot \mathbf{x} = \alpha f(\mathbf{x}) \tag{2.3.4}$$

*Proof.*

$$\begin{cases} \frac{\partial f(\lambda \mathbf{x})}{\partial \lambda} = \nabla f(\lambda \mathbf{x}) \cdot \mathbf{x} \\ \frac{\partial f(\lambda \mathbf{x})}{\partial \lambda} = \frac{\partial \lambda^\alpha f(\mathbf{x})}{\partial \lambda} = \alpha \lambda^{\alpha - 1} f(\mathbf{x}) \end{cases} \tag{2.3.5}$$

$$\implies \nabla f(\lambda \mathbf{x}) \cdot \mathbf{x} = \alpha \lambda^{\alpha - 1} f(\mathbf{x}) \tag{2.3.6}$$

$$\implies \nabla f(\mathbf{x}) \cdot \mathbf{x} = \alpha f(\mathbf{x}) \text{ evaluated at } \lambda = 1 \tag{2.3.7}$$

■

**Definition 2.3.2.** Let $C$ be the level set of $f : S \to \mathbb{R}$ at $\mathbf{a} \in S$ defined as

$$C := \{\mathbf{x} \in S : f(\mathbf{x}) = f(\mathbf{a})\} \tag{2.3.8}$$

and a vector $\mathbf{v}$ is **tangent to $C$ at a** if there exists a function $\gamma : I \to C$ defined on interval $I$ containing 0, such that

$$\gamma(0) = \mathbf{a} \tag{2.3.9}$$

and

$$\mathbf{v} = \gamma'(0) \tag{2.3.10}$$

[1] Also all functions $\mathbf{f}$ and $\mathbf{g}$ and $\mathbf{f} \circ \mathbf{g}$ are well-defined near $\mathbf{a}$ and $\mathbf{g}(\mathbf{a})$.

**Theorem 2.3.3.** Let $S \subset \mathbb{R}^n$ be an open set, and $f : S \to \mathbb{R}$ is differentiable at $\mathbf{a}$. Then $\nabla f(\mathbf{a})$ is orthogonal to the level set of $f$ passes through $\mathbf{a}$.

*Proof Idea.* Let $\mathbf{v}$ be an arbitrary tangent vector to $C$ at $\mathbf{a}$, there must exists a function $\gamma : I \to C$. And define

$$h(t) := f \circ \gamma(t) \tag{2.3.11}$$

by definition of $\gamma$, $h(I) = \{f(\mathbf{a})\}$. Thus

$$\begin{aligned} \frac{d}{dt}h(t) &= \frac{d}{dt}f \circ \gamma(t) && (2.3.12)\\ &= \nabla f(\gamma(0)) \cdot \gamma'(0) && (2.3.13)\\ &= \nabla f(\mathbf{a}) \cdot \gamma'(0) && (2.3.14)\\ &= \nabla f(\mathbf{a}) \cdot \mathbf{v} = 0 && (2.3.15) \end{aligned}$$

So $\nabla f(\mathbf{a})$ is orthogonal to any tangent vector of $C$ at $\mathbf{a}$, which means $\nabla f(\mathbf{a})$ is orthogonal to $C$. ■

## 2.4 the Mean Value Theorem

**Theorem 2.4.1** (the Mean Value Theorem)**.** Assume $f : S \to \mathbb{R}$, where $S$ is a convex and open subset of $\mathbb{R}^n$, of class $C^1$, then

$$\forall \mathbf{a}, \mathbf{b} \in S,\ \exists \lambda \in [0,1]\ s.t.\ \mathbf{c} = \lambda \mathbf{a} + (1-\lambda)\mathbf{b} \tag{2.4.1}$$

$$\nabla f(\mathbf{c}) \cdot (\mathbf{b} - \mathbf{a}) = f(\mathbf{b}) - f(\mathbf{a}) \tag{2.4.2}$$

*Proof Idea.* Define $\gamma(t) := t\mathbf{a} + (1-t)\mathbf{b}$. Construct $h : [0,1] \to \mathbb{R}$ defined as $h := f(\gamma(t))$ then apply one dimensional mean value theorem on $h$. ■

**Definition 2.4.1.** A set $S \subset \mathbb{R}^n$ is **convex** if

$$\forall \mathbf{a}, \mathbf{b} \in S, \lambda \in [0,1],\ \lambda \mathbf{a} + (1-\lambda)\mathbf{b} \in S \tag{2.4.3}$$

**Theorem 2.4.2.** Assume that $S$ is an open and convex subset of $\mathbb{R}^n$ and that $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable in $S$ such that

$$||\nabla f(\mathbf{x})|| \leq M\ \forall \mathbf{x} \in S \tag{2.4.4}$$

then for every $\mathbf{a}, \mathbf{b} \in S$,

$$|f(\mathbf{b}) - f(\mathbf{a})| \leq M||\mathbf{b} - \mathbf{a}|| \tag{2.4.5}$$

*Proof Idea.* Use *Cauchy's Inequality.* ■

**Theorem 2.4.3.** Assume that $S$ is an open and convex subset of $\mathbb{R}^n$, and $f : S \to \mathbb{R}$ is a function differentiable on $S$. If $\nabla f(\mathbf{x}) = \mathbf{0}\ \forall \mathbf{x} \in S$, then $f$ is constant on $S$.

*Proof Idea.* Take two arbitrary $\mathbf{a}, \mathbf{b} \in S$, then use mean value theorem to show $f(\mathbf{a}) = f(\mathbf{b})$. ■

**Theorem 2.4.4.** Assume that $S$ is an open and path-connected subset of $\mathbb{R}^n$, and $f : S \to \mathbb{R}$ is a function differentiable on $S$. If $\nabla f(\mathbf{x}) = \mathbf{0}\ \forall \mathbf{x} \in S$, then $f$ is constant on $S$.

*Proof.* Any path-connected set can be written as a countable union of convex sets $S = \cup_{i \in \mathcal{A}} C_i$ such that

$$\forall \alpha \subset \mathcal{A}\ s.t.\ \alpha \neq \varnothing,\ \cup_{i \in \alpha} C_i \cap \cup_{i \in \alpha^c} C_i \neq \varnothing \tag{2.4.6}$$

then apply the previous theorem. ■

## 2.5 Higher Order Derivatives

**Definition 2.5.1.** A function $f$ defined on $S$ is **of class** $C^k$ if all of its $k^{th}$ order partial derivatives exists and continuous everywhere in $S$.

**Theorem 2.5.1.** Assume that $S$ is an open subset of $\mathbb{R}^n$ and that $f : S \to \mathbb{R}$ is $C^k$. Let $\alpha \in [n]^k$, and let $\beta$ be any permutation of $\alpha$,

$$\partial^\alpha f = \partial^\beta f \tag{2.5.1}$$

**Definition 2.5.2.** A **multi-index** is an $n$-tuple of nonnegative integers. And we define

$$\partial^\alpha f := (\frac{\partial}{\partial x_1})^{\alpha_1}(\frac{\partial}{\partial x_2})^{\alpha_2}\cdots(\frac{\partial}{\partial x_n})^{\alpha_n} f, \quad \text{for } \alpha = (\alpha_1, \ldots, \alpha_n) \tag{2.5.2}$$

where the **order** of multi-index is defined as

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n \tag{2.5.3}$$

**Theorem 2.5.2** (the Multinomial Theorem)**.**

$$(x_1 + x_2 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!}\mathbf{x}^\alpha \tag{2.5.4}$$

*Proof Idea.* Prove by induction on $n$, with Binomial Theorem. ■

**Definition 2.5.3.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^2$ function, then its **Hessian matrix** is defined as

$$H_f := \begin{pmatrix} \partial_1\partial_1 f & \cdots & \partial_n\partial_1 f \\ \vdots & \ddots & \vdots \\ \partial_1\partial_n f & \cdots & \partial_n\partial_n f \end{pmatrix} \tag{2.5.5}$$

## 2.6 Taylor's Theorem

**Definition 2.6.1.** Let $f : I \to \mathbb{R}$, where $I$ is an open subset of $\mathbb{R}$, be $C^k$. Let $a \in I$. then the $k^{th}$ **order Taylor polynomial of** $f$ **at** $a$ is the unique polynomial of order at most $k$, denoted $P_{a,k}(h)$ such that

$$f^{(j)}(a) = P^{(j)}_{a,k}(0)\ \forall j \in \{0, 1, \ldots, k\} \tag{2.6.1}$$

Note

$$P^{(j)}_{a,k}(h) = \sum_{j=0}^{k} \frac{h^j}{j!} f^{(j)}(a) \tag{2.6.2}$$

**Theorem 2.6.1** (Taylor's Theorem in 1 Dimension)**.** Assume that $I \subset \mathbb{R}$ is an open interval and that $f : I \to \mathbb{R}$ is a function of class $C^k$ on $I$. For $a \in I$ and $h \in \mathbb{R}$ such that $a + h \in I$. Define the **reminder**

$$R_{a,k}(h) := f(a + h) - P_{a,k}(h) \tag{2.6.3}$$

Then

$$\lim_{h\to 0} \frac{R_{a,k}(h)}{h^k} = 0 \tag{2.6.4}$$

**Proposition 2.6.1.** Assume that $I \subset \mathbb{R}$ is an open interval and that $f : I \to \mathbb{R}$ is a function of class $C^k$ on $I$. For $a \in I$ and $h \in \mathbb{R}$ such that $a + h \in I$, there exists some $\theta \in (0, 1)$ such that

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2}f''(a) + \cdots + \frac{h^{k-1}}{(k-1)!}f^{(k-1)}(a) + \frac{h^k}{k!}f^{(k)}(a+\theta h) \tag{2.6.5}$$

**Definition 2.6.2.** Assume that $S \subset \mathbb{R}^n$ is an open interval and that $f : S \to \mathbb{R}$ is a function of class $C^k$ on $S$. For a point $\mathbf{a} \in S$, the $k^{th}$ **order Taylor polynomial of** $f : S \to \mathbb{R}$ is a polynomial of order at most $k$, denoted $P_{\mathbf{a},k}(\mathbf{h})$ satisfying

$$f(\mathbf{a}) = P_{\mathbf{a},k}(\mathbf{0}) \tag{2.6.6}$$

$$\partial^\alpha f(\mathbf{a}) = \partial^\alpha P_{\mathbf{a},k}(\mathbf{0}) \ \forall \alpha \ s.t. \ |\alpha| \le k \tag{2.6.7}$$

**Theorem 2.6.2** (Taylor's Theorem in $n$ Dimensions)**.** Assume that $S \subset \mathbb{R}^n$ is an open set and that $f : S \to \mathbb{R}$ is a function of class $C^k$ on $S$. For $\mathbf{a} \in S$ and $\mathbf{h} \in \mathbb{R}^n$ such that $\mathbf{a} + \mathbf{h} \in S$. Define the **reminder**

$$R_{\mathbf{a},k}(\mathbf{h}) := f(\mathbf{x} + \mathbf{h}) - P_{\mathbf{a},k}(\mathbf{h}) \tag{2.6.8}$$

Then

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{R_{\mathbf{a},k}(\mathbf{h})}{||\mathbf{h}||^k} = 0 \tag{2.6.9}$$

**Theorem 2.6.3** (the Quadratic Case)**.**

$$P_{\mathbf{a},2}(\mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + \frac{1}{2}\mathbf{h}^T H_f(\mathbf{a})\mathbf{h} \tag{2.6.10}$$

$$\exists \theta \in (0,1) \ s.t. \ f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + \frac{1}{2}\mathbf{h}^T H_f(\mathbf{a}+\theta\mathbf{h})\mathbf{h} \tag{2.6.11}$$

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{R_{\mathbf{a},2}(\mathbf{h})}{||\mathbf{h}||^2} = 0 \tag{2.6.12}$$

**Definition 2.6.3** (the General Taylor's Polynomial)**.**

$$P_{\mathbf{a},k}(\mathbf{h}) = \sum_{\{\alpha : |\alpha| \le k\}} \frac{\mathbf{h}^\alpha}{\alpha!}\partial^\alpha f(\mathbf{a}) \tag{2.6.13}$$

## 2.7 Critical Points

**Definition 2.7.1.** A symmetric $n \times n$ matrix $A$ is said to be

- **Positive definite** if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$.
- **Non-negative definite** if $\mathbf{x}^T A\mathbf{x} \ge 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
- **Negative definite** if $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$.
- **Non-positive definite** if $\mathbf{x}^T A\mathbf{x} \le 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

and **indefinite** otherwise.

**Theorem 2.7.1.** Assume $A$ is a symmetric matrix. Then

$$\begin{aligned} A \text{ is positive definite} &\iff \text{all its eigenvalues are positive} \\ &\iff \exists \lambda_i > 0 \text{ such that } \mathbf{x}^T A\mathbf{x} \geq \lambda_i ||\mathbf{x}||^2 \text{ for all } \mathbf{x} \in \mathbb{R}^n. \end{aligned}$$

and

$$A \text{ is nonnegative definite} \iff \text{all its eigenvalues are nonnegative.} \tag{2.7.1}$$

and

$$A \text{ is indefinite} \iff A \text{ has both positive and negative eigenvalues.} \tag{2.7.2}$$

**Lemma 2.7.1.** Let $A$ be a symmetric matrix, then

$$\text{the smallest eigenvalue of } A = \min_{\{\mathbf{u} \in \mathbb{R}^n : |\mathbf{u}|=1\}} \mathbf{u}^T A\mathbf{u} \tag{2.7.3}$$

**Definition 2.7.2.** A point $\mathbf{a} \in S$ is a **local minimum point** for $f : S \to \mathbb{R}$ if

$$\exists \varepsilon > 0 \; s.t. \; \forall \mathbf{x} \in \mathcal{B}(\varepsilon, \mathbf{a}) \; f(\mathbf{a}) \leq f(\mathbf{x}) \tag{2.7.4}$$

**Definition 2.7.3.** A point $\mathbf{a} \in S$ is a **local maximum point** for $f : S \to \mathbb{R}$ if

$$\exists \varepsilon > 0 \; s.t. \; \forall \mathbf{x} \in \mathcal{B}(\varepsilon, \mathbf{a}) \; f(\mathbf{a}) \geq f(\mathbf{x}) \tag{2.7.5}$$

**Definition 2.7.4.** Let $f : S \to \mathbb{R}$ is differentiable on the open sub $S \subset \mathbb{R}^n$, then a point $\mathbf{a} \in S$ is a **critical point** if

$$\nabla f(\mathbf{a}) = \mathbf{0} \tag{2.7.6}$$

**Definition 2.7.5.** Let $\mathbf{a} \in S$ be a critical point of $f$, then $\mathbf{a}$ is a **saddle point** if $H_f(\mathbf{a})$ is indefinite.

**Theorem 2.7.2** (First Derivative Test)**.** If $f : S \to R$ is differentiable, then

$$\text{local extremum} \implies \text{critical point} \tag{2.7.7}$$

**Theorem 2.7.3** (Necessary Condition for a Local Minimum)**.** If $f : S \to \mathbb{R}$ is $C^2$ and $\mathbf{a}$ is a local minimum point for $f$, then

(i) $\mathbf{a}$ is critical point of $f$;

(ii) $H_f(\mathbf{a})$ is positive semi-definite.

**Theorem 2.7.4** (Sufficient Condition for a Local Minimum)**.** If

(i) $\mathbf{a}$ is a critical point of $f$;

(ii) $H_f(\mathbf{a})$ is positive definite.

Then $\mathbf{a}$ is a local minimum for $f$.

**Corollary 2.7.1.** Assume $f$ is $C^2$ and $\nabla f(\mathbf{a}) = \mathbf{0}$, then

(i) If $H_f(\mathbf{a})$ is positive definite, then $\mathbf{a}$ is a local minimum;

(ii) If $H_f(\mathbf{a})$ is negative definite, then $\mathbf{a}$ is a local maximum;

(iii) If $H_f(\mathbf{a})$ is indefinite, then $\mathbf{a}$ is a saddle point.

*If none of the above hold, then we cannot determine the character of the critical point without further thought.*

**Definition 2.7.6.** A critical point $\mathbf{a}$ of $f$ is **degenerate** if $\det H_f(\mathbf{a}) = 0$, and **non-degenerate** if $\det H_f(\mathbf{a}) \neq 0$.

## 2.8 Optimization

**Theorem 2.8.1.** Let $S \subset \mathbb{R}^n$ be an open set and $f, g : S \to \mathbb{R}$ be $C^1$ functions. If $\mathbf{x}$ is a *local extremal* satisfying $g(\mathbf{x}) = 0$, and $\nabla g(\mathbf{x}) \neq 0$, then

$$\exists \lambda \in \mathbb{R} \; s.t. \begin{cases} \nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}) \\ g(\mathbf{x}) = 0 \end{cases} \tag{2.8.1}$$

**Lemma 2.8.1.** $\nabla g(\mathbf{x})$ is <u>orthogonal</u> to the constraint set $g^{-1}(0)$.

**Proposition 2.8.1.** Equations (2.8.1) $\implies \nabla f(\mathbf{x}) \perp g^{-1}(0)$ at $\mathbf{x}$.

**Theorem 2.8.2.** Let $S \subseteq \mathbb{R}^n$ be an open set, and $f, \{g_i\}_{i=1}^k : S \to \mathbb{R}$ be $C^1$ functions. Define $\mathbf{g}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^k \equiv (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_k(\mathbf{x}))$.
If $\mathbf{x} \in S$ is a *local extremal* of $f$ such that $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, and $\{\nabla g_i(\mathbf{x})\}$ are <u>linearly independent</u> (i.e. $rank(D\mathbf{g}(\mathbf{x})) = k$), then

$$\exists \boldsymbol{\lambda} \in \mathbb{R}^k \; s.t. \begin{cases} \nabla f(\mathbf{x}) = \boldsymbol{\lambda}^T D\mathbf{g}(\mathbf{x}) \\ \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{cases} \tag{2.8.2}$$

**Remark 2.8.1.** Procedure of optimization on *open sets*:

(i) Find all critical points.

(ii) Find optimizers among critical points.

**Remark 2.8.2.** Procedure of optimization with *inequality constraints*:

(i) Find critical points without the constraints.

(ii) Find critical points on the constraints.

(iii) Find optimizers among candidates.

# 3 The Implicit and Inverse Function Theorems

## 3.1 The Implicit Function Theorem I

**Theorem 3.1.1** (Implicit Function Theorem)**.** Let $S \subseteq \mathbb{R}^{n+k}$ be an open set, and function $F : S \to \mathbb{R}^k$ be a $C^1$ function. Suppose there exists point $\mathbf{a} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^k$ such that

$$F(\mathbf{a}, \mathbf{b}) = \mathbf{0} \tag{3.1.1}$$

If

$$det(D_{\mathbf{y}}(F(\mathbf{a}, \mathbf{b}))) \neq 0 \tag{3.1.2}$$

then there exists $r_0, r_1 > 0$ and a $C^1$ function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^k$ such that

$$\forall \mathbf{x} \in \mathcal{B}(r_0, \mathbf{a}), \;\; \mathbf{f}(\mathbf{x}) \in \mathcal{B}(r_1, \mathbf{b}) \wedge F(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{0} \tag{3.1.3}$$

and define $\mathbf{y} \equiv \mathbf{f}(\mathbf{x})$, the derivative of $\mathbf{f}$ can be found as

$$D\mathbf{f}(\mathbf{x}) = -[D_{\mathbf{y}} F(\mathbf{x}, \mathbf{y})]^{-1} D_{\mathbf{x}} F(\mathbf{x}, \mathbf{y}) \tag{3.1.4}$$

**Remark 3.1.1.** Procedure to prove solvability of non-linear equations

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0} \tag{3.1.5}$$

near $(\mathbf{a}, \mathbf{b})$.

(i) Verify $\mathbf{F}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$.

(ii) Assert

$$det(D_{\mathbf{y}}\mathbf{F}(\mathbf{a}, \mathbf{b})) \neq 0 \tag{3.1.6}$$

(iii) Approximate solution $\mathbf{y} = \mathbf{f}(\mathbf{x})$ using

$$\mathbf{f}(\mathbf{x} + \mathbf{h}) \approx \mathbf{a} + D\mathbf{f}(\mathbf{a})\mathbf{h} \tag{3.1.7}$$

$$= \mathbf{a} - [D_{\mathbf{y}}\mathbf{F}(\mathbf{a}, \mathbf{b})]^{-1} D_{\mathbf{x}}\mathbf{F}(\mathbf{a}, \mathbf{b}) \tag{3.1.8}$$

## 3.2 Geometric content of the Implicit Function Theorem

**Definition 3.2.1.** Let $S \subseteq \mathbb{R}^n$ and $\mathbf{a} \in S$. $S$ is **singular** at $\mathbf{a}$ if

$$\forall r > 0 \ S \cap \mathcal{B}(r, \mathbf{a}) \text{ cannot be represented as a } C^1 \text{ graph.} \tag{3.2.1}$$

$S$ is **regular** at $\mathbf{a}$ is its not singular there.

**Theorem 3.2.1** ($k$ dimensional manifold as level set)**.** Let $U \subseteq \mathbb{R}^n$ and let $\mathbf{F} : U \to \mathbb{R}^{n-k}$ be a $C^1$ function.

$$S \equiv \mathbf{F}^{-1}(\mathbf{0}) \tag{3.2.2}$$

Let $\mathbf{a} \in U$, if

$$rank(D\mathbf{F}(\mathbf{a})) = n - k \tag{3.2.3}$$

then $\exists r > 0$ such that the *level set of* $\boldsymbol{F}$ *near* $\boldsymbol{a}$

$$\mathcal{B}(r, \mathbf{a}) \cap S \tag{3.2.4}$$

can be represented as a $C^1$ graph.

**Theorem 3.2.2** ($k$ dimensional manifold as parameterization)**.** Let $T \subseteq \mathbb{R}^k$ and let $\mathbf{f} : U \to \mathbb{R}^n$ be a $C^1$ function.

$$S \equiv \mathbf{f}(T) \tag{3.2.5}$$

Let $\mathbf{t} \in T$, if

$$rank(\mathbf{f}(\mathbf{t})) = k \tag{3.2.6}$$

then $\exists r > 0$ such that the *parameterization of* $\boldsymbol{f}$ *near* $\boldsymbol{t}$

$$\mathbf{f}(T \cap \mathcal{B}(r, \mathbf{t})) \tag{3.2.7}$$

can be represented as a $C^1$ graph.

### 3.3 Transformations, and the Inverse Function Theorem

**Example 3.3.1** (Polar coordinate in $\mathbb{R}^2$)**.** Let

$$U \equiv \{(r, \theta) : r > 0 \wedge \theta \in (-\pi, \pi)\} \tag{3.3.1}$$

$$V \equiv \mathbb{R}^2 \backslash \{(x, 0) : x \leq 0\} \tag{3.3.2}$$

Define $\mathbf{f} : U \to V$ as

$$\mathbf{f}(r, \theta) \equiv \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix} \tag{3.3.3}$$

**Example 3.3.2** (Spherical coordinate in $\mathbb{R}^3$)**.** Define

$$\mathbf{f}(r, \theta, \varphi) = \begin{pmatrix} r\cos(\theta)\sin(\varphi) \\ r\sin(\theta)\sin(\varphi) \\ r\cos(\varphi) \end{pmatrix} \tag{3.3.4}$$

**Example 3.3.3** (Cylindrical coordinate in $\mathbb{R}^3$)**.** Define

$$\mathbf{f}(r, \theta, z) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ z \end{pmatrix} \tag{3.3.5}$$

**Theorem 3.3.1** (Inverse Function Theorem)**.** Let $U$ and $V$ be open subsets in $\mathbb{R}^n$, and $\mathbf{f} : U \to V$. Let $\mathbf{a} \in U$ and define $\mathbf{b} \equiv \mathbf{f}(\mathbf{a}) \in V$. If

$$det(D\mathbf{f}(\mathbf{a})) \neq 0 \tag{3.3.6}$$

then there exists $M \subseteq U$ and $N \subseteq V$ such that

(i) $\mathbf{a} \in M$ and $\mathbf{b} \in N$,

(ii) $\mathbf{f}$ is bijective between $M$ and $N$,

(iii) $\mathbf{f}^{-1} : N \to M$ is $C^1$,

and for all $\mathbf{x} \in M$ such $\mathbf{y} \equiv \mathbf{f}(\mathbf{x}) \in N$,

$$D\mathbf{f}^{-1}(\mathbf{y}) = [D\mathbf{f}(\mathbf{x})]^{-1} \tag{3.3.7}$$

# 4 Integration

## 4.1 Basics

**Theorem 4.1.1** (Properties of infimum and supremum)**.** Let $A \subseteq \mathbb{R}^n$ and $A \neq \emptyset$, and $f, g : A \to \mathbb{R}$ are bounded functions. Let $m$ and $M$ denote the infimum and supremum respectively, then

(i) $m_A f + m_A g \leq m_A(f+g) \leq M_A(f+g) \leq M_A f + M_A g$

(ii) If $A' \subseteq A$, then $m_A f \leq m_{A'} f \leq M_{A'} f \leq M_A f$

(iii) If $f(\mathbf{x}) \leq g(\mathbf{x})\ \forall \mathbf{x} \in A$, then $m_A f \leq m_A g$ and $M_A f \leq M_A g$

(iv) $|M_A f| \leq M_A |f|$

(v) $M_A|f| - m_A|f| \leq M_A f - m_A f$

(vi) $\forall c \in \mathbb{R}$, $M_A(cf) - m_A(cf) = |c|(M_A f - m_A f)$

(vii) $M_A f - m_A f = sup\{f(x) - f(y) : x, y \in A\}$

## 4.2 Integration on Higher Dimensions

**Definition 4.2.1.** A **rectangle** $\mathcal{R} \subseteq \mathbb{R}^n$ is defined as

$$\mathcal{R} \equiv \prod_{i=1}^{n} [a_i, b_i] \tag{4.2.1}$$

where $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$.

**Definition 4.2.2.** A **partition** $P$ of rectangle $\mathcal{R} = \prod_{i=1}^{n} [a_i, b_i]$ is a list of $n$ finite and increasing list of real numbers

$$P = \{L_1, L_2, \dots, L_n\} \tag{4.2.2}$$

where $L_i = \{e_j\}_{j=0}^{T_i}$ such that

$$a_i = e_0 < e_1 < \cdots < e_{T_i} = b_i \tag{4.2.3}$$

and such partition induces a set of rectangles(boxes) $\mathcal{B}(P) \equiv \{B_j\}_{j=1}^{J} \subseteq \mathcal{R}$.

**Definition 4.2.3.** Let $P$ and $P'$ be two partitions of $\mathcal{R}$. Then $P'$ is a **refinement** of $P$ if

$$\forall B_j \in \mathcal{B}(P), B'_j \in \mathcal{B}(P') \quad B'_j \subseteq B_j \vee B_j'^{int} \cap B_j^{int} = \emptyset \tag{4.2.4}$$

**Definition 4.2.4.** Define the **volume** of rectangle $\mathcal{R} = \prod_{i=1}^{n} [a_i, b_i]$ as

$$V^n(\mathcal{R}) \equiv \prod_{i=1}^{n} (b_i - a_i) \tag{4.2.5}$$

**Definition 4.2.5.** The **lower Riemann sum** of $f$ with partition $P$ on $\mathcal{R}$ is defined as

$$L_P f \equiv \sum_{B_j \in \mathcal{B}(P)} \inf_{\mathbf{x} \in B_j} f(\mathbf{x}) V^n(B_j) \tag{4.2.6}$$

and the **upper Riemann sum** is defined as

$$U_P f \equiv \sum_{B_j \in \mathcal{B}(P)} \sup_{\mathbf{x} \in B_j} f(\mathbf{x}) V^n(B_j) \tag{4.2.7}$$

**Definition 4.2.6.** The **upper integral** and **lower integral** of $f$ on $\mathcal{R}$ are defined as

$$\bar{I}_{\mathcal{R}} f \equiv \inf_P U_P f \tag{4.2.8}$$

$$\underline{I}_{\mathcal{R}} f \equiv \sup_P L_P f \tag{4.2.9}$$

**Definition 4.2.7.** A bounded real-valued function $f$ defined on $\mathcal{R}$ is **integrable** if

$$\underline{I}_{\mathcal{R}} f = \bar{I}_{\mathcal{R}} f \tag{4.2.10}$$

and the integral is defined as

$$\int \cdots \int_{\mathcal{R}} f \; dV^n \equiv \underline{I}_{\mathcal{R}} f = \bar{I}_{\mathcal{R}} f \tag{4.2.11}$$

**Lemma 4.2.1.** Let $f$ be a bounded real-valued function defined on $\mathcal{R}$, $f$ is integrable if and only if $\forall \epsilon > 0$, there exists a partition $P$ of $\mathcal{R}$ such that

$$U_P f - L_P f < \epsilon \tag{4.2.12}$$

**Theorem 4.2.1.** Let $f$ and $g$ be two integrable functions on $\mathcal{R} \subseteq \mathbb{R}^n$, let $c \in \mathbb{R}$,

(i) $f + g : \mathcal{R} \to \mathbb{R}$ is integrable and $\int_{\mathcal{R}} (f + g) = \int_{\mathcal{R}} f + \int_{\mathcal{R}} g$

(ii) $c \cdot f$ is integrable and $\int_{\mathcal{R}} c \cdot f = c \int_{\mathcal{R}} f$

(iii) $f(\mathbf{x}) \geq g(\mathbf{x})\ \forall \mathbf{x} \in \mathcal{R} \implies \int_{\mathcal{R}} f \geq \int_{\mathcal{R}} g$

(iv) $|f|$ is integrable and $|\int_R f| \leq \int_R |f|$

**Definition 4.2.8.** Let $S \subseteq \mathbb{R}^n$ be a bounded set, and there exists rectangle $\mathcal{R}$ covers $S$, the **indicator function** of $S$ is $\chi_S : \mathcal{R} \to \{0, 1\}$, defined as

$$\chi_S(\mathbf{x}) \equiv \mathbb{I}(\mathbf{x} \in S) \tag{4.2.13}$$

**Definition 4.2.9.** Let $S \subseteq \mathbb{R}^n$ be a bounded set, and there exists rectangle $\mathcal{R}$ covers $S$. Let $f : \mathcal{R} \to \mathbb{R}$ be a bounded function, then $f$ is **integrable on** $S$ if $\chi_S f$ is integrable on $\mathcal{R}$. And

$$\int \cdots \int_S f\ dV^n \equiv \int \cdots \int_{\mathcal{R}} \chi_S f\ dV^n \tag{4.2.14}$$

**Definition 4.2.10.** Let $Z \subseteq \mathbb{R}^n$, $Z$ has **zero content** if for all $\epsilon > 0$, there exists a <u>finite</u> set of rectangles $\{R_\ell\}_{\ell=1}^L$ covers $Z$ and

$$\sum_{\ell=1}^{L} V^n(R_\ell) < \epsilon \tag{4.2.15}$$

**Proposition 4.2.1.** Let $Z \subseteq \mathbb{R}^n$ has zero content, then

(i) For any $Z' \subseteq Z$, $Z'$ has zero content.

(ii) <u>Finite</u> union of content zero sets has zero content.

(iii) Let $f : [a, b] \to \mathbb{R}$ be an integrable function, it's graph $\{(x, f(x)) : x \in [a, b]\}$ has zero content.

(iv) Let $\mathbf{f} : [a, b] \to \mathbb{R}^2$ be a $C^1$ function, the parameterization $\mathbf{f}([a, b])$ has zero content.

**Theorem 4.2.2.** Let $\mathcal{R}$ be a rectangle in $\mathbb{R}^n$ and $f$ is integrable on $\mathcal{R}$ if

$$\{\mathbf{x} \in \mathcal{R} : f \text{ is discontinuous at } \mathbf{x}\} \tag{4.2.16}$$

has zero content.

**Proposition 4.2.2** (Folland 4.22)**.** Suppose $Z \subseteq \mathbb{R}^n$ has zero content. If $f : \mathbb{R}^n \to \mathbb{R}$ is bounded, then $f$ is integrable on $Z$ and $\int_Z f\ dV^n = 0$.

## 4.3 Iterated Integrals

**Theorem 4.3.1** (Fubini's Theorem)**.** Let $\mathcal{R} = [a, b] \times [c, d] \subseteq \mathbb{R}^2$ and $f : \mathcal{R} \to \mathbb{R}$ is bounded. Assuming that

(i) $f$ is integrable on $\mathcal{R}$.

(ii) for each $y \in [c, d]$, the function $f_y(x) \equiv f(x, y)$ is integrable on $[a, b]$.

(iii) Define $g(y) \equiv \int_a^b f(x, y) dy$ is integrable on $[c, d]$.

Then

$$\iint_{\mathcal{R}} f\ dA = \int_c^d \Big( \int_a^b f(x, y)\ dx \Big) dy \tag{4.3.1}$$

**Proposition 4.3.1.** Let $S \subseteq \mathbb{R}^n$ be an unbounded set, and $f : S \to \mathbb{R}$. Then improper integral $\int \cdots \int_S f\ d^n\mathbf{x}$ is absolutely convergent on $\mathbb{R}^n$ <u>if and only if</u> $\int \cdots \int_{\mathbb{R}^n} \chi_S f\ d^n\mathbf{x}$ is absolutely convergent.

## 4.4 Change of Variables

**Theorem 4.4.1** (Change of Variable)**.** Let $U$ and $V$ be two open subsets of $\mathbb{R}^n$, and let $\mathbf{G} : U \to V$ be a $C^1$ bijection. Let $T \subset U$ and $S \subset V$. Suppose $\mathbf{G}(T) = S$, then

$$\int \cdots \int_S f \ d\Omega = \int \cdots \int_T f \circ \mathbf{G} \ |\mathrm{det} D\mathbf{G}| \ d\Theta \tag{4.4.1}$$

**Corollary 4.4.1.** Let $S$ be a region in $\mathbb{R}^n$, suppose $S$ can be parameterized by $\mathbf{G} : T \to S$. By the change of variable formula, consider the special case $f(\mathbf{x}) = 1$,

$$|S| = \int \cdots \int_S 1 \ d\Omega = \int \cdots \int_T 1 \ |\mathrm{det} \ D\mathbf{G}(\mathbf{u})| \ d\Theta \tag{4.4.2}$$

**Example 4.4.1** (Polar Coordinate)**.** Define the coordinate transformation mapping from polar to Cartesian,

$$\mathbf{P}(r, \theta) \equiv (x, y) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}, \ \theta \in [0, 2\pi] \ r \in \mathbb{R}_+ \tag{4.4.3}$$

and $|\mathrm{det} \ D\mathbf{P}(r, \theta)| = r$.

**Example 4.4.2** (Cylindrical Coordinate)**.** Define the coordinate transformation mapping from cylindrical to Cartesian as

$$\mathbf{C}(r, \theta, z) \equiv (x, y, z) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ z \end{pmatrix}, \ \theta \in [0, 2\pi] \ r \in \mathbb{R}_+ \ z \in \mathbb{R} \tag{4.4.4}$$

and $|\mathrm{det} \ D\mathbf{C}(r, \theta, z)| = r$.

**Example 4.4.3** (Spherical Coordinate)**.** Define the coordinate transformation mapping from spherical to Cartesian as

$$\mathbf{S}(r, \theta, \varphi) = \begin{pmatrix} r\cos\theta\sin\varphi \\ r\sin\theta\sin\varphi \\ r\cos\varphi \end{pmatrix} \tag{4.4.5}$$

and $|\mathrm{det} \ D\mathbf{S}(r, \theta, \varphi)| = r^2 \sin\varphi$

## 4.5 Further Aspects

### 4.5.1 Exchanging Differentiation and Integration

**Theorem 4.5.1** (Exchanging Differentiation and Integration)**.** Let $f(\mathbf{x}, \mathbf{t}) : S \times T \to \mathbb{R}$ and define $F(\mathbf{x}) : S \to \mathbb{R}$ as

$$F(\mathbf{x}) \equiv \int \cdots \int_T f(\mathbf{x}, \mathbf{t}) \ d\Omega \tag{4.5.1}$$

If

(i) $S$ is open and $T$ is compact and bounded;

(ii) $f$ and $F$ are continuous on their domains;

(iii) and $\forall \ x_j \in \mathbf{x}$, $\frac{\partial f(\mathbf{x}, \mathbf{t})}{\partial x_j}$ is continuous, then $F$ is $C^1$ in $S$ and for every $j$,

$$\frac{\partial F(\mathbf{x})}{\partial x_j} = \int \cdots \int_T \frac{\partial f(\mathbf{x}, \mathbf{t})}{\partial x_j} \, d\Omega \tag{4.5.2}$$

**Corollary 4.5.1.** By the definition of partial derivative, above theorem is equivalent to

$$\lim_{h\to 0} \int \cdots \int_T \frac{f(\mathbf{x} + h\mathbf{e}_j, \mathbf{t})}{h} \, d\Omega = \int \cdots \int_T \lim_{h\to 0} \frac{f(\mathbf{x} + h\mathbf{e}_j, \mathbf{t})}{h} \, d\Omega \tag{4.5.3}$$

### 4.5.2 Improper Integrals

**Definition 4.5.1** (Unbounded Domains)**.** An **improper integral** with unbounded domain $\int \cdots \int_{\mathbb{R}^n} f \, d\Omega$ is **absolutely convergent** if there exists $L \in \mathbb{R}$ such that

$$\forall \varepsilon > 0 \; \exists R > 0 \; s.t. \; \forall S \subseteq \mathbb{R}^n \; B(R, \mathbf{0}) \subset S \implies \left| \int \cdots \int_S f \, d\Omega - L \right| < \varepsilon \tag{4.5.4}$$

**Theorem 4.5.2.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function, and that

$$\lim_{R\to\infty} \int \cdots \int_{B(R,\mathbf{0})} |f| \, d\Omega \text{ exists} \tag{4.5.5}$$

then $\int \cdots \int_{\mathbb{R}^n} f \, d\Omega$ is absolutely convergent.

**Corollary 4.5.2** (Equivalence)**.** Above improper integral $\int \cdots \int_{\mathbb{R}^n} f \, d\Omega$ is absolutely convergent if set

$$\{\int \cdots \int_{B(R,\mathbf{0})} |f| \, d\Omega : R \in \mathbb{R}_{++}\} \tag{4.5.6}$$

is bounded.

**Corollary 4.5.3.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be an continuous function, if

$$\exists p > n, \; C > 0 \; s.t. \; |f(\mathbf{x})| \leq \frac{1}{||\mathbf{x}||^p} \; \forall \mathbf{x} \in \mathbb{R}^n \tag{4.5.7}$$

then $\int \cdots \int_{\mathbb{R}^n} f \, d\Omega$ is absolutely convergent.

**Definition 4.5.2** (Unbounded Function)**.** Let $S \subset \mathbb{R}^n$, $\mathbf{a} \in \mathbb{R}^n$. Consider a function $f : S \backslash \{\mathbf{a}\} \to \mathbb{R}$. Then the improper integral $\int \cdots \int_S f d\Omega$ is absolutely convergent if

$$\exists L \in \mathbb{R} \; s.t \; \forall \varepsilon > 0 \; \exists r > 0 \; s.t. \; \forall U \subset S \; s.t. \; \mathbf{a} \in U^{int} \wedge U \subset B(r, \mathbf{a}), \; \left| \int \cdots \int_{S\backslash U} f \, d\Omega - L \right| < \varepsilon \tag{4.5.8}$$

**Theorem 4.5.3.** Let $f : S \backslash \{\mathbf{a}\} \to \mathbb{R}$, if

$$\lim_{r\to 0} \int \cdots \int_{S\backslash B(r,\mathbf{a})} |f| \, d\Omega \text{ exists} \tag{4.5.9}$$

then $\int \cdots \int_S f \, d\Omega$ is absolutely convergent.

**Corollary 4.5.4** (Equivalence)**.** If the set

$$\{\iint_{S\backslash B(r,\mathbf{a})} |f| \, d\Omega : r \in \mathbb{R}_{++}\} \tag{4.5.10}$$

is bounded, then $\int \cdots \int_S f \, d\Omega$ is absolutely convergent.

**Corollary 4.5.5.** Let $f : S \backslash \{\mathbf{a}\} \to \mathbb{R}$, if

$$\exists p < n, \; C > 0 \; s.t. |f(\mathbf{x})| \leq \frac{C}{||\mathbf{x} - \mathbf{a}||^p} \; \forall \mathbf{x} \in S \backslash \{\mathbf{a}\} \tag{4.5.11}$$

then the improper integral $\int \cdots \int_S f \, d\Omega$ is absolutely convergent.

# 5 Vector Calculus

## 5.1 Line Integrals

### 5.1.1 Arc Length

**Definition 5.1.1.** Let $C$ be a smooth curve in $\mathbb{R}^n$ parameterized by $C^1$ function $\mathbf{g}$ such that $\mathbf{g}'(t) \neq \mathbf{0}$ for every appropriate $t$.

$$C \equiv \{\mathbf{g}(t) : t \in [a, b]\} \tag{5.1.1}$$

and the **arc length** of $C$ is defined as

$$\int_C d^n\mathbf{x} \equiv \int_C ds \equiv \int_a^b ||\mathbf{g}'(t)||\ dt \tag{5.1.2}$$

**Proposition 5.1.1.** The arc length of a curve $C$ is an intrinsic property of the geometric object $C$ and should not depend on the particular parameterization we use.

*Proof.* Let $\varphi : [c, d] \to [a, b]$ be a bijection, so that $\mathbf{h} \equiv \mathbf{g} \circ \varphi$ is also a valid parameterization of $C$ such that

$$C \equiv \{\mathbf{h}(u) : u \in [c, d]\} \tag{5.1.3}$$

The arc length of $C$ can be computed using

$$\int_C ds = \int_c^d ||\mathbf{h}'(u)||\ du \tag{5.1.4}$$

$$= \int_c^d ||\mathbf{g}'(\varphi(u))|| \times ||\varphi'(u)||\ du \tag{5.1.5}$$

$$= \int_a^b ||\mathbf{g}'(t)||\ dt \text{ by change of variable formula.} \tag{5.1.6}$$

■

**Remark 5.1.1** (Interpretations)**.** Suppose $\mathbf{g}$ is a parameterization of $C$.

(i) $\int_a^b \mathbf{g}'(t)\ dt = \mathbf{g}(b) - \mathbf{g}(a)$ measures the distance between two endpoints of $C$.

(ii) Choosing a parameterization is effectively choosing an **orientation** for the curve $C$.

**Definition 5.1.2.** A function $\mathbf{g} : [a, b] \to \mathbb{R}^n$ is called **piecewise smooth** if

(i) it's *continuous*, and

(ii) it's derivate exists and is continuous except at finitely many points $t_j$, at which the one-sided limits exists.

### 5.1.2 Line Integrals of Scalar Functions

**Definition 5.1.3.** Let smooth curve $C \subseteq \mathbb{R}^n$, $f : C \to \mathbb{R}$ and $\mathbf{g}$ be a parameterization of $C$, then

$$\int_C f\ ds = \int_a^b f(\mathbf{g}(t))\ ||\mathbf{g}'(t)||\ dt \tag{5.1.7}$$

**Remark 5.1.2.** The line integrals of scalar functions are also independent from the choices of parameterizations.

**Definition 5.1.4.**

$$\text{Average of } f \text{ over } C \equiv \frac{\int_C f\ ds}{\int_C\ ds} \tag{5.1.8}$$

### 5.1.3 Line Integrals of Vector Fields

**Definition 5.1.5.** Let smooth $C \in \mathbb{R}^n$ with parameterization $\mathbf{g}$ and $\mathbf{F} : C \to \mathbb{R}^n$ defined on it, the **line integral** of $\mathbf{F}$ over $C$ is defined as

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_a^b \mathbf{F}(\mathbf{g}(t)) \cdot \mathbf{g}'(t)\ dt \tag{5.1.9}$$

**Proposition 5.1.2.** The line integral $\int_C \mathbf{F} \cdot d\mathbf{x}$ is independent of the parameterization *as long as the orientation is unchanged.*

**Theorem 5.1.1** (The Fundamental Theorem of Line Integral)**.** Let $f : C \to \mathbb{R}$ defined on smooth curve $C$ parameterized by $\mathbf{g} : [a, b] \to \mathbb{R}^n$, then

$$\int_C \nabla f(\mathbf{x}) \cdot d^n\mathbf{x} = f(\mathbf{g}(b)) - f(\mathbf{g}(a)) \tag{5.1.10}$$

*Proof.*

$$\int_C \nabla f(\mathbf{x}) \cdot d^n\mathbf{x} = \int_a^b \frac{\partial f(\mathbf{g}(t))}{\partial \mathbf{g}(t)} \cdot \mathbf{g}'(t)\ dt \tag{5.1.11}$$

$$= \int_a^b \frac{\partial f(\mathbf{g}(t))}{\partial t}\ dt = f(\mathbf{g}(b)) - f(\mathbf{g}(a)) \tag{5.1.12}$$

■

### 5.1.4 Rectifiable Curves

**Remark 5.1.3.** Let $C$ be a curve in $\mathbb{R}^n$ parameterized by injection $\mathbf{g} : [a, b] \to \mathbb{R}^n$ such that $\mathbf{g}'(t) \neq \mathbf{0}$. Let $P$ be a partition of $[a, b]$. Denote

$$L_P(C) \equiv \sum_j ||\mathbf{g}(t_j) - \mathbf{g}(t_{j-1})|| \tag{5.1.13}$$

**Definition 5.1.6.** A curve $C$ is **rectifiable** if the set $\{L_P(C) : P\}$ is bounded. And the arc length of $C$ s defined as

$$L(C) \equiv \sup\{L_P(C) : P\} \tag{5.1.14}$$

**Theorem 5.1.2.** The supremum found above, $L(C)$ is the precisely the arc length of $C$:

$$L(C) = \int_a^b ||\mathbf{g}'(t)||\ dt \tag{5.1.15}$$

## 5.2 Green's Theorem

### 5.2.1 Preliminary Definitions

**Definition 5.2.1.** A **simple closed curve** is a curve with parameterization $\mathbf{g} : [a, b] \to \mathbb{R}^n$ where

(i) $\mathbf{g}$ is continuous;

(ii) $\mathbf{g}(a) = \mathbf{g}(b)$;

(iii) $\mathbf{g}$ is injective with its domain restricted to $(a, b)$.

**Definition 5.2.2.** A *simple closed curve* is **piecewise smooth** if it has a parameterization $\mathbf{g}$ such that

(i) $\mathbf{g}$ is continuously differentiable with $\mathbf{g}'(t) \neq \mathbf{0}$ except finitely many breakpoints;

(ii) $\mathbf{g}'(t)$ is *one side continuous* at breakpoints of the curve.

**Definition 5.2.3.** A **regular region** $S \subseteq \mathbb{R}^n$ is a set satisfying both

(i) $S$ is compact;

(ii) $\overline{S^{int}} = S$.

**Definition 5.2.4.** Let $S \subseteq \mathbb{R}^2$, $S$ has **piecewise smooth boundary** if $\partial S$ consists of one or more *disjoint, piecewise smooth, simple closed curve.*

**Definition 5.2.5.** Let $S \subseteq \mathbb{R}^2$, then **positive orientation** on $\partial S$ is the orientation on each of the closed curves that make up the boundary such that the region is on the *left* with respect to the positive direction on the curve.

**Theorem 5.2.1** (Green's Theorem)**.** Suppose $S \subseteq \mathbb{R}^2$ is a regular region with piecewise smooth region $\partial S$. Suppose $\mathbf{F}$ is a $C^1$ vector field defined on $\overline{S}$, then

$$\int_{\partial S} \mathbf{F} \cdot d\mathbf{x} = \iint_S \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) dA \tag{5.2.1}$$

**Corollary 5.2.1.** Suppose $S$ is a regular region in $\mathbb{R}^2$ with piecewise smooth boundary $\partial S$, and let $\mathbf{n}(\mathbf{x})$ be the *unit outward normal* vector to $\partial S$ at $\mathbf{x} \in \partial S$. Suppose also that $\mathbf{F}$ is a vector field defined on $\overline{S}$, then

$$\int_{\partial S} \mathbf{F} \cdot \mathbf{n}\ ds = \iint_S \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2}\right) dA \tag{5.2.2}$$

*Proof.* Let $\mathbf{g}(t)$ be a parameterization of boundary $\partial S$. Then the tangent vector would be $\mathbf{g}'(t)$ and we can conclude the *outer normal vector* $\mathbf{n}$ is $\frac{(g_2'(t), -g_1'(t))}{||(g_2'(t), -g_1'(t))||}$. Then

$$\int_{\partial S} \mathbf{F} \cdot \mathbf{n}\ ds = \int_T \mathbf{F} \circ \mathbf{g} \cdot \frac{(g_2'(t), -g_1'(t))}{||(g_2'(t), -g_1'(t))||} ||\mathbf{g}'(t)||\ dt \tag{5.2.3}$$

$$= \int_T F_1 g_2'(t) - F_2 g_1'(t)\ dt \tag{5.2.4}$$

$$= \int_T \begin{pmatrix} -F_2 \\ F_1 \end{pmatrix} \cdot \begin{pmatrix} g_1'(t) \\ g_2'(t) \end{pmatrix}\ dt \tag{5.2.5}$$

$$= \int_{\partial S} \begin{pmatrix} -F_2 \\ F_1 \end{pmatrix} \cdot\ d^2\mathbf{x} \tag{5.2.6}$$

$$= \iint_S \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y}\ dA \text{ By Green's Theorem} \tag{5.2.7}$$

■

## 5.3 Surface Integrals

### 5.3.1 Surface Areas and Surface Integrals

**Definition 5.3.1.** Suppose $S$ is a surface in $\mathbb{R}^3$ and parameterized by

$$\mathbf{G}(\mathbf{u}) : R \to S \tag{5.3.1}$$

where $rank(D\mathbf{G}(\mathbf{u})) = 2$ for every $\mathbf{u} \in R \backslash Z$ where $Z$ is a probably empty set with zero content. If $||\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}||$ is integrable, then

$$Area(S) \equiv \iint_R ||\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}||\ d\Theta \tag{5.3.2}$$

**Definition 5.3.2.** Let $f : S \to \mathbb{R}$ be a real-valued continuous function defined on a super set of $S$, the **integral of a real-valued function on a surface** is defined as

$$\iint_S f(\mathbf{x})\ dA \equiv \iint_R f(\mathbf{G}(\mathbf{u}))||\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}||\ d\Theta \tag{5.3.3}$$

**Definition 5.3.3.** Let $\mathbf{F} : S \to \mathbb{R}^3$ be a continuous vector field defined on a super set of $S$, the **integral of vector field on a surface** is defined as

$$\iint_S \mathbf{F}(\mathbf{x}) \cdot \mathbf{n}\ dA \equiv \iint_R \mathbf{F}(\mathbf{G}(\mathbf{u})) \cdot (\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v})\ d\Theta \tag{5.3.4}$$

**Remark 5.3.1.** Surface integrals of real-valued functions are independent of the choice of parametrization.

**Remark 5.3.2.** But the choice of parameterization can change the sign of surface integrals of vector fields. We need to choose the direction of the normal, $\mathbf{n}$.

**Definition 5.3.4.** Let $S \subseteq \mathbb{R}^3$ be a two dimensional sub-manifold, and $f$ is a real-valued function defined on a super set of $S$. Define the **average of $f$ over $S$** as

$$\text{aver of } f \text{ over } S \equiv \frac{\iint_S f\ dA}{\iint_S 1\ dA} \tag{5.3.5}$$

**Remark 5.3.3.** A note on the relation between integrals of a vector field and a real-valued function. The surface of vector field $\mathbf{F}$ on $S$ is defined by *reducing $\boldsymbol{F}$ to a real-valued function $\boldsymbol{F} \cdot \boldsymbol{n}$* and then follow the definition of conventional real-valued function on $S$. Define $h \equiv \mathbf{F} \cdot \mathbf{n}$,

$$\iint_S \mathbf{F} \cdot \mathbf{n}\ dA = \iint_S h\ dA \tag{5.3.6}$$

$$\equiv \iint_R h(\mathbf{G}(\mathbf{u}))||\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}||\ d\Theta \tag{5.3.7}$$

$$= \iint_R \mathbf{F}(\mathbf{G}(\mathbf{u})) \cdot \mathbf{n}(\mathbf{G}(\mathbf{u}))||\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}||\ d\Theta \tag{5.3.8}$$

$$= \iint_R \mathbf{F}(\mathbf{G}(\mathbf{u})) \cdot \frac{\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}}{||\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}||}||\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}||\ d\Theta \tag{5.3.9}$$

$$= \iint_R \mathbf{F}(\mathbf{G}(\mathbf{u})) \cdot (\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v})\ d\Theta \tag{5.3.10}$$

### 5.3.2 An invariance property

**Remark 5.3.4.** As mentioned above, given $\mathbf{n}(\mathbf{x})$ fixed, we can define the surface integral of vector field as the surface integral of a real-valued function defined as $h(\mathbf{x}) \equiv \mathbf{F}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})$. And as argued before, one $\mathbf{n}$ is fixed (i.e. orientation is fixed), the value of integral is deterministic. Therefore we can conclude the integral of a vector field $\mathbf{F}$ over a surface $S$ depends on the **orientation** of $S$ but otherwise independent of the parameterization.

**Remark 5.3.5.** Let $S \subseteq \mathbb{R}^2$ be a two dimensional sub-manifold parameterized by $\mathbf{G} : R \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ such that $rank(\mathbf{G}(\mathbf{u})) = 2$ for all but zero-content sets on its domain.
Let $\varphi : W \subseteq \mathbb{R}^2 \to R$ be a bijection such that $\mathbf{H} \equiv \mathbf{G} \circ \varphi : W \to \mathbb{R}^3$ is another parameterization of $S$.
Now consider the integral of vector field $\mathbf{F}$ under parameterization $\mathbf{H}$,

$$\iint_S \mathbf{F} \cdot \mathbf{u}\, dA = \iint_W \mathbf{F}(\mathbf{H}) \cdot (\frac{\partial \mathbf{H}}{\partial s} \times \frac{\partial \mathbf{H}}{\partial t})\, d\Theta \tag{5.3.11}$$

$$= \iint_W \mathbf{F} \circ \mathbf{G} \circ \varphi \cdot (\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}) det\ D\varphi\, d\Theta \tag{5.3.12}$$

$$= \pm \iint_R \mathbf{F} \circ \mathbf{G} \cdot (\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v})\, d\Theta \text{ (change of variable)} \tag{5.3.13}$$

**Theorem 5.3.1** (Invariance)**.** Let $\mathbf{G} : R \to \mathbb{R}^3$ and $\mathbf{H} \equiv \mathbf{G} \circ \varphi : W \to \mathbb{R}^3$ be two parameterizations of $S$, then

$$\iint_R f \circ \mathbf{G} || \frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v} ||\, d\Theta = \iint_W f \circ \mathbf{H} || \frac{\partial \mathbf{H}}{\partial s} \times \frac{\partial \mathbf{H}}{\partial t} ||\, d\Theta \tag{5.3.14}$$

and

$$\iint_R \mathbf{F} \circ \mathbf{G} \cdot (\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v})\, d\Theta = \pm \iint_W \mathbf{F} \circ \mathbf{H} \cdot (\frac{\partial \mathbf{H}}{\partial u} \times \frac{\partial \mathbf{H}}{\partial v})\, d\Theta \tag{5.3.15}$$

### 5.3.3 Volume and Area

**Theorem 5.3.2.** Let $R$ be an arbitrary regular region in $\mathbb{R}^3$, and let $S$ be the boundary surface of $R$, define

$$S_h \equiv \{\mathbf{x} + \delta \mathbf{n} : \mathbf{x} \in S \wedge \delta \in [0, h]\} \tag{5.3.16}$$

where $S_h$ can be interpreted as *a shell of region $R$ with thickness $h$*. Then the surface area of $S$ is

$$\text{area}(S) = \lim_{h \to 0} \frac{|S_h|}{h} \tag{5.3.17}$$

## 5.4 Divergence, Gradient and Curl

**Definition 5.4.1.** Let $U \subseteq \mathbb{R}^n$ be an open set, and define real-valued function $f : U \to \mathbb{R}$ and vector field $\mathbf{F} : U \to \mathbb{R}^n$. Then we define

1. The **gradient** of $f$ as $\nabla f$;
2. The **divergence** of $\mathbf{F}$ as $\nabla \cdot \mathbf{F}$;
3. The **curl** of $\mathbf{F}$ as $\nabla \times \mathbf{F}$.

**Definition 5.4.2.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^1$ real-valued function, define the **Laplacian** of $f$ as a mapping from *real-valued functional space* to *real-valued functional space* defined as

$$\mathrm{div}(\mathrm{grad})f \equiv \sum_j \partial_j^2 f = \Delta f = \nabla^2 f \tag{5.4.1}$$

**Theorem 5.4.1.** For every $C^2$ real valued function $f : \mathbb{R}^3 \to \mathbb{R}$,

$$\mathrm{curl}(\mathrm{grad}f) = \mathbf{0} \tag{5.4.2}$$

For every $C^2$ vector field defined in $\mathbb{R}^3$ or a subset of it,

$$\mathrm{div}(\mathrm{curl}\mathbf{F}) = 0 \tag{5.4.3}$$

Note that the domain of $f$ and $\mathbf{F}$ must be $\mathbb{R}^3$ or a subset of it, otherwise the curl operation is not well-defined.

**Theorem 5.4.2** (Product rules)**.**

$$\mathrm{grad}(fg) = f\ \mathrm{grad}g + \mathrm{grad}f\ g \tag{5.4.4}$$

$$\mathrm{div}(f\mathbf{G}) = f\ \mathrm{div}G + \mathrm{grad}f \cdot \mathbf{G} \tag{5.4.5}$$

$$\mathrm{curl}(f\mathbf{G}) = f\ \mathrm{curl}G + \mathrm{grad}f \times \mathbf{G} \tag{5.4.6}$$

## 5.5 Divergence Theorem

**Remark 5.5.1.** vector field integral on boundary (2-dimensional sub-manifold) of region in $\mathbb{R}^3$ ($\mathbf{F} \cdot \mathbf{n}\ dA$ 2-form) and scalar valued function ($\mathrm{div}(\mathbf{F})\ dV$3-form) in a region (3-dimensional sub-manifold).

**Theorem 5.5.1** (Divergence Theorem)**.** Let $R \subseteq \mathbb{R}^3$ be a *regular region* with *piece-wise smooth boundary* $\partial S$. And $\mathbf{n}$ is the *outer normal vector* on $\partial S$, then,

$$\iint_{\partial S} \mathbf{F} \cdot \mathbf{n}\ dA = \iiint_S \mathrm{div}(\mathbf{F})\ dV \tag{5.5.1}$$

*Proof.*

**Definition 5.5.1.** A region $R \subseteq \mathbb{R}^3$ is said to be $xy$**-simple** if and only if it can be expressed as the following form

$$R = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in W, \varphi_1(x, y) \le z \le \varphi_2(x, y)\} \tag{5.5.2}$$

Suppose $S$ is simple in terms of all combinations of $x, y, z$.
Then

$$\iint_{\partial S} \mathbf{F} \cdot \mathbf{n}\ dA = \iint_{\partial S} F_1 n_1 + F_2 n_2 + F_3 n_2\ dA \tag{5.5.3}$$

Consider $\iint_{\partial S} F_3 n_3\ dA$, since $R$ is $xy$-simple,

$$\iint_{\partial S} F_3 n_3\ dA = \iint_{\partial S} F_3 \mathbf{k} \cdot \mathbf{n}\ dA \tag{5.5.4}$$

Note that except the bottom and top sides, which are parameterized by $\mathbf{G}_1(x,y)=(x,y,\varphi_1(x,y))$ and $\mathbf{G}_2(x,y)=(x,y,\varphi_2(x,y))$, the outer normal vector of those region has form $(\cdot,\cdot,0)$, and therefore $\mathbf{n}\cdot\mathbf{k}=0$ for every $\mathbf{x}$ on those regions, and contribute nothing to the integral.
Therefore, to evaluate $\iint_{\partial S} F_3\mathbf{k}\cdot\mathbf{n}\ dA$, we only need to consider the upper and bottom surfaces. Also note that $\mathbf{n}$ has opposite $z$ component on those two surfaces.
Moreover, the undirected $\mathbf{n}$ on those two surfaces is

$$\tilde{\mathbf{n}}=\begin{pmatrix}1\\0\\\partial_x\varphi_i\end{pmatrix}\times\begin{pmatrix}0\\1\\\partial_y\varphi_i\end{pmatrix}=\begin{pmatrix}-\partial_x\varphi_i\\\partial_x\varphi_i-\partial_y\varphi_i\\1\end{pmatrix} \tag{5.5.5}$$

$$\implies \iint_{\partial S} F_3\mathbf{k}\cdot\mathbf{n}\ dA=\iint_{\partial S} F_3\ dA \tag{5.5.6}$$

$$=\iint_{upper\ \partial S} F_3\ dA-\iint_{lower\ \partial S} F_3\ dA \tag{5.5.7}$$

$$=\iint_W F_3(x,y,\varphi_2(x,y))\ dxdy-\iint_W F_3(x,y,\varphi_1(x,y))\ dxdy \tag{5.5.8}$$

$$=\iint_W\int_{\varphi_1(x,y)}^{\varphi_2(x,y)}\partial_3F_3\ dzdxdy=\iiint_S\partial_3F_3\ dV \tag{5.5.9}$$

We can prove the equalities involving the other two components, and the proof is completed by the fact that any open set in $\mathbb{R}^n$ can be written as a countable union of *almost disjoint* cubes, which are simple and the boundary of $S$ has zero content. ■

**Proposition 5.5.1** (Geometric Interpretation of Divergence)**.** Let $S\subset\mathbb{R}^3$, $\mathbf{F}:S\to\mathbb{R}^3$, $\mathbf{a}\in S$,

$$\operatorname{div}(\mathbf{F})(\mathbf{a})=\lim_{r\to0}\frac{3}{4\pi r^2}\iiint_{\mathcal{B}(\mathbf{a},r)}\operatorname{div}(\mathbf{F})(\mathbf{x})\ d\mathbf{x} \tag{5.5.10}$$

$$=\lim_{r\to0}\frac{3}{4\pi r^2}\underbrace{\iint_{\partial\mathcal{B}(\mathbf{a},r)}\mathbf{F}\cdot\mathbf{n}\ d\mathbf{x}}_{\text{flux through boundary}} \tag{5.5.11}$$

thus $\operatorname{div}(\mathbf{F})(\mathbf{a})>0$ if and only if at point $\mathbf{a}$, matters are flowing away from this point.

**Corollary 5.5.1** (Green's Formula)**.** Suppose $R\subset\mathbb{R}^3$ and $f,g:R\to\mathbb{R}$ are $C^2$ functions, then

$$\iint_{\partial S} f\nabla g\cdot\mathbf{n}\ dA=\iiint_S\nabla f\cdot\nabla g+f\nabla^2g\ dV \tag{5.5.12}$$

$$\iint_{\partial S}(f\nabla g-g\nabla f)\ dA=\iiint_S(f\nabla^2g-g\nabla^2)\ dV \tag{5.5.13}$$

*Proof.*

$$\iint_{\partial S} f\nabla g\cdot\mathbf{n}\ dA=\iiint_S\operatorname{div}(f\nabla g)\ dA \tag{5.5.14}$$

$$=\iiint_S f\operatorname{div}(\nabla g)+\nabla f\cdot\nabla g\ dV=\iiint_S f\nabla f\cdot\nabla g+\nabla^2g\ dV \tag{5.5.15}$$

The second formula can be proved directly using divergence theorem the first formula. ■

## 5.6 Stokes Theorem

### 5.6.1 Stokes Theorem in $\mathbb{R}^3$

**Theorem 5.6.1** (Stokes Theorem, Special Case)**.** Let $S$ be a 2 dimensional sub-manifold in $\mathbb{R}^3$, and let $\mathbf{F}$ be a vector field defined on some neighbour of $S$, then

$$\int_{\partial S} \mathbf{F} \cdot d\mathbf{x} = \iint_S \operatorname{curl}(\mathbf{F}) \cdot \mathbf{n} \, dA \tag{5.6.1}$$

**Remark 5.6.1.** In above theorem, $\omega \equiv \mathbf{F} \cdot d\mathbf{x}$ is a 1-form in $\mathbb{R}^3$ and $d\omega \equiv \operatorname{curl}(\mathbf{F}) \cdot \mathbf{n} \, dA$ is a 2-form in $\mathbb{R}^3$.

**Corollary 5.6.1.** Let $S$ be a closed surface in $\mathbb{R}^3$, that's, $\partial S = \varnothing$, and let $\mathbf{n}$ denote the outer normal vector, and $\mathbf{F}$ is a $C^1$ vector field. Then

$$\iint_S \operatorname{curl}(\mathbf{F}) \cdot \mathbf{n} \, dA = 0 \tag{5.6.2}$$

*Proof.* We can construct a *small* simple closed curve $C$ on $S$ and divide $S$ into two regions sharing the same boundary. And note that given orientation fixed on $S$, the orientation on $\partial S_1$ and $\partial S_2$ are opposite. Then

$$\iint_S \operatorname{curl}(\mathbf{F}) \cdot \mathbf{n} \, dA = \iint_{S_1} \operatorname{curl}(\mathbf{F}) \cdot \mathbf{n} \, dA + \iint_{S_2} \operatorname{curl}(\mathbf{F}) \cdot \mathbf{n} \, dA \tag{5.6.3}$$

$$= \int_{\partial S_1} \mathbf{F} \cdot d\mathbf{x} - \int_{\partial S_2} \mathbf{F} \cdot d\mathbf{x} = \int_C \mathbf{F} \cdot d\mathbf{x} - \int_C \mathbf{F} \cdot d\mathbf{x} = 0 \tag{5.6.4}$$

■

**Proposition 5.6.1** (Geometric Interpretation of Curl)**.** Let $R \subset \mathbb{R}^3$ be a 2 dimensional sub-manifold with $\mathbf{n}$ as outer normal vector on it, and $\mathbf{a} \in R$,

$$\operatorname{curl}(F)(\mathbf{a}) \cdot \mathbf{n}(\mathbf{a}) = \lim_{r \to 0} \frac{1}{2\pi r^2} \iint_{\mathcal{D}(\mathbf{a},r)} \operatorname{curl}(F)(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, dA \tag{5.6.5}$$

$$= \lim_{r \to 0} \frac{1}{2\pi r^2} \int_{\partial \mathcal{D}(\mathbf{a},r)} \mathbf{F} \cdot d\mathbf{x} \tag{5.6.6}$$

$$\tag{5.6.7}$$

If we think of $\mathbf{F}$ as a force field, then $\int_{\partial \mathcal{D}(\mathbf{a},r)} \mathbf{F} \cdot d\mathbf{x}$ represents the work done by $\mathbf{F}$ on a particle moves around $\partial \mathcal{D}(\mathbf{a}, r)$. Thus $\operatorname{curl}(\mathbf{F}) \cdot \mathbf{u}$ represents the tendency of the force $\mathbf{F}$ to push the particle around $\partial \mathcal{D}(\mathbf{a}, r)$ in a direction compatible with $\mathbf{n}$.

### 5.6.2 The Generalization

**Proposition 5.6.2** (Properties of Exterior Products)**.** Let $\alpha_1, \alpha_2$ and $\beta$ be 1-forms on $\mathbb{R}^n$ and $f_1, f_2$ are continuous functions defined on $\mathbb{R}^n$,

1. **Distributive**

$$(f_1\alpha_1 + f_2\alpha_2) \wedge \beta = f_1(\alpha_1 \wedge \beta) + f_2(\alpha_2 \wedge \beta) \tag{5.6.8}$$

$$\beta \wedge (f_1\alpha_1 + f_2\alpha_2) = f_1(\beta \wedge \alpha_1) + f_2(\beta \wedge \alpha_2) \tag{5.6.9}$$

2. **Anti-commutative**

$$\beta \wedge \alpha = -\alpha \wedge \beta \tag{5.6.10}$$

**Theorem 5.6.2** (Divergence Theorem in $\mathbb{R}^n$)**.** Let $R$ be a regular region in $\mathbb{R}^n$ bounded by a piecewise smooth hyper-surface $\partial R$. Note here $R$ is a $n$ dimensional sub-manifold and $\partial R$ is a $n-1$ dimension sub-manifold. Then

$$\int \cdots \int_{\partial R} \mathbf{F} \cdot \mathbf{n} dV^{n-1} = \iint \cdots \int_R \operatorname{div}(\mathbf{F})\ dV^n \tag{5.6.11}$$

where if $\partial R$ is parameterized by $\mathbf{G}(u_1, \ldots, u_{n-1})$, then

$$\mathbf{n} dV^{n-1} = \det \begin{pmatrix} \mathbf{e}_1 & \ldots & \mathbf{e}_n \\ \partial_1 G_1 & \ldots & \partial_1 G_n \\ \vdots & & \vdots \\ \partial_{n-1} G_1 & \ldots & \partial_{n-1} G_n \end{pmatrix} \tag{5.6.12}$$

**Definition 5.6.1.** A 0-form on $\mathbb{R}^n$ is a real valued function $f$.

**Remark 5.6.2.** While writing the basis elements $dx_i \wedge dx_j$ with the variables in *cyclic order*. That's $dx$ before $dy$ before $dz$ before $dx$ in $\mathbb{R}^3$ case.

**Definition 5.6.2.** A $k$-form in $\mathbb{R}^n$ takes the expression of linear combination of $C(n,k)$ basis elements $\{\beta_i\}_i$.

**Example 5.6.1.** A 2-form $\omega$ in $\mathbb{R}^3$ can be expressed using a 3-element basis

$$\omega = \sum_{1 \le i < j \le 3} C_{ij}(\mathbf{x}) \beta_{ij} \tag{5.6.13}$$

$$\beta_{ij} \in \{dx \wedge dy, dy \wedge dz, dx \wedge dz\} \tag{5.6.14}$$

**Definition 5.6.3.** Let $\omega = \sum_{j=1}^{C(n,k)} f_j \beta_j$ be a $k$-form in $\mathbb{R}^n$, then it's **exterior derivative** is defined to be the $(k+1)$-form in $\mathbb{R}^n$ defined as

$$d\omega \equiv \sum_j df_j \wedge \beta_j \tag{5.6.15}$$

where $df_j$ can be computed using *total derivative*.

**Example 5.6.2.** In $\mathbb{R}^3$, the *exterior derivative* for a 0-form $f$ is its **gradient**, which is a 1-form. And the exterior derivate of a 1-form in $\mathbb{R}^3$ is its curl

$$\omega := F_1 dx + F_2 dy + F_3 dz \tag{5.6.16}$$

$$\implies d\omega = dF_1 \wedge dx + dF_2 \wedge dy + dF_3 \wedge dz \tag{5.6.17}$$

$$= (\partial_1 F_1 dx + \partial_2 F_1 dy + \partial_3 F_1 dz) \wedge dx \tag{5.6.18}$$

$$+ (\partial_1 F_2 dx + \partial_2 F_2 dy + \partial_3 F_2 dz) \wedge dy \tag{5.6.19}$$

$$+ (\partial_1 F_3 dx + \partial_2 F_3 dy + \partial_3 F_3 dz) \wedge dz \tag{5.6.20}$$

$$= (\partial_1 F_2 - \partial_2 F_1) dx \wedge dy + (\partial_2 F_3 - \partial_3 F_2) dy \wedge dz + (\partial_3 F_1 - \partial_1 F_3) dz \wedge dx \tag{5.6.21}$$

$$= \operatorname{curl}(\mathbf{F}) \tag{5.6.22}$$

The exterior derivate of a 2-form in $\mathbb{R}^3$ is its divergence

$$\omega := Ady \wedge dz + Bdz \wedge dx + Cdx \wedge dy \tag{5.6.23}$$
$$\implies d\omega = (\partial_1 Adx + \partial_2 Ady + \partial_3 Adz) \wedge dy \wedge dz \tag{5.6.24}$$
$$+(\partial_1 Bdx + \partial_2 Bdy + \partial_3 Bdz) \wedge dz \wedge dx \tag{5.6.25}$$
$$+(\partial_1 Cdx + \partial_2 Cdy + \partial_3 Cdz) \wedge dx \wedge dy \tag{5.6.26}$$
$$= (\partial_1 A + \partial_2 B + \partial_3 C) dx \wedge dy \wedge dz \tag{5.6.27}$$
$$= \mathrm{div}(\mathbf{F}) \tag{5.6.28}$$

**Theorem 5.6.3** (Stokes Theorem, 5.77)**.** Let $M$ be a smooth, oriented $k$ dimensional sub-manifold of $\mathbb{R}^n$ with a piecewise smooth boundary $\partial M$, and let $\partial M$ carry the orientation that is (in a suitable sense) compatible with the one on $M$. If $\omega$ is a $(k-1)$-form of class $C^1$ on an open set containing $M$, then

$$\int \cdots \int_{\partial M} \omega = \iint \cdots \int_M d\omega \tag{5.6.29}$$

**Theorem 5.6.4.** The *boundary* of a (smoothly bounded) region $M$ in a $k$ dimensional manifold is a $(k-1)$ dimensional manifold with no boundary.
That's let $M$ be a $k$ dimensional manifold with piecewise smooth boundary $\partial M$, then

$$\partial(\partial M) = \varnothing \tag{5.6.30}$$

**Theorem 5.6.5.** For any $k$-form $\omega$ on $\mathbb{R}^n$,

$$d(d\omega) = 0 \tag{5.6.31}$$

*Proof.* Let $M$ be a $k$ dimensional sub-manifold in $\mathbb{R}^n$ with piecewise smooth boundary, and $\omega$ is a $(k-2)$-form on $\mathbb{R}^n$, so $d(d\omega)$ is a $k$-form on $\mathbb{R}^n$. And

$$\iiint \cdots \int_M d(d\omega) = \iint \cdots \int_{\partial M} d\omega \tag{5.6.32}$$
$$= \int \cdots \int_{\partial(\partial M)} \omega = \int \cdots \int_{\varnothing} \omega = 0 \tag{5.6.33}$$

■

# 6 Further Topics

## 6.1 Fourier Analysis

**Definition 6.1.1.** A function $f : \mathbb{R} \to \mathbb{R}$ is **periodic with period** $p$ if[2]

$$f(x) = f(x+p)\ \forall x \in \mathbb{R} \tag{6.1.1}$$

**Assumption 6.1.** WLOG, we can assume $p = 2\pi$ for every periodic function encountered.

[2] Here $p$ is not necessarily the least period of $f$.

*Proof.* If $f$ is periodic with period $p$, we can transform it into a function with period $2\pi$ by defining $\tilde{f}(x) := f(x\frac{p}{2\pi})$ such that

$$\tilde{f}(x+2\pi) = f(x\frac{p}{2\pi} + p) \tag{6.1.2}$$

$$= f(x\frac{p}{2\pi}) = \tilde{f}(x) \tag{6.1.3}$$

■

**Theorem 6.1.1** (Fourier Decomposition)**.** Let $f : \mathbb{R} \to \mathbb{R}$ be a periodic function with period $2\pi$ continuous everywhere but finitely many points on its domain, then it can be expressed as

$$f(x) = a_0 + \sum_{n=1}^{\infty} \{a_n \cos(nx) + b_n \sin(nx)\} \tag{6.1.4}$$

**Lemma 6.1.1.**

$$\cos(a)\cos(b) = \frac{1}{2}[\cos(a+b) + \cos(a-b)] \tag{6.1.5}$$

*Proof.*

$$\cos(a)\cos(b) = \cos(a+b) + \sin(a)\sin(b) \tag{6.1.6}$$

$$\cos(a)\cos(b) = \cos(a-b) - \sin(a)\sin(b) \tag{6.1.7}$$

■

$$\sin(a)\sin(b) = \frac{1}{2}[\cos(a-b) - \cos(a+b)] \tag{6.1.8}$$

*Proof.*

$$\sin(a)\sin(b) = \cos(a)\cos(b) - \cos(a+b) \tag{6.1.9}$$

$$-\sin(a)\sin(b) = \sin(a)\sin(-b) = \cos(a)\cos(b) - \cos(a-b) \tag{6.1.10}$$

■

**Lemma 6.1.2.**

$$\int_{-\pi}^{\pi} \cos(nx)\sin(mx)\ dx = 0\ \forall n, m \in \mathbb{N} \tag{6.1.11}$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos(nx)\cos(mx)\ dx = \mathbb{1}\{n = m\} \tag{6.1.12}$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \sin(nx)\sin(mx)\ dx = \mathbb{1}\{n = m\} \tag{6.1.13}$$

$$\tag{6.1.14}$$

**Proposition 6.1.1** (Finding the Coefficients)**.**

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(mx)\ dx \tag{6.1.15}$$

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(mx)\ dx \tag{6.1.16}$$

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\ dx \tag{6.1.17}$$