Vector Analysis

向量分析

鄭經斆

Contents

Chapter 1

Linear Algebra

1.1 Vector Spaces

Definition 1.1 (Vector spaces)**.** A ***vector space*** $\mathcal{V}$ over a scalar field $\mathbb{F}$ is a set of elements called ***vectors***, together with two operations $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ and $\cdot : \mathbb{F} \times \mathcal{V} \to \mathcal{V}$, called the ***vector addition*** and ***scalar multiplication*** respectively, such that

1. $\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

2. $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

3. There is a zero vector $\mathbf{0}$ such that $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$.

4. For every $\boldsymbol{v}$ in $\mathcal{V}$, there is a vector $\boldsymbol{w}$ such that $\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}$.

5. $\alpha \cdot (\boldsymbol{v} + \boldsymbol{w}) = \alpha \cdot \boldsymbol{v} + \alpha \cdot \boldsymbol{w}$ for all $\alpha \in \mathbb{F}$ and $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

6. $\alpha \cdot (\beta \cdot \boldsymbol{v}) = (\alpha\beta) \cdot \boldsymbol{v}$ for all $\alpha, \beta \in \mathbb{F}$ and $\boldsymbol{v} \in \mathcal{V}$.

7. $(\alpha + \beta) \cdot \boldsymbol{v} = \alpha \cdot \boldsymbol{v} + \beta \cdot \boldsymbol{v}$ for all $\alpha, \beta \in \mathbb{F}$ and $\boldsymbol{v} \in \mathcal{V}$.

8. $1 \cdot \boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$.

For notational convenience, we often drop the $\cdot$ and write $\alpha\boldsymbol{v}$ instead of $\alpha \cdot \boldsymbol{v}$.

Remark 1.2. In property 4 of the definition above, it is easy to see that for each $\boldsymbol{v}$, there is only one vector $\boldsymbol{w}$ such that $\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}$. We often denote this $\boldsymbol{w}$ by $-\boldsymbol{v}$, and the vector substraction $- : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ is then defined (or understood) as $\boldsymbol{v} - \boldsymbol{w} = \boldsymbol{v} + (-\boldsymbol{w})$.

Example 1.3. Let $\mathbb{F}$ be a scalar field. The space $\mathbb{F}^{\mathrm{n}}$ is the collection of n-tuple $\boldsymbol{v} = (\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_{\mathrm{n}})$ with $\mathrm{v}_i \in \mathbb{F}$ with addition $+$ and scalar multiplication $\cdot$ defined by

$$
\begin{aligned}
(\mathrm{v}_1, \cdots, \mathrm{v}_{\mathrm{n}}) + (\mathrm{w}_1, \cdots, \mathrm{w}_{\mathrm{n}}) &\equiv (\mathrm{v}_1 + \mathrm{w}_1, \cdots, \mathrm{v}_{\mathrm{n}} + \mathrm{w}_{\mathrm{n}}) \,, \\
\alpha(\mathrm{v}_1, \cdots, \mathrm{v}_{\mathrm{n}}) &\equiv (\alpha\mathrm{v}_1, \cdots, \alpha\mathrm{v}_{\mathrm{n}}) \,.
\end{aligned}
$$

Then $\mathbb{F}^{\mathrm{n}}$ is a vector space.

Example 1.4. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\mathcal{V}$ be the collection of all $\mathbb{R}$-valued continuous functions on $[0, 1]$. The vector addition $+$ and scalar multiplication $\cdot$ is defined by

$$
\begin{aligned}
(f + g)(x) &= f(x) + g(x) & \forall\, f, g \in \mathcal{V} \,, \\
(\alpha \cdot f)(x) &= \alpha f(x) & \forall\, f \in \mathcal{V}, \alpha \in \mathbb{F} \,.
\end{aligned}
$$

Then $\mathcal{V}$ is a vector space, and is denoted by $\mathscr{C}([0, 1]; \mathbb{F})$. When the scalar field under consideration is clear, we simply use $\mathscr{C}([0, 1])$ to denote this vector space.

Definition 1.5 (Vector subspace)**.** Let $\mathcal{V}$ be a vector space over scalar field $\mathbb{F}$. A subset $\mathcal{W} \subseteq \mathcal{V}$ is called a vector subspace of $\mathcal{V}$ if itself is a vector space over $\mathbb{F}$.

1.1.1 The linear independence of vectors

Definition 1.6. Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$. k vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_k$ in $\mathcal{V}$ is said to be ***linearly dependent*** if there exists $(\alpha_1, \cdots, \alpha_k) \subseteq \mathbb{F}^k$, $(\alpha_1, \cdots, \alpha_k) \neq \mathbf{0}$ such that $\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_k\boldsymbol{v}_k = \mathbf{0}$. k vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_k$ in $\mathcal{V}$ is said to be ***linearly independent*** if they are not linearly dependent. In other words, $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$ are linearly independent if

$$
\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_k\boldsymbol{v}_k = \mathbf{0} \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \,.
$$

Example 1.7. The k vectors $\{1, x, x^2, \cdots, x^{k-1}\}$ are linearly independent in $\mathscr{C}([0, 1])$ for all $k \in \mathbb{N}$.

1.1.2 The dimension of a vector space

Definition 1.8. The ***dimension*** of a vector space $\mathcal{V}$ is the number of maximum linearly independent set in $\mathcal{V}$, and in such case $\mathcal{V}$ is called an n-dimensional vector space, where n the the dimension of $\mathcal{V}$. If for every number $\mathrm{n} \in \mathbb{N}$ there exists n linearly independent vectors in $\mathcal{V}$, the vector space $\mathcal{V}$ is said to be infinitely dimensional.

Example 1.9. The space $\mathbb{F}^{\mathrm{n}}$ is n-dimensional, and $\mathscr{C}([0,1])$ is infinitely dimensional (since $1, x, \cdots, x^{\mathrm{n}-1}$ are n linearly independent vectors in $\mathscr{C}([0,1])$).

1.1.3 Bases of a vector space

Definition 1.10 (Basis)**.** Let $\mathcal{V}$ be a vector space over $\mathbb{F}$. A set of vectors $\{\boldsymbol{v}_i\}_{i\in\mathcal{I}}$ in $\mathcal{V}$ is called a ***basis*** of $\mathcal{V}$ if for every $\boldsymbol{v} \in \mathcal{V}$, there exists a unique $\{\alpha_i\}_{i\in\mathcal{I}} \subseteq \mathbb{F}$ such that

$$\boldsymbol{v} = \sum_{\alpha\in\mathcal{I}} \alpha_i \boldsymbol{v}_i \,.$$

For a given basis $\mathcal{B} = \{\boldsymbol{v}_i\}_{i\in\mathcal{I}}$, the coefficients $\{\alpha_i\}_{i\in\mathcal{I}}$ given in the above relation is denoted by $[\boldsymbol{v}]_{\mathcal{B}}$.

Example 1.11 (Standard Basis of $\mathbb{F}^{\mathrm{n}}$)**.** Let $\mathrm{e}_i = (0, , \cdots, 0, 1, 0, \cdots, 0)$, where 1 locates at the i-th slot. Then the collection $\{\mathrm{e}_i\}_{i=1}^{\mathrm{n}}$ is a basis of the vector space $\mathbb{F}^{\mathrm{n}}$ over $\mathbb{F}$ since

$$(\alpha_1, \cdots, \alpha_{\mathrm{n}}) = \sum_{i=1}^{\mathrm{n}} \alpha_i \mathrm{e}_i \qquad \forall\, \alpha_i \in \mathbb{F}.$$

The collection $\{\mathrm{e}_i\}_{i=1}^{\mathrm{n}}$ is called the ***standard basis*** of $\mathbb{F}^{\mathrm{n}}$.

Example 1.12. Even though $\left\{1, x, \cdots, x^k, \cdots\right\}$ is a set of linearly independent vectors, it is not a basis of $\mathscr{C}([0,1])$. However, let $\mathscr{P}([0,1])$ be the collection of polynomials defined on $[0,1]$. Then $\mathscr{P}([0,1])$ is still a vector space, and $\left\{1, x, \cdots, x^k, \cdots\right\}$ is a basis of $\mathscr{P}([0,1])$.

1.2 Inner Products and Inner Product Spaces

Definition 1.13 (Inner product space)**.** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$ with a bilinear form $(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$ is called an ***inner product space*** if the bilinear form satisfies

1. $(\boldsymbol{v}, \boldsymbol{v}) \geqslant 0$ for all $\boldsymbol{v} \in \mathcal{V}$.

2. $(\boldsymbol{v}, \boldsymbol{v}) = 0$ if and only if $\boldsymbol{v} = 0$.

3. $(\boldsymbol{v}, \boldsymbol{w}) = \overline{(\boldsymbol{w}, \boldsymbol{v})}$ for all $\boldsymbol{v}, w \in \mathcal{V}$, where the bar over the scalar $(\boldsymbol{w}, \boldsymbol{v})$ is the complex conjugate.

4. $(\boldsymbol{v} + \boldsymbol{w}, \boldsymbol{u}) = (\boldsymbol{v}, \boldsymbol{u}) + (\boldsymbol{w}, \boldsymbol{u})$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

5. $(\alpha\boldsymbol{v}, \boldsymbol{w}) = \alpha(\boldsymbol{v}, \boldsymbol{w})$ for all $\alpha \in \mathbb{F}$ and $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

The bilinear form $(\cdot,\cdot)$ is called an *inner product* on $\mathcal{V}$.

Example 1.14 (Standard Inner Product on $\mathbb{F}^{\mathrm{n}}$)**.** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\mathbb{F}^{\mathrm{n}}$ be the vector space defined in Example 1.3. A special inner product on the vector space $\mathbb{F}^{\mathrm{n}}$ over $\mathbb{F}$, called the ***standard inner product*** on $\mathbb{F}^{\mathrm{n}}$, is defined by

$$(\boldsymbol{v}, \boldsymbol{w}) \equiv \sum_{i=1}^{\mathrm{n}} \mathrm{v}_i \overline{\mathrm{w}}_i \,,$$

where v_i and w_i are the i-th component of $\boldsymbol{v}$ and $\boldsymbol{w}$, respectively, and $\overline{\mathrm{w}}_i$ is the complex conjugate of w_i. We sometimes use $\boldsymbol{v} \cdot \boldsymbol{w}$ to denote $(\boldsymbol{v}, \boldsymbol{w})$.

Example 1.15. Let $\mathcal{V} = \mathscr{C}([0,1];\mathbb{R})$. Define

$$(f, g) = \int_0^1 f(x)g(x)dx \,.$$

Then $\big(\mathscr{C}([0,1];\mathbb{R}), (\cdot,\cdot)\big)$ is an inner product space. The norm induced by this inner product is given by

$$\|f\| = \Big[\int_0^1 |f(x)|^2 dx\Big]^{\frac{1}{2}} \,,$$

and is called the L^2-norm.

Proposition 1.16. *Let $\mathcal{V}$ be an inner product space with inner product $(\cdot,\cdot)$. The inner product $(\cdot,\cdot)$ on $\mathcal{V}$ induces a norm defined by*

$$\|\boldsymbol{v}\| \equiv \sqrt{(\boldsymbol{v}, \boldsymbol{v})}$$

satisfying

1. $\|\boldsymbol{v}\| \geqslant 0$ *for all* $\boldsymbol{v} \in \mathcal{V}$.
2. $\|\boldsymbol{v}\| = 0$ *if and only if* $\boldsymbol{v} = 0$.
3. $\|\alpha\boldsymbol{v}\| = |\alpha|\|\boldsymbol{v}\|$ *for all* $\alpha \in \mathbb{F}$ *and* $\boldsymbol{v} \in \mathcal{V}$.
4. $\|\boldsymbol{v} + \boldsymbol{w}\| \leqslant \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$ *for all* $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.
5. $|(\boldsymbol{v}, \boldsymbol{w})| \leqslant \|\boldsymbol{v}\|\|\boldsymbol{w}\|$ *for all* $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

Proof. Properties 1 through 3 are obvious. We focus on proving property 5 first, and as we will see, property 4 is a direct consequence of property 5.

Let $\alpha \in \mathbb{F}$ satisfy $\alpha(\boldsymbol{v}, \boldsymbol{w}) = |(\boldsymbol{v}, \boldsymbol{w})|$. Then $|\alpha| = 1$. For all $\lambda \in \mathbb{R}$,

$$\begin{aligned}(\lambda\alpha\boldsymbol{v} + \boldsymbol{w}, \lambda\alpha\boldsymbol{v} + \boldsymbol{w}) &= (\lambda\alpha\boldsymbol{v}, \lambda\alpha\boldsymbol{v}) + (\lambda\alpha\boldsymbol{v}, \boldsymbol{w}) + (\boldsymbol{w}, \lambda\alpha\boldsymbol{v}) + (\boldsymbol{w}, \boldsymbol{w}) \\ &= \lambda^2\|\boldsymbol{v}\|^2 + \lambda\alpha(\boldsymbol{v}, \boldsymbol{w}) + \overline{\lambda\alpha(\boldsymbol{v}, \boldsymbol{w})} + \|\boldsymbol{w}\|^2 \\ &= \lambda^2\|\boldsymbol{v}\|^2 + 2\lambda|(\boldsymbol{v}, \boldsymbol{w})| + \|\boldsymbol{w}\|^2 .\end{aligned}$$

Since the left-hand side of the quantity above is always non-negative for all $\lambda \in \mathbb{R}$, we must have

$$|(\boldsymbol{v}, \boldsymbol{w})|^2 - \|\boldsymbol{v}\|^2\|\boldsymbol{w}\|^2 \leqslant 0$$

which implies property 5. To prove property 4, we note that

$$\begin{aligned}\|\boldsymbol{v} + \boldsymbol{w}\| \leqslant \|\boldsymbol{v}\| + \|\boldsymbol{w}\| &\Leftrightarrow \|\boldsymbol{v} + \boldsymbol{w}\|^2 \leqslant (\|\boldsymbol{v}\| + \|\boldsymbol{w}\|)^2 \\ &\Leftrightarrow (\boldsymbol{v} + \boldsymbol{w}, \boldsymbol{v} + \boldsymbol{w}) \leqslant \|\boldsymbol{v}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\| + \|\boldsymbol{w}\|^2 \\ &\Leftrightarrow \mathrm{Re}(\boldsymbol{v}, \boldsymbol{w}) \leqslant \|\boldsymbol{v}\|\|\boldsymbol{w}\|\end{aligned}$$

while the last inequality is valid because of property 5. □

Remark 1.17. The inequality in property 5 is called the ***Cauchy-Schwarz inequality***.

Definition 1.18. Let $\big(\mathcal{V}, (\cdot,\cdot)\big)$ be an inner product space. A basis $\mathcal{B}$ of $\mathcal{V}$ is called ***orthogonal*** if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ if $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{B}$ and $\boldsymbol{u} \neq \boldsymbol{v}$, and is called ***orthonormal*** if it is an orthogonal basis such that $\|\boldsymbol{v}\| = 1$ for all $\boldsymbol{v} \in \mathcal{B}$.

Definition 1.19 (Orthogoanl complement)**.** Let $\big(\mathcal{V}, (\cdot,\cdot)\big)$ be an inner product space over scalar field $\mathbb{F}$, and $\mathcal{W} \subseteq \mathcal{V}$ be a vector subspace of $\mathcal{V}$. The ***orthogonal complement*** of $\mathcal{W}$, denoted by $\mathcal{W}^\perp$, is the set

$$\mathcal{W}^\perp = \big\{\boldsymbol{v} \in \mathcal{V} \,\big|\, (\boldsymbol{v}, \boldsymbol{w}) = 0 \text{ for all } \boldsymbol{w} \in \mathcal{W}\big\} .$$

Proposition 1.20. *Let $\big(\mathcal{V}, (\cdot,\cdot)\big)$ be an inner product space over scalar field $\mathbb{F}$, and $\mathcal{W}$ be a vector subspace of $\mathcal{V}$. Then $\mathcal{W}^\perp$ is a vector subspace of $\mathcal{V}$.*

1.3 Normed Vector Spaces

The norm introduced in Proposition 1.16 is a good way of measure the magnitude of vectors. In general if a real-valued function can be used as a measurement of the magnitude of vectors if certain properties are satisfied.

Definition 1.21. Let $\mathcal{V}$ be a vector space over scalar field $\mathbb{F}$. A real-valued function $\|\cdot\|$: $\mathcal{V} \to \mathbb{R}$ is said to be a norm of $\mathcal{V}$ if

1. $\|\boldsymbol{v}\| \geqslant 0$ for all $\boldsymbol{v} \in \mathcal{V}$.

2. $\|\boldsymbol{v}\| = 0$ if and only if $\boldsymbol{v} = 0$.

3. $\|\alpha\boldsymbol{v}\| = |\alpha|\|\boldsymbol{v}\|$ for all $\boldsymbol{v} \in \mathcal{V}$ and $\alpha \in \mathbb{F}$.

4. $\|\boldsymbol{v} + \boldsymbol{w}\| \leqslant \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

The pair $(\mathcal{V}, \|\cdot\|)$ is called a normed vector space.

Example 1.22. Let $\mathcal{V} = \mathbb{F}^{\mathrm{n}}$, and $\|\cdot\|_p$ be defined by

$$\|\boldsymbol{x}\|_p = \begin{cases} \Big[\sum\limits_{i=1}^{\mathrm{n}} |x_i|^p\Big]^{\frac{1}{p}} & \text{if} \quad 1 \leqslant p < \infty\,, \\ \max\limits_{1\leqslant i\leqslant \mathrm{n}} |x_i| & \text{if} \quad p = \infty\,, \end{cases}$$

where $\boldsymbol{x} = (x_1, \cdots, x_{\mathrm{n}})$. The function $\|\cdot\|_p$ is a norm of $\mathbb{F}^{\mathrm{n}}$, and is called the p-norm of $\mathbb{F}^{\mathrm{n}}$.

Theorem 1.23 (Hölder's inequality)**.** *Let $1 \leqslant p \leqslant \infty$. Then*

$$\big|(\boldsymbol{x}, \boldsymbol{y})\big| \leqslant \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_{p'} \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^{\mathrm{n}}\,, \tag{1.1}$$

where $(\cdot,\cdot)$ is the standard inner product on $\mathbb{F}^{\mathrm{n}}$ and p' is the conjugate of p satisfying $\dfrac{1}{p} + \dfrac{1}{p'} = 1$.

Proof. Let $\boldsymbol{x} = (x_1, \cdots, x_{\mathrm{n}})$ and $\boldsymbol{y} = (y_1, \cdots, y_{\mathrm{n}})$ be given. Without loss of generality we can assume that $\boldsymbol{x} \neq \mathbf{0}$ and $\boldsymbol{y} \neq \mathbf{0}$. Define $\widetilde{\boldsymbol{x}} = \boldsymbol{x}/\|\boldsymbol{x}\|_p$ and $\widetilde{\boldsymbol{y}} = \boldsymbol{y}/\|\boldsymbol{y}\|_{p'}$. Then $\|\widetilde{\boldsymbol{x}}\|_p = 1$ and $\|\widetilde{\boldsymbol{y}}\|_{p'} = 1$. By Young's inequality

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{p'}b^{p'} \qquad \forall\, a, b \geqslant 0\,,$$

we find that for $1 < p < \infty$,

$$\begin{aligned}\big|(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{y}})\big| &= \Big|\sum_{k=1}^{\mathrm{n}} \frac{x_k}{\|\boldsymbol{x}\|_p} \frac{y_k}{\|\boldsymbol{y}\|_{p'}}\Big| \leqslant \sum_{k=1}^{\mathrm{n}} \frac{|x_k|}{\|\boldsymbol{x}\|_p} \frac{|y_k|}{\|\boldsymbol{y}\|_{p'}} \leqslant \sum_{k=1}^{\mathrm{n}} \Big(\frac{1}{p} \frac{|x_k|^p}{\|\boldsymbol{x}\|_p^p} + \frac{1}{p'} \frac{|y_k|^{p'}}{\|\boldsymbol{y}\|_{p'}^{p'}}\Big) \\ &= \frac{1}{p\|\boldsymbol{x}\|_p^p} \sum_{k=1}^{\mathrm{n}} |x_k|^p + \frac{1}{p'\|\boldsymbol{y}\|_{p'}^{p'}} \sum_{k=1}^{\mathrm{n}} |y_k|^{p'} = \frac{\|\boldsymbol{x}\|_p^p}{p\|\boldsymbol{x}\|_p^p} + \frac{\|\boldsymbol{y}\|^{p'}}{p'\|\boldsymbol{y}\|_{p'}^{p'}} = 1\end{aligned}$$

which conclude the case for $1 < p < \infty$. The proof for the case that $p = 1$ or $p = \infty$ is trivial, and is left to the reader. □

Corollary 1.24 (Minkowski inequality)**.** *Let $1 \leqslant p \leqslant \infty$. Then*

$$\|\boldsymbol{x} + \boldsymbol{y}\|_p \leqslant \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^{\mathrm{n}}\,.$$

Proof. We only prove the case that $1 < p < \infty$. First we note that

$$\begin{aligned}\|\boldsymbol{x} + \boldsymbol{y}\|_p^p &= \sum_{k=1}^{\mathrm{n}} |x_k + y_k|^p \leqslant \sum_{k=1}^{\mathrm{n}} |x_k + y_k|^{p-1}\big(|x_k| + |y_k|\big) \\ &= \sum_{k=1}^{\mathrm{n}} |x_k + y_k|^{p-1}|x_k| + \sum_{k=1}^{\mathrm{n}} |x_k + y_k|^{p-1}|y_k|\,.\end{aligned}$$

Let $\boldsymbol{u} = \big(|x_1|, |x_2|, \cdots, |x_{\mathrm{n}}|\big)$ and $\boldsymbol{v} = \big(|x_1 + y_1|^{p-1}, |x_2 + y_2|^{p-1}, \cdots, |x_{\mathrm{n}} + y_{\mathrm{n}}|^{p-1}\big)$. By Hölder's inequality,

$$\begin{aligned}\sum_{k=1}^{\mathrm{n}} |x_k + y_k|^{p-1}|x_k| &= (\boldsymbol{u}, \boldsymbol{v}) \leqslant \|\boldsymbol{u}\|_p \|\boldsymbol{v}\|_{p'} = \|\boldsymbol{x}\|_p \Big(\sum_{k=1}^{\mathrm{n}} |x_k + y_k|^{(p-1)p'}\Big)^{\frac{1}{p'}} \\ &= \|\boldsymbol{x}\|_p \Big(\sum_{k=1}^{\mathrm{n}} |x_k + y_k|^p\Big)^{\frac{p-1}{p}} = \|\boldsymbol{x}\|_p \|\boldsymbol{x} + \boldsymbol{y}\|_p^{p-1}\,.\end{aligned}$$

Similarly, we have $\sum\limits_{k=1}^{\mathrm{n}} |x_k + y_k|^{p-1}|y_k| \leqslant \|\boldsymbol{y}\|_p \|\boldsymbol{x} + \boldsymbol{y}\|_p^{p-1}$; thus

$$\|\boldsymbol{x} + \boldsymbol{y}\|_p^p \leqslant \big(\|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p\big)\|\boldsymbol{x} + \boldsymbol{y}\|_p^{p-1}$$

which concludes the Minkowski inequality. □

Theorem 1.25. *Let $1 \leqslant p \leqslant \infty$, and p' be the conjugate of p; that is, $\dfrac{1}{p} + \dfrac{1}{p'} = 1$. Then*

$$\|\boldsymbol{x}\|_p = \sup_{\|\boldsymbol{y}\|_{p'} = 1} \big|(\boldsymbol{x}, \boldsymbol{y})\big| \qquad \forall\, \boldsymbol{x} \in \mathbb{F}^{\mathrm{n}}\,.$$

Proof. By Hölder's inequality, it is clear that $\|\boldsymbol{x}\|_p \geqslant \sup\limits_{\|\boldsymbol{y}\|_{p'}=1} \big|(\boldsymbol{x},\boldsymbol{y})\big|$ for all $\boldsymbol{x} \in \mathbb{F}^{\mathrm{n}}$. On the other hand, note that $|x_k|^p = x_k \cdot \overline{x_k}|x_k|^{p-2}$; thus letting $y_k = \dfrac{\overline{x_k}|x_k|^{p-2}}{\|\boldsymbol{x}\|_p^{p-1}}$ we find that $\|\boldsymbol{y}\|_{p'} = 1$ which implies that

$$\big|(\boldsymbol{x},\boldsymbol{y})\big| = \frac{1}{\|\boldsymbol{x}\|_p^{p-1}} \sum_{k=1}^{\mathrm{n}} |x_k|^p = \|\boldsymbol{x}\|_p$$

which implies that $\sup\limits_{\|\boldsymbol{y}\|_{p'}=1} \big|(\boldsymbol{x},\boldsymbol{y})\big| \geqslant \|\boldsymbol{x}\|_p$. ◻

Making use of Hölder's inequality (1.1) and the Riemann sum approximation of the Riemann integral, we can conclude the following

Theorem 1.26. *Let $1 \leqslant p \leqslant \infty$. If p' is the conjugate of p; that is, $\dfrac{1}{p} + \dfrac{1}{p'} = 1$, then*

$$\Big|\int_0^1 f(x)g(x)\,dx\Big| \leqslant \|f\|_p\|g\|_{p'} \qquad \forall\, f,g \in \mathscr{C}([0,1];\mathbb{R})\,,$$

where

$$\|f\|_p = \begin{cases} \Big(\displaystyle\int_0^1 |f(x)|^p dx\Big)^{\frac{1}{p}} & \text{if } 1 \leqslant p < \infty\,, \\ \displaystyle\max_{x\in[0,1]} |f(x)| & \text{if } p = \infty\,. \end{cases}$$

Remark 1.27. The Minkowski inequality implies that

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p \qquad \forall\, f,g \in \mathscr{C}([0,1];\mathbb{R})\,.$$

In other words, the function $\|\cdot\|_p : \mathscr{C}([0,1];\mathbb{R}) \to \mathbb{R}$ is a norm on $\mathscr{C}([0,1];\mathbb{R})$, and is called the L^p-norm.

1.4 Matrices

Definition 1.28 (Matrix)**.** Let $\mathbb{F}$ be a scalar field. The space $\mathbb{M}(\mathrm{m},\mathrm{n};\mathbb{F})$ is the collection of elements, called an *m*-by-*n* matrix or $\mathrm{m} \times \mathrm{n}$ matrix over $\mathbb{F}$, of the form

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

where $a_{ij} \in \mathbb{F}$ is called the (i,j)-th entry of A, and is denoted by $[\mathrm{A}]_{ij}$. We write $\mathrm{A} = [a_{ij}]_{1\leqslant i\leqslant m;1\leqslant j\leqslant n}$ or simply $\mathrm{A} = [a_{ij}]_{\mathrm{m}\times\mathrm{n}}$ to denote that A is an $\mathrm{m}\times\mathrm{n}$ matrix whose (i,j)-th entry is a_{ij}. A is called a ***square matrix*** if $\mathrm{m} = \mathrm{n}$. The $1\times\mathrm{m}$ matrix

$$a_{i*} = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{i\mathrm{n}} \end{bmatrix}$$

is called the i-th row of A, and the $m\times 1$ matrix

$$a_{*j} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{\mathrm{m}j} \end{bmatrix}$$

is called the j-th column of A.

Definition 1.29 (Matrix addition)**.** Let $\mathrm{A} = [a_{ij}]_{\mathrm{m}\times\mathrm{n}}$ and $\mathrm{B} = [b_{ij}]_{\mathrm{m}\times\mathrm{n}}$ be two $\mathrm{m}\times\mathrm{n}$ matrices over a scalar field $\mathbb{F}$. The sum of A and B, denoted by $\mathrm{A}+\mathrm{B}$, is another $\mathrm{m}\times\mathrm{n}$ matrix defined by $\mathrm{A}+\mathrm{B} = [a_{ij}+b_{ij}]_{\mathrm{m}\times\mathrm{n}}$ or more precisely,

$$\mathrm{A}+\mathrm{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}.$$

Definition 1.30 (Scalar multiplication)**.** Let $\mathrm{A} = [a_{ij}]_{\mathrm{m}\times\mathrm{n}}$ be an $\mathrm{m}\times\mathrm{n}$ matrix over a scalar field $\mathbb{F}$, and $\alpha \in \mathbb{F}$. The scalar multiplication of α and A, denoted by $\alpha\mathrm{A}$, is an $\mathrm{m}\times\mathrm{n}$ matrix defined by $\alpha\mathrm{A} = [\alpha a_{ij}]_{\mathrm{m}\times\mathrm{n}}$ or more precisely,

$$\alpha\mathrm{A} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}.$$

Proposition 1.31. *The space $\mathbb{M}(m,n;\mathbb{F})$ is a vector space over $\mathbb{F}$ under the matrix addition and scalar multiplication defined in previous two definitions.*

Definition 1.32 (Matrix product)**.** Let $\mathrm{A} \in \mathbb{M}(m,n;\mathbb{F})$ and $\mathrm{B} \in \mathbb{M}(n,\ell;\mathbb{F})$ be two matrices over a scalar field $\mathbb{F}$. The matrix product of A and B, denoted by AB, is an $m\times\ell$ matrix given by $\mathrm{AB} = [c_{ij}]_{\mathrm{m}\times\mathrm{n}}$ with $c_{ij} = \sum\limits_{k=1}^{\mathrm{n}} a_{ik}b_{kj}$. In other words, the (i,j)-th entry of the product AB is the inner product of the i-th row of A and the j-th column of B.

Remark 1.33. The matrix product AB is only defined if the number of columns of A is the same as the number of rows of B. Therefore, even if AB is defined, BA might not make sense. When A and B are both $n \times \mathrm{n}$ square matrix, AB and BA are both defined; however, in general AB $\neq$ BA.

Remark 1.34. Let $\boldsymbol{v} \in \mathbb{F}^{\mathrm{n}}$ be a vector such that the k-th component of $\boldsymbol{v}$ is the same as the (i,k)-th entry of $\mathrm{A} \in \mathbb{M}(m,n;\mathbb{F})$, and $\boldsymbol{w} \in \mathbb{F}^{\mathrm{n}}$ be a vector such that the k-th component of $\boldsymbol{w}$ is the same as the (k,j)-th entry of $\mathrm{B} \in \mathbb{M}(n,\ell;\mathbb{F})$. Then the (i,j)-th entry of AB is simply the inner product of $\boldsymbol{v}$ and $\boldsymbol{w}$ in $\mathbb{F}^{\mathrm{n}}$.

Example 1.35. Let $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -1 & 1 \\ 3 & 0 & 2 \\ -1 & 1 & 0 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 0 & 1 & 1 \\ -4 & 1 & -2 \end{bmatrix}$$

but BA is not defined.

Proposition 1.36. *Let* $\mathrm{A} \in \mathbb{M}(\mathrm{m},\mathrm{n};\mathbb{F})$, $\mathrm{B} \in \mathbb{M}(\mathrm{n},\ell;\mathbb{F})$ *and* $\mathrm{C} \in \mathbb{M}(\ell,k;\mathbb{F})$. *Then*

$$\mathrm{A}(\mathrm{BC}) = (\mathrm{AB})\mathrm{C}\,.$$

Definition 1.37 (The range and the null space of matrices)**.** Let $\mathrm{A} \in \mathbb{M}(\mathrm{m},\mathrm{n};\mathbb{F})$. The ***range*** of A, denoted by $R(\mathrm{A})$, is the subset of $\mathbb{F}^{\mathrm{m}}$ given by

$$R(\mathrm{A}) = \left\{ \mathrm{A}\boldsymbol{x} \in \mathbb{F}^{\mathrm{m}} \,\middle|\, \boldsymbol{x} \in \mathbb{F}^{\mathrm{n}} \right\},$$

and the ***null space*** of A, denoted by null(A), is the subset of $\mathbb{F}^{\mathrm{n}}$ given by

$$\mathrm{null}(\mathrm{A}) = \left\{ \boldsymbol{x} \in \mathbb{F}^{\mathrm{n}} \,\middle|\, \mathrm{A}\boldsymbol{x} = \mathbf{0} \right\}.$$

Proposition 1.38. *Let* $\mathrm{A} \in \mathbb{M}(\mathrm{m},\mathrm{n};\mathbb{F})$. *Then* $R(\mathrm{A})$ *and* $\mathrm{null}(\mathrm{A})$ *are vector subspaces of* $\mathbb{F}^{\mathrm{n}}$ *and* $\mathbb{F}^{\mathrm{m}}$, *respectively.*

Definition 1.39 (Kronecker's delta)**.** The Kronecker delta is a function, denoted by δ, of two variables (usually positive integers) such that the function is 1 if the two variables are equal, and 0 otherwise. When the two variables are i and j, the value $\delta(i,j)$ is usually written as δ_{ij}; that is,

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j\,, \\ 1 & \text{if } i = j\,. \end{cases}$$

Definition 1.40 (Identity matrix)**.** The identity matrix of size n, denoted by $\mathrm{I_n}$, is the $\mathrm{n}\times\mathrm{n}$ square matrix with ones on the main diagonal and zeros elsewhere. In other words,

$$\mathrm{I_n} = [\delta_{ij}]_{\mathrm{n}\times\mathrm{n}}\,,$$

where δ_{ij} is the Kronecker delta.

When the size is clear from the context, $\mathrm{I_n}$ is sometimes denoted by I.

Definition 1.41 (Transpose)**.** Let $\mathrm{A} = [a_{ij}]_{\mathrm{m}\times\mathrm{n}}$ be a $\mathrm{m}\times\mathrm{n}$ matrix over scalar field $\mathbb{F}$. The transpose of A, denoted by $\mathrm{A^T}$, is the $\mathrm{n}\times\mathrm{m}$ matrix given by $[\mathrm{A^T}]_{ij} = a_{ji}$.

By the definition of product of matrices, we can easily derive the following two propositions.

Proposition 1.42. *Let* $\mathrm{A}\in\mathbb{M}(\mathrm{m},\mathrm{n};\mathbb{F})$ *and* $\mathrm{B}\in\mathbb{M}(\mathrm{n},\ell;\mathbb{F})$*. Then* $(\mathrm{AB})^{\mathrm{T}} = \mathrm{B^T A^T}$*.*

Proposition 1.43. *Let* $\mathrm{A} = [a_{ij}]_{\mathrm{m}\times\mathrm{n}}$ *be a* $\mathrm{m}\times\mathrm{n}$ *matrix over scalar field* $\mathbb{F}$*, and* $(\cdot,\cdot)_{\mathbb{F}^\mathrm{n}}$ *and* $(\cdot,\cdot)_{\mathbb{F}^\mathrm{m}}$ *be the standard inner products on* $\mathbb{F}^\mathrm{n}$ *and* $\mathbb{F}^\mathrm{m}$*, respectively. Then*

$$(\mathrm{A}\boldsymbol{x},\boldsymbol{y})_{\mathbb{F}^\mathrm{m}} = (\boldsymbol{x},\overline{\mathrm{A^T}}\boldsymbol{y})_{\mathbb{F}^\mathrm{n}} \qquad \forall\,\boldsymbol{x}\in\mathbb{F}^\mathrm{n},\boldsymbol{y}\in\mathbb{F}^\mathrm{m}\,.$$

Definition 1.44 (Rank and nullity of matrices)**.** The ***rank*** of a matrix A, denoted by rank(A), is the dimension of the vector space generated (or spanned) by its columns. The ***nullity*** of a matrix A, denoted by nullity(A), is the dimension of the null space of A.

Remark 1.45. The matrix $\overline{\mathrm{A^T}}$ is often called the ***conjugate transpose*** of the matrix A.

Remark 1.46. The rank defined above is also referred to the ***column rank***, and the ***row rank*** of a matrix is the dimension of the vector space spanned by its rows. One should immediately notice that the column rank of A equals the dimension of $R(\mathrm{A})$ and the row rank of A equalis the dimension of $R(\mathrm{A^T})$.

Theorem 1.47. *Let* $\mathrm{A}\in\mathbb{M}(\mathrm{m},\mathrm{n};\mathbb{F})$*. Then* $\mathrm{rank(A)} + \mathrm{nullity(A)} = \mathrm{n}$.

Proof. Without loss of generality, we assume that $\mathrm{nulltiy(A)} = k < \mathrm{n}$, and $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ be a basis of null(A). Then there exists $\mathrm{n}-k$ vectors $\{\boldsymbol{v}_{k+1},\cdots,\boldsymbol{v}_\mathrm{n}\}$ such that $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_\mathrm{n}\}$ is a basis of $\mathbb{F}^\mathrm{n}$. We conclude the theorem by showing that $\{\mathrm{A}\boldsymbol{v}_{k+1},\cdots,\mathrm{A}\boldsymbol{v}_\mathrm{n}\}$ is a basis of $R(\mathrm{A})$.

First, we claim that $\{A\boldsymbol{v}_{k+1}, \cdots, A\boldsymbol{v}_n\}$ is a linearly independent set of vectors. To see this, suppose that $\alpha_{k+1}, \cdots, \alpha_n \in \mathbb{F}$ such that

$$\alpha_{k+1}A\boldsymbol{v}_{k+1} + \cdots + \alpha_n A\boldsymbol{v}_n = \boldsymbol{0}\,.$$

Then $A(\alpha_{k+1}\boldsymbol{v}_{k+1} + \cdots + \alpha_n\boldsymbol{v}_n) = \boldsymbol{0}$ which implies that $\alpha_{k+1}\boldsymbol{v}_{k+1} + \cdots + \alpha_n\boldsymbol{v}_n \in \mathrm{null}(A)$. Since $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$ is a basis of null(A), there exist $\alpha_1, \cdots, \alpha_k \in \mathbb{F}$ such that

$$\alpha_1\boldsymbol{v}_1 + \alpha_k\boldsymbol{v}_k = \alpha_{k+1}\boldsymbol{v}_{k+1} + \cdots + \alpha_n\boldsymbol{v}_n\,.$$

By the linear independence of $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$, we must have $\alpha_1 = \cdots = \alpha_n = 0$ which shows the linear independence of $\{A\boldsymbol{v}_{k+1}, \cdots, A\boldsymbol{v}_n\}$.

Let $\boldsymbol{w} \in R(A)$. Then $\boldsymbol{w} = A\boldsymbol{v}$ for some $\boldsymbol{v} \in \mathbb{F}^n$. Since $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is a basis of $\mathbb{F}^n$, there exist $\beta_1, \cdots, \beta_n \in \mathbb{F}$ such that $\boldsymbol{v} = \beta_1\boldsymbol{v}_1 + \cdots + \beta_n\boldsymbol{v}_n$. As a consequence, by the fact that $A\boldsymbol{v}_j = \boldsymbol{0}$ for $1 \leqslant j \leqslant k$,

$$\boldsymbol{w} = A\boldsymbol{v} = A(\beta_1\boldsymbol{v}_1 + \cdots + \beta_n\boldsymbol{v}_n) = \beta_1 A\boldsymbol{v}_1 + \cdots \beta_n A\boldsymbol{v}_n = \beta_{k+1}A\boldsymbol{v}_{k+1} + \cdots + \beta_n A\boldsymbol{v}_n\,;$$

thus $\boldsymbol{w}$ can be written as a linear combination of $\{A\boldsymbol{v}_{k+1}, \cdots, A\boldsymbol{v}_n\}$. □

Theorem 1.48. *The rank of a matrix is the same as the rank of its transpose. In other words, for a given matrix the row rank equals the column rank.*

Proof. Let A be a m × n matrix, and $(\cdot,\cdot)_{\mathbb{F}^n}$, $(\cdot,\cdot)_{\mathbb{F}^m}$ be the standard inner products on $\mathbb{F}^n$, $\mathbb{F}^m$, respectively. Then Proposition 1.43 implies that

$$\begin{aligned} \boldsymbol{y} \in R(A)^\perp \;&\Leftrightarrow\; (\boldsymbol{y}, A\boldsymbol{x})_{\mathbb{F}^m} = 0 \text{ for all } \boldsymbol{x} \in \mathbb{F}^n \;\Leftrightarrow\; (\overline{A^T}\boldsymbol{y}, \boldsymbol{x})_{\mathbb{F}^n} = 0 \text{ for all } \boldsymbol{x} \in \mathbb{F}^n \\ &\Leftrightarrow\; \overline{A^T}\boldsymbol{y} = \boldsymbol{0} \;\Leftrightarrow\; \boldsymbol{y} \in \mathrm{null}(\overline{A^T})\,. \end{aligned}$$

In other words, $R(A)^\perp = \mathrm{null}(\overline{A^T})$. Since the column rank of A is the dimension of $R(A)$, we must have

$$\mathrm{nullity}(A^T) = \mathrm{nullity}(\overline{A^T}) = \dim\big(R(A)^\perp\big) = m - \text{the column rank of } A\,.$$

On the other hand, Theorem 1.47 implies that

$$\mathrm{rank}(A^T) + \mathrm{nullity}(A^T) = m\,;$$

thus the column rank of A is the same as the row rank of A. □

Definition 1.49. Let $\mathrm{A} \in \mathbb{M}(\mathrm{n}, \mathrm{n}; \mathbb{F})$ be a square matrix. A is said to be ***invertible*** if there exists $\mathrm{B} \in \mathbb{M}(\mathrm{n}, \mathrm{n}; \mathbb{F})$ such that $\mathrm{AB} = \mathrm{I_n}$. The matrix B is called the ***inverse matrix*** of A, and is usually denoted by A^{-1}.

Proposition 1.50. *Let* $\mathrm{A} \in \mathbb{M}(\mathrm{n}, \mathrm{n}; \mathbb{F})$ *be invertible. Then* $\mathit{rank}(\mathrm{A}) = \mathit{rank}(\mathrm{A}^{-1}) = \mathrm{n}$.

Proof. Since $\mathrm{A}(\mathrm{A}^{-1}\boldsymbol{b}) = (\mathrm{AA}^{-1})\boldsymbol{b} = \boldsymbol{b}$ for all $\boldsymbol{b} \in \mathbb{F}^{\mathrm{n}}$, $R(\mathrm{A}) = \mathbb{F}^{\mathrm{n}}$ which implies that $\mathrm{rank}(\mathrm{A}) = \mathrm{n}$. We next show that $R(\mathrm{A}^{-1}) = \mathbb{F}^{\mathrm{n}}$. Denote A^{-1} by B, and let $\boldsymbol{b} \in \mathbb{F}^{\mathrm{n}}$. Then $\mathrm{B^T}(\mathrm{A^T}\boldsymbol{b}) = (\mathrm{B^T A^T})\boldsymbol{b} = \boldsymbol{b}$ since $\mathrm{B^T A^T} = (\mathrm{AB})^{\mathrm{T}} = \mathrm{I_n}$. This observation implies that $R(\mathrm{B^T}) = \mathbb{F}^{\mathrm{n}}$, and the theorem is then concluded by Theorem 1.48. □

Proposition 1.51. *Let* $\mathrm{A} \in \mathbb{M}(\mathrm{n}, \mathrm{n}; \mathbb{F})$ *be invertible. Then* $\mathrm{A}^{-1}\mathrm{A} = \mathrm{AA}^{-1} = \mathrm{I_n}$.

Proof. We show that for all $\boldsymbol{b} \in \mathbb{F}^{\mathrm{n}}$, $\mathrm{A}^{-1}\mathrm{A}\boldsymbol{b} = \boldsymbol{b}$. Since A is invertible, $\mathrm{rank}(\mathrm{A}^{-1}) = \mathrm{n}$; thus $R(\mathrm{A}^{-1}) = \mathbb{F}^{\mathrm{n}}$ which implies that for each $\boldsymbol{b} \in \mathbb{F}^{\mathrm{n}}$, there exists $\boldsymbol{x} \in \mathbb{F}$ such that $\mathrm{A}^{-1}\boldsymbol{x} = \boldsymbol{b}$. As a consequence,

$$(\mathrm{A}^{-1}\mathrm{A})\boldsymbol{b} = (\mathrm{A}^{-1}\mathrm{A})(\mathrm{A}^{-1}\boldsymbol{x}) = \mathrm{A}^{-1}(\mathrm{AA}^{-1})\boldsymbol{x} = \mathrm{A}^{-1}\boldsymbol{x} = \boldsymbol{b}\,.$$ □

1.4.1 Elementary Row Operations and Elementary Matrices

Definition 1.52 (Elementary row operations)**.** For an $\mathrm{n} \times \mathrm{m}$ matrix A, three types of elementary row operations can be performed on A:

1. The first type of row operation on A switches all matrix elements on the i-th row with their counterparts on j-th row.

2. The second type of row operation on A multiplies all elements on the i-th row by a non-zero scalar λ.

3. The third type of row operation on A adds j-th row multiplied by a scalar μ to the i-th row.

The elementary row operation on an $\mathrm{n} \times \mathrm{m}$ matrix A can be done by multiplying A by an $\mathrm{n} \times \mathrm{n}$ matrix, called an elementary matrix, on the left. The elementary matrices are defined in the following

Definition 1.53 (Elementary matrices)**.** An elementary matrix is a matrix which differs from the identity matrix by one single elementary row operation.

1. Switching the i_0-th and j_0-th rows of A, where $i_0 \neq j_0$, is done by left multiplied A by the matrix $\mathrm{E} = [e_{ij}]_{\mathrm{n}\times\mathrm{n}}$ given by

$$e_{ij} = \begin{cases} 1 & \text{if } (i,j) = (i_0, j_0) \text{ or } (i,j) = (j_0, i_0) \text{ or } i = j = k_0 \text{ for some } k_0 \neq i_0, j_0, \\ 0 & \text{otherwise,} \end{cases}$$

or in the matrix form,

$$\mathrm{E} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & & & & & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & & & & & \vdots \\ \vdots & & 0 & 0 & 0 & & & 1 & & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & & & & & \vdots \\ \vdots & & & & 0 & \ddots & 0 & & & & \vdots \\ \vdots & & & & & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & & 1 & & & 0 & 0 & 0 & & \vdots \\ 0 & & & & & & & \ddots & 1 & \ddots & \vdots \\ 0 & & & & & & & & 0 & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{the } i_0\text{-th row} \\ \\ \\ \\ \leftarrow \text{the } j_0\text{-th row} \\ \\ \\ \\ \end{array}$$

$$\begin{array}{cc} \uparrow & \uparrow \\ \text{the } i_0\text{-th column} & \text{the } j_0\text{-th column} \end{array}$$

2. Multiplying the k_0-th row of A by a non-zero scalar λ is done by left multiplied A by the matrix $\mathrm{E} = [e_{ij}]_{\mathrm{n}\times\mathrm{n}}$ given by

$$e_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ \lambda & \text{if } i = j = k_0, \\ 1 & \text{otherwise,} \end{cases}$$

or in the matrix form,

$$\mathrm{E} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & & \vdots \\ \vdots & & 0 & 1 & 0 & & & & \vdots \\ \vdots & & & 0 & \lambda & 0 & & & \vdots \\ \vdots & & & & 0 & 1 & 0 & & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & & 0 & 1 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \leftarrow \text{the } k_0\text{-th row} \\ \\ \\ \\ \\ \end{array}$$

3. Adding the j_0-th row of A multiplied by a scalar μ to the i_0-th row, where $i_0 \neq j_0$, is done by left multiplied A by the matrix $\mathrm{E} = [e_{ij}]_{\mathrm{n}\times\mathrm{n}}$ given by

$$e_{ij} = \begin{cases} 1 & \text{if } i = j, \\ \mu & \text{if } (i,j) = (i_0, j_0), \\ 0 & \text{otherwise,} \end{cases}$$

or in the matrix form,

$$\mathrm{E} = \begin{array}{cl} \left[\begin{array}{ccccccccc} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & & & & & & 0 \\ \vdots & \ddots & \ddots & \ddots & & & \mu & & 0 \\ \vdots & & \ddots & \ddots & \ddots & & & & 0 \\ \vdots & & & 0 & 1 & 0 & & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & & 0 & 1 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{array}\right] & \begin{array}{l} \\ \\ \leftarrow \text{the } i_0\text{-th row} \\ \\ \\ \\ \\ \\ \\ \end{array} \\ \begin{array}{ccccccccc} & & & & & & \uparrow & & \\ & & & & & & \text{the } j_0\text{-th column} & & \end{array} & \end{array}$$

Proposition 1.54. *Every elementary matrix is invertible.*

Theorem 1.55. *Let* $\mathrm{A} \in \mathbb{M}(\mathrm{n}, \mathrm{n}; \mathbb{F})$ *be a square matrix. The following statements are equivalent:*

1. $R(\mathrm{A}) = \mathbb{F}^{\mathrm{n}}$.

2. $rank(\mathrm{A}) = \mathrm{n}$.

3. $\mathrm{A}\boldsymbol{x} = \boldsymbol{b}$ *has a unique solution* $\boldsymbol{x}$ *for all* $\boldsymbol{b} \in \mathbb{F}^{\mathrm{n}}$.

4. A *is invertible.*

5. $\mathrm{A} = \mathrm{E}_k\mathrm{E}_{k-1}\cdots\mathrm{E}_2\mathrm{E}_1$ *for some elementary matrices* $\mathrm{E}_1, \cdots, \mathrm{E}_k$.

Proof. Note that by definition 1,2,3 are equivalent, and Proposition 1.50 shows that $4 \Rightarrow 2$. The implication from 3 to 4 is due to the fact that the map $\boldsymbol{b} \mapsto \boldsymbol{x}$, where $\boldsymbol{x}$ is the unique solution to $\mathrm{A}\boldsymbol{x} = \boldsymbol{b}$, is the inverse of A. Proposition 1.54 provides that $5 \Rightarrow 4$. That $3 \Rightarrow 5$ follows from that at most $\mathrm{n}(\mathrm{n}+1)$ elementary row operations has to be applied on A to reach the identity matrix. □

1.5 Determinants

In order to introduce the notion of the determinant of square matrices, we need to talk about permutations first. Note that there are many other ways of defining determinants, but it is quite elegant to use the notion of permutations, and we can derive a lot of useful results via this definition.

Definition 1.56 (Permutations)**.** A sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ of positive integers not exceeding n, with the property that no two of the k_i are equal, is called a ***permutation of degree*** n. The collection of all permutations of degree n is denoted by $\mathbb{P}(\mathrm{n})$.

A sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ can be obtained from the sequence $(1, 2, \cdots, \mathrm{n})$ by a finite number of interchanges of pairs of elements. For example, if $k_1 \neq 1$, we can transpose 1 and k_1, obtaining $(k_1, \cdots, 1, \cdots)$. Proceeding in this way we shall arrive at the sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ after n or less such interchanges of pairs.

In general, a permutation $(k_1, k_2, \cdots, k_\mathrm{n})$ can be expressed as

$$\tau_{(i_N, j_N)} \cdots \tau_{(i_2, j_2)} \tau_{(i_1, j_1)} (1, 2, \cdots, \mathrm{n}) = (k_1, k_2, \cdots, k_\mathrm{n}),$$

where $\tau_{(i,j)}$ is a "pair-interchange operator" which swaps the i-th and the j-th elements (of the object fed into), and N is the number of pair interchanges. We call such pair-interchange operators the permutation operator. Since $\tau_{(i,j)}$ is the inverse operator of itself, we also have

$$\tau_{(i_1, j_1)} \tau_{(i_2, j_2)} \cdots \tau_{(i_N, j_N)} (k_1, k_2, \cdots, k_\mathrm{n}) = (1, 2, \cdots, \mathrm{n}).$$

We remark here that the number of pair interchanges (from $(1, 2, \cdots, \mathrm{n})$ to $(k_1, k_2, \cdots, k_\mathrm{n})$) is not unique; nevertheless, if two processes of pair interchanges lead to the same permutation, then the numbers of interchanges differ by an even number. This leads to the following

Definition 1.57 (Even and odd permutations)**.** A permutation $(k_1, \cdots, k_\mathrm{n})$ is called an ***even*** (***odd***) ***permutation*** of degree n if the number required to interchange pairs of $(1, 2, \cdots, \mathrm{n})$ in order to obtain $(k_1, k_2, \cdots, k_\mathrm{n})$ is even (odd).

Example 1.58. If $\mathrm{n} = 3$, the permutation $(3, 1, 2)$ can be obtained by interchanging pairs of $(1, 2, 3)$ twice:

$$(1, 2, 3) \xrightarrow{\tau_{(1,3)}} (3, 2, 1) \xrightarrow{\tau_{(2,3)}} (3, 1, 2);$$

thus $(3,1,2)$ is an even permutation of $(1,2,3)$. On the other hand, $(1,3,2)$ is obtained by interchanging pairs of $(1,2,3)$ once:

$$(1,2,3) \xrightarrow{\tau_{(2,3)}} (1,3,2);$$

thus $(1,3,2)$ is an odd permutation of $(1,2,3)$.

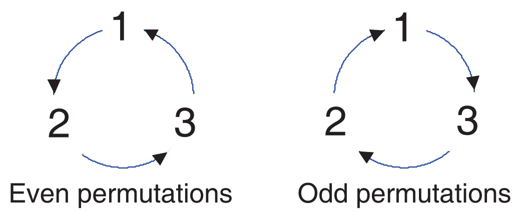


Figure 1.1: Even and odd permutations of degree 3

For $\mathrm{n} = 3$, the even and odd permutations can also be viewed as the orientation of the permutation (k_1, k_2, k_3). To be more precise, if $(1,2,3)$ is arranged in a counter-clockwise orientation (see Figure 1.1), then an even permutation of degree 3 is a permutation in the counter-clockwise orientation, while an odd permutation of degree 3 is a permutation in the clockwise orientation. From figure 1.1, it is easy to see that $(3,1,2)$ is an even permutation of degree 3 and $(1,3,2)$ is an odd permutation of degree 3.

Definition 1.59 (The permutation symbol)**.** The permutation symbol $\varepsilon_{k_1 k_2 \cdots k_\mathrm{n}}$ is a function of permutations of degree n defined by

$$\varepsilon_{k_1 k_2 \cdots k_\mathrm{n}} = \begin{cases} 1 & \text{if } (k_1, k_2, \cdots, k_\mathrm{n}) \text{ is an even permutation of degree n,} \\ -1 & \text{if } (k_1, k_2, \cdots, k_\mathrm{n}) \text{ is an odd permutation of degree n.} \end{cases}$$

Remark 1.60. One can extend the domain the permutation symbol to all the sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ by defining that $\varepsilon_{k_1 k_2 \cdots k_\mathrm{n}} = 0$ if $(k_1, k_2, \cdots, k_\mathrm{n})$ is not a permutation of degree n.

Definition 1.61 (Determinants)**.** Given an $\mathrm{n} \times \mathrm{n}$ matrix $\mathrm{A} = [a_{ij}]$, the determinants of A, denoted by $\det(\mathrm{A})$, is defined by

$$\det(\mathrm{A}) = \sum_{(k_1, \cdots, k_\mathrm{n}) \in \mathbb{P}(\mathrm{n})} \varepsilon_{k_1 k_2 \cdots k_\mathrm{n}} \prod_{\ell=1}^{\mathrm{n}} a_{\ell k_\ell}.$$

We note that the product $\prod\limits_{\ell=1}^{\mathrm{n}} a_{\ell k_\ell}$ in the definition of the determinant is formed by multiplying n-elements which appears exactly once in each row and column.

Proposition 1.62. *Let* E *be an elementary matrix. Then*

1. $\det(\mathrm{E}) \neq 0$.
2. $\det(\mathrm{E}) = \det(\mathrm{E}^{\mathrm{T}})$.
3. *If* A *is an* $\mathrm{n}\times\mathrm{n}$ *matrix, then* $\det(\mathrm{EA}) = \det(\mathrm{E})\det(\mathrm{A})$.

The proof of the proposition above is not difficult, and is left as an exercise.

Corollary 1.63. *Let* $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{\mathrm{n}} \in \mathbb{R}^{\mathrm{n}}$ *be (column) vectors,* $c \in \mathbb{R}$, *and*

$$\begin{aligned}
\mathrm{A} &= \left[\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{\mathrm{n}}\right], \\
\mathrm{B} &= \left[\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{j-1} \vdots \lambda\boldsymbol{v}_j \vdots \boldsymbol{v}_{j+1} \vdots \cdots \vdots \boldsymbol{v}_{\mathrm{n}}\right], \\
\mathrm{C} &= \left[\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{j-1} \vdots \boldsymbol{v}_j + \mu\boldsymbol{v}_i \vdots \boldsymbol{v}_{j+1} \vdots \cdots \vdots \boldsymbol{v}_{\mathrm{n}}\right] \qquad \textit{for some } i \neq j\,.
\end{aligned}$$

Then $\det(\mathrm{B}) = \lambda\det(\mathrm{A})$, *and* $\det(\mathrm{C}) = \det(\mathrm{A})$.

Proof. The corollary is easily concluded since $\mathrm{B} = \mathrm{E}_1\mathrm{A}$ and $\mathrm{C} = \mathrm{E}_2\mathrm{A}$ for some elementary matrices E_1 and E_2 with $\det(\mathrm{E}_1) = c$ and $\det(\mathrm{E}_2) = 1$. □

Corollary 1.64. *Let* A *be an* $\mathrm{n}\times\mathrm{n}$ *matrix. Then* A *is invertible if and only if* $\det(\mathrm{A}) \neq 0$.

Proof. ($\Rightarrow$) Since A is invertible, Theorem 1.55 implies that

$$\mathrm{A} = \mathrm{E}_k\mathrm{E}_{k-1}\cdots\mathrm{E}_2\mathrm{E}_1$$

for some elementary matrices $\mathrm{E}_1, \cdots, \mathrm{E}_k$, and this corollary follows from Proposition 1.62.

($\Leftarrow$) Note that A is invertible if and only if $\mathrm{rank}(\mathrm{A}) = \mathrm{rank}(\mathrm{A}^{\mathrm{T}}) = \mathrm{n}$. Therefore, if A is not invertible, the row vectors of A are linearly dependent; thus there exists a non-zero vectors $(\alpha_1, \cdots, \alpha_{\mathrm{n}}) \in \mathbb{F}^{\mathrm{n}}$ such that

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots \alpha_{\mathrm{n}}\boldsymbol{v}_{\mathrm{n}} = \mathbf{0}\,,$$

where $\mathrm{A}^{\mathrm{T}} = \left[\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{\mathrm{n}}\right]$. Suppose that $\alpha_j \neq 0$. Then

$$\boldsymbol{v}_j = \beta_1\boldsymbol{v}_1 + \cdots \beta_{j-1}\boldsymbol{v}_{j-1} + \beta_{j+1}\boldsymbol{v}_{j+1} + \cdots \beta_{\mathrm{n}}\boldsymbol{v}_{\mathrm{n}}\,;$$

thus after applying $(\mathrm{n}-1)$-times elementary row operations of the third type (adding some multiple of certain row to another row) on A we reach a matrix whose j-th row is a zero (row) vector. Thereofre, for some elementary matrices $\mathrm{E}_1, \cdots, \mathrm{E}_{\mathrm{n}-1}$ we have

$$\det(\mathrm{E}_{\mathrm{n}-1}\cdots \mathrm{E}_1\mathrm{A}) = 0$$

which implies that $\det(\mathrm{A}) = 0$. □

Corollary 1.65. *Let* A *be an* $\mathrm{n}\times\mathrm{n}$ *matrix. Then the determinant of* A *and* A^{T}*, the transpose of* A*, are the same; that is,*

$$\det(\mathrm{A}) = \det(\mathrm{A}^{\mathrm{T}}).$$

Proof. If A is not invertible, then A^{T} is not invertible either because of Theorem 1.48. Therefore, $\det(\mathrm{A}) = 0 = \det(\mathrm{A}^{\mathrm{T}})$.

Now suppose that A is invertible. Then Theorem 1.55 implies that

$$\mathrm{A} = \mathrm{E}_k\mathrm{E}_{k-1}\cdots\mathrm{E}_2\mathrm{E}_1$$

for some elementary matrices $\mathrm{E}_1, \cdots, \mathrm{E}_k$. Since all $\mathrm{E}_j^{\mathrm{T}}$'s are also elementary matrices, by Proposition 1.62 we conclude that

$$\begin{aligned}\det(\mathrm{A}^{\mathrm{T}}) &= \det(\mathrm{E}_1^{\mathrm{T}}\cdots\mathrm{E}_k^{\mathrm{T}}) = \det(\mathrm{E}_1^{\mathrm{T}})\cdots\det(\mathrm{E}_k^{\mathrm{T}})\\ &= \det(\mathrm{E}_k^{\mathrm{T}})\cdots\det(\mathrm{E}_1^{\mathrm{T}})\\ &= \det(\mathrm{E}_k)\cdots\det(\mathrm{E}_1) = \det(\mathrm{E}_k\cdots\mathrm{E}_1) = \det(\mathrm{A})\,.\end{aligned}$$

□

Corollary 1.66. *Let* A, B *be* $\mathrm{n}\times\mathrm{n}$ *matrices. Then* $\det(\mathrm{AB}) = \det(\mathrm{A})\det(\mathrm{B})$.

Proof. If A is not invertible, then AB is not invertible either; thus in this case $\det(\mathrm{A})\det(\mathrm{B}) = 0 = \det(\mathrm{AB})$.

Now suppose that A is invertible. Then Theorem 1.55 implies that

$$\mathrm{A} = \mathrm{E}_k\mathrm{E}_{k-1}\cdots\mathrm{E}_2\mathrm{E}_1$$

for some elementary matrices $\mathrm{E}_1, \cdots, \mathrm{E}_k$. As a consequence, Proposition 1.62 implies that

$$\begin{aligned}\det(\mathrm{AB}) &= \det(\mathrm{E}_k\cdots\mathrm{E}_1\mathrm{B}) = \det(\mathrm{E}_k)\det(\mathrm{E}_{k-1}\cdots\mathrm{E}_1\mathrm{B})\\ &= \cdots = \det(\mathrm{E}_k)\cdots\det(\mathrm{E}_1)\det(\mathrm{B})\\ &= \det(\mathrm{E}_k\cdots\mathrm{E}_1)\det(\mathrm{B}) = \det(\mathrm{A})\det(\mathrm{B})\,.\end{aligned}$$

□

Definition 1.67 (Minor, Cofactor, and Adjoint matrices)**.** Let A be an $\mathrm{n} \times \mathrm{n}$ matrix, and $\mathrm{A}(\hat{i}, \hat{j})$ be the $(\mathrm{n}-1) \times (\mathrm{n}-1)$ matrix obtained by eliminating the i-th row and j-th column of A; that is,

$$\mathrm{A}(\hat{i}, \hat{j}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1\mathrm{n}} \\ \vdots & \ddots & & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & a_{(i-1)2} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)\mathrm{n}} \\ a_{(i+1)1} & a_{(i+1)2} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)\mathrm{n}} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{\mathrm{n}1} & a_{\mathrm{n}2} & \cdots & a_{\mathrm{n}(j-1)} & a_{\mathrm{n}(j+1)} & \cdots & a_{\mathrm{nn}} \end{bmatrix}.$$

The (i, j)-th ***minor*** of A is the determinant of $A(\hat{i}, \hat{j})$, and the (i, j)-th ***cofactor***, is the (i, j)-th minor of A multiplied by $(-1)^{i+j}$. The ***adjoint matrix*** of A, denoted by Adj(A), is the transpose of the cofactor matrix; that is,

$$\big[\mathrm{Adj}(\mathrm{A})\big]_{ij} = (-1)^{i+j} \det \big(\mathrm{A}(\hat{j}, \hat{i})\big)\,.$$

Example 1.68. Let $\mathrm{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & -1 & 2 \\ 0 & 2 & -1 \end{bmatrix}$. Then the minor matrix of A is $\begin{bmatrix} -3 & -3 & 6 \\ -8 & -1 & 2 \\ 7 & -7 & -7 \end{bmatrix}$, the cofactor matrix of A is $\begin{bmatrix} -3 & 3 & 6 \\ 8 & -1 & -2 \\ 7 & 7 & -7 \end{bmatrix}$, and the adjoint matrix of A is $\begin{bmatrix} -3 & 8 & 7 \\ 3 & -1 & 7 \\ 6 & -2 & -7 \end{bmatrix}$.

The following lemma provides a way of computing the minors of a matrix.

Lemma 1.69. *Let* A *be an* $\mathrm{n} \times \mathrm{n}$ *matrix. Then*

$$\det \big(\mathrm{A}(\hat{i}, \hat{j})\big) = (-1)^{i+j} \sum_{(k_1, \cdots, k_\mathrm{n}) \in \mathbb{P}(\mathrm{n}),\, k_i = j} \varepsilon_{k_1 k_2 \cdots k_\mathrm{n}} \prod_{\substack{1 \leqslant \ell \leqslant \mathrm{n} \\ \ell \neq i}} a_{\ell k_\ell}\,.$$

Proof. Fix $(i, j) \in \{1, 2, \cdots, \mathrm{n}\} \times \{1, 2, \cdots, \mathrm{n}\}$. The matrix $\mathrm{A}(\hat{i}, \hat{j})$ is given by $\mathrm{A}(\hat{i}, \hat{j}) = [b_{\alpha\beta}]$, where $\alpha, \beta = 1, 2, \cdots, \mathrm{n}-1$, and

$$b_{\alpha\beta} = \begin{cases} a_{\alpha\beta} & \text{if } \alpha < i \text{ and } \beta < j, \\ a_{(\alpha+1)\beta} & \text{if } \alpha > i \text{ and } \beta < j, \\ a_{\alpha(\beta+1)} & \text{if } \alpha < i \text{ and } \beta > j, \\ a_{(\alpha+1)(\beta+1)} & \text{if } \alpha > i \text{ and } \beta > j. \end{cases}$$

Each permutation $(\sigma_1, \sigma_2, \cdots, \sigma_{\mathrm{n}-1})$ of degree $\mathrm{n}-1$ corresponds a unique permutation $(k_1, k_2, \cdots, k_\mathrm{n})$ of degree n such that

1. $k_i = j$;

2. for each $\tau \in \{1, \cdots, i-1\}$ and $\iota \in \{i, i+1, \cdots, \mathrm{n}-1\}$,
$$k_\tau = \begin{cases} \sigma_\tau & \text{if } \sigma_\tau < j\,, \\ \sigma_\tau + 1 & \text{if } \sigma_\tau \geqslant j\,, \end{cases} \qquad \text{and} \qquad k_{\iota+1} = \begin{cases} \sigma_\iota & \text{if } \sigma_\iota < j\,, \\ \sigma_\iota + 1 & \text{if } \sigma_\iota \geqslant j\,. \end{cases}$$

We now determine the sign of $\varepsilon_{\sigma_1\sigma_2\cdots\sigma_{\mathrm{n}-1}}$ and $\varepsilon_{k_1k_2\cdots k_{\mathrm{n}}}$. Note that if a process of pair interchanges of the permutation $(\sigma_1, \sigma_2, \cdots, \sigma_{\mathrm{n}-1})$ leads to $(1, 2, \cdots, \mathrm{n}-1)$, then similar process of pair interchanges of the permutation $(k_1, k_2, \cdots, k_{i-1}, j, k_{i+1}, \cdots, k_{\mathrm{n}})$, by leaving the i-th slot fixed, leads to the permutation of degree n
$$\begin{cases} (1, 2, \cdots, j-1, j+1, \cdots, i-1, j, i, \cdots, \mathrm{n}) & \text{if } \; i > j, \\ (1, 2, \cdots, i-1, j, i, \cdots, j-1, j+1, \cdots, \mathrm{n}) & \text{if } \; i < j, \\ (1, 2, \cdots, \mathrm{n}) & \text{if } \; j = i. \end{cases}$$

For the case that $i \neq j$, another $|i-j|$-times of pair interchanges leads to $(1, 2, \cdots, \mathrm{n})$. To be more precise, suppose that $i > j$. We first interchange the $(i-2)$-th and the $(i-1)$-th components, and then interchange that $(i-3)$-th and the $(i-2)$-th components, and so on. After $(i-j)$-times of pair interchanges, we reach $(1, 2, \cdots, \mathrm{n})$. Symbolically,
$$\begin{array}{c} (1, 2, \cdots, j-1, j+1, \cdots, i-1, j, i, \cdots, \mathrm{n}) \\ \downarrow \tau_{(i\,-\,2,\,i\,-\,1)} \\ (1, 2, \cdots, j-1, j+1, \cdots, i-2, j, i-1, \cdots, \mathrm{n}) \\ \downarrow \tau_{(i\,-\,3,\,i\,-\,2)} \\ (1, 2, \cdots, j-1, j+1, \cdots, i-3, j, i-2 \cdots, \mathrm{n}) \\ \downarrow \\ \vdots \\ \downarrow \\ (1, 2, \cdots, \mathrm{n}). \end{array}$$

Similar argument applies to the case $i < j$; thus
$$\varepsilon_{\sigma_1\sigma_2\cdots\sigma_{\mathrm{n}-1}} = (-1)^{|i-j|}\varepsilon_{k_1k_2\cdots k_{\mathrm{n}}} = (-1)^{i+j}\varepsilon_{k_1k_2\cdots k_{\mathrm{n}}}.$$

As a consequence,
$$\begin{aligned} \det\big(\mathrm{A}(\hat{i}, \hat{j})\big) &= \sum_{(\sigma_1, \sigma_2, \cdots, \sigma_{\mathrm{n}-1}) \in \mathbb{P}(\mathrm{n}-1)} \varepsilon_{\sigma_1\sigma_2\cdots\sigma_{\mathrm{n}-1}} \prod_{\tau=1}^{\mathrm{n}-1} b_{\tau\sigma_\tau} \\ &= (-1)^{i+j} \sum_{(k_1, \cdots, k_{\mathrm{n}}) \in \mathbb{P}(\mathrm{n}),\, k_i = j} \varepsilon_{k_1k_2\cdots k_{\mathrm{n}}} \prod_{\substack{1 \leqslant \ell \leqslant \mathrm{n} \\ \ell \neq i}} a_{\ell k_\ell}. \end{aligned}$$
$\square$

Theorem 1.70. *Let* A *be an* $n \times n$ *matrix. Then*

$$\mathrm{Adj}(A)A = A\mathrm{Adj}(A) = \det(A)I_n.$$

Proof. Let $A = [a_{ij}]$. By definition of matrix multiplications,

$$\begin{aligned}\big(\mathrm{Adj}(A)A\big)_{ij} &= \sum_{m=1}^{n} \big(\mathrm{Adj}(A)\big)_{im} a_{mj} = \sum_{m=1}^{n} \Big[\sum_{(k_1,\cdots,k_n)\in\mathbb{P}(n),\, k_m=i} \varepsilon_{k_1k_2\cdots k_n} \prod_{\substack{1\leqslant \ell\leqslant n\\ \ell\neq m}} a_{\ell k_\ell}\Big] a_{mj} \\ &= \begin{cases} \displaystyle\sum_{(k_1,\cdots,k_n)\in\mathbb{P}(n)} \varepsilon_{k_1k_2\cdots k_n} \prod_{\ell=1}^{n} a_{\ell k_\ell} & \text{if } \ i=j, \\ 0 & \text{if } \ i\neq j. \end{cases}\end{aligned}$$

The conclusion then follows from the definition of the determinant. □

Corollary 1.71. *Let* $A = [a_{ij}]$ *be an* $n \times n$ *matrix, and* $C = [c_{ij}]$ *be the adjoint matrix of* A. *Then*

$$\det(A) = \sum_{j=1}^{n} a_{ij}c_{ji} = \sum_{j=1}^{n} a_{ji}c_{ij} \quad \forall\, 1 \leqslant i \leqslant n\,.$$

Corollary 1.72. *Let* A *be an* $n \times n$ *matrix and* $\det(A) \neq 0$. *Then the matrix* $\dfrac{\mathrm{Adj}(A)}{\det(A)}$ *is the inverse matrix of* A, *or equivalently,*

$$\mathrm{Adj}(A) = \det(A)A^{-1}. \tag{1.2}$$

1.5.1 Variations of determinants

Let δ be an operator satisfying the "product rule" $\delta(fg) = f\delta g + (\delta f)g$. Typically δ will be differential operators. By the definition of the determinant,

$$\begin{aligned}\delta \det(A) &= \sum_{(k_1,\cdots,k_n)\in\mathbb{P}(n)} \varepsilon_{k_1k_2\cdots k_n} \delta \prod_{\ell=1}^{n} a_{\ell k_\ell} \\ &= \sum_{i=1}^{n} \Big[\sum_{(k_1,\cdots,k_n)\in\mathbb{P}(n)} \varepsilon_{k_1k_2\cdots k_n} \delta a_{ik_i} \prod_{\substack{1\leqslant \ell\leqslant n\\ \ell\neq i}} a_{\ell k_\ell}\Big] \\ &= \sum_{i,j=1}^{n} \Big[\sum_{(k_1,\cdots,k_n)\in\mathbb{P}(n),\, k_i=j} \varepsilon_{k_1k_2\cdots k_n} \delta a_{ik_i} \prod_{\substack{1\leqslant \ell\leqslant n\\ \ell\neq i}} a_{\ell k_\ell}\Big] \\ &= \sum_{i,j=1}^{n} (-1)^{i+j} \det\big(A(\hat{i},\hat{j})\big)\delta a_{ij}\,.\end{aligned}$$

Therefore, we obtain the following

Theorem 1.73. *Let* A *be an* n×n *matrix, and* δ *be an operator satisfying* $\delta(fg) = f\delta g + (\delta f)g$ *whenever the product makes sense. Then*

$$\delta \det(\mathrm{A}) = \mathrm{tr}\big(\mathrm{Adj}(\mathrm{A})\delta\mathrm{A}\big)\,, \tag{1.3}$$

where $\delta\mathrm{A} \equiv [\delta a_{ij}]_{\mathrm{n}\times\mathrm{n}}$ *if* $\mathrm{A} = [a_{ij}]_{\mathrm{n}\times\mathrm{n}}$. *In particular, if* A *is invertible,*

$$\delta \det(\mathrm{A}) = \det(\mathrm{A})\mathrm{tr}(\big(\mathrm{A}^{-1}\delta\mathrm{A}\big)\,.$$

Example 1.74. Let $\mathrm{A}(x) = \begin{bmatrix} f(x) & g(x) \\ h(x) & k(x) \end{bmatrix}$ and $\delta = \dfrac{d}{dx}$. Then

$$\delta \det(\mathrm{A}) = \mathrm{tr}\Big(\begin{bmatrix} k & -g \\ -h & f \end{bmatrix} \begin{bmatrix} f' & g' \\ h' & k' \end{bmatrix} \Big) = kf' + fk' - gk' - hg'\,.$$

1.6 Bounded Linear Maps

Definition 1.75 (Linear map)**.** Let $\mathcal{V}$ and $\mathcal{W}$ be two vector spaces over a scalar field $\mathbb{F}$. A map $L : \mathcal{V} \to \mathcal{W}$ is called a ***linear map*** from $\mathcal{V}$ into $\mathcal{W}$ if

$$L(\alpha\boldsymbol{v} + \boldsymbol{w}) = \alpha L(\boldsymbol{v}) + L(\boldsymbol{w}) \qquad \forall\, \alpha \in \mathbb{F} \text{ and } \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}.$$

For notational convenience, we often write $L\boldsymbol{v}$ instead of $L(\boldsymbol{v})$. When $\mathcal{V}$ and $\mathcal{W}$ are finite dimensional, linear maps (from $\mathcal{V}$ into $\mathcal{W}$) are sometimes called ***linear transformations*** (from $\mathcal{V}$ into $\mathcal{W}$).

Let $L_1, L_2 : \mathcal{V} \to \mathcal{W}$ be two linear maps, and $\alpha \in \mathbb{F}$ be a scalar. It is easy to see that $\alpha L_1 + L_2 : \mathcal{V} \to \mathcal{W}$ is also a linear map. This is equivalent to say that the collection of linear maps is a vector space, and this induces the following

Definition 1.76. The vector space $\mathscr{L}(\mathcal{V}, \mathcal{W})$ is the collection of linear maps from $\mathcal{V}$ to $\mathcal{W}$.

Definition 1.77 (Boundedness of linear maps)**.** Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ and $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be two normed vector spaces over a scalar field $\mathbb{F}$. A linear map $L : \mathcal{V} \to \mathcal{W}$ is said to be bounded if the number

$$\|L\|_{\mathscr{B}(\mathcal{V},\mathcal{W})} \equiv \sup_{\|\boldsymbol{v}\|_{\mathcal{V}}=1} \|L\boldsymbol{v}\|_{\mathcal{W}} = \sup_{\boldsymbol{v}\neq 0} \frac{\|L\boldsymbol{v}\|_{\mathcal{W}}}{\|\boldsymbol{v}\|_{\mathcal{V}}} \tag{1.4}$$

is finite. The collection of all bounded linear map from $\mathcal{V}$ to $\mathcal{W}$ is denoted by $\mathscr{B}(\mathcal{V}, \mathcal{W})$, and $\mathscr{B}(\mathcal{V}, \mathcal{V})$ is also denoted by $\mathscr{B}(\mathcal{V})$ for simplicity.

Remark 1.78. When the domain $\mathcal{V}$ and the target $\mathcal{W}$ under consideration are clear, we use $\|\cdot\|$ instead of $\|\cdot\|_{\mathscr{B}(\mathcal{V},\mathcal{W})}$ to simplify the notation of operator norm.

Remark 1.79. If $\mathcal{V}$ is finite dimensional, then $\mathscr{L}(\mathcal{V},\mathcal{W}) = \mathscr{B}(\mathcal{V},\mathcal{W})$.

Proposition 1.80. *Let $(\mathcal{V},\|\cdot\|_{\mathcal{V}})$ and $(\mathcal{W},\|\cdot\|_{\mathcal{W}})$ be two normed vector spaces over a scalar field $\mathbb{F}$. Then $\big(\mathscr{B}(\mathcal{V},\mathcal{W}),\|\cdot\|\big)$ with $\|\cdot\|$ defined by* (1.4) *is a normed vector space.* (*Therefore, $\|\cdot\|$ is also called an operator norm*).

Definition 1.81 (Dual space)**.** Let $(\mathcal{V},\|\cdot\|)$ be a normed vector space over field $\mathbb{F}$. An element in $\mathscr{B}(\mathcal{V},\mathbb{F})$ is called a bounded linear functional on $\mathcal{V}$, and the space $\big(\mathscr{B}(\mathcal{V},\mathbb{F}),\|\cdot\|_{\mathscr{B}(\mathcal{V},\mathbb{F})}\big)$ is called the dual space of $(\mathcal{V},\|\cdot\|)$, and is usually denoted by $\mathcal{V}'$.

Definition 1.82. Let $(\mathcal{V},\|\cdot\|_{\mathcal{V}})$ and $(\mathcal{W},\|\cdot\|_{\mathcal{W}})$ be two normed vector spaces over a scalar field $\mathbb{F}$, and $L \in \mathscr{B}(\mathcal{V},\mathcal{W})$. The collection of all elements $\boldsymbol{v} \in \mathcal{V}$ such that $L\boldsymbol{v} = \mathbf{0}$ is called the kernel (or the null space) of L and is denoted by $\ker(L)$ or $\mathrm{Null}(L)$. In other words,

$$\ker(L) = \big\{\boldsymbol{v} \in \mathcal{V} \,\big|\, L\boldsymbol{v} = \mathbf{0}\big\}\,.$$

Theorem 1.83 (Riesz Representation Theorem)**.** *Let $(\mathcal{V},(\cdot,\cdot)_{\mathcal{V}})$ be an inner product space, and $f:\mathcal{V}\to\mathbb{R}$ be a bounded linear map. Then there exists a unique $\boldsymbol{w} \in \mathcal{V}$ such that $f(\boldsymbol{v}) = (\boldsymbol{v},\boldsymbol{w})_{\mathcal{V}}$ for all $\boldsymbol{v} \in \mathcal{V}$.*

Proof. The uniqueness for such a vector $\boldsymbol{w}$ is simply due to the fact that there is no non-trivial vector which is orthogonal to itself.

Now we show the existence of $\boldsymbol{w}$. If $f(\boldsymbol{v}) = 0$ for all $\boldsymbol{v} \in \mathcal{V}$, then $\boldsymbol{w} = \mathbf{0}$ does the job. Now suppose that $\ker(f) \subsetneq \mathcal{V}$. Then there exists $\boldsymbol{u} \in \ker(f)^{\perp}$ such that $\|\boldsymbol{u}\|_{\mathcal{V}} = 1$.

For $\boldsymbol{v} \in \mathcal{V}$, consider the vector $\boldsymbol{y} = f(\boldsymbol{v})\boldsymbol{u} - f(\boldsymbol{u})\boldsymbol{v}$. Then $\boldsymbol{y} \in \ker(f)$; thus $\boldsymbol{y}\cdot\boldsymbol{u} = 0$. Therefore,

$$0 = f(\boldsymbol{v})\|\boldsymbol{u}\|_{\mathcal{V}}^2 - f(\boldsymbol{u})(\boldsymbol{v},\boldsymbol{u})_{\mathcal{V}} = f(\boldsymbol{v}) - (\boldsymbol{v},\overline{f(\boldsymbol{u})}\boldsymbol{u})_{\mathcal{V}}$$

which implies that $f(\boldsymbol{v}) = (\boldsymbol{v},\boldsymbol{w})_{\mathcal{V}}$ with $\boldsymbol{w} = \overline{f(\boldsymbol{u})}\boldsymbol{u}$. □

By the Riesz representation theorem, we conclude the following

Theorem 1.84. *Let $(\mathcal{V},(\cdot,\cdot)_{\mathcal{V}})$ and $(\mathcal{W},(\cdot,\cdot)_{\mathcal{W}})$ be two inner product spaces. Then for all $L \in \mathscr{B}(\mathcal{V},\mathcal{W})$, there exists a unique $L^* \in \mathscr{B}(\mathcal{W},\mathcal{V})$ such that*

$$(L\boldsymbol{v},\boldsymbol{w})_{\mathcal{W}} = (\boldsymbol{v},L^*\boldsymbol{w})_{\mathcal{V}} \qquad \forall\,\boldsymbol{v} \in \mathcal{V},\boldsymbol{w} \in \mathcal{W}\,.$$

Definition 1.85 (Dual operator)**.** Let $\mathcal{V}$ and $\mathcal{W}$ be two inner product spaces, and $L : \mathcal{V} \to \mathcal{W}$ be a bounded linear map. The ***dual operator*** of L, denoted by L^*, is the unique linear map from $\mathcal{W}$ into $\mathcal{V}$ satisfying

$$(L\boldsymbol{v}, \boldsymbol{w})_{\mathcal{W}} = (\boldsymbol{v}, L^*\boldsymbol{w})_{\mathcal{V}} \qquad \forall\, \boldsymbol{v} \in \mathcal{V}, \boldsymbol{w} \in \mathcal{W},$$

where $(\cdot,\cdot)_{\mathcal{V}}$ and $(\cdot,\cdot)_{\mathcal{W}}$ are inner products on $\mathcal{V}$ and $\mathcal{W}$, respectively.

Definition 1.86 (Symmetry of linear maps)**.** An linear map $L \in \mathscr{B}(\mathcal{H})$ is said to be ***symmetric*** if $L = L^*$.

The last part of this section contributes to the following theorem which states that *every bounded linear maps near by* (*measured by the operator norm*) *an invertible bounded linear map is also invertible*.

Theorem 1.87. *Let* $\mathrm{GL}(n)$ *be the set of all invertible linear maps on* $(\mathbb{R}^n, \|\cdot\|_2)$*; that is,*

$$\mathrm{GL}(n) = \big\{ L \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n) \,\big|\, L \textit{ is one-to-one (and onto)} \big\} .$$

1. *If* $L \in \mathrm{GL}(n)$ *and* $K \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ *satisfying* $\|K - L\|\|L^{-1}\| < 1$, *then* $K \in \mathrm{GL}(n)$.

2. *The mapping* $L \mapsto L^{-1}$ *is continuous on* $\mathrm{GL}(n)$*; that is,*

$$\forall\, \varepsilon > 0\,, \exists\, \delta > 0 \ni \|K^{-1} - L^{-1}\| < \varepsilon \quad \textit{whenever} \quad \|K - L\| < \delta\,.$$

Proof. 1. Let $\|L^{-1}\| = \dfrac{1}{\alpha}$ and $\|K - L\| = \beta$. Then $\beta < \alpha$; thus for every $x \in \mathbb{R}^n$,

$$\begin{aligned} \alpha\|x\|_{\mathbb{R}^n} &= \alpha\|L^{-1}Lx\|_{\mathbb{R}^n} \leqslant \alpha\|L^{-1}\|\|Lx\|_{\mathbb{R}^n} = \|Lx\|_{\mathbb{R}^n} \leqslant \|(L-K)x\|_{\mathbb{R}^n} + \|Kx\|_{\mathbb{R}^n} \\ &\leqslant \beta\|x\|_{\mathbb{R}^n} + \|Kx\|_{\mathbb{R}^n}\,. \end{aligned}$$

As a consequence, $(\alpha - \beta)\|x\|_{\mathbb{R}^n} \leqslant \|Kx\|_{\mathbb{R}^n}$ and this implies that $K : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one hence invertible.

2. Let $L \in \mathrm{GL}(n)$ and $\varepsilon > 0$ be given. Choose $\delta = \min\left\{\dfrac{1}{2\|L^{-1}\|}, \dfrac{\varepsilon}{2\|L^{-1}\|^2}\right\}$. If $\|K-L\| < \delta$, then $K \in \mathrm{GL}(n)$. Since $L^{-1} - K^{-1} = K^{-1}(K-L)L^{-1}$, we find that if $\|K - L\| < \delta$,

$$\|K^{-1}\| - \|L^{-1}\| \leqslant \|K^{-1} - L^{-1}\| \leqslant \|K^{-1}\|\|K - L\|\|L^{-1}\| < \frac{1}{2}\|K^{-1}\|$$

which implies that $\|K^{-1}\| < 2\|L^{-1}\|$. Therefore, if $\|K - L\| < \delta$,

$$\|L^{-1} - K^{-1}\| \leqslant \|K^{-1}\|\|K - L\|\|L^{-1}\| < 2\|L^{-1}\|^2\delta < \varepsilon\,. \qquad \square$$

1.6.1 Matrix norms

Each $\mathrm{m} \times \mathrm{n}$ matrix $\mathrm{A} \in \mathbb{M}(\mathrm{m}, \mathrm{n}; \mathbb{F})$ induces a linear map $L : \mathbb{F}^{\mathrm{n}} \to \mathbb{F}^{\mathrm{m}}$ in a natural way: let $\mathrm{A} = [a_{ij}]_{\mathrm{m}\times\mathrm{n}}$ be a $\mathrm{m} \times \mathrm{n}$ matrix, $\mathcal{B} = \{\mathrm{e}_j\}_{j=1}^{\mathrm{n}}$ and $\widetilde{\mathcal{B}} = \{\widetilde{\mathrm{e}}_k\}_{k=1}^{\mathrm{m}}$ be the standard basis of $\mathbb{F}^{\mathrm{n}}$ and $\mathbb{F}^{\mathrm{m}}$, respectively. We define the linear map $L : \mathbb{F}^{\mathrm{n}} \to \mathbb{F}^{\mathrm{m}}$ by

$$Lx = \sum_{i=1}^{\mathrm{m}} \sum_{j=1}^{\mathrm{n}} a_{ij} x_j \widetilde{\mathrm{e}}_i \in \mathbb{F}^{\mathrm{m}}, \quad \text{where } \; x = \sum_{j=1}^{\mathrm{n}} x_j \mathrm{e}_j \in \mathbb{F}^{\mathrm{n}}\,,$$

or equivalently, $[Lx]_{\widetilde{\mathcal{B}}} = \mathrm{A}[x]_{\mathcal{B}}$. The linear map L is called the ***linear map induced by the matrix*** A.

By matrix norms it means the operator norm of the induced linear map. However, as introduced in Section 1.6, the operator norm of a linear map depends on the norms equipped on the vector spaces. In particular, we have introduced p-norm on $\mathbb{F}^{\mathrm{n}}$, and we have the following

Definition 1.88. Let $A \in \mathbb{M}(\mathrm{m}, \mathrm{n}; \mathbb{F})$ with induced linear map $L : \mathbb{F}^{\mathrm{n}} \to \mathbb{F}^{\mathrm{m}}$. The p-norm of A, denoted by $\|A\|_p$, is the operator norm of $L : (\mathbb{F}^{\mathrm{n}}, \|\cdot\|_p) \to (\mathbb{F}^{\mathrm{m}}, \|\cdot\|_p)$ given by

$$\|A\|_p = \sup_{\|x\|_p=1} \|Lx\|_p = \sup_{x\neq 0} \frac{\|Lx\|_p}{\|x\|_p}\,.$$

Remark 1.89. We can also choose different p in the domain and the co-domain. In other words, the (p, q)-norm of $A \in \mathbb{M}(\mathrm{m}, \mathrm{n}, \mathbb{F})$ is the operator norm of the induced linear map $L : (\mathbb{F}^{\mathrm{n}}, \|\cdot\|_p) \to (\mathbb{F}^{\mathrm{m}}, \|\cdot\|_q)$ given by

$$\|A\|_{(p,q)} = \sup_{\|x\|_p=1} \|Lx\|_q = \sup_{x\neq 0} \frac{\|Lx\|_q}{\|x\|_p}\,.$$

From now on, for notational simplicity we use Ax to denote $[Lx]_{\widetilde{\mathcal{B}}}$ if $\widetilde{\mathcal{B}}$ is the standard basis of the co-domain.

Example 1.90. Consider the case $p = 1$ and $p = \infty$, respectively.

1. $p = \infty$: $\|A\|_\infty = \sup\limits_{\|x\|_\infty=1} \|Ax\|_\infty = \max\Big\{\sum\limits_{j=1}^{\mathrm{m}} |a_{1j}|, \sum\limits_{j=1}^{\mathrm{m}} |a_{2j}|, \cdots \sum\limits_{j=1}^{\mathrm{m}} |a_{nj}|\Big\}$.

   Reason: Let $x = (x_1, x_2, \cdots, x_{\mathrm{n}})^{\mathrm{T}}$ and $A = \big[a_{ij}\big]_{n\times\mathrm{m}}$. Then

$$Ax = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ a_{21}x_1 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{bmatrix}$$

Assume $\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{m}|a_{ij}|=\sum\limits_{j=1}^{m}|a_{kj}|$ for some $1\leqslant k\leqslant n$. Let

$$x=\left(\operatorname{sgn}(a_{k1}),\operatorname{sgn}(a_{k2}),\cdots,\operatorname{sgn}(a_{kn})\right)\,.$$

Then $\|x\|_\infty=1$, and $\|Ax\|_\infty=\sum\limits_{j=1}^{m}|a_{kj}|$.

On the other hand, if $\|x\|_\infty=1$, then

$$|a_{i1}x_1+a_{i2}x_2+\cdots a_{im}x_m|\leqslant\sum_{j=1}^{\mathrm{m}}|a_{ij}|\leqslant\max\Big\{\sum_{j=1}^{\mathrm{m}}|a_{1j}|,\sum_{j=1}^{\mathrm{m}}|a_{2j}|,\cdots\sum_{j=1}^{\mathrm{m}}|a_{nj}|\Big\};$$

thus $\|A\|_\infty=\max\Big\{\sum\limits_{j=1}^{\mathrm{m}}|a_{1j}|,\sum\limits_{j=1}^{\mathrm{m}}|a_{2j}|,\cdots\sum\limits_{j=1}^{\mathrm{m}}|a_{nj}|\Big\}$. In other words, $\|A\|_\infty$ is the largest sum of the absolute value of row entries.

2. $p=1$: $\|A\|_1=\max\Big\{\sum\limits_{i=1}^{n}|a_{i1}|,\sum\limits_{i=1}^{n}|a_{i2}|,\cdots,\sum\limits_{i=1}^{n}|a_{im}|\Big\}$.

   Reason: Let $(\cdot,\cdot)$ denote the inner product in $\mathbb{R}^{\mathrm{m}}$. Then for $x\in\mathbb{R}^{\mathrm{n}}$ with $\|x\|_1=1$, by Hölder's inequality (1.1) and Theorem 1.25 we have

$$\begin{aligned}\|Ax\|_1&=\sup_{\|y\|_\infty=1}(Ax,y)=\sup_{\|y\|_\infty=1}(x,A^{\mathrm{T}}y)\leqslant\sup_{\|y\|_\infty=1}\|x\|_1\|A^{\mathrm{T}}y\|_\infty\\&=\sup_{\|y\|_\infty=1}\|A^{\mathrm{T}}y\|_\infty=\|A^{\mathrm{T}}\|_\infty\,;\end{aligned}$$

   thus $\|A\|_1=\sup\limits_{\|x\|_1=1}\|Ax\|_1\leqslant\|A^{\mathrm{T}}\|_\infty$. Similarly, if $y\in\mathbb{R}^{\mathrm{m}}$ and $\|y\|_\infty=1$, then

$$\begin{aligned}\|A^{\mathrm{T}}y\|_\infty&=\sup_{\|x\|_1=1}(x,A^{\mathrm{T}}y)=\sup_{\|x\|_1=1}(Ax,y)\leqslant\sup_{\|x\|_1=1}\|Ax\|_1\|y\|_\infty\\&=\sup_{\|x\|_1=1}\|Ax\|_1=\|A\|_1\end{aligned}$$

   which implies that $\|A^{\mathrm{T}}\|_\infty=\sup\limits_{\|y\|_\infty=1}\|A^{\mathrm{T}}y\|_\infty\leqslant\|A\|_1$. As a consequence,

$$\|A\|_1=\|A^{\mathrm{T}}\|_\infty=\max\Big\{\sum_{i=1}^{n}|a_{i1}|,\sum_{i=1}^{n}|a_{i2}|,\cdots,\sum_{i=1}^{n}|a_{im}|\Big\}\,.$$

1.7 Representation of Linear Transformations

In Section 1.6.1, we see that any $\mathrm{m} \times \mathrm{n}$ matrix is associated with a linear map. On the other hand, suppose that $\mathcal{V}$ is a n-dimensional vector space with basis $\mathcal{B} = \{\boldsymbol{v}_j\}_{j=1}^{\mathrm{n}}$, and $\mathcal{W}$ is a m-dimensional vector space with basis $\widetilde{\mathcal{B}} = \{\boldsymbol{w}_i\}_{i=1}^{\mathrm{m}}$. Define $\mathbf{V} = \big[\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{\mathrm{n}}\big]$ and $\mathbf{W} = \big[\boldsymbol{w}_1 \vdots \cdots \vdots \boldsymbol{w}_{\mathrm{m}}\big]$, and let $L \in \mathscr{L}(\mathcal{V}, \mathcal{W})$. Since $L\boldsymbol{v}_j \in \mathcal{W}$, for each $1 \leqslant j \leqslant \mathrm{n}$ we can write $L\boldsymbol{v}_j = \sum\limits_{i=1}^{\mathrm{m}} a_{ij} \boldsymbol{w}_i$ for some coefficients a_{ij}. Moreover, if $\boldsymbol{u} \in \mathcal{V}$, then

$$\boldsymbol{u} = \sum_{j=1}^{\mathrm{n}} c_j \boldsymbol{v}_j \qquad \text{or} \qquad \boldsymbol{c} = [\boldsymbol{u}]_{\mathcal{B}} \qquad \text{or} \qquad \boldsymbol{u} = \mathbf{V}\boldsymbol{c}\,,$$

and by the linearity of L,

$$L\boldsymbol{u} = L\Big(\sum_{j=1}^{\mathrm{n}} c_j \boldsymbol{v}_j\Big) = \sum_{j=1}^{\mathrm{n}} c_j L\boldsymbol{v}_j = \sum_{j=1}^{\mathrm{n}}\sum_{i=1}^{\mathrm{m}} c_j a_{ij} \boldsymbol{w}_i = \sum_{i=1}^{\mathrm{m}} \Big(\sum_{j=1}^{\mathrm{n}} a_{ij} c_j\Big) \boldsymbol{w}_i\,.$$

Let $b_i = \sum\limits_{j=1}^{\mathrm{n}} a_{ij} c_j$, and $\boldsymbol{b} = [b_1, \cdots, b_{\mathrm{m}}]^{\mathrm{T}}$. Then

$$[L\boldsymbol{u}]_{\widetilde{\mathcal{B}}} = \boldsymbol{b} = A\boldsymbol{c} = A[\boldsymbol{u}]_{\mathcal{B}}\,.$$

The discussion above induces the following

Definition 1.91. Let $\mathcal{V}, \mathcal{W}$ be two vector spaces, $\dim(\mathcal{V}) = \mathrm{n}$ and $\dim(\mathcal{W}) = \mathrm{m}$, and $\mathcal{B}, \widetilde{\mathcal{B}}$ are basis of $\mathcal{V}, \mathcal{W}$, respectively. For $L \in \mathscr{L}(\mathcal{V}, \mathcal{W})$, the ***matrix representation*** of L relative to bases $\mathcal{B}$ and $\widetilde{\mathcal{B}}$, denoted by $[L]_{\widetilde{\mathcal{B}},\mathcal{B}}$, is the matrix satisfying

$$[L\boldsymbol{u}]_{\widetilde{\mathcal{B}}} = [L]_{\widetilde{\mathcal{B}},\mathcal{B}} [\boldsymbol{u}]_{\mathcal{B}} \qquad \forall\, \boldsymbol{u} \in \mathcal{V}\,.$$

If $L \in \mathscr{L}(\mathcal{V}, \mathcal{V})$, we simply use $[L]_{\mathcal{B}}$ to denote $[L]_{\mathcal{B},\mathcal{B}}$.

1.8 Matrix Diagonalization

Definition 1.92 (Eigenvalues and Eigenvectors)**.** Let $\mathcal{V}$ be a finite dimensional vector spaces over a scalar field $\mathbb{F}$, and $L \in \mathscr{B}(\mathcal{V})$. A scalar $\lambda \in \mathbb{F}$ is said to be an ***eigenvalue*** of L if there is a non-zero vector $\boldsymbol{v} \in \mathcal{V}$ such that $L\boldsymbol{v} = \lambda\boldsymbol{v}$. The collection of all eigenvalues of L is denoted by $\sigma(L)$.

For an eigenvalue $\lambda \in \mathbb{F}$ of L, a non-zero vector $\boldsymbol{v} \in \mathcal{V}$ satisfying $L\boldsymbol{v} = \lambda\boldsymbol{v}$ is called an ***eigenvector*** associated with the eigenvalue λ, and the collection of all $\boldsymbol{v} \in \mathcal{V}$ such that $L\boldsymbol{v} = \lambda\boldsymbol{v}$ is called the ***eigenspace*** associated with λ.

Let $\dim(\mathcal{V}) = \mathrm{n}$ and $\mathcal{B}$ be a basis of $\mathcal{V}$. Then if $\lambda \in \mathbb{F}$ is an eigenvalue of $L \in \mathscr{B}(\mathcal{V})$, there exists non-zero vector $\boldsymbol{v} \in \mathcal{V}$ such that

$$[L]_{\mathcal{B}}[\boldsymbol{v}]_{\mathcal{V}} = [L\boldsymbol{v}]_{\mathcal{B}} = \lambda[\boldsymbol{v}]_{\mathcal{B}}\,;$$

thus the matrix representation $[L]_{\mathcal{B}}$ of L satisfies that $[L]_{\mathcal{B}} - \lambda \mathrm{I}_{\mathrm{n}}$ is singular (not invertible). Therefore, $\det([L]_{\mathcal{B}} - \lambda \mathrm{I}_{\mathrm{n}}) = 0$ which motivates the following

Definition 1.93. Let $A \in \mathbb{M}(\mathrm{n}, \mathrm{n}; \mathbb{F})$ be a $\mathrm{n} \times \mathrm{n}$ matrix over scalar field $\mathbb{F}$. An ***eigenvalue*** of A is a scalar $\lambda \in \mathbb{F}$ such that $\det(A - \lambda \mathrm{I}_{\mathrm{n}}) = 0$.

Theorem 1.94. *Let $L \in \mathscr{B}(\mathbb{F}^{\mathrm{n}})$ be symmetric. Then $\sigma(L) \subseteq \mathbb{R}$.*

Proof. Let $\lambda \in \sigma(L)$, and $\boldsymbol{v}$ be an eigenvector associated with λ. Then

$$\lambda(\boldsymbol{v}, \boldsymbol{v})_{\mathbb{F}^{\mathrm{n}}} = (\lambda\boldsymbol{v}, \boldsymbol{v})_{\mathbb{F}^{\mathrm{n}}} = (L\boldsymbol{v}, \boldsymbol{v})_{\mathbb{F}^{\mathrm{n}}} = (\boldsymbol{v}, L^*\boldsymbol{v})_{\mathbb{F}^{\mathrm{n}}} = (\boldsymbol{v}, L\boldsymbol{v})_{\mathbb{F}^{\mathrm{n}}} = (\boldsymbol{v}, \lambda\boldsymbol{v}) = \bar{\lambda}(\boldsymbol{v}, \boldsymbol{v})_{\mathbb{F}^{\mathrm{n}}}$$

which implies that $\lambda \in \mathbb{R}$. □

Lemma 1.95. *Let $L \in \mathscr{B}(\mathbb{F}^{\mathrm{n}})$ be symmetric, and $(\cdot,\cdot)_{\mathbb{F}^{\mathrm{n}}}$ be the standard inner product on $\mathbb{F}^{\mathrm{n}}$. Then the two numbers*

$$m \equiv \inf_{\|u\|_{\mathbb{F}^{\mathrm{n}}}=1} (Lu, u)_{\mathbb{F}^{\mathrm{n}}} \quad \textit{and} \quad M \equiv \sup_{\|u\|_{\mathbb{F}^{\mathrm{n}}}=1} (Lu, u)_{\mathbb{F}^{\mathrm{n}}}$$

belong to $\sigma(L)$.

Proof. Suppose that $M \notin \sigma(L)$. Let $[u, v] = (Mu - Lu, v)_{\mathbb{F}^{\mathrm{n}}}$. Then $[\cdot,\cdot]$ is an inner product on $\mathbb{F}^{\mathrm{n}}$; thus the Cauchy-Schwarz inequality (Proposition 1.16) implies that

$$\big|[u, v]\big| \leqslant \big|[u, u]\big|^{1/2}\big|[v, v]\big|^{1/2}.$$

By Theorem 1.25, we find that

$$\begin{aligned}\|Mu - Lu\|_{\mathbb{F}^{\mathrm{n}}} &= \sup_{\|v\|_{\mathbb{F}^{\mathrm{n}}}=1} \big|(Mu - Lu, v)_{\mathbb{F}^{\mathrm{n}}}\big| = \sup_{\|v\|_{\mathbb{F}^{\mathrm{n}}}=1} \big|[u, v]\big| \leqslant \sup_{\|v\|_{\mathbb{F}^{\mathrm{n}}}=1} \big|[u, u]\big|^{1/2}\big|[v, v]\big|^{1/2} \\ &\leqslant (M - m)^{1/2}(Mu - Lu, u)_{\mathbb{F}^{\mathrm{n}}}^{1/2} \qquad \forall\, u \in \mathbb{F}^{\mathrm{n}}\,,\end{aligned} \tag{1.5}$$

where we use the fact that $\sup\limits_{\|v\|_{\mathbb{F}^{\mathrm{n}}}=1} \big|[v, v]\big|^{1/2} = (M - m)^{1/2}$ to conclude the last inequality.

Let $\mathcal{B}$ be the standard basis of $\mathbb{F}^{\rm n}$, and $\{\boldsymbol{u}_k\}_{k=1}^{\infty}$ be a sequence of vectors in $\mathbb{F}^{\rm n}$ such that $\|\boldsymbol{u}_k\|_{\mathbb{F}^{\rm n}} = 1$, and $\lim\limits_{k\to\infty}(L\boldsymbol{u}_k, \boldsymbol{u}_k)_{\mathbb{F}^{\rm n}} = M$. Then (1.5) implies $\|M\boldsymbol{u}_k - L\boldsymbol{u}_k\|_{\mathbb{F}^{\rm n}} \to 0$ as $k \to \infty$. Since $M \notin \sigma(L)$, $M\mathrm{I}_{\rm n} - [L]_{\mathcal{B}}$ is invertible; thus

$$[\boldsymbol{u}_k]_{\mathcal{B}} = (M\mathrm{I}_{\rm n} - [L]_{\mathcal{B}})^{-1}(M[\boldsymbol{u}_k]_{\mathcal{B}} - [L]_{\mathcal{B}}[\boldsymbol{u}_k]_{\mathcal{B}}) \to \mathbf{0} \quad \text{in} \quad \mathbb{F}^{\rm n}$$

which contradicts to $\|u_k\|_{\mathbb{F}^{\rm n}} = 1$ for all $k \in \mathbb{N}$. Hence $M \in \sigma(L)$. Similarly, $m \in \sigma(L)$. □

Definition 1.96 (Diagonalizable linear maps)**.** Let $\mathcal{V}$ be a finite dimensional vector spaces over a scalar field $\mathbb{F}$. A linear map $L : \mathcal{V} \to \mathcal{V}$ is said to be ***diagonalizable*** if there is a basis $\mathcal{B}$ of $\mathcal{V}$ such that each $\boldsymbol{v} \in \mathcal{B}$ is an eigenvector of L.

Theorem 1.97. *Let $L \in \mathscr{B}(\mathbb{R}^{\rm n})$ be symmetric. Then there exists an orthonormal basis of $\mathbb{R}^{\rm n}$ consisting of eigenvectors of L.*

Example 1.98 (The 2-norm of matrices)**.** Let $(\cdot,\cdot)_{\mathbb{R}^k}$ denote the inner product in Euclidean space $\mathbb{R}^k$, and $A \in \mathbb{M}(\mathrm{m}, \mathrm{n}; \mathbb{R})$. Since $A^{\mathrm{T}}A$ is a symmetric $\mathrm{n} \times \mathrm{n}$ matrix, it is diagonalizable by an orthonormal matrix P; that is, $A^{\mathrm{T}}A = P\Lambda P^{\mathrm{T}}$ for some orthonormal $\mathrm{n} \times \mathrm{n}$ matrix P and diagonal $\mathrm{n} \times \mathrm{n}$ matrix $\Lambda = [\lambda_i \delta_{ij}]$. Therefore,

$$\|Ax\|_2^2 = (Ax, Ax)_{\mathbb{R}^{\rm m}} = (x, A^{\mathrm{T}}Ax)_{\mathbb{R}^{\rm n}} = (x, P\Lambda P^{\mathrm{T}}x)_{\mathbb{R}^{\rm n}} = (P^{\mathrm{T}}x, \Lambda P^{\mathrm{T}}x)_{\mathbb{R}^{\rm n}}$$

which implies that

$$\begin{aligned}
\sup_{\|x\|_2=1} \|Ax\|_2^2 &= \sup_{\|x\|_2=1} (P^{\mathrm{T}}x, \Lambda P^{\mathrm{T}}x)_{\mathbb{R}^{\rm n}} = \sup_{\|y\|_2=1} (y, \Lambda y)_{\mathbb{R}^{\rm n}} \quad (\text{Let } y = P^{\mathrm{T}}x, \text{ then } \|y\|_2 = 1) \\
&= \sup_{\|y\|_2=1} (\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_{\rm n} y_{\rm n}^2) \\
&= \max\big\{\lambda_1, \cdots, \lambda_{\rm n}\big\} = \text{maximum eigenvalue of } A^{\mathrm{T}}A\,.
\end{aligned}$$

As a consequence, $\|A\|_2 = \sqrt{\text{maximum eigenvalue of } A^{\mathrm{T}}A}$.

1.9 The Einstein Summation Convention

In mathematics, especially in applications of linear algebra to physics, the Einstein summation convention is a notational convention that implies summation over a set of indexed terms in a formula, thus achieving notational brevity. According to this convention, when an index variable appears twice in a single term it implies summation of that term over all the values of the index. For example, with this convention, the inner product $\boldsymbol{u} \cdot \boldsymbol{v}$ of two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{\mathrm{n}}$, where $\boldsymbol{u} = (u_1, \cdots, u_{\mathrm{n}})$ and $\boldsymbol{v} = (v_1, \cdots, v_{\mathrm{n}})$, can be expressed as $u_i v_i$, and the i-th component of the cross product $\boldsymbol{u} \times \boldsymbol{v}$ of two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ can be expressed as $\varepsilon_{ijk} u^j v^k$.

In this book, we make a further convention that repeated Latin indices are summed from 1 to n, and repeated Greek indices are summed from 1 to n $-$ 1, where n is the space dimension. In other words, we use the symbol $f_i g_i$ to denote the sum $\sum\limits_{i=1}^{\mathrm{n}} f_i g_i$, and the symbol $f_\alpha g_\alpha$ to denote the sum $\sum\limits_{i=1}^{\mathrm{n}-1} f_\alpha g_\alpha$. Starting from the next Chapter, we use such summation convention for notational simplicity.

Chapter 2

Differentiation of Functions of Several Variables

2.1 Functions of Several Variables

Definition 2.1. Let $\mathcal{V}$ be a vector space (over a scalar field $\mathbb{F}$). A ***$\mathcal{V}$-valued function*** f of n real variables is a rule that assigns a unique vector $f(x_1, \cdots, x_n) \in \mathcal{V}$ to each point $(x_1, \cdots, x_n)$ in some subset A of $\mathbb{R}^n$. The set A is called the ***domain*** of f, and usually is denoted by $\mathrm{Dom}(f)$. The set of vectors $f(x_1, \cdots, x_n)$ obtained from points in the domain is called the ***range*** of f and is denoted by $\mathrm{Ran}(f)$. We write $f : A \to \mathcal{V}$ if f is a $\mathcal{V}$-valued function defined on $A \subseteq \mathbb{R}^n$.

If $\mathcal{V} = \mathbb{R}$, we simply call $f : \mathrm{Dom}(f) \to \mathbb{R}$ a ***real-valued function***, while if $\mathcal{V} = \mathbb{R}^m$, we simply call $f : \mathrm{Dom}(f) \to \mathcal{V}$ as a ***vector-valued function***.

A ***vector field*** is a vector-valued function $f : \mathrm{Dom}(f) \to \mathcal{V}$ such that $\mathrm{Dom}(f) \subseteq \mathcal{V} = \mathbb{R}^n$ for some $n \in \mathbb{N}$.

Definition 2.2. Let $\mathcal{V}$ be a vector space over $\mathbb{R}$, $A \subseteq \mathbb{R}^n$ be a set, and $f, g : A \to \mathcal{V}$ be $\mathcal{V}$-valued functions, $h : A \to \mathbb{R}$ be a real-valued function. The functions $f + g$, $f - g$ and hf, mapping from A to $\mathcal{V}$, are defined by

$$\begin{aligned}
(f+g)(x) &= f(x) + g(x) \qquad \forall\, x \in A\,, \\
(f-g)(x) &= f(x) - g(x) \qquad \forall\, x \in A\,, \\
(hf)(x) &= h(x)f(x) \qquad \forall\, x \in A\,.
\end{aligned}$$

The map $\dfrac{f}{h} : A \backslash \{x \in A \,|\, h(x) = 0\} \to \mathcal{V}$ is defined by

$$\Big(\frac{f}{h}\Big)(x) = \frac{f(x)}{h(x)} \qquad \forall\, x \in A \backslash \{x \in A \,|\, h(x) = 0\}\,.$$

Definition 2.3. A set $\mathcal{U} \subseteq \mathbb{R}^n$ is said to be open in $\mathbb{R}^n$ if for each $x \in \mathcal{U}$, there exists $r > 0$ such that $B(x, r)$, the ball centered at x with radius r given by

$$B(x, r) = \big\{ y \in \mathbb{R}^n \,\big|\, \|x - y\|_{\mathbb{R}^n} < r \big\}\,,$$

is contained in $\mathcal{U}$. A set $\mathcal{F} \subseteq \mathbb{R}^n$ is said to be closed in $\mathbb{R}^n$ if $\mathcal{F}^\complement$, the complement of $\mathcal{F}$, is open in $\mathbb{R}^n$.

Let $A \subseteq \mathbb{R}^n$ be a set. A point x_0 is said to be

1. an ***interior point*** of A if there exists $r > 0$ such that $B(x_0, r) \subseteq A$;

2. an ***isolated point*** of A if there exists $r > 0$ such that $B(x_0, r) \cap A = \{x_0\}$;

3. an ***exterior point*** of A if there exists $r > 0$ such that $B(x_0, r) \subseteq A^\complement$;

4. a ***boundary point*** of A if for each $r > 0$, $B(x_0, r) \cap A \neq \varnothing$ and $B(x_0, r) \cap A^\complement \neq \varnothing$.

The collection of all interior points of A is called the interior of A and is denoted by $\mathring{A}$. The collection of all exterior points of A is called the exterior of A, and the collection of all boundary point of A is called the boundary of A. The boundary of A is denoted by ∂A. The closure of A is defined as $A \cup \partial A$ and is denoted by $\overline{A}$. The derived set of A, denoted by A', is the collection of all points in $\overline{A}$ that are not isolated points.

A is said to be bounded in $\mathbb{R}^n$ if there exists a constant $M > 0$ such that

$$\|x\|_{\mathbb{R}^n} < M \quad \forall\, x \in A \qquad \big(\Leftrightarrow A \subseteq B(0, M)\big)\,.$$

A is said to be unbounded if A is not bounded.

The following theorem is a fundamental result in point-set topology. We omit the proof since it is not the main concern in vector analysis; however, the result should look intuitive and the proof of this theorem is not difficult. Interested readers can try to establish this result by yourselves.

Theorem 2.4. *Let $A \subseteq \mathbb{R}^n$ be a set. Then*

1. *A is open if and only if $A = \mathring{A}$;*

2. *A is closed if and only if $A = \overline{A}$;*

3. *A is closed if and only if $\partial A \subseteq A$.*

Definition 2.5 (Level Sets, and Graphs)**.** Let $A \to \mathbb{R}^{\mathrm{n}}$ be a set, and $f : A \to \mathbb{R}$ be a real-valued function. The collection of points in A where f has a constant value is called a ***level set*** of f. The collection of all points $\big(x, f(x)\big)$ is called the ***graph*** of f.

Remark 2.6. A level surface is conventionally called a level curve when n = 2.

2.2 Limits and Continuity

Definition 2.7. Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a set, and $f : A \to \mathbb{R}^{\mathrm{m}}$ be a vector-valued function. For a given $x_0 \in A'$, we say that $b \in \mathbb{R}^{\mathrm{m}}$ is the limit of f at x_0, written

$$\lim_{x \to x_0} f(x) = b \qquad \text{or} \qquad f(x) \to b \text{ as } x \to x_0 \,,$$

if for each $\varepsilon > 0$, there exists $\delta = \delta(x_0, \varepsilon) > 0$ such that

$$\|f(x) - b\|_{\mathbb{R}^{\mathrm{m}}} < \varepsilon \text{ whenever } 0 < \|x - x_0\|_{\mathbb{R}^{\mathrm{n}}} < \delta \text{ and } x \in A \,.$$

By the definition above, it is easy to see the following

Proposition 2.8. *Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a set, and $f, g : A \to \mathbb{R}^{\mathrm{m}}$ be a vector-valued functions. Suppose that $x_0 \in A'$, $f(x) = g(x)$ for all $x \in A \backslash \{x_0\}$, and $\lim\limits_{x \to x_0} f(x)$ exists. Then $\lim\limits_{x \to x_0} g(x)$ exists and*

$$\lim_{x \to x_0} g(x) = \lim_{x \to x_0} f(x) \,.$$

The following proposition is standard, and we omit the proof.

Proposition 2.9. *Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a set, and $f, g : A \to \mathbb{R}^{\mathrm{m}}$ be vector-valued functions, $h : A \to \mathbb{R}$ be a real-valued function. Suppose that $x_0 \in A'$, and $\lim\limits_{x \to x_0} f(x) = a$, $\lim\limits_{x \to x_0} g(x) = b$, $\lim\limits_{x \to x_0} h(x) = c$. Then*

$$\begin{aligned} \lim_{x \to x_0} (f + g)(x) &= a + b \,, & \lim_{x \to x_0} (f - g)(x) &= a - b \,, \\ \lim_{x \to x_0} (hf)(x) &= ca \,, & \lim_{x \to x_0} (f \cdot g)(x) &= a \cdot b \,, \\ \lim_{x \to x_0} \Big(\frac{f}{h}\Big) &= \frac{a}{c} \quad \textit{if } c \neq 0 \,. \end{aligned}$$

Example 2.10. By Proposition 2.9,

$$\lim_{(x,y)\to(0,1)} \frac{x - xy + 3}{x^2y + 5xy - y^3} = \frac{0 - (0)(1) + 3}{(0)^2(1) + 5(0)(1) - (1)^3} = -3\,.$$

Example 2.11. Let $f : (0,\infty) \times (0,\infty) \to \mathbb{R}$ be given by $f(x,y) = \dfrac{x^2 - xy}{\sqrt{x} - \sqrt{y}}$. We cannot apply Proposition 2.9 to compute the limit $\lim\limits_{(x,y)\to(0,0)} f(x,y)$, if the limit exists, since $\lim\limits_{(x,y)\to(0,0)} (\sqrt{x} - \sqrt{y}) = 0$. Nevertheless, if $(x,y) \neq (0,0)$,

$$f(x,y) = \frac{x^2 - xy}{\sqrt{x} - \sqrt{y}} = \frac{x(x-y)(\sqrt{x} + \sqrt{y})}{(\sqrt{x} - \sqrt{y})(\sqrt{x} + \sqrt{y})} = x(\sqrt{x} + \sqrt{y})\,;$$

thus Proposition 2.8 and 2.9 imply that

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} x(\sqrt{x} + \sqrt{y}) = 0\,.$$

Definition 2.12. Let $A \subseteq \mathbb{R}^n$ be a set, and $f : A \to \mathbb{R}^m$ be a vector-valued function. The function f is said to be continuous at $x_0 \in A \cap A'$ if $\lim\limits_{x\to x_0} f(x) = f(x_0)$. In other words, f is continuous at x_0 if

$$\forall\, \varepsilon > 0, \exists\, \delta = \delta(x_0, \varepsilon) > 0 \ni \|f(x) - f(x_0)\|_{\mathbb{R}^m} < \varepsilon \text{ whenever } \|x - x_0\|_{\mathbb{R}^n} < \delta \text{ and } x \in A\,.$$

If f is continuous at each point of $B \subseteq A \cap A'$, then f is said to be continuous on B.

Remark 2.13. 1. The notation $\delta = \delta(x_0, \varepsilon)$ means that the number δ could depend on x_0 and ε.

2. Another way of interpreting the continuity of f at x_0 is as follows: $f : A \to \mathbb{R}^m$ is continuous at $x_0 \in \mathcal{U}$ if

$$\forall\, \varepsilon > 0, \exists\, \delta = \delta(x_0, \varepsilon) > 0 \ni f\big(B(x_0, \delta) \cap A\big) \subseteq B(f(x_0), \varepsilon)\,.$$

3. If $A = \mathcal{U}$ is an open set, we can assume that δ is chosen small enough so that $B(x_0, \delta) \subseteq \mathcal{U}$ in both Definition 2.7 and 2.12. In other words, $\lim\limits_{x\to x_0} f(x) = b$ if

$$\forall\, \varepsilon > 0, \exists\, \delta = \delta(x_0, \varepsilon) > 0 \ni \|f(x) - b\|_{\mathbb{R}^m} < \varepsilon \text{ whenever } 0 < \|x - x_0\|_{\mathbb{R}^n} < \delta\,,$$

and $f : \mathcal{U} \to \mathbb{R}^m$ is continuous at $x_0 \in \mathcal{U}$ if

$$\forall\, \varepsilon > 0, \exists\, \delta = \delta(x_0, \varepsilon) > 0 \ni \|f(x) - f(x_0)\|_{\mathbb{R}^m} < \varepsilon \text{ whenever } \|x - x_0\|_{\mathbb{R}^n} < \delta\,.$$

4. If $A \subseteq \mathbb{R}^n$ is closed and bounded, and $f : A \to \mathbb{R}^m$ is continuous, then for each $\varepsilon > 0$ we can choose δ depending only on ε such that
$$\|f(x) - f(y)\|_{\mathbb{R}^m} < \varepsilon \text{ whenever } \|x - y\|_{\mathbb{R}^n} < \delta \text{ and } x, y \in A\,.$$
The property (that δ can be chosen independent of the point x_0) is called ***uniform continuity***.

Theorem 2.14. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$ be a vector-valued function. Then the following assertions are equivalent:*

1. *f is continuous on $\mathcal{U}$.*
2. *For each open set $\mathcal{V} \subseteq \mathbb{R}^m$, $f^{-1}(\mathcal{V}) \subseteq \mathcal{U}$ is open, where $f^{-1}(\mathcal{V})$ is the pre-image of $\mathcal{V}$ under f defined by*
$$f^{-1}(\mathcal{V}) \equiv \big\{x \in \mathcal{U} \,\big|\, f(x) \in \mathcal{V}\big\}\,.$$

Proof. Before proceeding, we recall that $B \subseteq f^{-1}(f(B))$ for all $B \subseteq \mathcal{U}$ and $f(f^{-1}(B)) \subseteq B$ for all $B \subseteq \mathbb{R}^m$.

"1 $\Rightarrow$ 2" Let $a \in f^{-1}(\mathcal{V})$. Then $f(a) \in \mathcal{V}$. Since $\mathcal{V}$ is open in $\mathbb{R}^m$, $\exists\, \varepsilon_{f(a)} > 0$ such that $B(f(a), \varepsilon_{f(a)}) \subseteq \mathcal{V}$. By continuity of f (and Remark 2.13), there exists $\delta_a > 0$ such that
$$f\big(B(a, \delta_a)\big) \subseteq B\big(f(a), \varepsilon_{f(a)}\big)\,.$$
Therefore, for each $a \in f^{-1}(\mathcal{V})$, $\exists\, \delta_a > 0$ such that
$$B(a, \delta_a) \subseteq f^{-1}\big(f\big(B(a, \delta_a)\big)\big) \subseteq f^{-1}\big(B\big(f(a), \varepsilon_{f(a)}\big)\big) \subseteq f^{-1}(\mathcal{V})\,.$$
Therefore, $f^{-1}(\mathcal{V})$ is open.

"2 $\Rightarrow$ 1" Let $a \in \mathcal{U}$ and $\varepsilon > 0$ be given. Define $\mathcal{V} = B(f(a), \varepsilon)$, then $\mathcal{V}$ is open. Since $a \in f^{-1}(\mathcal{V})$ and $f^{-1}(\mathcal{V})$ is open by assumption, there exists $\delta > 0$ such that $B(a, \delta) \subseteq f^{-1}(\mathcal{V})$. Therefore,
$$f\big(B(a, \delta)\big) \subseteq f(f^{-1}(\mathcal{V})) \subseteq \mathcal{V} = B(f(a), \varepsilon)$$
which (with the help of Remark 2.13) implies that f is continuous at a. $\square$

2.3 Definition of Derivatives and the Matrix Representation of Derivatives

Definition 2.15. Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{m}}$ be an open set. A function $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ is said to be ***differentiable*** at $x_0 \in A$ if there is a linear transformation from $\mathbb{R}^{\mathrm{n}}$ to $\mathbb{R}^{\mathrm{m}}$, denoted by $(Df)(x_0)$ and called the ***derivative*** of f at x_0, such that

$$\lim_{x \to x_0} \frac{\big\|f(x) - f(x_0) - (Df)(x_0)(x - x_0)\big\|_{\mathbb{R}^{\mathrm{m}}}}{\|x - x_0\|_{\mathbb{R}^{\mathrm{n}}}} = 0\,,$$

where $(Df)(x_0)(x - x_0)$ denotes the value of the linear transformation $(Df)(x_0)$ applied to the vector $x - x_0$. In other words, f is differentiable at $x_0 \in \mathcal{U}$ if there exists $L \in \mathscr{B}(\mathbb{R}^{\mathrm{n}}, \mathbb{R}^{\mathrm{m}})$ such that

$$\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni \|f(x) - f(x_0) - L(x - x_0)\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \varepsilon \|x - x_0\|_{\mathbb{R}^{\mathrm{n}}} \text{ whenever } \|x - x_0\|_{\mathbb{R}^{\mathrm{n}}} < \delta\,.$$

If f is differentiable at each point of $\mathcal{U}$, we say that f is differentiable on $\mathcal{U}$.

Example 2.16. Let $L : \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{m}}$ be a linear transformation; that is, there is a matrix $[L]_{m \times n}$ such that $L(x) = [L]_{m \times n}[x]_n$ for all $x \in \mathbb{R}^{\mathrm{n}}$. Then L is differentiable. In fact, $(DL)(x_0) = L$ for all $x_0 \in X$ since

$$\lim_{x \to x_0} \frac{\|Lx - Lx_0 - L(x - x_0)\|_{\mathbb{R}^{\mathrm{m}}}}{\|x - x_0\|_{\mathbb{R}^{\mathrm{n}}}} = 0\,.$$

Example 2.17. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x^2 + 2y$. Define $L_{(a,b)}(x, y) = 2ax + 2y$. Then $L_{(a,b)}$ is a linear transformation (from $\mathbb{R}^2$ to $\mathbb{R}$) and

$$\begin{aligned}
&\frac{\big|x^2 + 2y - a^2 - 2b - L_{(a,b)}(x - a, y - b)\big|}{\sqrt{(x - a)^2 + (y - b)^2}} \\
&\qquad = \frac{\big|x^2 + 2y - a^2 - 2b - 2a(x - a) - 2(y - b)\big|}{\sqrt{(x - a)^2 + (y - b)^2}} \\
&\qquad = \frac{(x - a)^2}{\sqrt{(x - a)^2 + (y - b)^2}} \leqslant |x - a|\,;
\end{aligned}$$

thus

$$\lim_{(x,y) \to (a,b)} \frac{\big|x^2 + 2y - a^2 - 2b - L_{(a,b)}(x - a, y - b)\big|}{\sqrt{(x - a)^2 + (y - b)^2}} = 0\,.$$

Therefore, f is differentiable at (a, b) and $(Df)(a, b) = L_{(a,b)}$.

Remark 2.18. Adopting the standard basis of $\mathbb{R}^{\mathrm{n}}$ and $\mathbb{R}^{\mathrm{m}}$, a linear transformation $L : \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{m}}$ has a matrix representation $[L]_{\mathrm{m}\times\mathrm{n}}$ such that $L(x) = [L]_{\mathrm{m}\times\mathrm{n}}[x]_{\mathrm{n}}$ for all $x \in \mathbb{R}^{\mathrm{n}}$. In the following, we will always use the standard basis for $\mathbb{R}^{\mathrm{n}}$ and $\mathbb{R}^{\mathrm{m}}$ and use L and $L(x)$ to denote $[L]_{\mathrm{m}\times\mathrm{m}}$ and $[L]_{\mathrm{m}\times\mathrm{n}}[x]_{\mathrm{n}}$, respectively, if L is a linear transformation from $\mathbb{R}^{\mathrm{n}}$ to $\mathbb{R}^{\mathrm{m}}$ and $x \in \mathbb{R}^{\mathrm{n}}$.

Proposition 2.19. *Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be an open set, and $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ be differentiable at $x_0 \in \mathcal{U}$. Then $(Df)(x_0)$, the derivative of f at x_0, is uniquely determined by f.*

Proof. Suppose $L_1, L_2 \in \mathscr{B}(\mathbb{R}^{\mathrm{n}}, \mathbb{R}^{\mathrm{m}})$ are derivatives of f at x_0. Let $\varepsilon > 0$ be given and $\mathrm{e} \in \mathbb{R}^{\mathrm{n}}$ be a unit vector; that is, $\|\mathrm{e}\|_{\mathbb{R}^{\mathrm{n}}} = 1$. Since $\mathcal{U}$ is open, there exists $r > 0$ such that $B(x_0, r) \subseteq \mathcal{U}$. By Definition 2.15, there exists $0 < \delta < r$ such that

$$\frac{\|f(x) - f(x_0) - L_1(x - x_0)\|_{\mathbb{R}^{\mathrm{m}}}}{\|x - x_0\|_{\mathbb{R}^{\mathrm{n}}}} < \frac{\varepsilon}{2} \quad \text{and} \quad \frac{\|f(x) - f(x_0) - L_2(x - x_0)\|_{\mathbb{R}^{\mathrm{m}}}}{\|x - x_0\|_{\mathbb{R}^{\mathrm{n}}}} < \frac{\varepsilon}{2}$$

if $0 < \|x - x_0\|_{\mathbb{R}^{\mathrm{n}}} < \delta$. Letting $x = x_0 + \lambda\mathrm{e}$ with $0 < |\lambda| < \delta$, we have

$$\begin{aligned}
\|L_1\mathrm{e} - L_2\mathrm{e}\|_{\mathbb{R}^{\mathrm{m}}} &= \frac{1}{|\lambda|}\|L_1(x - x_0) - L_2(x - x_0)\|_{\mathbb{R}^{\mathrm{m}}} \\
&\leqslant \frac{1}{|\lambda|}\Big(\big\|f(x) - f(x_0) - L_1(x - x_0)\big\|_{\mathbb{R}^{\mathrm{m}}} + \big\|f(x) - f(x_0) - L_2(x - x_2)\big\|_{\mathbb{R}^{\mathrm{m}}}\Big) \\
&= \frac{\big\|f(x) - f(x_0) - L_1(x - x_0)\big\|_{\mathbb{R}^{\mathrm{m}}}}{\|x - x_0\|_{\mathbb{R}^{\mathrm{n}}}} + \frac{\big\|f(x) - f(x_0) - L_2(x - x_0)\big\|_{\mathbb{R}^{\mathrm{m}}}}{\|x - x_0\|_{\mathbb{R}^{\mathrm{n}}}} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $L_1\mathrm{e} = L_2\mathrm{e}$ for all unit vectors $\mathrm{e} \in \mathbb{R}^{\mathrm{n}}$ which guarantees that $L_1 = L_2$ (since if $x \neq 0$, $L_1 x = \|x\|_{\mathbb{R}^{\mathrm{n}}} L_1\big(\frac{x}{\|x\|_{\mathbb{R}^{\mathrm{n}}}}\big) = \|x\|_{\mathbb{R}^{\mathrm{n}}} L_2\big(\frac{x}{\|x\|_{\mathbb{R}^{\mathrm{n}}}}\big) = L_2 x$). ▫

Example 2.20. $(Df)(x_0)$ may not be unique if the domain of f is not open. For example, let $A = \big\{(x, y) \,\big|\, 0 \leqslant x \leqslant 1, y = 0\big\}$ be a subset of $\mathbb{R}^2$, and $f : A \to \mathbb{R}$ be given by $f(x, y) = 0$. Fix $x_0 = (a, 0) \in A$, then both of the linear maps

$$L_1(x, y) = 0 \quad \text{and} \quad L_2(x, y) = ay \qquad \forall\,(x, y) \in \mathbb{R}^2$$

satisfy Definition 2.15 since

$$\lim_{(x,0)\to(a,0)} \frac{\big|f(x, 0) - f(a, 0) - L_1(x - a, 0)\big|}{\big\|(x, 0) - (a, 0)\big\|_{\mathbb{R}^2}} = \lim_{(x,0)\to(a,0)} \frac{\big|f(x, 0) - f(a, 0) - L_2(x - a, 0)\big|}{\big\|(x, 0) - (a, 0)\big\|_{\mathbb{R}^2}} = 0\,.$$

Definition 2.21. Let $\{\mathrm{e}_k\}_{k=1}^{\mathrm{n}}$ be the standard basis of $\mathbb{R}^{\mathrm{n}}$, $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be an open set, $a \in \mathcal{U}$ and $f : \mathcal{U} \to \mathbb{R}$ be a function. The partial derivative of f at a with respect to x_j, denoted by $\dfrac{\partial f}{\partial x_j}(a)$, is the limit

$$\lim_{h \to 0} \frac{f(a + h\mathrm{e}_j) - f(a)}{h}$$

if it exists. In other words, if $a = (a_1, \cdots, a_n)$, then

$$\frac{\partial f}{\partial x_j}(a) = \lim_{h \to 0} \frac{f(a_1, \cdots, a_{j-1}, a_j + h, a_{j+1}, \cdots, a_n) - f(a_1, \cdots, a_n)}{h}.$$

Theorem 2.22. *Suppose $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ is an open set and $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ is differentiable at $a \in \mathcal{U}$. Then the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(a)$ exists for all $i = 1, \cdots m$ and $j = 1, \cdots n$, and the matrix representation of the linear transformation $Df(a)$ (with respect to the standard basis of $\mathbb{R}^{\mathrm{n}}$ and $\mathbb{R}^{\mathrm{m}}$) is given by*

$$\big[Df(a)\big] = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_{\mathrm{n}}}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_{\mathrm{m}}}{\partial x_1}(a) & \cdots & \dfrac{\partial f_{\mathrm{m}}}{\partial x_{\mathrm{n}}}(a) \end{bmatrix} \quad \textit{or} \quad \big[Df(a)\big]_{ij} = \frac{\partial f_i}{\partial x_j}(a).$$

Proof. Since $\mathcal{U}$ is open and $a \in \mathcal{U}$, there exists $r > 0$ such that $B(a, r) \subseteq \mathcal{U}$. By the differentiability of f at a, there is $L \in \mathscr{B}(\mathbb{R}^{\mathrm{n}}, \mathbb{R}^{\mathrm{m}})$ such that for any given $\varepsilon > 0$, there exists $0 < \delta < r$ such that

$$\|f(x) - f(a) - L(x - a)\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \varepsilon \|x - a\|_{\mathbb{R}^{\mathrm{n}}} \text{ whenever } x \in B(a, \delta).$$

In particular, for each $i = 1, \cdots, m$,

$$\Big|\frac{f_i(a + h\mathrm{e}_j) - f_i(a)}{h} - (L\mathrm{e}_j)_i\Big| \leqslant \Big\|\frac{f(a + h\mathrm{e}_j) - f(a)}{h} - L\mathrm{e}_j\Big\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \varepsilon \quad \forall\, 0 < |h| < \delta, h \in \mathbb{R},$$

where $(L\mathrm{e}_j)_i$ denotes the i-th component of $L\mathrm{e}_j$ in the standard basis. As a consequence, for each $i = 1, \cdots, \mathrm{m}$,

$$\lim_{h \to 0} \frac{f_i(a + h\mathrm{e}_j) - f_i(a)}{h} = (L\mathrm{e}_j)_i \text{ exists}$$

and by definition, we must have $(L\mathrm{e}_j)_i = \dfrac{\partial f_i}{\partial x_j}(a)$. Therefore, $L_{ij} = \dfrac{\partial f_i}{\partial x_j}(a)$. □

Definition 2.23. Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be an open set, and $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$. The matrix

$$(Jf)(x) \equiv \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_{\mathrm{n}}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_{\mathrm{m}}}{\partial x_1} & \cdots & \dfrac{\partial f_{\mathrm{m}}}{\partial x_{\mathrm{n}}} \end{bmatrix}(x) \equiv \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(x) & \cdots & \dfrac{\partial f_1}{\partial x_{\mathrm{n}}}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_{\mathrm{m}}}{\partial x_1}(x) & \cdots & \dfrac{\partial f_{\mathrm{m}}}{\partial x_{\mathrm{n}}}(x) \end{bmatrix}$$

is called the ***Jacobian matrix*** of f at x (if each entry exists).

Remark 2.24. A function f might not be differential even if the Jacobian matrix Jf exists; however, if f is differentiable at x_0, then $(Df)(x)$ can be represented by $(Jf)(x)$; that is, $[(Df)(x)] = (Jf)(x)$.

Example 2.25. Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be given by $f(x_1, x_2) = (x_1^2, x_1^3 x_2, x_1^4 x_2^2)$. Suppose that f is differentiable at $x = (x_1, x_2)$, then

$$\big[(Df)(x)\big] = \begin{bmatrix} 2x_1 & 0 \\ 3x_1^2 x_2 & x_1^3 \\ 4x_1^3 x_2^2 & 2x_1^4 x_2 \end{bmatrix}.$$

Remark 2.26. For each $x \in A$, $Df(x)$ is a linear transformation, but Df in general is not linear in x.

Example 2.27. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

Then $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{\partial f}{\partial y}(0,0) = 0$; thus if f is differentiable at $(0,0)$, then $(Df)(0,0) = \begin{bmatrix} 0 & 0 \end{bmatrix}$. However,

$$\Big| f(x,y) - f(0,0) - \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Big| = \frac{|xy|}{x^2+y^2} = \frac{|xy|}{(x^2+y^2)^{\frac{3}{2}}} \sqrt{x^2+y^2}\,;$$

thus f is not differentiable at $(0,0)$ since $\dfrac{|xy|}{(x^2+y^2)^{\frac{3}{2}}}$ cannot be arbitrarily small even if x^2+y^2 is small.

Example 2.28. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} x & \text{if } y = 0\,, \\ y & \text{if } x = 0\,, \\ 1 & \text{otherwise}\,. \end{cases}$$

Then $\dfrac{\partial f}{\partial x}(0,0) = \lim\limits_{h\to 0} \dfrac{f(h,0)-f(0,0)}{h} = \lim\limits_{h\to 0} \dfrac{h}{h} = 1$. Similarly, $\dfrac{\partial f}{\partial y}(0,0) = 1$; thus if f is differentiable at $(0,0)$, then $(Df)(0,0) = \begin{bmatrix} 1 & 1 \end{bmatrix}$. However,

$$\left| f(x,y) - f(0,0) - \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right| = \big| f(x,y) - (x+y) \big| ;$$

thus if $xy \neq 0$,

$$\big| f(x,y) - (x+y) \big| = |1-x-y| \nrightarrow 0 \text{ as } (x,y) \to (0,0), xy \neq 0.$$

Therefore, f is not differentiable at $(0,0)$.

2.4 Conditions for Differentiability

Proposition 2.29. *Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be open, $a \in \mathcal{U}$, and $f = (f_1, \cdots, f_m) : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$. Then f is differentiable at a if and only if f_i is differentiable at a for all $i = 1, \cdots, m$. In other words, for vector-valued functions defined on an open subset of $\mathbb{R}^{\mathrm{n}}$,*

Componentwise differentiable $\Leftrightarrow$ Differentiable.

Proof. "$\Rightarrow$" Let $(Df)(a)$ be the Jacobian matrix of f at a. Then

$$\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni \big\| f(x) - f(a) - (Df)(a)(x-a) \big\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \varepsilon \|x-a\|_{\mathbb{R}^{\mathrm{n}}} \text{ if } \|x-a\|_{\mathbb{R}^{\mathrm{n}}} < \delta\,.$$

Let $\{\mathrm{e}_j\}_{j=1}^m$ be the standard basis of $\mathbb{R}^{\mathrm{m}}$, and $L_i \in \mathscr{L}(\mathbb{R}^{\mathrm{n}}, \mathbb{R})$ be given by $L_i(h) = \mathrm{e}_i^{\mathrm{T}}[(Df)(a)]h$. Then $L_i \in \mathscr{B}(\mathbb{R}^{\mathrm{n}}, \mathbb{R})$ by Remark 1.79, and if $\|x-a\|_{\mathbb{R}^{\mathrm{n}}} < \delta$,

$$\begin{aligned} \big| f_i(x) - f_i(a) - L_i(x-a) \big| &= \big| \mathrm{e}_i \cdot \big(f(x) - f(a) - (Df)(a)(x-a) \big) \big| \\ &\leqslant \big\| f(x) - f(a) - (Df)(a)(x-a) \big\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \varepsilon \|x-a\|_{\mathbb{R}^{\mathrm{n}}}\,; \end{aligned}$$

thus f_i is differentiable at a with derivatives L_i.

"$\Leftarrow$" Suppose that $f_i : \mathcal{U} \to \mathbb{R}$ is differentiable at a for each $i = 1, \cdots, m$. Then there exists $L_i \in \mathscr{B}(\mathbb{R}^{\mathrm{n}}, \mathbb{R})$ such that

$$\forall\, \varepsilon > 0, \exists\, \delta_i > 0 \ni \big| f_i(x) - f_i(a) - L_i(x-a) \big| \leqslant \frac{\varepsilon}{m} \|x-a\|_{\mathbb{R}^{\mathrm{n}}} \text{ if } \|x-a\|_{\mathbb{R}^{\mathrm{n}}} < \delta_i\,.$$

Let $L \in \mathscr{L}(\mathbb{R}^{\mathrm{n}}, \mathbb{R}^{\mathrm{m}})$ be given by $Lx = (L_1 x, L_2 x, \cdots, L_m x) \in \mathbb{R}^{\mathrm{m}}$ if $x \in \mathbb{R}^{\mathrm{n}}$. Then $L \in \mathscr{B}(\mathbb{R}^{\mathrm{n}}, \mathbb{R}^{\mathrm{m}})$ by Remark 1.79, and

$$\big\| f(x) - f(a) - L(x-a) \big\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \sum_{i=1}^m \big| f_i(x) - f_i(a) - L_i(x-a) \big| \leqslant \varepsilon \|x-a\|_{\mathbb{R}^{\mathrm{n}}}$$

if $\|x-a\|_{\mathbb{R}^{\mathrm{n}}} < \delta = \min\big\{\delta_1, \cdots, \delta_m\big\}$. □

Theorem 2.30. *Let* $\mathcal{U} \subseteq \mathbb{R}^{\rm n}$ *be open,* $a \in \mathcal{U}$*, and* $f : \mathcal{U} \to \mathbb{R}$*. If*

1. *the Jacobian matrix of* f *exists in a neighborhood of* a*, and*

2. *at least* $({\rm n}-1)$ *entries of the Jacobian matrix of* f *are continuous at* a*,*

then f *is differentiable at* a*.*

Proof. W.L.O.G. we can assume that $\dfrac{\partial f}{\partial x_1}, \dfrac{\partial f}{\partial x_2}, \cdots, \dfrac{\partial f}{\partial x_{\rm n-1}}$ are continuous at a. Let $\{{\rm e}_j\}_{j=1}^{\rm n}$ be the standard basis of $\mathbb{R}^{\rm n}$, and $\varepsilon > 0$ be given. Since $\dfrac{\partial f}{\partial x_i}$ is continuous at a for $i = 1, \cdots, n-1$,

$$\exists\, \delta_i > 0 \ni \Big|\frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(a)\Big| < \frac{\varepsilon}{\sqrt{n}} \text{ whenever } \|x-a\|_{\mathbb{R}^{\rm n}} < \delta_i\,.$$

On the other hand, by the definition of the partial derivatives,

$$\exists\, \delta_n > 0 \ni \Big|\frac{f(a+h{\rm e}_n)-f(a)}{h} - \frac{\partial f}{\partial x_n}(a)\Big| < \frac{\varepsilon}{\sqrt{n}} \text{ whenever } 0 < |h| < \delta_n\,.$$

Let $k = x-a$ and $\delta = \min\big\{\delta_1, \cdots, \delta_n\big\}$. Then

$$\begin{aligned}
&\Big|f(x) - f(a) - \Big[\frac{\partial f}{\partial x_1}(a)(x_1-a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n-a_n)\Big]\Big| \\
&\quad = \Big|f(a+k) - f(a) - \frac{\partial f}{\partial x_1}(a)k_1 - \cdots - \frac{\partial f}{\partial x_n}(a)k_n\Big| \\
&\quad = \Big|f(a_1+k_1, \cdots, a_n+k_n) - f(a_1, \cdots, a_n) - \frac{\partial f}{\partial x_1}(a)k_1 - \cdots - \frac{\partial f}{\partial x_n}(a)k_n\Big| \\
&\quad \leqslant \Big|f(a_1+k_1, \cdots, a_n+k_n) - f(a_1, a_2+k_2, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_1}(a)k_1\Big| \\
&\qquad + \Big|f(a_1, a_2+k_2, \cdots, a_n+k_n) - f(a_1, a_2, a_3+k_3, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_2}(a)k_2\Big| \\
&\qquad + \cdots + \Big|f(a_1, \cdots, a_{n-1}, a_n+k_n) - f(a_1, \cdots, a_n) - \frac{\partial f}{\partial x_n}(a)k_n\Big|\,.
\end{aligned}$$

By the mean value theorem,

$$\begin{aligned}
&f(a_1, \cdots, a_{j-1}, a_j+k_j, \cdots, a_n+k_n) - f(a_1, \cdots, a_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n) \\
&\qquad = k_j \frac{\partial f}{\partial x_j}(a_1, \cdots, a_{j-1}, a_j+\theta_j k_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n)
\end{aligned}$$

for some $0 < \theta_j < 1$; thus for $j = 1, \cdots, n-1$, if $\|x-a\|_{\mathbb{R}^{\rm n}} = \|k\|_{\mathbb{R}^{\rm n}} < \delta$,

$$\begin{aligned}
&\Big|f(a_1, \cdots, a_{j-1}, a_j+k_j, \cdots, a_n+k_n) - f(a_1, \cdots, a_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_j}(a)k_j\Big| \\
&\quad = \Big|\frac{\partial f}{\partial x_j}(a_1, \cdots, a_{j-1}, a_j+\theta_j k_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_j}(a)\Big||k_j| \leqslant \frac{\varepsilon}{\sqrt{n}}|k_j|\,.
\end{aligned}$$

Moreover, if $\|x-a\|_{\mathbb{R}^n} < \delta$, then $|k_n| \leqslant \|k\|_{\mathbb{R}^n} = \|x-a\|_{\mathbb{R}^n} < \delta \leqslant \delta_n$; thus

$$\Big| f(a_1, \cdots, a_{n-1}, a_n + k_n) - f(a_1, \cdots, a_n) - \frac{\partial f}{\partial x_n}(a) k_n \Big| \leqslant \frac{\varepsilon}{\sqrt{n}} |k_n| \,.$$

As a consequence, if $\|x-a\|_{\mathbb{R}^n} < \delta$, by Cauchy's inequality,

$$\begin{aligned} &\Big| f(x) - f(a) - \Big[\frac{\partial f}{\partial x_1}(a)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n - a_n)\Big]\Big| \\ &\qquad \leqslant \frac{\varepsilon}{\sqrt{n}} \sum_{j=1}^{n} |k_j| \leqslant \varepsilon \|k\|_{\mathbb{R}^n} = \varepsilon \|x-a\|_{\mathbb{R}^n} \end{aligned}$$

which implies that f is differentiable at a. □

Remark 2.31. When two or more components of the Jacobian matrix $\Big[\frac{\partial f}{\partial x_1} \ \cdots \ \frac{\partial f}{\partial x_n}\Big]$ of a scalar function f are discontinuous at a point $x_0 \in \mathcal{U}$, in general f is not differentiable at x_0. For example, both components of the Jacobian matrix of the functions given in Example 2.27, 2.28, 2.44 are discontinuous at $(0,0)$, and these functions are not differentiable at $(0,0)$.

Example 2.32. Let $\mathcal{U} = \mathbb{R}^2 \backslash \big\{(x,0) \in \mathbb{R}^2 \,\big|\, x \geqslant 0\big\}$, and $f : \mathcal{U} \to \mathbb{R}$ be given by

$$f(x,y) = \arg(x+iy) = \left\{ \begin{array}{cl} \cos^{-1} \dfrac{x}{\sqrt{x^2+y^2}} & \text{if } y > 0\,, \\ \pi & \text{if } y = 0\,, \\ 2\pi - \cos^{-1} \dfrac{x}{\sqrt{x^2+y^2}} & \text{if } y < 0\,. \end{array} \right.$$

Then

$$\frac{\partial f}{\partial x}(x,y) = \left\{ \begin{array}{cl} -\dfrac{y}{x^2+y^2} & \text{if } y \neq 0\,, \\ 0 & \text{if } y = 0\,, \end{array} \right. \quad \text{and} \quad \frac{\partial f}{\partial y}(x,y) = \left\{ \begin{array}{cl} \dfrac{x}{x^2+y^2} & \text{if } y \neq 0\,, \\ \dfrac{1}{x} & \text{if } y = 0\,. \end{array} \right.$$

Since $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are both continuous on $\mathcal{U}$, f is differentiable on $\mathcal{U}$.

Definition 2.33. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$ be differentiable on $\mathcal{U}$. f is said to be ***continuously differentiable*** on $\mathcal{U}$ if the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ exist and are continuous on $\mathcal{U}$ for $i = 1, \cdots, m$ and $j = 1, \cdots, n$. The collection of all continuously differentiable functions from $\mathcal{U}$ to $\mathbb{R}^m$ is denoted by $\mathscr{C}^1(\mathcal{U};\mathbb{R}^m)$. The collection of all bounded differentiable functions from $\mathcal{U}$ to $\mathbb{R}^m$ whose partial derivatives are continuous and bounded is denoted by $\mathscr{C}^1_b(\mathcal{U};\mathbb{R}^m)$.

Example 2.34. If $f : \mathbb{R} \to \mathbb{R}$ is differentiable at x_0, must f' be continuous at x_0? In other words, is it always true that $\lim\limits_{x\to x_0} f'(x) = f'(x_0)$?

Answer: No! For example, take

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

1° Show $f(x)$ is differentiable at $x = 0$:

$$f'(0) = \lim_{h\to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h\to 0} \frac{h^2 \sin \frac{1}{h}}{h} = \lim_{h\to 0} h \sin \frac{1}{h} = 0\,.$$

2° We compute the derivative of f and find that

$$f'(x) = \begin{cases} 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

However, $\lim\limits_{x\to 0} f'(x)$ does not exist.

Definition 2.35. Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be open, and $f : \mathcal{U} \to \mathbb{R}$ be a function. If the partial derivative $\dfrac{\partial f}{\partial x_j}$ exists in $\mathcal{U}$ and has partial derivatives (at every point in $\mathcal{U}$) with respect to x_i, then the second-order partial derivatives $\dfrac{\partial}{\partial x_i}\Big(\dfrac{\partial f}{\partial x_j}\Big)$ is denoted by $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

In general, if the k-th order partial derivatives $\dfrac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \cdots \partial x_{i_1}}$ exists in $\mathcal{U}$ and has partial derivatives (at every point in $\mathcal{U}$) with respect to $x_{i_{k+1}}$, then the $(k+1)$-th order partial derivatives $\dfrac{\partial}{\partial x_{i_{k+1}}}\Big(\dfrac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \cdots \partial x_{i_1}}\Big)$ is denoted by $\dfrac{\partial^{k+1} f}{\partial x_{i_{k+1}} \partial x_{i_k} \cdots \partial x_{i_1}}$; that is,

$$\frac{\partial^{k+1} f}{\partial x_{i_{k+1}} \partial x_{i_k} \cdots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_{k+1}}}\Big(\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \cdots \partial x_{i_1}}\Big)\,.$$

Theorem 2.36. *Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be open, $a \in \mathcal{U}$, and $f : \mathcal{U} \to \mathbb{R}$ be a real-valued function. Suppose that for some $1 \leqslant i, j \leqslant \mathrm{n}$, $\dfrac{\partial f}{\partial x_i}$, $\dfrac{\partial f}{\partial x_j}$, $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ and $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ exist in a neighborhood of a and are continuous at a. Then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)\,.$$

Proof. W.L.O.G., we assume that f is a function of two variables; that is, n $= 2$. For fixed $h, k \in \mathbb{R}$, define $\varphi(x,y) = f(x,y+k) - f(x,y)$ and $\psi(x,y) = f(x+h,y) - f(x,y)$. Then

$$\begin{aligned}\varphi(a+h,b) - \varphi(a,b) &= f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b)\\ &= \psi(a,b+k) - \psi(a,b)\,.\end{aligned}$$

By the mean value theorem (Theorem A.9), for $h, k \neq 0$ and sufficiently small,

$$\begin{aligned}\varphi(a+h,b) - \varphi(a,b) = \varphi_x(a+\theta_1 h,b)h &= \big[f_x(a+\theta_1 h,b+k) - f_x(a+\theta_1 h,b)\big]h\\ &= (f_x)_y(a+\theta_1 h,b+\theta_2 k)hk\end{aligned}$$

for some $\theta_1,\theta_2 \in (0,1)$, and similarly, for some $\theta_3,\theta_4 \in (0,1)$,

$$\psi(a,b+k) - \psi(a,b) = (f_y)_x(a+\theta_3 h,b+\theta_4 k)hk\,.$$

Therefore, for $h, k \neq 0$ and sufficiently small, there exist $\theta_1,\theta_2,\theta_3,\theta_4 \in (0,1)$ such that

$$(f_x)_y(a+\theta_1 h,b+\theta_2 k) = (f_y)_x(a+\theta_3 h,b+\theta_4 k)\,. \tag{2.1}$$

Let $\varepsilon > 0$ be given. Since $(f_x)_y$ and $(f_y)_x$ are continuous at (a,b), there exist $\delta_1,\delta_2 > 0$ such that

$$\begin{aligned}\big|(f_x)_y(x,y) - (f_x)_y(a,b)\big| &< \frac{\varepsilon}{2} \;\text{ if }\; \sqrt{(x-a)^2+(y-b)^2} < \delta_1\,,\\ \big|(f_x)_y(x,y) - (f_x)_y(a,b)\big| &< \frac{\varepsilon}{2} \;\text{ if }\; \sqrt{(x-a)^2+(y-b)^2} < \delta_2\,.\end{aligned}$$

In particular, if $\delta = \min\{\delta_1,\delta_2\}$ and $h, k \neq 0$ satisfying $\sqrt{h^2+k^2} < \delta$,

$$\big|(f_x)_y(a+\theta_1 h,b+\theta_2 k) - (f_x)_y(a,b)\big| + \big|(f_x)_y(a+\theta_3 h,b+\theta_4 k) - (f_x)_y(a,b)\big| < \varepsilon\,,$$

where $\theta_1,\theta_2,\theta_3,\theta_4 \in (0,1)$ are chosen to validate (2.1). As a consequence,

$$\begin{aligned}&\big|(f_x)_y(a,b) - (f_y)_x(a,b)\big|\\ &\quad= \big|(f_x)_y(a,b) - (f_x)_y(a+\theta_1 h,b+\theta_2 k) + (f_x)_y(a+\theta_3 h,b+\theta_4 k) - (f_x)_y(a,b)\big|\\ &\quad\leqslant \big|(f_x)_y(a+\theta_1 h,b+\theta_2 k) - (f_x)_y(a,b)\big| + \big|(f_x)_y(a+\theta_3 h,b+\theta_4 k) - (f_x)_y(a,b)\big| < \varepsilon\end{aligned}$$

which concludes the theorem (since $\varepsilon > 0$ is given arbitrarily). □

Example 2.37. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

Then

$$f_x(x,y) = \begin{cases} \dfrac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,, \end{cases}$$

and

$$f_y(x,y) = \begin{cases} \dfrac{x^5-4x^3y^2-xy^4}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,, \end{cases}$$

It is clear that f_x and f_y are continuous on $\mathbb{R}^2$; thus f is differentiable on $\mathbb{R}^2$. However,

$$f_{xy}(0,0) = \lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = -1\,,$$

while

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = 1\,;$$

thus the Hessian matrix of f at the origin is not symmetric.

Definition 2.38. Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be open, and $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ be a vector-valued function. The function f is said to be of class $\mathscr{C}^2$ if $f \in \mathscr{C}^1(\mathcal{U};\mathbb{R}^{\mathrm{m}})$ and the second partial derivatives $\dfrac{\partial^2 f_i}{\partial x_j \partial x_k}$ exists and is continuous in $\mathcal{U}$ for all $1 \leqslant i \leqslant \mathrm{m}$ and $1 \leqslant j,k \leqslant \mathrm{n}$. The collection of all $\mathscr{C}^2$-functions $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ is denoted by $\mathscr{C}^2(\mathcal{U};\mathbb{R}^{\mathrm{m}})$.

In general, the function f is said to be of class $\mathscr{C}^k$ if $f \in \mathscr{C}^{k-1}(\mathcal{U};\mathbb{R}^{\mathrm{m}})$ and the k-th order partial derivatives $\dfrac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} \cdots \partial x_{i_1}}$ exists and is continuous in $\mathcal{U}$ for all $1 \leqslant i \leqslant \mathrm{m}$ and $1 \leqslant i_1, \cdots, i_k \leqslant \mathrm{n}$. The collection of all $\mathscr{C}^k$-functions $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ is denoted by $\mathscr{C}^k(\mathcal{U};\mathbb{R}^{\mathrm{m}})$.

A function is said to be ***smooth*** or ***of class*** $\mathscr{C}^\infty$ if it is of class $\mathscr{C}^k$ for all positive integer k.

Corollary 2.39. *Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be open, and $f \in \mathscr{C}^2(\mathcal{U};\mathbb{R})$. Then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) \qquad \forall\, a \in \mathcal{U} \text{ and } 1 \leqslant i,j \leqslant \mathrm{n}\,.$$

2.5 Properties of Differentiable Functions

2.5.1 Continuity of Differentiable Functions

Theorem 2.40. *Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be open, and $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ be differentiable at $x_0 \in \mathcal{U}$. Then f is continuous at x_0.*

Proof. Since f is differentiable at x_0, there exists $L \in \mathscr{B}(\mathbb{R}^{\mathrm{n}}, \mathbb{R}^{\mathrm{m}})$ such that

$$\exists\, \delta_1 > 0 \ni \big\| f(x) - f(x_0) - L(x - x_0) \big\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \|x - x_0\|_{\mathbb{R}^{\mathrm{n}}} \quad \forall\, x \in B(x_0, \delta_1)\,.$$

As a consequence,

$$\big\| f(x) - f(x_0) \big\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \big(\|L\| + 1\big) \|x - x_0\|_{\mathbb{R}^{\mathrm{n}}} \quad \forall\, x \in B(x_0, \delta_1)\,. \tag{2.2}$$

For a given $\varepsilon > 0$, let $\delta = \min\Big\{\delta_1, \dfrac{\varepsilon}{2(\|L\| + 1)}\Big\}$. Then $\delta > 0$, and if $x \in B(x_0, \delta)$,

$$\big\| f(x) - f(x_0) \big\|_{\mathbb{R}^{\mathrm{m}}} \leqslant \frac{\varepsilon}{2} < \varepsilon\,. \qquad \square$$

Remark 2.41. In fact, if f is differentiable at x_0, then f satisfies the "local Lipschitz property"; that is,

$$\exists\, M = M(x_0) > 0 \text{ and } \delta = \delta(x_0) > 0 \ni \text{if } \|x - x_0\|_X < \delta, \text{ then } \|f(x) - f(x_0)\|_Y \leqslant M \|x - x_0\|_X$$

since we can choose $M = \|L\| + 1$ and $\delta = \delta_1$ (see (2.2)).

Example 2.42. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given in Example 2.27. We have shown that f is not differentiable at $(0, 0)$. In fact, f is not even continuous at $(0, 0)$ since when approaching the origin along the straight line $x_2 = m x_1$,

$$\lim_{(x_1, m x_1) \to (0,0)} f(x_1, m x_1) = \lim_{x_1 \to 0} \frac{m x_1^2}{(m^2 + 1) x_1^2} = \frac{m^2}{m^2 + 1} \neq f(0, 0) \text{ if } m \neq 0\,.$$

Example 2.43. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given in Example 2.28. Then f is not continuous at $(0, 0)$; thus not differentiable at $(0, 0)$.

Example 2.44. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = \begin{cases} \dfrac{x^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0)\,, \\ 0 & \text{if } (x, y) = (0, 0)\,. \end{cases}$$

Then $f_x(0,0) = 1$ and $f_y(0,0) = 0$. However,

$$\frac{\left|f(x,y) - f(0,0) - \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}\right|}{\sqrt{x^2+y^2}} = \frac{|x|y^2}{(x^2+y^2)^{\frac{3}{2}}} \nrightarrow 0 \text{ as } (x,y) \to (0,0).$$

Therefore, f is not differentiable at $(0,0)$. On the other hand, f is continuous at $(0,0)$ since

$$\big|f(x,y) - f(0,0)\big| = \big|f(x,y)\big| \leqslant |x| \to 0 \text{ as } (x,y) \to (0,0).$$

2.5.2 The Product Rules

Proposition 2.45. *Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be an open set, and $f : \mathcal{U} \to \mathbb{R}^{\mathrm{m}}$ and $g : \mathcal{U} \to \mathbb{R}$ be differentiable at $x_0 \in A$. Then $gf : A \to \mathbb{R}^{\mathrm{m}}$ is differentiable at x_0, and*

$$D(gf)(x_0)(v) = g(x_0)(Df)(x_0)(v) + (Dg)(x_0)(v)f(x_0)\,. \tag{2.3}$$

Moreover, if $g(x_0) \neq 0$, then $\dfrac{f}{g} : A \to \mathbb{R}^{\mathrm{m}}$ is also differentiable at x_0, and $D(\dfrac{f}{g})(x_0) : \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{m}}$ is given by

$$D\big(\frac{f}{g}\big)(x_0)(v) = \frac{g(x_0)\big((Df)(x_0)(v)\big) - (Dg)(x_0)(v)f(x_0)}{g^2(x_0)}\,. \tag{2.4}$$

Proof. We only prove (2.3), and (2.4) is left as an exercise.

Let A be the Jacobian matrix of gf at x_0; that is, the (i,j)-th entry of A is

$$\frac{\partial(gf_i)}{\partial x_j}(x_0) = g(x_0)\frac{\partial f_i}{\partial x_j}(x_0) + \frac{\partial g}{\partial x_j}(x_0)f_i(x_0)\,.$$

Then $Av = g(x_0)(Df)(x_0)(v) + (Dg)(x_0)(v)f(x_0)$; thus

$$\begin{aligned}(gf)(x) - (gf)(x_0) - A(x-x_0) &= g(x_0)\big(f(x) - f(x_0) - (Df)(x_0)(x-x_0)\big)\\ &\quad + \big(g(x) - g(x_0) - (Dg)(x_0)(x-x_0)\big)f(x)\\ &\quad + \big((Dg)(x_0)(x-x_0)\big)\big(f(x) - f(x_0)\big)\,.\end{aligned}$$

Since $(Dg)(x_0) \in \mathscr{B}(\mathbb{R}^{\mathrm{n}},\mathbb{R})$, $\|(Dg)(x_0)\|_{\mathscr{B}(\mathbb{R}^{\mathrm{n}},\mathbb{R})} < \infty$; thus using the inequality

$$\big|(Dg)(x_0)(x-x_0)\big| \leqslant \big\|(Dg)(x_0)\big\|_{\mathscr{B}(\mathbb{R}^{\mathrm{n}},\mathbb{R})}\|x-x_0\|_{\mathbb{R}^{\mathrm{n}}}$$

and the continuity of f at x_0 (due to Theorem 2.40), we find that

$$\lim_{x\to x_0}\left|\frac{\big|(Dg)(x_0)(x-x_0)\big|}{\|x-x_0\|_{\mathbb{R}^{\mathrm{n}}}}\big\|f(x) - f(x_0)\big\|_{\mathbb{R}^{\mathrm{m}}}\right| \leqslant \lim_{x\to x_0}\big\|(Dg)(x_0)\big\|_{\mathscr{B}(\mathbb{R}^{\mathrm{n}},\mathbb{R})}\big\|f(x) - f(x_0)\big\|_{\mathbb{R}^{\mathrm{m}}} = 0\,.$$

As a consequence,

$$\begin{aligned}
&\lim_{x \to x_0} \frac{\big\|(gf)(x) - (gf)(x_0) - A(x - x_0)\big\|_{\mathbb{R}^m}}{\|x - x_0\|_{\mathbb{R}^n}} \\
&\qquad \leqslant \big|g(x_0)\big| \lim_{x \to x_0} \frac{\big\|f(x) - f(x_0) - (Df)(x_0)(x - x_0)\big\|_{\mathbb{R}^m}}{\|x - x_0\|_{\mathbb{R}^n}} \\
&\qquad\quad + \lim_{x \to x_0} \Big[\frac{\big|g(x) - g(x_0) - (Dg)(x_0)(x - x_0)\big|}{\|x - x_0\|_{\mathbb{R}^n}} \|f(x)\|_{\mathbb{R}^m}\Big] \\
&\qquad\quad + \lim_{x \to x_0} \Big[\frac{\big|(Dg)(x_0)(x - x_0)\big|}{\|x - x_0\|_{\mathbb{R}^n}} \big\|f(x) - f(x_0)\big\|_{\mathbb{R}^m}\Big] = 0
\end{aligned}$$

which implies that gf is differentiable at x_0 with derivative $D(gf)(x_0)$ given by (2.3). □

• The differentiation of the Jacobian

Before going into the next section, we study the differentiation of a special determinant, the Jacobian.

Example 2.46. Suppose that $\psi : \Omega \subseteq \mathbb{R}^n \to \psi(\Omega) \subseteq \mathbb{R}^n$ is a given diffeomorphism (thus $\det(\nabla\psi) \neq 0$). Let $\mathrm{M} = \nabla\psi$, and $\mathrm{J} = \det(\mathrm{M})$. By Corollary 1.72, the adjoint matrix of M is $\mathrm{J}\mathrm{M}^{-1}$. Letting δ be a (first order) partial differential operator which satisfies $\delta(fg) = f\delta g + (\delta f)g$, by Theorem 1.73 we find that

$$\delta\mathrm{J} = \mathrm{tr}(\mathrm{J}\mathrm{M}^{-1}\delta\mathrm{M}) = \sum_{i,j=1}^{n} \mathrm{J}\mathrm{A}_i^j \delta\psi^i_{,j} \overset{\substack{\text{Einstein's summation}\\ \text{convention}}}{\equiv} \mathrm{J}\mathrm{A}_i^j \delta\psi^i_{,j}\,, \tag{2.5}$$

where $\mathrm{A}_i^j = a_{ji}$ with $\mathrm{M}^{-1} = [a_{ji}]_{n \times n}$, and $f_{,j} \equiv \dfrac{\partial f}{\partial x_j}$.

Remark 2.47. From now on we sometimes write the row index of a matrix as a super-script for the following reason: if $\psi : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a differentiable vector-valued function, then $\nabla\psi$ is usually expressed by

$$\nabla\psi = \begin{bmatrix} \dfrac{\partial\psi_1}{\partial x_1} & \dfrac{\partial\psi_1}{\partial x_2} & \cdots & \dfrac{\partial\psi_1}{\partial x_n} \\ \dfrac{\partial\psi_2}{\partial x_1} & \dfrac{\partial\psi_2}{\partial x_2} & \cdots & \dfrac{\partial\psi_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial\psi_m}{\partial x_1} & \dfrac{\partial\psi_m}{\partial x_2} & \cdots & \dfrac{\partial\psi_m}{\partial x_n} \end{bmatrix};$$

thus the (i, j) element of $\nabla\psi$ is $\dfrac{\partial\psi_i}{\partial x_j}$, and the row index i appears "above" the column index j.

Theorem 2.48 (Piola's identity)**.** *Let $\psi : \Omega \subseteq \mathbb{R}^n \to \psi(\Omega) \subseteq \mathbb{R}^n$ be a $\mathscr{C}^2$-diffeomorphism, and $[a_{ij}]_{n\times n}$ be the adjoint matrix of $\nabla\psi$. Then*

$$a_{ji},_j \overset{\text{Einstein's summation convention}}{\equiv} \sum_{j=1}^{n} \frac{\partial}{\partial x_j} a_{ji} = 0. \tag{2.6}$$

In other words, each column of the adjoint matrix of the Jacobian matrix of ψ is divergence-free (see Definition 4.74).

Proof. Let $\mathrm{J} = \det(\nabla\psi)$ and $\mathrm{A} = (\nabla\psi)^{-1}$. Then $a_{ji} = \mathrm{J}\mathrm{A}^j_i$. Moreover, since $\mathrm{A}\nabla\psi = \mathrm{I}_n$, $\sum\limits_{r=1}^{n} \mathrm{A}^j_r \psi^r_{,s} = \delta_{js}$; thus

$$0 = \Big[\sum_{r=1}^{n} \mathrm{A}^j_r \psi^r_{,s}\Big]_{,k} = \sum_{r=1}^{n} \big[\mathrm{A}^j_{r,k}\psi^r_{,s} + \mathrm{A}^j_r \psi^r_{,sk}\big]$$

which, after multiplying the equality above by A^s_i and then summing over s, implies that

$$\mathrm{A}^j_{i,k} = -\sum_{r,s=1}^{n} \mathrm{A}^j_r \psi^r_{,sk} \mathrm{A}^s_i. \tag{2.7}$$

As a consequence, by Theorem 2.36 we conclude that

$$\sum_{j=1}^{n} \frac{\partial}{\partial x_j} (\mathrm{J}\mathrm{A}^j_i) = \sum_{j=1}^{n}\sum_{r,s=1}^{n} \Big[\mathrm{J}\mathrm{A}^r_s \psi^s_{,rj} \mathrm{A}^j_i - \mathrm{J}\mathrm{A}^j_r \psi^r_{,sj} \mathrm{A}^s_i\Big] = 0\,.$$

□

2.5.3 The Chain Rule

Theorem 2.49. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ and $\mathcal{V} \subseteq \mathbb{R}^m$ be open sets, $f : \mathcal{U} \to \mathbb{R}^m$ and $g : \mathcal{V} \to \mathbb{R}^\ell$ be vector-valued functions, and $f(\mathcal{U}) \subseteq \mathcal{V}$. If f is differentiable at $x_0 \in \mathcal{U}$ and g is differentiable at $f(x_0)$, then the map $F = g \circ f$ defined by*

$$F(x) = g\big(f(x)\big) \qquad \forall\, x \in \mathcal{U}$$

is differentiable at x_0, and

$$(DF)(x_0)(h) = (Dg)\big(f(x_0)\big)\big((Df)(x_0)(h)\big)$$

or in component,

$$\big[(DF)(x_0)\big]_{ij} = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}\big(f(x_0)\big)\frac{\partial f_k}{\partial x_j}(x_0)\,.$$

Proof. To simplify the notation, let $y_0 = f(x_0)$, $A = (Df)(x_0) \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$, and $B = (Dg)(y_0) \in \mathscr{B}(\mathbb{R}^m, \mathbb{R}^\ell)$. Let $\varepsilon > 0$ be given. By the differentiability of f and g at x_0 and y_0, there exists $\delta_1, \delta_2 > 0$ such that if $\|x - x_0\|_{\mathbb{R}^n} < \delta_1$ and $\|y - y_0\|_{\mathbb{R}^m} < \delta_2$, we have

$$\begin{aligned}\|f(x) - f(x_0) - A(x - x_0)\|_{\mathbb{R}^m} &\leqslant \min\Big\{1, \frac{\varepsilon}{2(\|B\| + 1)}\Big\}\|x - x_0\|_{\mathbb{R}^n}\,,\\ \|g(y) - g(y_0) - B(y - y_0)\|_{\mathbb{R}^\ell} &\leqslant \frac{\varepsilon}{2(\|A\| + 1)}\|y - y_0\|_{\mathbb{R}^m}\,.\end{aligned}$$

Define

$$\begin{aligned}u(h) &= f(x_0 + h) - f(x_0) - Ah \qquad \forall\, \|h\|_{\mathbb{R}^n} < \delta_1\,,\\ v(k) &= g(y_0 + k) - g(y_0) - Bk \qquad \forall\, \|k\|_{\mathbb{R}^m} < \delta_2\,.\end{aligned}$$

Then if $\|h\|_{\mathbb{R}^n} < \delta_1$ and $\|k\|_{\mathbb{R}^m} < \delta_2$,

$$\|u(h)\|_{\mathbb{R}^m} \leqslant \|h\|_{\mathbb{R}^n}\,, \quad \|u(h)\|_{\mathbb{R}^m} \leqslant \frac{\varepsilon}{2(\|B\| + 1)}\|h\|_{\mathbb{R}^n} \quad \text{and} \quad \|v(k)\|_{\mathbb{R}^\ell} \leqslant \frac{\varepsilon}{2(\|A\| + 1)}\|k\|_{\mathbb{R}^m}\,.$$

Let $k = f(x_0 + h) - f(x_0) = Ah + u(h)$. Then $\lim\limits_{h \to 0} k = 0$; thus there exists $\delta_3 > 0$ such that

$$\|k\|_{\mathbb{R}^m} < \delta_2 \quad \text{whenever} \quad \|h\|_{\mathbb{R}^n} < \delta_3\,.$$

Since

$$\begin{aligned}F(x_0 + h) - F(x_0) &= g(y_0 + k) - g(y_0) = Bk + v(k) = B\big(Ah + u(h)\big) + v(k)\\ &= BAh + Bu(h) + v(k)\,,\end{aligned}$$

we conclude that if $\|h\|_{\mathbb{R}^n} < \delta = \min\{\delta_1, \delta_3\}$,

$$\begin{aligned}\|F(x_0 + h) - F(x_0) - BAh\|_{\mathbb{R}^\ell} &\leqslant \|Bu(h)\|_{\mathbb{R}^\ell} + \|v(k)\|_{\mathbb{R}^\ell} \leqslant \|B\|\|u(h)\|_{\mathbb{R}^m} + \frac{\varepsilon}{2(\|A\| + 1)}\|k\|_{\mathbb{R}^m}\\ &\leqslant \frac{\varepsilon}{2}\|h\|_{\mathbb{R}^n} + \frac{\varepsilon}{2(\|A\| + 1)}\big(\|A\|\|h\|_{\mathbb{R}^n} + \|u(h)\|_{\mathbb{R}^m}\big) \leqslant \frac{\varepsilon}{2}\|h\|_{\mathbb{R}^n} + \frac{\varepsilon}{2}\|h\|_{\mathbb{R}^n} = \varepsilon\|h\|_{\mathbb{R}^n}\end{aligned}$$

which implies that F is differentiable at x_0 and $\big[(DF)(x_0)\big] = BA$. □

Example 2.50. Consider the polar coordinate $x = r\cos\theta$, $y = r\sin\theta$. Then every function $f : \mathbb{R}^2 \to \mathbb{R}$ is associated with a function $F : [0, \infty) \times [0, 2\pi) \to \mathbb{R}$ satisfying

$$F(r, \theta) = f(r\cos\theta, r\sin\theta)\,.$$

Suppose that f is differentiable. Then F is differentiable, and the chain rule implies that

$$\begin{bmatrix} \dfrac{\partial F}{\partial r} & \dfrac{\partial F}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

Therefore, we arrive at the following form of chain rule

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y} \qquad \text{and} \qquad \frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y}$$

which is commonly seen in Calculus textbook.

Example 2.51. Let $f : \mathbb{R} \to \mathbb{R}$ and $F : \mathbb{R}^2 \to \mathbb{R}$ be differentiable, and $F\big(x, f(x)\big) = 0$ and $\dfrac{\partial F}{\partial y} \neq 0$. Then $f'(x) = -\dfrac{F_x\big(x, f(x)\big)}{F_y\big(x, f(x)\big)}$, where $F_x = \dfrac{\partial F}{\partial x}$ and $F_y = \dfrac{\partial F}{\partial y}$.

Example 2.52. Let $\gamma : (0,1) \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. Let $F(t) = f\big(\gamma(t)\big)$. Then $F'(t) = (Df)\big(\gamma(t)\big)\gamma'(t)$.

Example 2.53. Let $f(u,v,w) = u^2v + wv^2$ and $g(x,y) = (xy, \sin x, e^x)$. Let $h = f \circ g : \mathbb{R}^2 \to \mathbb{R}$. Find $\dfrac{\partial h}{\partial x}$.

Way I: Compute $\dfrac{\partial h}{\partial x}$ directly: Since

$$h(x,y) = f(g(x,y)) = f(xy, \sin x, e^x) = x^2y^2\sin x + e^x\sin^2 x\,,$$

we have

$$\frac{\partial h}{\partial x} = 2xy^2\sin x + x^2y^2\cos x + e^x\sin^2 x + 2e^x\sin x\cos x\,.$$

Way II: Use the chain rule:

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial g_1}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial g_2}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial g_3}{\partial x} = 2uv \cdot y + (u^2 + 2wv)\cdot\cos x + v^2 \cdot e^x \\ &= 2xy^2\sin x + (x^2y^2 + 2e^x\sin x)\cos x + e^x\sin^2 x. \end{aligned}$$

Example 2.54. Let $F(x,y) = f(x^2+y^2)$, $f : \mathbb{R} \to \mathbb{R}$, $F : \mathbb{R}^2 \to \mathbb{R}$. Show that $x\dfrac{\partial F}{\partial y} = y\dfrac{\partial F}{\partial x}$.

Proof: Let $g(x,y) = x^2 + y^2$, $g : \mathbb{R}^2 \to \mathbb{R}$, then $F(x,y) = (f \circ g)(x,y)$. By the chain rule,

$$\begin{bmatrix} \dfrac{\partial F}{\partial x} & \dfrac{\partial F}{\partial y} \end{bmatrix} = f'(g(x,y)) \cdot \begin{bmatrix} \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{bmatrix} = f'(g(x,y)) \begin{bmatrix} 2x & 2y \end{bmatrix}$$

which implies that

$$\frac{\partial F}{\partial x} = 2xf'(g(x,y)), \qquad \frac{\partial F}{\partial y} = 2yf'(g(x,y))\,.$$

So $y\dfrac{\partial F}{\partial x} = f'(g(x,y))2xy = x\dfrac{\partial F}{\partial y}$.

2.5.4 The Mean Value Theorem

Theorem 2.55. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$ with $f = (f_1, \cdots, f_m)$. Suppose that f is differentiable on $\mathcal{U}$ and the line segment joining x and y lies in $\mathcal{U}$. Then there exist points $c_1, \cdots, c_m$ on that segment such that*

$$f_i(y) - f_i(x) = (Df_i)(c_i)(y - x) \qquad \forall\, i = 1, \cdots, m.$$

Moreover, if $\mathcal{U}$ is convex and $\sup\limits_{x \in \mathcal{U}} \|(Df)(x)\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)} \leqslant M$, then

$$\|f(x) - f(y)\|_{\mathbb{R}^m} \leqslant M\|x - y\|_{\mathbb{R}^n} \qquad \forall\, x, y \in \mathcal{U}\,.$$

Proof. Let $\gamma : [0,1] \to \mathbb{R}^n$ be given by $\gamma(t) = (1-t)x + ty$. Then by Theorem 2.49, for each $i = 1, \cdots, m$, $(f_i \circ \gamma) : [0,1] \to \mathbb{R}$ is differentiable on $(0,1)$; thus the mean value theorem (Theorem A.9) implies that there exists $t_i \in (0,1)$ such that

$$f_i(y) - f_i(x) = (f_i \circ \gamma)(1) - (f_i \circ \gamma)(0) = (f_i \circ \gamma)'(t_i) = (Df_i)(c_i)\big(\gamma'(t_i)\big)\,,$$

where $c_i = \gamma(t_i)$. On the other hand, $\gamma'(t_i) = y - x$.

Let $g(t) = (f \circ \gamma)(t)$. Then the chain rule implies that $g'(t) = (Df)(\gamma(t))(y - x)$; thus

$$\|g'(t)\|_{\mathbb{R}^m} \leqslant \|(Df)(\gamma(t))\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)} \|y - x\|_{\mathbb{R}^m} \leqslant M\|x - y\|_{\mathbb{R}^n}\,.$$

Define $h(t) = \big(g(1) - g(0)\big) \cdot g(t)$. Then $h : [0,1] \to \mathbb{R}$ is differentiable; thus by the mean value theorem (Theorem A.9) we find that there exists $\xi \in (0,1)$ such that

$$h(1) - h(0) = h'(\xi) = \big(g(1) - g(0)\big) \cdot g'(\xi)\,;$$

thus by the fact that $g(0) = f(x)$ and $g(1) = f(y)$,

$$\begin{aligned}\|f(x) - f(y)\|^2_{\mathbb{R}^m} &= h(1) - h(0) \leqslant \|g(1) - g(0)\|_{\mathbb{R}^m} \|g'(\xi)\|_{\mathbb{R}^m} \\ &\leqslant M\|f(x) - f(y)\|_{\mathbb{R}^m} \|x - y\|_{\mathbb{R}^n}\end{aligned}$$

which concludes the theorem. □

Example 2.56. Let $f : [0,1] \to \mathbb{R}^2$ be given by $f(t) = (t^2, t^3)$. Then there is no $s \in (0,1)$ such that

$$(1,1) = f(1) - f(0) = f'(s)(1 - 0) = f'(s)$$

since $f'(s) = (2s, 3s^2) \neq (1,1)$ for all $s \in (0,1)$.

Example 2.57. Let $f:\mathbb{R}\to\mathbb{R}^2$ be given by $f(x)=(\cos x,\sin x)$. Then $f(2\pi)-f(0)=(0,0)$; however, $f'(x)=(-\sin x,\cos x)$ which cannot be a zero vector.

Example 2.58. Let f be given in Example 2.32, and $\mathcal{U}$ be a small neighborhood of the curve

$$\mathcal{C}=\big\{(x,y)\,\big|\,x^2+y^2=1,x\leqslant 0\big\}\cup\big\{(x,\pm 1)\,|\,0\leqslant x\leqslant 1\big\}.$$

Then

$$f(1,-1)-f(1,1)=\frac{3\pi}{2}.$$

On the other hand,

$$(Df)(x,y)(0,-2)=\begin{bmatrix}\dfrac{-y}{x^2+y^2} & \dfrac{x}{x^2+y^2}\end{bmatrix}\begin{bmatrix}0\\-2\end{bmatrix}=-\frac{2x}{x^2+y^2}$$

which can never be $\dfrac{3\pi}{2}$ since $\Big|\dfrac{2x}{x^2+y^2}\Big|\leqslant 3$ if $(x,y)\in\mathcal{U}$ while $\dfrac{3\pi}{2}>3$. Therefore, no point (x,y) in $\mathcal{U}$ validates

$$(Df)(x,y)\big((1,-1)-(1,1)\big)=f(1,-1)-f(1,1).$$

Example 2.59. Suppose that $\mathcal{U}\subseteq\mathbb{R}^n$ is an open convex set, and $f:\mathcal{U}\to\mathbb{R}^m$ is differentiable and $Df(x)=0$ for all $x\in\mathcal{U}$. Then f is a constant; that is, for some $\alpha\in\mathbb{R}^m$ we have $f(x)=\alpha$ for all $x\in\mathcal{U}$.
Reason: Since $\mathcal{U}$ is convex, then the Mean Value Theorem can be applied to any $x,y\in\mathcal{U}$ such that $f_i(x)-f_i(y)=Df_i(c_i)(x-y)=0$ ($\because Df_i=0$) for $i=1,2,\cdots,m$; thus $f(x)=f(y)$ for any $x,y\in\mathcal{U}$. Let $\alpha=f(x)\in\mathbb{R}^m$, then we reach the conclusion.

2.6 The Inverse Function Theorem（反函數定理）

反函數定理是用來探討一個函數的反函數是否存在的問題。只要一個函數不是一對一的，一般來說都不能定義其反函數，例如三角函數中，正弦、餘弦及正切函數都是周期函數，所以全域的反函數不存在。但是我們也知道有所謂的反三角函數 $\sin^{-1}$（或 arcsin），$\cos^{-1}$（或 arctan）及 $\tan^{-1}$（或 arctan），這是因為我們限制了原三角函數的定義域使其在新的定義域上是一對一的（因此反函數存在）。因此，要討論一個定在某一個（大範圍的）定義域的函數的反函數，常常我們最多只能說反函數只在某一小塊區域上存在。

如何知道一個函數在一小塊區域上的反函數存在，我們首先該問的是在定義域是一維（或是指單變數函數）的情況下發生什麼事？由一維的反函數定理 (Theorem A.10) 我們知

道首先應該要保留的條件是類似於微分不為零的這個條件。但是在多變數函數之下，微分不為零的條件該怎麼呈現，這是第一個問題。而當我們觀察 (A.1)，應該可以猜出在多變數版本裡面所該對應到的條件，即是 $(Df)(x)$ 這個 bounded linear map 的可逆性。

另外，假設 $f \in \mathscr{C}^1$，那麼由 Theorem 1.87 我們知道在一個點 x_0 如果 $(Df)(x_0)$ 可逆的話，那麼在一個鄰域裡 $(Df)(x)$ 都可逆。所以下面這個反函數定理的條件中只有 (Df) 在一個點可逆這個條件，因為我們暫時也只能討論在小區域的反函數存不存在。

Before proceeding, we first prove the following important proposition which is used crucially in the proof of the inverse function theorem.

Proposition 2.60 (Contraction Mapping Principle)**.** *Let $F \subseteq \mathbb{R}^n$ be a closed subset (on which every Cauchy sequence converges), and $\Phi : F \to F$ be a **contraction mapping**; that is, there is a constant $\theta \in [0,1)$ such that*

$$\big\|\Phi(x) - \Phi(y)\big\|_{\mathbb{R}^n} \leqslant \theta \|x - y\|_{\mathbb{R}^n} .$$

*Then there exists a unique point $x \in F$, called the **fixed-point** of Φ, such that $\Phi(x) = x$.*

Proof. Let $x_0 \in F$, and define $x_{k+1} = \Phi(x_k)$ for all $k \in \mathbb{N} \cup \{0\}$. Then

$$\|x_{k+1} - x_k\|_{\mathbb{R}^n} = \big\|\Phi(x_k) - \Phi(x_{k-1})\big\|_{\mathbb{R}^n} \leqslant \theta \|x_k - x_{k-1}\|_{\mathbb{R}^n} \leqslant \cdots \leqslant \theta^k \|x_1 - x_0\|_{\mathbb{R}^n} ;$$

thus if $\ell > k$,

$$\begin{aligned} \|x_\ell - x_k\|_{\mathbb{R}^n} &\leqslant \|x_k - x_{k+1}\|_{\mathbb{R}^n} + \|x_{k+1} - x_{k+2}\|_{\mathbb{R}^n} + \cdots + \|x_{\ell-1} - x_\ell\|_{\mathbb{R}^n} \\ &\leqslant (\theta^k + \theta^{k+1} + \cdots + \theta^{\ell-1}) \|x_1 - x_0\|_{\mathbb{R}^n} \\ &\leqslant \theta^k (1 + \theta + \theta^2 + \cdots) \|x_1 - x_0\|_{\mathbb{R}^n} = \frac{\theta^k}{1-\theta} \|x_1 - x_0\|_{\mathbb{R}^n} . \end{aligned} \tag{2.8}$$

Since $\theta \in [0,1)$, $\lim\limits_{k\to\infty} \dfrac{\theta^k}{1-\theta} \|x_1 - x_0\|_{\mathbb{R}^n} = 0$; thus

$$\forall\, \varepsilon > 0, \exists\, N > 0 \ni \|x_k - x_\ell\|_{\mathbb{R}^n} < \varepsilon \quad \forall\, k, \ell \geqslant N .$$

In other words, $\{x_k\}_{k=1}^\infty$ is a Cauchy sequence in F. By assumption, $x_k \to x$ as $k \to \infty$ for some $x \in F$. Finally, since $\Phi(x_k) = x_{k+1}$ for all $k \in \mathbb{N}$, by the continuity of Φ we obtain that

$$\Phi(x) = \lim_{k\to\infty} \Phi(x_k) = \lim_{k\to\infty} x_{k+1} = x$$

which guarantees the existence of a fixed-point.

Suppose that for some $x, y \in M$, $\Phi(x) = x$ and $\Phi(y) = y$. Then

$$\|x - y\|_{\mathbb{R}^n} = \big\|\Phi(x) - \Phi(y)\big\|_{\mathbb{R}^n} \leqslant \theta\|x - y\|_{\mathbb{R}^n}$$

which suggests that $\|x - y\|_{\mathbb{R}^n} = 0$ or $x = y$. Therefore, the fixed-point of Φ is unique. □

Now we state and prove the inverse function theorem.

Theorem 2.61 (Inverse Function Theorem). *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $x_0 \in \mathcal{D}$, $f : \mathcal{D} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x_0)$ be invertible. Then there exist an open neighborhood $\mathcal{U}$ of x_0 and an open neighborhood $\mathcal{V}$ of $f(x_0)$ such that*

1. *$f : \mathcal{U} \to \mathcal{V}$ is one-to-one and onto;*
2. *The inverse function $f^{-1} : \mathcal{V} \to \mathcal{U}$ is of class $\mathscr{C}^1$;*
3. *If $x = f^{-1}(y)$, then $(Df^{-1})(y) = \big((Df)(x)\big)^{-1}$;*
4. *If f is of class $\mathscr{C}^r$ for some $r > 1$, so is f^{-1}.*

Proof. We will omit the proof of 4 since it requires more knowledge about differentiation.

Assume that $A = (Df)(x_0)$. Then $\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} \neq 0$. Choose $\lambda > 0$ such that $2\lambda\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} = 1$. Since $f \in \mathscr{C}^1$, there exists $\delta > 0$ such that

$$\big\|(Df)(x) - A\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} = \big\|(Df)(x) - (Df)(x_0)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} < \lambda \quad \text{whenever } x \in B(x_0, \delta) \cap \mathcal{D}\,.$$

By choosing δ even smaller if necessary, we can assume that $B(x_0, \delta) \subseteq \mathcal{D}$. Let $\mathcal{U} = B(x_0, \delta)$.
Claim: $f : \mathcal{U} \to \mathbb{R}^n$ is one-to-one (hence $f : \mathcal{U} \to f(\mathcal{U})$ is one-to-one and onto).
Proof of claim: For each $y \in \mathbb{R}^n$, define $\varphi_y(x) = x + A^{-1}\big(y - f(x)\big)$ (and we note that every fixed-point of φ_y corresponds to a solution to $f(x) = y$). Then

$$(D\varphi_y)(x) = \mathrm{Id} - A^{-1}(Df)(x) = A^{-1}\big(A - (Df)(x)\big)\,,$$

where Id is the identity map on $\mathbb{R}^n$. Therefore,

$$\big\|(D\varphi_y)(x)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} \leqslant \|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\big\|A - (Df)(x)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} < \frac{1}{2} \quad \forall\, x \in B(x_0, \delta)\,.$$

By the mean value theorem (Theorem 2.55),

$$\big\|\varphi_y(x_1) - \varphi_y(x_2)\big\|_{\mathbb{R}^n} \leqslant \frac{1}{2}\|x_1 - x_2\|_{\mathbb{R}^n} \qquad \forall\, x_1, x_2 \in B(x_0, \delta),\ x_1 \neq x_2\,; \tag{2.9}$$

thus at most one x satisfies $\varphi_y(x) = x$; that is, φ_y has at most one fixed-point. As a consequence, $f : B(x_0, \delta) \to \mathbb{R}^n$ is one-to-one.
Claim: The set $\mathcal{V} = f(\mathcal{U})$ is open.
Proof of claim: Let $b \in \mathcal{V}$. Then there is $a \in \mathcal{V}$ with $f(a) = b$. Choose $r > 0$ such that $\overline{B(a,r)} \subseteq \mathcal{U}$. We observe that if $y \in B(b, \lambda r)$, then

$$\|\varphi_y(a) - a\|_{\mathbb{R}^n} \leqslant \|A^{-1}\big(y - f(a)\big)\|_{\mathbb{R}^n} \leqslant \|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\|y - b\|_{\mathbb{R}^n} < \lambda\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}r = \frac{r}{2}\,;$$

thus if $y \in B(b, \lambda r)$ and $x \in B(a, r)$,

$$\|\varphi_y(x) - a\|_{\mathbb{R}^n} \leqslant \big\|\varphi_y(x) - \varphi_y(a)\big\|_{\mathbb{R}^n} + \|\varphi_y(a) - a\|_{\mathbb{R}^n} < \frac{1}{2}\|x - a\|_{\mathbb{R}^n} + \frac{r}{2} < r\,.$$

Therefore, if $y \in B(b, \lambda r)$, then $\varphi_y : B(a, r) \to B(a, r)$. By the continuity of φ_y,

$$\varphi_y : \overline{B(a,r)} \to \overline{B(a,r)}\,.$$

On the other hand, (2.9) implies that φ_y is a contraction mapping if $y \in B(b, \lambda r)$; thus by the contraction mapping principle (Proposition 2.60) φ_y has a unique fixed-point $x \in B(a, r)$. As a result, every $y \in B(b, \lambda r)$ corresponds to a unique $x \in B(a, r)$ such that $\varphi_y(x) = x$ or equivalently, $f(x) = y$. Therefore,

$$B(b, \lambda r) \subseteq f\big(B(a, r)\big) \subseteq f(\mathcal{U}) = \mathcal{V}\,.$$

Next we show that $f^{-1} : \mathcal{V} \to \mathcal{U}$ is differentiable. We note that if $x \in B(x_0, \delta)$,

$$\|(Df)(x_0) - (Df)(x)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} < \lambda\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} = \frac{1}{2}\,;$$

thus Theorem 1.87 implies that $(Df)(x)$ is invertible if $x \in B(x_0, \delta)$.

Let $b \in \mathcal{V}$ and $k \in \mathbb{R}^n$ such that $b + k \in \mathcal{V}$. Then there exists a unique $a \in \mathcal{U}$ and $h = h(k) \in \mathbb{R}^n$ such that $a + h \in \mathcal{U}$, $b = f(a)$ and $b + k = f(a + h)$. By the mean value theorem and (2.9),

$$\big\|\varphi_y(a + h) - \varphi_y(a)\big\|_{\mathbb{R}^n} < \frac{1}{2}\|h\|_{\mathbb{R}^n}\,;$$

thus the fact that $f(a + h) - f(a) = k$ implies that

$$\|h - A^{-1}k\|_{\mathbb{R}^n} < \frac{1}{2}\|h\|_{\mathbb{R}^n}$$

which further suggests that

$$\frac{1}{2}\|h\|_{\mathbb{R}^n} \leqslant \|A^{-1}k\|_{\mathbb{R}^n} \leqslant \|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\|k\|_{\mathbb{R}^n} \leqslant \frac{1}{2\lambda}\|k\|_{\mathbb{R}^n}\,. \tag{2.10}$$

As a consequence, if k is such that $b+k \in \mathcal{V}$,

$$\begin{aligned}
\frac{\left\|f^{-1}(b+k)-f^{-1}(b)-\big((Df)(a)\big)^{-1}k\right\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} &= \frac{\left\|a+h-a-\big((Df)(a)\big)^{-1}k\right\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} \\
&\leqslant \left\|\big((Df)(a)\big)^{-1}\right\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} \frac{\left\|k-(Df)(a)(h)\right\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} \\
&\leqslant \left\|\big((Df)(a)\big)^{-1}\right\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} \frac{\left\|f(a+h)-f(a)-(Df)(a)(h)\right\|_{\mathbb{R}^n}}{\|h\|_{\mathbb{R}^n}} \frac{\|h\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} \\
&\leqslant \frac{\left\|\big((Df)(a)\big)^{-1}\right\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}}{\lambda} \frac{\left\|f(a+h)-f(a)-(Df)(a)(h)\right\|_{\mathbb{R}^n}}{\|h\|_{\mathbb{R}^n}}\,.
\end{aligned}$$

Using (2.10), $h \to 0$ as $k \to 0$; thus passing $k \to 0$ on the left-hand side of the inequality above, by the differentiability of f we conclude that

$$\lim_{k\to 0} \frac{\left\|f^{-1}(b+k)-f^{-1}(b)-\big((Df)(a)\big)^{-1}k\right\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} = 0\,.$$

This proves 3. ▫

Remark 2.62. Since $f^{-1}: \mathcal{V} \to \mathcal{U}$ is continuous, for any open subset $\mathcal{W}$ of $\mathcal{U}$ $f(\mathcal{W}) = (f^{-1})^{-1}(\mathcal{W})$ is open relative to $\mathcal{V}$, or $f(\mathcal{W}) = \mathcal{O} \cap \mathcal{V}$ for some open set $\mathcal{O} \subseteq \mathbb{R}^n$. In other words, if $\mathcal{U}$ is an open neighborhood of x_0 given by the inverse function theorem, then $f(\mathcal{W})$ is also open for all open subsets $\mathcal{W}$ of $\mathcal{U}$. We call this property as f is a ***local open mapping*** at x_0.

Remark 2.63. Since $(Df)(x_0) \in \mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$, the condition that $(Df)(x_0)$ is invertible can be replaced by that the determinant of the Jacobian matrix of f at x_0 is not zero; that is,

$$\det\big([(Df)(x_0)]\big) \neq 0\,.$$

The determinant of the Jacobian matrix of f at x_0 is called the ***Jacobian*** of f at x_0. The Jacobian of f at x sometimes is denoted by $\mathbf{J}_f(x)$ or $\dfrac{\partial(f_1,\cdots,f_n)}{\partial(x_1,\cdots,x_n)}$.

Example 2.64. Let $f:\mathbb{R}\to\mathbb{R}$ be given by

$$f(x) = \begin{cases} x+2x^2\sin\dfrac{1}{x} & \text{if } x\neq 0\,, \\ 0 & \text{if } x = 0\,. \end{cases}$$

Let $0 \in (a,b)$ for some (small) open interval (a,b). Since $f'(x) = 1-2\cos\dfrac{1}{x}+4x\sin\dfrac{1}{x}$ for $x \neq 0$, f has infinitely many critical points in (a,b), and (for whatever reasons) these critical points are local maximum points or local minimum points of f which implies that f is **not** locally invertible even though we have $f'(0) = 1 \neq 0$. One cannot apply the inverse function theorem in this case since f is not $\mathscr{C}^1$.

Corollary 2.65. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, $f : \mathcal{U} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x)$ be invertible for all $x \in \mathcal{U}$. Then $f(\mathcal{W})$ is open for every open set $\mathcal{W} \subseteq \mathcal{U}$.*

在證明了小區域的（local）反函數定理 (Theorem 2.61) 之後，我們接下來要問的是全域的（global）反函數在什麼條件之下會存在。如果照一維的反函數定理，我們會猜測是不是只要 $(Df)(x)$ 在整個區域都可逆就能得到在全域的反函數都存在。以下給個反例說單單在這個條件之下，函數不一定會有一對一的性質。

Example 2.66. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$f(x, y) = (e^x \cos y, e^x \sin y) .$$

Then

$$\big[(Df)(x, y)\big] = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix} .$$

It is easy to see that the Jacobian of f at any point is not zero (thus $(Df)(x)$ is invertible for all $x \in \mathbb{R}^2$), and f is not globally one-to-one (thus the inverse of f does not exist globally) since for example, $f(x, y) = f(x, y + 2\pi)$.

要再加什麼條件進來才能得到反函數在全域都存在是個不容易的問題。在一維的情況下，導數是 sign definite 就表示函數在全域是嚴格單調的。在高維度的情況，即使是 $(Df)(x)$ 到處都可逆，仍然有很多情況可能發生（如上例）。下面這個定理（全域的反函數存在定理），從某種角度來說並沒有真的加了什麼條件以確保全域的反函數存在，只是多要求了在所考慮的區域邊界上函數是一對一的。這個條件在一維的情況之下是自動成立的：因為如果一單變數函數的導數是 sign definite，那麼函數在邊界上必定是一對一的（因為嚴格單調的關係）。

Theorem 2.67 (Global Existence of Inverse Function)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $f : \mathcal{D} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x)$ be invertible for all $x \in K$. Suppose that K is a connected*（連通，即只有一塊）*, closed and bounded subset of $\mathcal{D}$, and $f : \partial K \to \mathbb{R}^n$ is one-to-one. Then $f : K \to \mathbb{R}^n$ is one-to-one.*

全域的反函數定理的證明需要更多關於點拓的知識，所以不在這門課中證明。

2.7 The Implicit Function Theorem（隱函數定理）

Theorem 2.68 (Implicit Function Theorem)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open, and $F : \mathcal{D} \to \mathbb{R}^m$ be a function of class $\mathscr{C}^1$. Suppose that for some $(x_0, y_0) \in \mathcal{D}$, where $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$, $F(x_0, y_0) = 0$ and*

$$\big[(D_yF)(x_0,y_0)\big] = \begin{bmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_m} \end{bmatrix} (x_0, y_0)$$

is invertible. Then there exists an open neighborhood $\mathcal{U} \subseteq \mathbb{R}^n$ of x_0, an open neighborhood $\mathcal{V} \subseteq \mathbb{R}^m$ of y_0, and $f : \mathcal{U} \to \mathcal{V}$ such that

1. *$F\big(x, f(x)\big) = 0$ for all $x \in \mathcal{U}$;*

2. *$y_0 = f(x_0)$;*

3. *$(Df)(x) = -\big((D_yF)(x, f(x))\big)^{-1}(D_xF)\big(x, f(x)\big)$ for all $x \in \mathcal{U}$;*

4. *f is of class $\mathscr{C}^1$;*

5. *If F is of class $\mathscr{C}^r$ for some $r > 1$, so is f.*

Proof. Let $z = (x, y)$ and $w = (u, v)$, where $x, u \in \mathbb{R}^n$ and $y, v \in \mathbb{R}^m$. Define $w = G(z)$, where G is given by $G(x, y) = \big(x, F(x, y)\big)$. Then $G : \mathcal{D} \to \mathbb{R}^{n+m}$, and

$$\big[(DG)(x,y)\big] = \begin{bmatrix} \mathbb{I}_n & 0 \\ (D_xF)(x,y) & (D_yF)(x,y) \end{bmatrix},$$

where $\mathbb{I}_n$ is the $n \times n$ identity matrix and $(D_xF)(x, y) \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$ whose matrix representation is given by

$$\big[(D_xF)(x,y)\big] = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & \cdots & \dfrac{\partial F_m}{\partial x_n} \end{bmatrix} (x, y)\,.$$

We note that the Jacobian of G at (x_0, y_0) is $\det\big([(D_yF)(x_0, y_0)]\big)$ which does not vanish since $(D_yF)(x_0, y_0)$ is invertible, so the inverse function theorem implies that there exists open neighborhoods $\mathcal{O}$ of (x_0, y_0) and $\mathcal{W}$ of $\big(x_0, F(x_0, y_0)\big) = (x_0, 0)$ such that

(a) $G : \mathcal{O} \to \mathcal{W}$ is one-to-one and onto;

(b) the inverse function $G^{-1} : \mathcal{W} \to \mathcal{O}$ is of class $\mathscr{C}^r$;

(c) $(DG^{-1})\big(x, F(x,y)\big) = \big((DG)(x,y)\big)^{-1}$.

By Remark 2.62, W.L.O.G. we can assume that $\mathcal{O} = \mathcal{U} \times \mathcal{V}$, where $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ and $\mathcal{V} \subseteq \mathbb{R}^{\mathrm{m}}$ are open, and $x_0 \in \mathcal{U}$, $y_0 \in \mathcal{V}$.

Write $G^{-1}(u,v) = \big(\varphi(u,v), \psi(u,v)\big)$, where $\varphi : \mathcal{W} \to \mathcal{U}$ and $\psi : \mathcal{W} \to \mathcal{V}$. Then

$$(u,v) = G\big(\varphi(u,v), \psi(u,v)\big) = \big(\varphi(u,v), F(u, \psi(u,v))\big)$$

which implies that $\varphi(u,v) = u$ and $v = F(u, \psi(u,v))$. Let $f(x) = \psi(x,0)$. Then $\big(u, f(u)\big) \in \mathcal{U} \times \mathcal{V}$ is the unique point satisfying $F\big(u, f(u)\big) = 0$ if $u \in \mathcal{U}$. Therefore, $f : \mathcal{U} \to \mathcal{V}$, and

$$F\big(x, f(x)\big) = 0 \qquad \forall\, x \in \mathcal{U}\,.$$

Since $G(x_0, y_0) = (x_0, 0) = G\big(x_0, f(x_0)\big)$, $(x_0, y_0), \big(x_0, f(x_0)\big) \in \mathcal{O}$, and $G : \mathcal{O} \to \mathcal{W}$ is one-to-one, we must have $y_0 = f(x_0)$.

By (b) and (c), we have G^{-1} is of class $\mathscr{C}^1$, and

$$(DG^{-1})(u,v) = \big((DG)(x,y)\big)^{-1}.$$

As a consequence, $\psi \in \mathscr{C}^1$, and

$$\begin{aligned}\begin{bmatrix}(D_u\varphi)(u,v) & (D_v\varphi)(u,v)\\ (D_u\psi)(u,v) & (D_v\psi)(u,v)\end{bmatrix} &= \begin{bmatrix}\mathbb{I}_n & 0\\ (D_xF)(x,y) & (D_yF)(x,y)\end{bmatrix}^{-1}\\ &= \begin{bmatrix}\mathbb{I}_n & 0\\ -\big((D_yF)(x,y)\big)^{-1}(D_xF)(x,y) & \big((D_yF)(x,y)\big)^{-1}\end{bmatrix}.\end{aligned}$$

Evaluating the equation above at $v = 0$, we conclude that

$$(Df)(u) = (D_u\psi)(u,0) = -\big((D_yF)(u, f(u))\big)^{-1}(D_xF)\big(u, f(u)\big)$$

which implies 3. We also note that 4 follows from (b) and 5 follows from 3. □

Example 2.69. Let $F(x,y) = x^2 + y^2 - 1$.

1. If $(x_0, y_0) = (1, 0)$, then $F_x(x_0, y_0) = 2 \neq 0$; thus the implicit function theorem implies that locally x can be expressed as a function of y.

2. If $(x_0, y_0) = (0, -1)$, then $F_y(x_0, y_0) = -2 \neq 0$; thus the implicit function theorem implies that locally y can be expressed as a function of x.

3. If $(x_0, y_0) = \big(-\frac{1}{2}, \frac{\sqrt{3}}{2}\big)$, then $F_x(x_0, y_0) = -1 \neq 0$ and $F_y(x_0, y_0) = \sqrt{3} \neq 0$; thus the implicit function theorem implies that locally x can be expressed as a function of y and locally y can be expressed as a function of x.

Example 2.70. Suppose that (x, y, u, v) satisfies the equation

$$\left\{\begin{aligned} xu + yv^2 &= 0 \\ xv^3 + y^2u^6 &= 0 \end{aligned}\right.$$

and $(x_0, y_0, u_0, v_0) = (1, -1, 1, -1)$. Let $F(x, y, u, v) = (xu + yv^2, xv^3 + y^2u^6)$. Then $F(x_0, y_0, u_0, v_0) = 0$.

1. Since $(D_{x,y}F)(x_0, y_0, u_0, v_0) = \begin{bmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{bmatrix}(x_0, y_0, u_0, v_0) = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ is invertible, locally (x, y) can be expressed in terms of u, v; that is, locally $x = x(u, v)$ and $y = y(u, v)$.

2. Since $(D_{y,u}F)(x_0, y_0, u_0, v_0) = \begin{bmatrix} \dfrac{\partial F_1}{\partial y} & \dfrac{\partial F_1}{\partial u} \\ \dfrac{\partial F_2}{\partial y} & \dfrac{\partial F_2}{\partial u} \end{bmatrix}(x_0, y_0, u_0, v_0) = \begin{bmatrix} 1 & 1 \\ -2 & 6 \end{bmatrix}$ is invertible, locally (y, u) can be expressed in terms of x, v.

Example 2.71. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$f(x, y, z) = (xe^y + ye^z, xe^z + ze^y)\,.$$

Then f is of class $\mathscr{C}^1$, $f(-1, 1, 1) = (0, 0)$ and

$$\big[(Df)(x, y, z)\big] = \begin{bmatrix} e^y & xe^y + e^z & ye^z \\ e^z & ze^y & xe^z + e^y \end{bmatrix}.$$

Since $(D_{y,z}f)(-1, 1, 1) = \begin{bmatrix} 0 & e \\ e & 0 \end{bmatrix}$ is invertible, the implicit function theorem implies that the system

$$\left\{\begin{aligned} xe^y + ye^z &= 0 \\ xe^z + ze^y &= 0 \end{aligned}\right.$$

can be solved for y and z as continuously differentiable function of x for x near -1 and (y,z) near $(1,1)$. Furthermore, if we write $(y,z) = g(x)$ for x near -1, then

$$g'(x) = \begin{bmatrix} xe^y + e^z & ye^z & ye^z \\ ze^y & xe^z + e^y & \end{bmatrix}^{-1} \begin{bmatrix} e^y \\ e^z \end{bmatrix}.$$

2.8 Directional Derivatives and Gradient Vectors

Definition 2.72 (Directional Derivatives)**.** Let f be real-valued and defined on a neighborhood of $x_0 \in \mathbb{R}^{\rm n}$, and let ${\rm v} \in \mathbb{R}^{\rm n}$ be a unit vector. Then

$$(D_{\rm v} f)(x_0) \equiv \frac{d}{dt}\Big|_{t=0} f(x_0 + t{\rm v}) = \lim_{t\to 0} \frac{f(x_0 + t{\rm v}) - f(x_0)}{t}$$

is called the ***directional derivative***（方向導數）of f at x_0 in the direction v.

Remark 2.73. Let $\{{\rm e}_j\}_{j=1}^{\rm n}$ be the standard basis of $\mathbb{R}^{\rm n}$. Then the partial derivative $\dfrac{\partial f}{\partial x_j}(x_0)$ (if it exists) is the directional derivative of f at x_0 in the direction ${\rm e}_j$.

Remark 2.74. Let f be a real-valued differentiable function defined on a neighborhood of $x_0 \in \mathbb{R}^{\rm n}$, and let ${\rm v} \in \mathbb{R}^{\rm n}$ be a unit vector. For a curve $\gamma : (-\delta, \delta) \to \mathbb{R}^{\rm n}$ satisfying that $\gamma(0) = x_0$ and $\gamma'(0) = {\rm v}$, the chain rule shows that

$$\frac{d}{dt}\Big|_{t=0} (f \circ \gamma)(t) = (Df)(x_0)({\rm v}) = (D_{\rm v} f)(x_0)\,.$$

In other words, for a differentiable function f in a neighborhood of x_0, the derivative $\dfrac{d}{dt}\Big|_{t=0}(f\circ\gamma)$ is independent of γ as long as $\gamma(0) = x_0$ and $\gamma'(0) = {\rm v}$. Therefore, directional derivative of a differential function f at x_0 in the direction v can also be defined by the value $\dfrac{d}{dt}\Big|_{t=0}(f\circ\gamma)(t)$, where $\gamma : (-\delta,\delta) \to \mathbb{R}^{\rm n}$ is any curve satisfying $\gamma(0) = x_0$ and $\gamma'(0) = {\rm v}$.

Theorem 2.75. *Let $\mathcal{U} \subseteq \mathbb{R}^{\rm n}$ be open, and $f : \mathcal{U} \to \mathbb{R}$ be differentiable at x_0. Then the directional derivative of f at x_0 in the direction* v *is $(Df)(x_0)({\rm v})$.*

Proof. Since f is differentiable at x_0, $\forall\, \varepsilon > 0$, $\ni \delta > 0$ such that

$$\big|f(x) - f(x_0) - (Df)(x_0)(x - x_0)\big| \leqslant \frac{\varepsilon}{2}\|x - x_0\|_{\mathbb{R}^{\rm n}} \text{ whenever } \|x - x_0\|_{\mathbb{R}^{\rm n}} < \delta\,.$$

In particular, if $x = x_0 + t\mathrm{v}$ with v being a unit vector in $\mathbb{R}^n$ and $0 < |t| < \delta$, then

$$\Big|\frac{f(x_0 + t\mathrm{v}) - f(x_0)}{t} - (Df)(x_0)(\mathrm{v})\Big| = \frac{\big|f(x_0 + t\mathrm{v}) - f(x_0) - (Df)(x_0)(t\mathrm{v})\big|}{|t|}$$
$$= \frac{\big|f(x) - f(x_0) - (Df)(x_0)(x - x_0)\big|}{|t|} \leqslant \frac{\varepsilon}{2} < \varepsilon\,;$$

thus $(D_\mathrm{v} f)(x_0) = (Df)(x_0)(\mathrm{v})$. □

Remark 2.76. When $v \in \mathbb{R}^n$ but $0 < \|v\|_{\mathbb{R}^n} \neq 1$, we let $\mathrm{v} = \dfrac{v}{\|v\|_{\mathbb{R}^n}}$. Then the direction derivatives of a function $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$ at $a \in \mathcal{U}$ in the direction v is

$$(D_\mathrm{v} f)(a) = \lim_{t \to 0} \frac{f(a + t\mathrm{v}) - f(a)}{t}\,.$$

Making a change of variable $s = \dfrac{t}{\|v\|_{\mathbb{R}^n}}$. Then

$$(Df)(x_0)(v) = \|v\|_{\mathbb{R}^n}(Df)(x_0)(\mathrm{v}) = \|v\|_{\mathbb{R}^n} \lim_{t \to 0} \frac{f(a + t\mathrm{v}) - f(a)}{t} = \lim_{s \to 0} \frac{f(a + sv) - f(a)}{s}\,.$$

We sometimes also call the value $(Df)(x_0)(v)$ the "directional derivative" of f in the "direction" v.

Example 2.77. The existence of directional derivatives of a function f at x_0 in all directions does not guarantee the differentiability of f at x_0. For example, let $f : \mathbb{R}^2 \to \mathbb{R}$ be given as in Example 2.44, and $\mathrm{v} = (\mathrm{v}_1, \mathrm{v}_2) \in \mathbb{R}^2$ be a unit vector. Then

$$(D_\mathrm{v} f)(0) = \lim_{t \to 0} \frac{f(t\mathrm{v}_1, t\mathrm{v}_2) - f(0,0)}{t} = \mathrm{v}_1^3\,.$$

However, f is not differentiable at $(0,0)$. We also note that in this example, $(D_\mathrm{v} f)(0) \neq (Jf)(0)\mathrm{v}$, where $(Jf)(0) = \begin{bmatrix} \dfrac{\partial f}{\partial x}(0,0) & \dfrac{\partial f}{\partial y}(0,0) \end{bmatrix}$ is the Jacobian matrix of f at $(0,0)$.

Example 2.78. The existence of directional derivatives of a function f at x_0 in all directions does not even guarantee the continuity of f at x_0. For example, let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,, \end{cases}$$

and $\mathrm{v} = (\mathrm{v}_1, \mathrm{v}_2) \in \mathbb{R}^2$ be a unit vector. Then if $\mathrm{v}_1 \neq 0$,

$$(D_\mathrm{v} f)(0) = \lim_{t \to 0} \frac{f(t\mathrm{v}_1, t\mathrm{v}_2) - f(0,0)}{t} = \lim_{t \to 0} \frac{t^3 \mathrm{v}_1 \mathrm{v}_2^2}{t(t^2\mathrm{v}_1^2 + t^4\mathrm{v}_2^4)} = \frac{\mathrm{v}_2^2}{\mathrm{v}_1}$$

while if $v_1 = 0$,

$$(D_{\mathrm{v}} f)(0) = \lim_{t\to 0} \frac{f(t\mathrm{v}_1, t\mathrm{v}_2) - f(0,0)}{t} = 0\,.$$

However, f is not continuous at $(0,0)$ since if (x, y) approaches $(0,0)$ along the curve $x = my^2$ with $m \neq 0$, we have

$$\lim_{y\to 0} f(my^2, y) = \lim_{y\to 0} \frac{my^4}{m^2y^4 + y^4} = \frac{m}{m^2+1}$$

which depends on m. Therefore, f is not continuous at $(0,0)$.

Example 2.79. Here comes another example showing that a function having directional derivative in all directions might not be continuous. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} \dfrac{xy}{x+y^2} & \text{if } x + y^2 \neq 0\,, \\ 0 & \text{if } x + y^2 = 0\,, \end{cases}$$

and $\mathrm{v} = (\mathrm{v}_1, \mathrm{v}_2) \in \mathbb{R}^2$ be a unit vector. Then if $\mathrm{v}_1 \neq 0$,

$$(D_{\mathrm{v}} f)(0) = \lim_{t\to 0} \frac{f(t\mathrm{v}_1, t\mathrm{v}_2) - f(0,0)}{t} = \lim_{t\to 0} \frac{t^2\mathrm{v}_1\mathrm{v}_2}{t(t\mathrm{v}_1 + t^2\mathrm{v}_2^2)} = \mathrm{v}_2$$

while if $\mathrm{v}_1 = 0$,

$$(D_{\mathrm{v}} f)(0) = \lim_{t\to 0} \frac{f(t\mathrm{v}_1, t\mathrm{v}_2) - f(0,0)}{t} = 0\,.$$

However, f is not continuous at $(0,0)$ since if (x, y) approaches $(0,0)$ along the polar curve

$$\theta(r) = \frac{\pi}{2} + \sin^{-1}(r - mr^2) \qquad 0 < r \ll 1\,,$$

we have

$$\begin{aligned}\lim_{\substack{(x,y)\to(0,0)\\ x=r\cos\theta(r),\, y=r\sin\theta(r)}} f(x,y) &= \lim_{r\to 0^+} \frac{r^2\cos\theta(r)\sin\theta(r)}{r^2\sin^2\theta(r) + r\cos\theta(r)} = \lim_{r\to 0^+} \frac{r(-r+mr^2)\sin\theta(r)}{r\sin^2\theta(r) - r + mr^2} \\ &= \lim_{r\to 0^+} \frac{(-r+mr^2)\sin\theta(r)}{\sin^2\theta(r) - 1 + mr} = \frac{-1}{m}\end{aligned}$$

which depends on m. Therefore, f is not continuous at $(0,0)$.

Definition 2.80. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set. The derivative of a scalar function $f : \mathcal{U} \to \mathbb{R}$ is called the ***gradient*** of f and is denoted by $\mathrm{grad} f$ or ∇f.

Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be an open set, $a \in \mathcal{U}$ and $f : \mathcal{U} \to \mathbb{R}$ be a real-valued function. Suppose that $f \in \mathscr{C}^1(\mathcal{U};\mathbb{R})$ and $(\nabla f)(a) \neq 0$. Then $\dfrac{\partial f}{\partial x_k}(a) \neq 0$ for some $1 \leqslant k \leqslant \mathrm{n}$. W.L.O.G., we can assume that $\dfrac{\partial f}{\partial x_{\mathrm{n}}}(a) \neq 0$. By the implicit function theorem, there exists an open neighborhood $\mathcal{V} \subseteq \mathbb{R}^{\mathrm{n}-1}$ of $(a_1, \cdots, a_{\mathrm{n}-1})$ and an open neighborhood $\mathcal{W} \subseteq \mathbb{R}$ of a_{n}, as well as a $\mathscr{C}^1$-function $\varphi : \mathcal{V} \to \mathbb{R}$ such that in a neighborhood of a the level set $\big\{x \in \mathcal{U} \,\big|\, f(x) = f(a)\big\}$ can be represented by $x_{\mathrm{n}} = \varphi(x_1, \cdots, x_{\mathrm{n}-1})$; that is,

$$f\big(x_1, \cdots, x_{\mathrm{n}-1}, \varphi(x_1, \cdots, x_{\mathrm{n}-1})\big) = f(a) \qquad \forall\, (x_1, \cdots, x_{\mathrm{n}-1}) \in \mathcal{V}\,.$$

Moreover,

$$\varphi_{x_j}(x_1, \cdots, x_{\mathrm{n}-1}) = -\frac{f_{x_j}\big(x_1, \cdots, x_{\mathrm{n}-1}, \varphi(x_1, \cdots, x_{\mathrm{n}-1})\big)}{f_{x_n}\big(x_1, \cdots, x_{\mathrm{n}-1}, \varphi(x_1, \cdots, x_{\mathrm{n}-1})\big)} \qquad \forall\, (x_1, \cdots, x_{\mathrm{n}-1}) \in \mathcal{V}\,.$$

Consider the collection of vectors $\{v_j\}_{j=1}^{\mathrm{n}-1}$ given by

$$v_j = \frac{\partial}{\partial x_j}\Big|_{x=a}\big(x_1, \cdots, x_{\mathrm{n}-1}, \varphi(x_1, \cdots, x_{\mathrm{n}-1})\big) \qquad (x_1, \cdots, x_{\mathrm{n}-1}) \in \mathcal{V}\,.$$

Then $v_j's$ are tangent vectors of the level surface. If $\{\mathrm{e}_j\}_{j=1}^{\mathrm{n}}$ is the standard basis of $\mathbb{R}^{\mathrm{n}}$, then

$$v_j = e_j + \big(0, \cdots, 0, \varphi_{x_j}(a_1, \cdots, a_{\mathrm{n}-1})\big) = e_j - \Big(0, \cdots, 0, \frac{f_{x_j}(a)}{f_{x_n}(a)}\Big)\,.$$

Therefore, the gradient vector $(\nabla f)(a)$ is perpendicular to v_j for all $1 \leqslant j \leqslant \mathrm{n}-1$ which conclude the following

Proposition 2.81. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open and $f \in \mathscr{C}^1(\mathcal{U};\mathbb{R})$; that is, $f : \mathcal{U} \to \mathbb{R}$ is continuously differentiable. Then if $(\nabla f)(x_0) \neq 0$, the vector $\dfrac{(\nabla f)(x_0)}{\|(\nabla f)(x_0)\|_{\mathbb{R}^n}}$ is the unit normal to the level set $\big\{x \in \mathcal{U} \,\big|\, f(x) = f(x_0)\big\}$ at x_0.*

Example 2.82. Find the normal to $\mathcal{S} = \big\{(x, y, z) \,\big|\, x^2 + y^2 + z^2 = 3\big\}$ at $(1, 1, 1) \in \mathcal{S}$.
Solution: Take $f(x, y, z) = x^2 + y^2 + z^2 - 3$. Then $(\nabla f)(x, y, z) = (2x, 2y, 2z)$; thus $(\nabla f)(1, 1, 1) = (2, 2, 2)$ is normal to S at $(1, 1, 1)$.

Example 2.83. Consider the surface

$$\mathcal{S} = \big\{(x, y, z) \in \mathbb{R}^3 \,\big|\, x^2 - y^2 + xyz = 1\big\}.$$

Find the tangent plane of $\mathcal{S}$ at $(1, 0, 1)$.

Solution: Let $f(x,y,z) = x^2 - y^2 + xyz$. Then

$$\mathcal{S} = \big\{(x,y,z) \in \mathbb{R}^3 \,|\, f(x,y,z) = f(1,0,1)\big\}\,;$$

that is, $\mathcal{S}$ is a level set of f. Since $(\nabla f)(1,0,1) = (2,1,0) \neq (0,0,0)$, $(2,1,0)$ is normal to $\mathcal{S}$ at $(1,0,1)$; thus the tangent plane of S at $(1,0,1)$ is $2(x-1)+y=0$. ▫

Proposition 2.84. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. If $(\nabla f)(x_0) \neq 0$, then $\pm\dfrac{(\nabla f)(x_0)}{\|(\nabla f)(x_0)\|_{\mathbb{R}^n}}$ is the direction in which the function f increases/decreases most rapidly*（最速上升／下降方向）*at x_0.*

Proof. Let $x_0 \in \mathbb{R}^n$ be given. Suppose that f increases most rapidly in the direction v, then $(D_{\rm v} f)(x_0) = \sup\limits_{\|w\|_{\mathbb{R}^n}=1} (D_w f)(x_0)$. Since f is differentiable, $(D_w f)(x_0) = (Df)(x_0)(w) = (\nabla f)(x_0)\cdot w$ which is maximized in the direction $\dfrac{(\nabla f)(x_0)}{\|(\nabla f)(x_0)\|_{\mathbb{R}^n}}$. ▫

Example 2.85. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by $f(x,y,z) = x^2 y \sin z$. Find the direction of the greatest rate of change at $(3,2,0)$.
Solution: We compute the gradient of f at $(3,2,0)$ as follows:

$$\begin{aligned}(\nabla f)(3,2,0) &= \Big(\frac{\partial f}{\partial x}(3,2,0), \frac{\partial f}{\partial y}(3,2,0), \frac{\partial f}{\partial z}(3,2,0)\Big)\\ &= (2xy\sin z, x^2\sin z, x^2 y\cos z)\big|_{(x,y,z)=(3,,2,0)} = (0,0,18).\end{aligned}$$

Therefore, the direction of the greatest rate of change of f at $(3,2,0)$ is $(0,0,1)$.

Chapter 3

Multiple Integrals

3.1 Integrable Functions

Let us start our discussion on the integrability of functions of two variables.

Definition 3.1. Let $A \subseteq \mathbb{R}^2$ be a bounded set. Define

$$\begin{aligned}
a_1 &= \inf\left\{x \in \mathbb{R} \,\middle|\, (x, y) \in A \text{ for some } y \in \mathbb{R}\right\}, \\
b_1 &= \sup\left\{x \in \mathbb{R} \,\middle|\, (x, y) \in A \text{ for some } y \in \mathbb{R}\right\}, \\
a_2 &= \inf\left\{y \in \mathbb{R} \,\middle|\, (x, y) \in A \text{ for some } x \in \mathbb{R}\right\}, \\
b_2 &= \sup\left\{y \in \mathbb{R} \,\middle|\, (x, y) \in A \text{ for some } x \in \mathbb{R}\right\}.
\end{aligned}$$

A collection of rectangles $\mathcal{P}$ is called a ***partition*** of A if there exists a partition $\mathcal{P}_x$ of $[a_1, b_1]$ and a partition P_y of $[a_2, b_2]$,

$$\mathcal{P}_x = \left\{a_1 = x_0 < x_1 < \cdots < x_{\mathrm{n}} = b_1\right\} \quad \text{and} \quad \mathcal{P}_y = \left\{a_2 = y_0 < y_1 < \cdots < y_m = b_2\right\},$$

such that

$$\mathcal{P} = \left\{\Delta_{ij} \,\middle|\, \Delta_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}] \text{ for } i = 0, 1, \cdots, n-1 \text{ and } j = 0, 1, \cdots, m-1\right\}.$$

The ***mesh size*** of the partition $\mathcal{P}$, denoted by $\|\mathcal{P}\|$ and also called the norm of $\mathcal{P}$, is defined by

$$\|\mathcal{P}\| = \max\left\{\sqrt{(x_{i+1} - x_i)^2 + (y_{j+1} - y_j)^2} \,\middle|\, i = 0, 1, \cdots, n-1, j = 0, 1, \cdots, m-1\right\}.$$

The number $\sqrt{(x_{i+1} - x_i)^2 + (y_{j+1} - y_j)^2}$ is often denoted by $\operatorname{diam}(\Delta_{ij})$, and is called the ***diameter*** of Δ_{ij}.

Definition 3.2. Let $A \subseteq \mathbb{R}^2$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. For any partition $\mathcal{P} = \{\Delta_{ij} \,|\, \Delta_{ij} = (x_i, x_{i+1}) \times (y_j, y_{j+1}), i = 0, \cdots, n-1, j = 0, \cdots, m-1\}$, the ***upper sum*** and the ***lower sum*** of f with respect to the partition $\mathcal{P}$, denoted by $U(f, \mathcal{P})$ and $L(f, \mathcal{P})$ respectively, are numbers defined by

$$U(f, \mathcal{P}) = \sum_{\substack{0 \leqslant i \leqslant n-1 \\ 0 \leqslant j \leqslant m-1}} \sup_{(x,y) \in \Delta_{ij}} \overline{f}^A(x, y) \mathbb{A}(\Delta_{ij}) \,,$$
$$L(f, \mathcal{P}) = \sum_{\substack{0 \leqslant i \leqslant n-1 \\ 0 \leqslant j \leqslant m-1}} \inf_{(x,y) \in \Delta_{ij}} \overline{f}^A(x, y) \mathbb{A}(\Delta_{ij}) \,,$$

where $\mathbb{A}(\Delta_{ij}) = (x_{i+1} - x_i)(y_{j+1} - y_j)$ is the area of the rectangle Δ_{ij}, and $\overline{f}^A$ is an extension of f, called the extension of f by zero outside A, given by

$$\overline{f}^A(x) = \begin{cases} f(x) & x \in A \,, \\ 0 & x \notin A \,. \end{cases}$$

The two numbers

$$\overline{\int_A} f(x, y) \, d\mathbb{A} \equiv \inf \{ U(f, \mathcal{P}) \,|\, \mathcal{P} \text{ is a partition of } A \}$$

and

$$\underline{\int_A} f(x, y) \, d\mathbb{A} \equiv \sup \{ L(f, \mathcal{P}) \,|\, \mathcal{P} \text{ is a partition of } A \}$$

are called the ***upper integral*** and ***lower integral*** of f over A, respectively. The function f is said to be ***Riemann*** (***Darboux***) ***integrable*** (over A) if $\overline{\int_A} f(x, y) d\mathbb{A} = \underline{\int_A} f(x, y) d\mathbb{A}$, and in this case, we express the upper and lower integral as $\int_A f(x, y) d\mathbb{A}$, called the ***double integral*** of f over A.

Similar to the case of double integrals, we can consider the integrability of a bounded function f defined on a bounded set $A \subseteq \mathbb{R}^{\mathrm{n}}$ as follows

Definition 3.3. Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded set. Define the numbers $a_1, a_2, \cdots, a_n$ and $b_1, b_2, \cdots, b_n$ by

$$a_k = \inf \{ x_k \in \mathbb{R} \,|\, x = (x_1, \cdots, x_{\mathrm{n}}) \in A \text{ for some } x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_{\mathrm{n}} \in \mathbb{R} \},$$
$$b_k = \sup \{ x_k \in \mathbb{R} \,|\, x = (x_1, \cdots, x_{\mathrm{n}}) \in A \text{ for some } x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_{\mathrm{n}} \in \mathbb{R} \}.$$

A collection of rectangles $\mathcal{P}$ is called a ***partition*** of A if there exists partitions $\mathcal{P}^{(k)}$ of $[a_k, b_k]$, $k = 1, \cdots, n$, $\mathcal{P}^{(k)} = \big\{a_k = x_0^{(k)} < x_1^{(k)} < \cdots < x_{N_k}^{(k)} = b_k\big\}$, such that

$$\mathcal{P} = \Big\{\Delta_{i_1 i_2 \cdots i_n} \,\Big|\, \Delta_{i_1 i_2 \cdots i_n} = \big[x_{i_1}^{(1)}, x_{i_1+1}^{(1)}\big] \times \big[x_{i_2}^{(2)}, x_{i_2+1}^{(2)}\big] \times \cdots \times \big[x_{i_n}^{(\mathrm{n})}, x_{i_{n+1}}^{(\mathrm{n+1})}\big], \\ i_k = 0, 1, \cdots, N_k - 1, k = 1, \cdots, n\Big\}.$$

The ***mesh size*** of the partition $\mathcal{P}$, denoted by $\|\mathcal{P}\|$, is defined by

$$\|\mathcal{P}\| = \max\Big\{\sqrt{\sum_{k=1}^{\mathrm{n}} (x_{i_k+1}^{(k)} - x_{i_k}^{(k)})^2} \,\Big|\, i_k = 0, 1, \cdots, N_k - 1, k = 1, \cdots, n\Big\}.$$

The number $\sqrt{\sum\limits_{k=1}^{\mathrm{n}} (x_{i_k+1}^{(k)} - x_{i_k}^{(k)})^2}$ is often denoted by $\mathrm{diam}(\Delta_{i_1 i_2 \cdots i_n})$, and is called the ***diameter*** of the rectangle $\Delta_{i_1 i_2 \cdots i_n}$.

Definition 3.4. Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. For any partition

$$\mathcal{P} = \Big\{\Delta_{i_1 i_2 \cdots i_n} \,\Big|\, \Delta_{i_1 i_2 \cdots i_n} = \big[x_{i_1}^{(1)}, x_{i_1+1}^{(1)}\big] \times \big[x_{i_2}^{(2)}, x_{i_2+1}^{(2)}\big] \times \cdots \times \big[x_{i_n}^{(\mathrm{n})}, x_{i_{n+1}}^{(\mathrm{n+1})}\big], \\ i_k = 0, 1, \cdots, N_k - 1, k = 1, \cdots, n\Big\},$$

the ***upper sum*** and the ***lower sum*** of f with respect to the partition $\mathcal{P}$, denoted by $U(f, \mathcal{P})$ and $L(f, \mathcal{P})$ respectively, are numbers defined by

$$U(f, \mathcal{P}) = \sum_{\Delta \in \mathcal{P}} \sup_{x \in \Delta} \overline{f}^A(x) \nu_{\mathrm{n}}(\Delta)\,,$$
$$L(f, \mathcal{P}) = \sum_{\Delta \in \mathcal{P}} \inf_{x \in \Delta} \overline{f}^A(x) \nu_{\mathrm{n}}(\Delta)\,,$$

where $\nu_{\mathrm{n}}(\Delta)$ is the n-dimensional ***volume*** of the rectangle Δ given by

$$\nu_{\mathrm{n}}(\Delta) = (x_{i_1+1}^{(1)} - x_{i_1}^{(1)})(x_{i_2+1}^{(2)} - x_{i_2}^{(2)}) \cdots (x_{i_n+1}^{(\mathrm{n})} - x_{i_n}^{(\mathrm{n})})$$

if $\Delta = \big[x_{i_1}^{(1)} - x_{i_1+1}^{(1)}\big] \times \big[x_{i_2}^{(2)} - x_{i_2+1}^{(2)}\big] \times \cdots \times \big[x_{i_n}^{(\mathrm{n})} - x_{i_n+1}^{(\mathrm{n})}\big]$, and $\overline{f}^A$ is the extension of f by zero outside A given by

$$\overline{f}^A(x) = \begin{cases} f(x) & x \in A\,, \\ 0 & x \notin A\,. \end{cases} \tag{3.1}$$

The two numbers

$$\overline{\int_A} f(x)dx \equiv \inf\big\{U(f,\mathcal{P})\,\big|\,\mathcal{P} \text{ is a partition of } A\big\}\,,$$

and

$$\underline{\int_A} f(x)dx \equiv \sup\big\{L(f,\mathcal{P})\,\big|\,\mathcal{P} \text{ is a partition of } A\big\}$$

are called the ***upper integral*** and ***lower integral*** of f over A, respective. The function f is said to be ***Riemann*** (***Darboux***) ***integrable*** (over A) if $\overline{\int_A} f(x)dx = \underline{\int_A} f(x)dx$, and in this case, we express the upper and lower integral as $\int_A f(x)dx$, called the ***n-tuple integral*** of f over A.

Definition 3.5. A partition $\mathcal{P}'$ of a bounded set $A \subseteq \mathbb{R}^n$ is said to be a ***refinement*** of another partition $\mathcal{P}$ of A if for any $\Delta' \in \mathcal{P}'$, there is $\Delta \in \mathcal{P}$ such that $\Delta' \subseteq \Delta$. A partition $\mathcal{P}$ of a bounded set $A \subseteq \mathbb{R}^n$ is said to be the ***common refinement*** of another partitions $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ of A if

1. $\mathcal{P}$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$.

2. If $\mathcal{P}'$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$, then $\mathcal{P}'$ is also a refinement of $\mathcal{P}$.

In other words, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if it is the coarsest refinement.

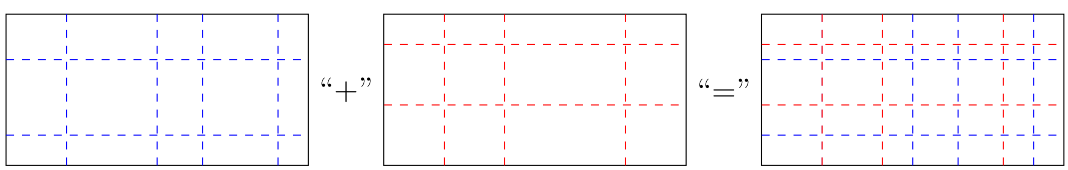


Figure 3.1: The common refinement of two partitions

Qualitatively speaking, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if for each $j = 1, \cdots n$, the j-th component c_j of the vertex $(c_1, \cdots, c_n)$ of each rectangle $\Delta \in \mathcal{P}$ belongs to $\mathcal{P}_i^{(j)}$ for some $i = 1, \cdots, k$.

Proposition 3.6. *Let $A \subseteq \mathbb{R}^n$ be a bounded subset, and $f : A \to \mathbb{R}$ be a bounded function. If $\mathcal{P}$ and $\mathcal{P}'$ are partitions of A and $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then*

$$L(f,\mathcal{P}) \leqslant L(f,\mathcal{P}') \leqslant U(f,\mathcal{P}') \leqslant U(f,\mathcal{P})\,.$$

Corollary 3.7. *Let $A \subseteq \mathbb{R}^n$ be a bounded subset, and $f : A \to \mathbb{R}$ be a bounded function. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are partitions of A, then*

$$L(f, \mathcal{P}_1) \leqslant U(f, \mathcal{P}_2)\,.$$

Proof. Let $\mathcal{P}$ be the common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then Proposition 3.6 implies that

$$L(f, \mathcal{P}_1) \leqslant L(f, \mathcal{P}) \leqslant U(f, \mathcal{P}) \leqslant U(f, \mathcal{P}_2)\,. \qquad \square$$

Corollary 3.8. *Let $A \subseteq \mathbb{R}^n$ be a bounded subset, and $f : A \to \mathbb{R}$ be a bounded function. Then*

$$\underline{\int_A} f(x)dx \leqslant \overline{\int_A} f(x)dx\,.$$

Proof. Noting that for each given partition $\mathcal{P}$ of A, $L(f, \mathcal{P})$ is a lower bounded for all possible upper sum; thus

$$L(f, \mathcal{P}) \leqslant \overline{\int_A} f(x)dx \qquad \text{for all partitions } \mathcal{P} \text{ of } A$$

which further implies that $\underline{\int_A} f(x)dx \leqslant \overline{\int_A} f(x)dx\,.$ $\square$

Proposition 3.9 (Riemann's condition)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over A if and only if*

$$\forall\, \varepsilon > 0, \exists \text{ a partition } \mathcal{P} \text{ of } A \ni U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon\,.$$

Proof. "$\Rightarrow$" Let $\varepsilon > 0$ be given. By the definition of infimum and supremum, there exist partition $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int_A} f(x)\,dx - \frac{\varepsilon}{2} < L(f, \mathcal{P}_2) \quad \text{and} \quad \overline{\int_A} f(x)\,dx + \frac{\varepsilon}{2} > U(f, \mathcal{P}_1)\,.$$

Let $\mathcal{P}$ be a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Since f is Riemann integrable over A, $\underline{\int_A} f(x)dx = \overline{\int_A} f(x)dx$; thus Proposition 3.6 implies that

$$\begin{aligned} U(f, \mathcal{P}) - L(f, \mathcal{P}) &\leqslant U(f, \mathcal{P}_1) - L(f, \mathcal{P}_2) \\ &< \overline{\int_A} f(x)\,dx + \frac{\varepsilon}{2} - \Big(\underline{\int_A} f(x)\,dx - \frac{\varepsilon}{2}\Big) = \varepsilon\,. \end{aligned}$$

"$\Leftarrow$" Let $\varepsilon > 0$ be given. By assumption there exists a partition $\mathcal{P}$ of A such that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$. Then

$$0 \leqslant \overline{\int_A} f(x)\, dx - \underline{\int_A} f(x)\, dx \leqslant U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon\,.$$

Since $\varepsilon > 0$ is given arbitrary, we must have $\underline{\int_A} f(x)dx = \overline{\int_A} f(x)dx$; thus f is Riemann integrable over A. $\square$

Definition 3.10. Let $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ be a partition of a bounded set $A \subseteq \mathbb{R}^n$. A collection of N points $\{\xi_1, \cdots, \xi_N\}$ is called a ***sample set*** for the partition $\mathcal{P}$ if $\xi_k \in \Delta_k$ for all $k = 1, \cdots, N$. Points in a sample set are called sample points for the partition $\mathcal{P}$.

Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. A ***Riemann sum*** of f for the the partition $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ of A is a sum which takes the form

$$\sum_{k=1}^{N} \overline{f}^A(\xi_i)\nu_n(\Delta_k)\,,$$

where the set $\Xi = \{\xi_1, \xi_2, \cdots, \xi_N\}$ is a sample set for the partition $\mathcal{P}$.

Theorem 3.11 (Darboux). *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function with extension $\overline{f}^A$ given by* (3.1)*. Then f is Riemann integrable over A if and only if there exists $\mathrm{I} \in \mathbb{R}$ such that for every given $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums for the partition $\mathcal{P}$ belongs to the interval $(\mathrm{I} - \varepsilon, \mathrm{I} + \varepsilon)$. In other words, f is Riemann integrable over A if and only if there exists $\mathrm{I} \in \mathbb{R}$ such that for every given $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\Big| \sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k) - \mathrm{I} \Big| < \varepsilon \tag{3.2}$$

whenever $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$ and $\{\xi_1, \xi_2, \cdots, \xi_N\}$ is a sample set for $\mathcal{P}$.

Proof. The boundedness of A guarantees that $A \subseteq \big[-\frac{r}{2}, \frac{r}{2}\big]^n$ for some $r > 0$. Let $R = \big[-\frac{r}{2}, \frac{r}{2}\big]^n$.

"$\Leftarrow$" Suppose the right-hand side statement is true. Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that if $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$, then for all sets of sample points $\{\xi_1, \cdots, \xi_N\}$ for $\mathcal{P}$, we must have

$$\Big|\sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k) - \mathrm{I}\Big| < \frac{\varepsilon}{4}\,.$$

Let $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ be a partition of A with $\|\mathcal{P}\| < \delta$. Choose two sample sets $\{\xi_1, \cdots, \xi_N\}$ and $\{\eta_1, \cdots, \eta_N\}$ for $\mathcal{P}$ such that

(a) $\displaystyle\sup_{x\in\Delta_k} \overline{f}^A(x) - \frac{\varepsilon}{4\nu(R)} < \overline{f}^A(\xi_k) \leqslant \sup_{x\in\Delta_k} \overline{f}^A(x)$;

(b) $\displaystyle\inf_{x\in\Delta_k} \overline{f}^A(x) + \frac{\varepsilon}{4\nu(R)} > \overline{f}^A(\eta_k) \geqslant \inf_{x\in\Delta_k} \overline{f}^A(x)$.

Then

$$\begin{aligned} U(f,\mathcal{P}) &= \sum_{k=1}^{N} \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) < \sum_{k=1}^{N} \big[\overline{f}^A(\xi_k) + \frac{\varepsilon}{4\nu(R)}\big]\nu(\Delta_k) \\ &= \sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k) + \frac{\varepsilon}{4\nu(R)}\sum_{k=1}^{N}\nu(\Delta_k) < \mathrm{I} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \mathrm{I} + \frac{\varepsilon}{2} \end{aligned}$$

and

$$\begin{aligned} L(f,\mathcal{P}) &= \sum_{k=1}^{N} \inf_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) > \sum_{k=1}^{N} \big[\overline{f}^A(\eta_k) - \frac{\varepsilon}{4\nu(R)}\big]\nu(\Delta_k) \\ &= \sum_{k=1}^{N} \overline{f}^A(\eta_k)\nu(\Delta_k) - \frac{\varepsilon}{4\nu(R)}\sum_{k=1}^{N}\nu(\Delta_k) > \mathrm{I} - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} = \mathrm{I} - \frac{\varepsilon}{2}\,. \end{aligned}$$

As a consequence, $\mathrm{I} - \frac{\varepsilon}{2} < L(f,\mathcal{P}) \leqslant U(f,\mathcal{P}) < \mathrm{I} + \frac{\varepsilon}{2}$; thus $U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon$.

"$\Rightarrow$" Let $\mathrm{I} = \displaystyle\int_A f(x)dx$, and $\varepsilon > 0$ be given. Since f is Riemann integrable over A, there exists a partition $\mathcal{P}_1$ of A such that $U(f,\mathcal{P}_1) - L(f,\mathcal{P}_1) < \frac{\varepsilon}{2}$. Suppose that $\mathcal{P}_1^{(i)} = \big\{y_0^{(i)}, y_1^{(i)}, \cdots, y_{m_i}^{(i)}\big\}$ for $1 \leqslant i \leqslant n$. With M denoting the number $m_1 + m_2 + \cdots + m_n$, we define

$$\delta = \frac{\varepsilon}{4r^{\mathrm{n}-1}(M+\mathrm{n})\big(\sup \overline{f}^A(R) - \inf \overline{f}^A(R) + 1\big)}\,.$$

Then $\delta > 0$. Our goal is to show that if $\mathcal{P}$ is a partition of A with $\|\mathcal{P}\| < \delta$ and $\{\xi_1, \cdots, \xi_N\}$ is a set of sample points for $\mathcal{P}$, then (3.2) holds.

Assume that $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ is a given partition of A with $\|\mathcal{P}\| < \delta$. Let $\mathcal{P}'$ be the common refinement of $\mathcal{P}$ and $\mathcal{P}_1$. Write $\mathcal{P}' = \{\Delta_1', \Delta_2', \cdots, \Delta_{N'}'\}$ and

$\Delta_k = \Delta_k^{(1)} \times \Delta_k^{(2)} \times \cdots \times \Delta_k^{(\mathrm{n})}$ as well as $\Delta'_k = \Delta_k'^{(1)} \times \Delta_k'^{(2)} \times \cdots \times \Delta_k'^{(\mathrm{n})}$. By the definition of the upper sum,

$$\begin{aligned} U(f,\mathcal{P}) &= \sum_{k=1}^{N} \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) \\ &= \sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\notin\Delta_k^{(i)} \text{for all } i,j}} \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) + \sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\in\Delta_k^{(i)} \text{for some } i,j}} \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) \end{aligned}$$

and similarly,

$$U(f,\mathcal{P}') = \sum_{\substack{1\leqslant k\leqslant N' \text{ with} \\ y_j^{(i)}\notin\Delta_k'^{(i)} \text{for all } i,j}} \sup_{x\in\Delta'_k} \overline{f}^A(x)\nu(\Delta'_k) + \sum_{\substack{1\leqslant k\leqslant N' \text{ with} \\ y_j^{(i)}\in\Delta_k'^{(i)} \text{for some } i,j}} \sup_{x\in\Delta'_k} \overline{f}^A(x)\nu(\Delta'_k)\,.$$

By the fact that $\Delta_k \in \mathcal{P}'$ if $y_j^{(i)} \notin \Delta_k'^{(i)}$ for all i, j, we must have

$$\sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\in\Delta_k^{(i)} \text{for some } i,j}} \nu(\Delta_k) = \sum_{\substack{1\leqslant k\leqslant N' \text{ with} \\ y_j^{(i)}\in\Delta_k'^{(i)} \text{for some } i,j}} \nu(\Delta'_k)\,.$$

The equality above further implies that

$$\begin{aligned} U(f,\mathcal{P})-U(f,\mathcal{P}') &= \sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\in\Delta_k^{(i)} \text{for some } i,j}} \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) - \sum_{\substack{1\leqslant k\leqslant N' \text{ with} \\ y_j^{(i)}\in\Delta_k'^{(i)} \text{for some } i,j}} \sup_{x\in\Delta'_k} \overline{f}^A(x)\nu(\Delta'_k) \\ &\leqslant \big(\sup \overline{f}^A(R) - \inf \overline{f}^A(R)\big) \sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\in\Delta_k^{(i)} \text{for some } i,j}} \nu(\Delta_k)\,. \end{aligned}$$

Moreover, for each fixed i, j,

$$\bigcup_{\substack{1\leqslant k\leqslant N \\ y_j^{(i)}\in\Delta_k^{(i)}}} \Delta_k \subseteq \big[-\frac{r}{2}, \frac{r}{2}\big]^{i-1} \times \big[y_j^{(i)} - \delta, y_j^{(i)} + \delta\big] \times \big[-\frac{r}{2}, \frac{r}{2}\big]^{\mathrm{n}-i};$$

thus

$$\sum_{\substack{1\leqslant k\leqslant N \text{ with} \\ y_j^{(i)}\in\Delta_k^{(i)}}} \nu(\Delta_k) \leqslant 2\delta r^{\mathrm{n}-1} \qquad \forall\, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m_i\,.$$

Therefore,

$$\begin{aligned}
&U(f,\mathcal{P})-U(f,\mathcal{P}')\\
&\qquad\leqslant \big(\sup \overline{f}^{A}(R)-\inf \overline{f}^{A}(R)\big)\sum_{i=1}^{\mathrm{n}}\sum_{j=0}^{m_i}\sum_{\substack{1\leqslant k\leqslant N \text{ with}\\ y_j^{(i)}\in\Delta_k^{(i)}}}\nu(\Delta_k)\\
&\qquad\leqslant \big(\sup \overline{f}^{A}(R)-\inf \overline{f}^{A}(R)\big)\sum_{i=1}^{\mathrm{n}}\sum_{j=0}^{m_i}2\delta r^{\mathrm{n}-1}\\
&\qquad\leqslant 2\delta r^{\mathrm{n}-1}(m_1+m_2+\cdots+m_n+n)\big(\sup \overline{f}^{A}(R)-\inf \overline{f}^{R}(A)\big)<\frac{\varepsilon}{2}\,,
\end{aligned}$$

and the fact that $U(f,\mathcal{P}_1)-L(f,\mathcal{P}_1)<\dfrac{\varepsilon}{2}$ shows that

$$\begin{aligned}
U(f,\mathcal{P})-\mathrm{I}&\leqslant U(f,\mathcal{P})-\mathrm{I}+U(f,\mathcal{P}_1)-U(f,\mathcal{P}_1)\\
&\leqslant U(f,\mathcal{P})-L(f,\mathcal{P}_1)+U(f,\mathcal{P}_1)-U(f,\mathcal{P}')<\varepsilon\,.
\end{aligned}$$

Therefore, for any sample set $\{\xi_1,\cdots,\xi_N\}$ for $\mathcal{P}$,

$$\sum_{k=1}^{N}\overline{f}^{A}(\xi_k)\nu(\Delta_k)\leqslant U(f,\mathcal{P})<\mathrm{I}+\varepsilon\,.$$

Similar argument can be used to show that

$$\sum_{k=1}^{N}\overline{f}^{A}(\xi_k)\nu(\Delta_k)\geqslant L(f,\mathcal{P})>\mathrm{I}-\varepsilon$$

which concludes the Theorem. □

Definition 3.12. A bounded set $A\subseteq\mathbb{R}^{\mathrm{n}}$ is said to ***have volume*** if the characteristic function of A, denoted by $\mathbf{1}_A$ and given by

$$\mathbf{1}_A(x)=\begin{cases}1 & \text{if } x\in A\,,\\ 0 & \text{otherwise}\,,\end{cases}$$

is Riemann integrable over A, and the number $\displaystyle\int_A\mathbf{1}_A(x)\,dx$ is called the ***volume*** of A and is denoted by $\nu(A)$. If $\nu(A)=0$, then A is said to have volume zero.

Remark 3.13. Having defined the indicator function, then for a bounded function $f:A\to\mathbb{R}$ with bounded domain A, any given partition $\mathcal{P}$ of A we have $\overline{f}^{A}=f\mathbf{1}_A$; thus

$$U(f,\mathcal{P})=\sum_{\Delta\in\mathcal{P}}\sup_{x\in\Delta}(f\mathbf{1}_A)(x)\nu(\Delta)\qquad\text{and}\qquad L(f,\mathcal{P})=\sum_{\Delta\in\mathcal{P}}\inf_{x\in\Delta}(f\mathbf{1}_A)(x)\nu(\Delta)\,.$$

3.2 Properties of the Integrals

Proposition 3.14. *Let $A \subseteq \mathbb{R}^n$ be bounded, and $f, g : A \to \mathbb{R}$ be bounded. Then*

(a) *If $B \subseteq A$, then $\underline{\int}_A (f\mathbf{1}_B)(x)\,dx = \underline{\int}_B f(x)\,dx$ and $\overline{\int}_A (f\mathbf{1}_B)(x)\,dx = \overline{\int}_B f(x)\,dx$.*

(b) $\underline{\int}_A f(x)\,dx + \underline{\int}_A g(x)\,dx \leqslant \underline{\int}_A (f+g)(x)\,dx \leqslant \overline{\int}_A (f+g)(x)\,dx \leqslant \overline{\int}_A f(x)\,dx + \overline{\int}_A g(x)\,dx$.

(c) *If $c \geqslant 0$, then $\underline{\int}_A (cf)(x)\,dx = c\underline{\int}_A f(x)\,dx$ and $\overline{\int}_A (cf)(x)\,dx = c\overline{\int}_A f(x)\,dx$. If $c < 0$, then $\underline{\int}_A (cf)(x)\,dx = c\overline{\int}_A f(x)\,dx$ and $\overline{\int}_A (cf)(x)\,dx = c\underline{\int}_A f(x)\,dx$.*

(d) *If $f \leqslant g$ on A, then $\underline{\int}_A f(x)\,dx \leqslant \underline{\int}_A g(x)\,dx$ and $\overline{\int}_A f(x)\,dx \leqslant \overline{\int}_A g(x)\,dx$.*

(e) *If A has volume zero, then f is Riemann integrable over A, and $\int_A f(x)\,dx = 0$.*

Proof. We only prove (a), (b), (c) and (e) since (d) are trivial.

(a) Let $\varepsilon > 0$ be given. By the definition of the lower integral, there exist partition $\mathcal{P}_A$ of A and $\mathcal{P}_B$ of B such that

$$\underline{\int}_A (f\mathbf{1}_B)(x)\,dx - \varepsilon < L(f\mathbf{1}_B, \mathcal{P}_A) = \sum_{\Delta \in \mathcal{P}_A} \inf_{x\in\Delta} \overline{f\mathbf{1}_B}^A(x)\nu(\Delta)$$

and

$$\underline{\int}_B f(x)\,dx - \frac{\varepsilon}{2} < L(f, \mathcal{P}_B) = \sum_{\Delta \in \mathcal{P}_B} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta)\,.$$

Let $\mathcal{P}'_A$ be a refinement of $\mathcal{P}_A$ such that some collection of rectangles in $\mathcal{P}'_A$ forms a partition of B. Denote this partition of B by $\mathcal{P}'_B$. Since $\inf\limits_{x\in\Delta} \overline{f}^B(x) \leqslant 0$ if $\Delta \in \mathcal{P}'_A \backslash \mathcal{P}'_B$, Proposition 3.6 implies that

$$\begin{aligned}
\underline{\int}_A (f\mathbf{1}_B)(x)\,dx - \varepsilon < L(f\mathbf{1}_B, \mathcal{P}_A) \leqslant L(f\mathbf{1}_B, \mathcal{P}'_A) &= \sum_{\Delta \in \mathcal{P}'_A} \inf_{x\in\Delta} \overline{f\mathbf{1}_B}^A(x)\nu(\Delta) \\
&= \Big(\sum_{\Delta \in \mathcal{P}'_A \backslash \mathcal{P}'_B} + \sum_{\Delta \in \mathcal{P}'_B}\Big) \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) \\
&\leqslant \sum_{\Delta \in \mathcal{P}'_B} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) = L(f, \mathcal{P}'_B) \leqslant \underline{\int}_B f(x)\,dx\,.
\end{aligned}$$

On the other hand, let $\widetilde{\mathcal{P}}_A$ be a partition of A such that $\mathcal{P}_B \subseteq \widetilde{\mathcal{P}}_A$ and

$$\sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\,\Delta\cap B\neq\varnothing} \nu(\Delta) \leqslant \frac{\varepsilon}{2(M+1)}\,,$$

where $M > 0$ is an upper bound of $|f|$. Then

$$\sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\,\Delta\cap B\neq\varnothing} \inf_{x\in\Delta} \overline{f}^{B}(x)\nu(\Delta) \geqslant -M \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\,\Delta\cap B\neq\varnothing} \nu(\Delta) \geqslant -\frac{\varepsilon}{2}$$

which further implies that

$$\begin{aligned}
\underline{\int_A} (f\mathbf{1}_B)(x)\,dx \geqslant L(f\mathbf{1}_B, \widetilde{\mathcal{P}}_A) &= \sum_{\Delta\in\widetilde{\mathcal{P}}_A} \inf_{x\in\Delta} \overline{f\mathbf{1}_B}^{A}(x)\nu(\Delta) \\
&= \Big(\sum_{\Delta\in\mathcal{P}_B} + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\,\Delta\cap B\neq\varnothing} + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\,\Delta\cap B=\varnothing}\Big) \inf_{x\in\Delta} \overline{f}^{B}(x)\nu(\Delta) \\
&= L(f, \mathcal{P}_B) + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\,\Delta\cap B\neq\varnothing} \inf_{x\in\Delta} \overline{f}^{B}(x)\nu(\Delta) > \underline{\int_B} f(x)\,dx - \varepsilon\,.
\end{aligned}$$

Therefore, we establish that

$$\underline{\int_B} f(x)\,dx - \varepsilon < \underline{\int_A} (f\mathbf{1}_B)(x)\,dx < \underline{\int_B} f(x)\,dx + \varepsilon\,.$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that $\underline{\int_A} (f\mathbf{1}_B)(x)\,dx = \underline{\int_B} f(x)\,dx$. Similar argument can be applied to conclude that $\overline{\int_A} (f\mathbf{1}_B)(x)\,dx = \overline{\int_B} f(x)\,dx$.

(b) Let $\varepsilon > 0$ be given. By the definition of the lower integral, there exist partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int_A} f(x)\,dx - \frac{\varepsilon}{2} < L(f, \mathcal{P}_1) \quad\text{and}\quad \underline{\int_A} g(x)\,dx - \frac{\varepsilon}{2} < L(g, \mathcal{P}_2)\,.$$

Let $\mathcal{P}$ be a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then

$$\begin{aligned}
\underline{\int_A} f(x)\,dx + \underline{\int_A} g(x)\,dx - \varepsilon < L(f, \mathcal{P}_1) + L(f, \mathcal{P}_2) &\leqslant L(f, \mathcal{P}) + L(g, \mathcal{P}) \\
&= \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} \overline{f}(x)\nu(\Delta) + \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} \overline{g}(x)\nu(\Delta) \\
&\leqslant \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} (\overline{f} + \overline{g})(x)\nu(\Delta) = L(f+g, \mathcal{P}) \leqslant \underline{\int_A} (f+g)(x)\,dx\,.
\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\underline{\int_A} f(x)\,dx + \underline{\int_A} g(x)\,dx \leqslant \underline{\int_A} (f+g)(x)\,dx\,.$$

Similarly, we have $\overline{\int_A} (f+g)(x)\,dx \leqslant \overline{\int_A} f(x)\,dx + \overline{\int_A} g(x)\,dx$; thus (b) is established.

(c) It suffices to show the case $c = -1$. Let $\varepsilon > 0$ be given. Then there exist partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int_A} -f(x)\,dx - \varepsilon < L(-f, \mathcal{P}_1) \qquad \text{and} \qquad U(f, \mathcal{P}_2) < \overline{\int_A} f(x)\,dx + \varepsilon\,.$$

Let $\mathcal{P}$ be the common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then

$$\underline{\int_A} -f(x)\,dx - \varepsilon < L(-f, \mathcal{P}_1) \leqslant L(-f, \mathcal{P}) \leqslant \underline{\int_A} -f(x)\,dx$$

and

$$\overline{\int_A} f(x)\,dx \leqslant U(f, \mathcal{P}) \leqslant U(f, \mathcal{P}_2) < \overline{\int_A} f(x)\,dx + \varepsilon\,.$$

By the fact that

$$L(-f, \mathcal{P}) = \sum_{\Delta \in \mathcal{P}} \inf_{x \in \Delta} \overline{(-f)}^A(x)\nu(\Delta) = -\sum_{\Delta \in \mathcal{P}} \sup_{x \in \Delta} \overline{f}^A(x)\nu(\Delta) = -U(f, \mathcal{P})\,,$$

we find that

$$\underline{\int_A} -f(x)\,dx - \varepsilon < L(-f, \mathcal{P}) = -U(f, \mathcal{P}) \leqslant -\overline{\int_A} f(x)\,dx$$

and

$$\underline{\int_A} -f(x)\,dx \geqslant L(-f, \mathcal{P}) = -U(f, \mathcal{P}) > -\overline{\int_A} f(x)\,dx - \varepsilon\,.$$

Therefore,

$$\underline{\int_A} -f(x)\,dx - \varepsilon < -\overline{\int_A} f(x)\,dx < \underline{\int_A} -f(x)\,dx + \varepsilon\,.$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude (c).

(e) Since f is bounded on A, there exist $M > 0$ such that $-M \leqslant f(x) \leqslant M$ for all $x \in A$. Therefore, $-1_A \leqslant \dfrac{f}{M} \leqslant 1_A$ on A; thus (c) and (d) imply that

$$0 = \int_A \mathbf{1}_A(x)\,dx = \overline{\int_A} \mathbf{1}_A(x)\,dx \geqslant \overline{\int_A} \frac{f(x)}{M}\,dx = \frac{1}{M}\overline{\int_A} f(x)\,dx$$

which implies that $\overline{\int_A} f(x)\,dx \leqslant 0$. Similarly, $\overline{\int_A} -f(x)\,dx \leqslant 0$ which further implies that $\underline{\int_A} f(x)\,dx \geqslant 0$. Therefore, by Corollary 3.8 we conclude that

$$0 \leqslant \underline{\int_A} f(x)\,dx \leqslant \overline{\int_A} f(x)\,dx \leqslant 0$$

which implies that f is Riemann integrable over A and $\int_A f(x)\,dx = 0$. ▫

Remark 3.15. Let $A \subseteq \mathbb{R}^n$ be bounded and $f, g : A \to \mathbb{R}$ be bounded. Then (b) of Proposition 3.14 also implies that

$$\underline{\int_A}(f-g)(x)\,dx \leqslant \underline{\int_A} f(x)\,dx - \underline{\int_A} g(x)\,dx \text{ and } \overline{\int_A} f(x)\,dx - \overline{\int_A} g(x)\,dx \leqslant \overline{\int_A}(f-g)(x)\,dx\,.$$

Corollary 3.16. *Let $A, B \subseteq \mathbb{R}^n$ be bounded such that $A \cap B$ has volume zero, and $f : A \cup B \to \mathbb{R}$ be bounded. Then*

$$\underline{\int_A} f(x)\,dx + \underline{\int_B} f(x)\,dx \leqslant \underline{\int_{A\cup B}} f(x)\,dx \leqslant \overline{\int_{A\cup B}} f(x)\,dx \leqslant \overline{\int_A} f(x)\,dx + \overline{\int_B} f(x)\,dx\,.$$

Proof. Note that $f\mathbf{1}_A + f\mathbf{1}_B = f\mathbf{1}_{A\cup B} + f\mathbf{1}_{A\cap B}$ on $A \cup B$. Therefore, (a), (b) of Proposition 3.14 and Remark 3.15 implies that

$$\begin{aligned}
\underline{\int_A} f(x)\,dx + \underline{\int_B} f(x)\,dx &= \underline{\int_{A\cup B}}(f\mathbf{1}_A)(x)\,dx + \underline{\int_{A\cup B}}(f\mathbf{1}_B)(x)\,dx \leqslant \underline{\int_{A\cup B}}(f\mathbf{1}_A + f\mathbf{1}_B)(x)\,dx \\
&= \underline{\int_{A\cup B}} \big(f\mathbf{1}_{A\cup B} - (-f\mathbf{1}_{A\cap B})\big)(x)\,dx \\
&\leqslant \underline{\int_{A\cup B}} f\mathbf{1}_{A\cup B}(x)\,dx - \underline{\int_{A\cup B}}(-f\mathbf{1}_{A\cap B})(x)\,dx \\
&= \underline{\int_{A\cup B}} f(x)\,dx - \underline{\int_{A\cap B}}(-f)(x)\,dx
\end{aligned}$$

which, with the help of Proposition 3.14 (e), further implies that

$$\underline{\int_A} f(x)\,dx + \underline{\int_B} f(x)\,dx \leqslant \underline{\int_{A\cup B}} f(x)\,dx\,.$$

The case of the upper integral can be proved in a similar fashion. ▫

Having established Proposition 3.14, it is easy to see the following theorem (except (c)). The proof is left as an exercise.

Theorem 3.17. *Let $A \subseteq \mathbb{R}^n$ be bounded, $c \in \mathbb{R}$, and $f, g : A \to \mathbb{R}$ be Riemann integrable. Then*

(a) *$f \pm g$ is Riemann integrable, and* $\displaystyle\int_A (f \pm g)(x)\,dx = \int_A f(x)\,dx \pm \int_A g(x)\,dx$.

(b) *cf is Riemann integrable, and* $\displaystyle\int_A (cf)(x)\,dx = c\int_A f(x)\,dx$.

(c) *$|f|$ is Riemann integrable, and* $\displaystyle\Big|\int_A f(x)\,dx\Big| \leqslant \int_A |f(x)|dx$.

(d) *If $f \leqslant g$, then* $\displaystyle\int_A f(x)\,dx \leqslant \int_A g(x)\,dx$.

(e) *If A has volume and $|f| \leqslant M$, then* $\displaystyle\Big|\int_A f(x)\,dx\Big| \leqslant M\nu(A)$.

Definition 3.18. Let $A \subseteq \mathbb{R}^n$ be a set and $f : A \to \mathbb{R}$ be a function. For $B \subseteq A$, the ***restriction of*** f ***to*** B is the function $f\big|_B : A \to \mathbb{R}$ given by $f|_B = f\mathbf{1}_B$. In other words,

$$f\big|_B(x) = \begin{cases} f(x) & \text{if } x \in B\,, \\ 0 & \text{if } x \in A\backslash B\,. \end{cases}$$

The following two theorems are direct consequences of (a) of Proposition 3.14 and Corollary 3.16.

Theorem 3.19. *Let $A, B \subseteq \mathbb{R}^n$ be bounded, $B \subseteq A$, and $f : A \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over B if and only if $f|_B$ is Riemann integrable over A. In either cases,*

$$\int_A f\big|_B(x)\,dx = \int_B f(x)\,dx\,.$$

Theorem 3.20. *Let A, B be bounded subsets of $\mathbb{R}^n$ be such that $A \cap B$ has volume zero, and $f : A \cup B \to \mathbb{R}$ be bounded such that $f\big|_A$ and $f\big|_B$ are all Riemann integrable over $A \cup B$. Then f is Riemann integrable over $A \cup B$, and*

$$\int_{A\cup B} f(x)\,dx = \int_A f(x)\,dx + \int_B f(x)\,dx\,.$$

3.3 Integrability for Almost Continuous Functions

Lemma 3.21. *Let $A \subseteq \mathbb{R}^n$ be a bounded set of volume zero. If $B \subseteq A$, then B has volume zero.*

Proof. By (a), (d) and (e) of Proposition 3.14,

$$0 = \int_A \mathbf{1}_B(x)\,dx = \underline{\int}_A \mathbf{1}_B(x)\,dx = \underline{\int}_B \mathbf{1}_B(x)\,dx$$

and

$$0 = \int_A \mathbf{1}_B(x)\,dx = \overline{\int}_A \mathbf{1}_B(x)\,dx = \overline{\int}_B \mathbf{1}_B(x)\,dx\,.$$

Therefore, $\displaystyle\int_B \mathbf{1}_B(x)\,dx = 0$ which implies that B has volume zero. □

Lemma 3.22. *Let $A_1, \cdots, A_k \subseteq \mathbb{R}^n$ be bounded sets of volume zero. Then $\bigcup\limits_{j=1}^{k} A_j$ has volume zero.*

Proof. It suffices to prove the case for $k = 2$. Suppose that A_1 and A_2 are bounded sets of volume zero, and $A = A_1 \cup A_2$. By Lemma 3.21, $A_1 \cap A_2$ has volume zero; thus (e) of Proposition 3.14 and Corollary 3.16 imply that

$$\underline{\int}_A \mathbf{1}_A(x)\,dx = \underline{\int}_{A_1 \cup A_2} \mathbf{1}_A(x)\,dx \geqslant \underline{\int}_{A_1} \mathbf{1}_A(x)\,dx + \underline{\int}_{A_2} \mathbf{1}_A(x)\,dx = 0$$

and

$$\overline{\int}_A \mathbf{1}_A(x)\,dx = \overline{\int}_{A_1 \cup A_2} \mathbf{1}_A(x)\,dx \leqslant \overline{\int}_{A_1} \mathbf{1}_A(x)\,dx + \overline{\int}_{A_2} \mathbf{1}_A(x)\,dx = 0\,.$$

Therefore, $\displaystyle\int_A \mathbf{1}_A(x)\,dx = 0$ which implies that A has volume zero. □

Theorem 3.23. *Let $A \subseteq \mathbb{R}^n$ be a bounded set such that ∂A has volume zero, and $f : A \to \mathbb{R}$ be a bounded function. If f is continuous except perhaps on a set of volume zero, then f is Riemann integrable over A.*

Proof. Let R be a closed cube such that $A \subseteq R$ and $\partial A \cap \partial R = \varnothing$. We show that $\overline{f}^A = f\mathbf{1}_A$ is Riemann integrable over R and by (a) of Proposition 3.14, we then obtain that

$$\underline{\int}_A f(x)\,dx = \underline{\int}_R (f\mathbf{1}_A)(x)\,dx = \int_R (f\mathbf{1}_A)(x)\,dx = \overline{\int}_R (f\mathbf{1}_A)(x)\,dx = \overline{\int}_A f(x)\,dx$$

which implies that f is Riemann integrable over A.

Let $\varepsilon > 0$ be given. Suppose that the collection of discontinuities of f is D, and $B = \partial A \cup D$. Since ∂A and D has volume zero, Lemma 3.22 implies that B has volume zero; thus (a) of Proposition 3.14 then implies (with $B \subseteq R$ in mind) that

$$\underline{\int}_R \mathbf{1}_B(x)\,dx = \underline{\int}_B \mathbf{1}_B(x)\,dx = 0 \qquad \text{and} \qquad \overline{\int}_R \mathbf{1}_B(x)\,dx = \overline{\int}_B \mathbf{1}_B(x)\,dx = 0\,.$$

Therefore, $\displaystyle\int_R \mathbf{1}_B(x)\,dx = 0$, so there exists a partition $\mathcal{P}_1$ of R such that

$$\sum_{\Delta\in\mathcal{P}_1,\Delta\cap B\neq\varnothing} \nu(\Delta) = U(\mathbf{1}_B,\mathcal{P}_1) < \frac{\varepsilon}{2\big[\sup \overline{f}^A(R) - \inf \overline{f}^A(R) + 1\big]}\,.$$

Let $\mathcal{U} \equiv \operatorname{int}\Big(\displaystyle\bigcup_{\Delta\in\mathcal{P}_1,\Delta\cap B\neq\varnothing}\Delta\Big)$. Then $B \subseteq \mathcal{U}$. Since the discontinuity of $\overline{f}^A$ is a subset of B, $\overline{f}^A : R\cap\mathcal{U}^\complement \to \mathbb{R}$ is continuous. Since $R\cap\mathcal{U}^\complement$ is closed and bounded, $\overline{f}^A$ is uniformly continuous; thus there exists $\delta > 0$ such that

$$\big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| < \frac{\varepsilon}{2\nu(R)} \qquad \text{if } x_1, x_2 \in R\cap\mathcal{U}^\complement \text{ and } \|x_1 - x_2\| < \delta\,.$$

Let $\mathcal{P}$ be a refinement of $\mathcal{P}_1$ such that $\|\mathcal{P}\| < \delta$, and define two classes C_1, C_2 of rectangles in $\mathcal{P}$ by $C_1 = \big\{\Delta' \in \mathcal{P} \,\big|\, \Delta' \nsubseteq \Delta \text{ for all } \Delta\in\mathcal{P}_1 \text{ satisfying } \Delta\cap B \neq \varnothing\big\}$ and $C_2 = \big\{\Delta' \in \mathcal{P} \,\big|\, \Delta' \notin C_1\big\}$. Then if $\Delta' \in C_1$, then $\Delta' \subseteq R\backslash\mathcal{U}^\complement$; thus

$$\begin{aligned}
U(\overline{f}^A,\mathcal{P}) - L(\overline{f}^A,\mathcal{P}) &= \sum_{\Delta'\in\mathcal{P}} \Big[\sup_{x\in\Delta'}(\overline{f}^A\mathbf{1}_R)(x) - \inf_{x\in\Delta'}(\overline{f}^A\mathbf{1}_R)(x)\Big]\nu(\Delta') \\
&= \Big(\sum_{\Delta'\in C_1} + \sum_{\Delta'\in C_2}\Big)\Big[\sup_{x\in\Delta'}\overline{f}^A(x) - \inf_{x\in\Delta'}\overline{f}^A(x)\Big]\nu(\Delta') \\
&\leqslant \frac{\varepsilon}{2\nu(R)}\sum_{\Delta'\in C_1}\nu(\Delta') + \Big[\sup\overline{f}^A(R) - \inf\overline{f}^A(R)\Big]\sum_{\Delta'\in C_2}\nu(\Delta') \\
&= \frac{\varepsilon}{2\nu(R)}\nu(R) + \Big[\sup\overline{f}^A(R) - \inf\overline{f}^A(R)\Big]\sum_{\Delta\in\mathcal{P}_1,\Delta\cap B\neq\varnothing}\nu(\Delta) \\
&< \frac{\varepsilon}{2} + \frac{\big[\sup\overline{f}^A(R) - \inf\overline{f}^A(R)\big]\varepsilon}{2\big[\sup\overline{f}^A(R) - \inf\overline{f}^A(R) + 1\big]} < \varepsilon\,,
\end{aligned}$$

and we conclude that f is Riemann integrable over A by Riemann's condition. □

3.4 The Fubini theorem

If $f : [a,b] \to \mathbb{R}$ is continuous, the fundamental theorem of Calculus can be applied to computed the integral of f over $[a,b]$. In the following two sections, we focus on how the integral of f over $A \subseteq \mathbb{R}^{\mathrm{n}}$, where $\mathrm{n} \geqslant 2$, can be computed if the integral exists.

Definition 3.24. Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ and $B \subseteq \mathbb{R}^{\mathrm{m}}$ be bounded sets, $S = A\times B$ be a product set in $\mathbb{R}^{\mathrm{n+m}}$, and $f : S \to \mathbb{R}$ be bounded. For each fixed $x \in A$, the lower integral of the function

$f(x,\cdot):B\to\mathbb{R}$ is denoted by $\underline{\int}_B f(x,y)\,dy$, and the upper integral of $f(x,\cdot):B\to\mathbb{R}$ is denoted by $\overline{\int}_B f(x,y)\,dy$. If for each $x\in A$ the upper integral and the lower integral of $f(x,\cdot):B\to\mathbb{R}$ are the same, we simply write $\int_B f(x,y)\,dy$ for the integrals of $f(x,\cdot)$ over B. The integrals $\overline{\int}_A f(x,y)\,dx$, $\underline{\int}_A f(x,y)\,dx$ and $\int_A f(x,y)\,dx$ are defined in a similar way.

Theorem 3.25 (Fubini's Theorem)**.** *Let $A\subseteq\mathbb{R}^{\rm n}$ and $B\subseteq\mathbb{R}^{\rm m}$ be bounded sets, and $f:A\times B\to\mathbb{R}$ be bounded. For $x\in\mathbb{R}^{\rm n}$ and $y\in\mathbb{R}^{\rm m}$, write $z=(x,y)$. Then*

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_A\Big(\underline{\int}_B f(x,y)dy\Big)dx \leqslant \overline{\int}_A\Big(\overline{\int}_B f(x,y)dy\Big)dx \leqslant \overline{\int}_{A\times B} f(z)\,dz \tag{3.3}$$

and

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_B\Big(\underline{\int}_A f(x,y)dx\Big)dy \leqslant \overline{\int}_B\Big(\overline{\int}_A f(x,y)dx\Big)dy \leqslant \overline{\int}_{A\times B} f(z)\,dz\,. \tag{3.4}$$

In particular, if $f:A\times B\to\mathbb{R}$ is Riemann integrable, then

$$\begin{aligned}\int_{A\times B} f(z)\,dz &= \int_A\Big(\underline{\int}_B f(x,y)dy\Big)dx = \int_A\Big(\overline{\int}_B f(x,y)dy\Big)dx\\ &= \int_B\Big(\underline{\int}_A f(x,y)dx\Big)dy = \int_B\Big(\overline{\int}_A f(x,y)dx\Big)dy\,.\end{aligned}$$

Proof. It suffices to prove (3.3). Let $\varepsilon>0$ be given. Choose a partition $\mathcal{P}$ of $A\times B$ such that $L(f,\mathcal{P})>\underline{\int}_{A\times B} f(z)\,dz-\varepsilon$. Since $\mathcal{P}$ is a partition of $A\times B$, there exist partition $\mathcal{P}_x$ of A and partition $\mathcal{P}_y$ of B such that $\mathcal{P}=\big\{\Delta=R\times S\,\big|\,R\in\mathcal{P}_x, S\in\mathcal{P}_y\big\}$. By Proposition 3.14 and Corollary 3.16, we find that

$$\begin{aligned}\underline{\int}_A\Big(\underline{\int}_B f(x,y)\,dy\Big)dx &= \underline{\int}_{\bigcup_{R\in\mathcal{P}_x}R}\mathbf{1}_A(x)\Big(\underline{\int}_{\bigcup_{S\in\mathcal{P}_y}S} f(x,y)\mathbf{1}_B(y)\,dy\Big)dx\\ &\geqslant \sum_{R\in\mathcal{P}_x}\underline{\int}_R\Big(\sum_{S\in\mathcal{P}_y}\underline{\int}_S \overline{f}^{A\times B}(x,y)\,dy\Big)dx\\ &\geqslant \sum_{R\in\mathcal{P}_x}\sum_{S\in\mathcal{P}_y}\underline{\int}_R\Big(\underline{\int}_S \overline{f}^{A\times B}(x,y)\,dy\Big)dx\\ &\geqslant \sum_{R\in\mathcal{P}_x,S\in\mathcal{P}_y}\inf_{(x,y)\in R\times S}\overline{f}^{A\times B}(x,y)\nu_{\rm m}(S)\nu_{\rm n}(R)\\ &= \sum_{\Delta\in\mathcal{P}}\inf_{(x,y)\in\Delta}\overline{f}^{A\times B}(x,y)\nu_{\rm n+m}(\Delta) = L(f,\mathcal{P}) > \underline{\int}_{A\times B} f(z)dz-\varepsilon\,.\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_{B}\Big(\underline{\int}_{A} f(x,y)dx\Big)dy\,.$$

Similarly, $\overline{\int}_{A}\Big(\overline{\int}_{B} f(x,y)dy\Big)dx \leqslant \overline{\int}_{A\times B} f(z)\,dz$; thus (3.3) is concluded. □

Corollary 3.26. *Let $S \subseteq \mathbb{R}^{\mathrm{n}}$ be a closed and bounded set such that ∂S has volume zero, $\varphi_1, \varphi_2 : S \to \mathbb{R}$ be continuous maps such that $\varphi_1(x) \leqslant \varphi_2(x)$ for all $x \in S$, $A = \big\{(x,y) \in \mathbb{R}^{\mathrm{n}} \times \mathbb{R}\,\big|\, x \in S, \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\big\}$, and $f : A \to \mathbb{R}$ be continuous. Then f is Riemann integrable over A, and*

$$\int_A f(x,y)\,d(x,y) = \int_S \Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\Big)dx\,. \tag{3.5}$$

Proof. To establish that f is Riemann integrable over A, by Theorem 3.23 it suffices to show that ∂A has volume zero. Let $m = \min\limits_{x\in S} \varphi_1(x)$ and $M = \max\limits_{x\in S} \varphi_2(x)$. Since

$$\partial A \subseteq \big\{(x,\varphi_1(x))\,\big|\, x \in S\big\} \cup \big\{(x,\varphi_2(x))\,\big|\, x \in S\big\} \cup \big(\partial S \times [m,M]\big)\,,$$

to see ∂A has volume zero it suffices to show that $\partial S \times [m,M]$, $\big\{(x,\varphi_1(x))\,\big|\, x \in S\big\}$ and $\big\{(x,\varphi_2(x))\,\big|\, x \in S\big\}$ have volume zero because of Lemma 3.21 and 3.22. Note that Theorem 3.23 implies that φ_1 is Riemann integrable over S; thus for a given $\varepsilon > 0$ there exists partition $\mathcal{P}$ of S such that

$$U(\varphi_1,\mathcal{P}) - L(\varphi_1,\mathcal{P}) < \varepsilon\,.$$

Let $B = \bigcup\limits_{\Delta\in\mathcal{P},\Delta\cap S\neq\varnothing} \Delta \times \big[\inf_{x\in\Delta} \overline{\varphi_1}^S(x), \sup_{x\in\Delta} \overline{\varphi_1}^S(x)\big]$. Then $C \equiv \big\{(x,\varphi_1(x))\,\big|\, x \in S\big\} \subseteq B$ and

$$\begin{aligned} 0 &\leqslant \overline{\int}_C \mathbf{1}_C(z)\,dz \leqslant \overline{\int}_B \mathbf{1}_B(z)\,dz \leqslant \sum_{\Delta\in\mathcal{P},\Delta\cap S\neq\varnothing} \big(\sup_{x\in\Delta}\overline{\varphi_1}^S(x) - \inf_{x\in\Delta}\overline{\varphi_1}^S(x)\big) \times \nu_{\mathrm{n}}(\Delta) \\ &\leqslant U(\varphi_1,\mathcal{P}) - L(\varphi_1,\mathcal{P}) < \varepsilon\,. \end{aligned}$$

Therefore, $C = \big\{(x,\varphi_1(x))\,\big|\, x \in S\big\}$ has volume zero and similarly, $\big\{(x,\varphi_2(x))\,\big|\, x \in S\big\}$ has volume zero.

Now we show that $\partial S \times [m,M]$ has volume zero. Since ∂S has volume zero in $\mathbb{R}^{\mathrm{n}}$, for a given $\varepsilon > 0$ there exists a partition $\mathcal{P}$ of ∂S such that

$$U(\mathbf{1}_S,\mathcal{P}) < \frac{\varepsilon}{M-m+1}\,.$$

Then $\partial S \times [m, M] \subseteq \bigcup\limits_{\Delta \in \mathcal{P}, \Delta \cap \partial S \neq \varnothing} \Delta \times [m, M]$, and as above

$$\overline{\int}_{\partial S \times [m,M]} \mathbf{1}_{\partial S \times [m,M]}(z)\, dz \leqslant \sum_{\Delta \in \mathcal{P}, \Delta \cap \partial S \neq \varnothing} \nu_{\rm n}(\Delta) \times (M-m) \leqslant (M-m) U(\mathbf{1}_S, \mathcal{P}) < \varepsilon\,.$$

Therefore, $\partial S \times [m, M]$ has volume zero; thus we establish that f is Riemann integrable over A.

Next we prove (3.5). Note that $A \subseteq S \times [m, M]$; thus Theorem 3.20 and the Fubini Theorem imply that

$$\begin{aligned}\int_A f(x,y)\, d(x,y) &= \int_{S \times [m,M]} \overline{f}^A(x,y)\, d(x,y) = \underline{\int}_S \Big(\underline{\int}_m^M \overline{f}^A(x,y)\, dy \Big) dx \\ &= \overline{\int}_S \Big(\overline{\int}_m^M \overline{f}^A(x,y)\, dy \Big) dx\,.\end{aligned}$$

Noting that $[m, M]$ has a boundary of volume zero in $\mathbb{R}$, and for each $x \in S$, $\overline{f}^A(x, \cdot)$ is continuous except perhaps at $y = \varphi_1(x)$ and $y = \varphi_2(x)$, Theorem 3.23 implies that $\overline{f}^A(x, \cdot)$ is Riemann integrable over $[m, M]$ for each $x \in S$; thus $\underline{\int}_m^M \overline{f}^A(x,y)\, dy = \overline{\int}_m^M \overline{f}^A(x,y)\, dy$ which further implies that

$$\int_A f(x,y)\, d(x,y) = \int_S \Big(\int_m^M \overline{f}^A(x,y)\, dy \Big) dx\,. \tag{3.6}$$

For each fixed $x \in S$, let $A_x = \big\{ y \in \mathbb{R} \,\big|\, \varphi_1(x) \leqslant y \leqslant \varphi_2(x) \big\}$. Then $\overline{f}^A(x,y) = f(x,y)\mathbf{1}_{A_x}(y)$ for all $(x,y) \in S \times [m, M]$ or equivalently, $\overline{f}^A(x, \cdot) = f(x, \cdot)|_{A_x}$ for all $x \in S$; thus Proposition 3.14 (a) implies that

$$\int_m^M \overline{f}^A(x,y)\, dy = \int_{A_x} f(x,y)\, dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy \qquad \forall\, x \in S\,. \tag{3.7}$$

Combining (3.6) and (3.7), we conclude (3.5). □

Example 3.27. Let $A = \big\{ (x,y) \in \mathbb{R}^2 \,\big|\, 0 \leqslant x \leqslant 1, x \leqslant y \leqslant 1 \big\}$, and $f : A \to \mathbb{R}$ be given by $f(x,y) = xy$. Then Corollary 3.26 implies that

$$\int_A f(x,y)\, d\mathbb{A} = \int_0^1 \Big(\int_x^1 xy\, dy \Big) dx = \int_0^1 \frac{xy^2}{2}\Big|_{y=x}^{y=1} dx = \int_0^1 \Big(\frac{x}{2} - \frac{x^3}{2} \Big) dx = \frac{1}{4} - \frac{1}{8} = \frac{1}{8}\,.$$

On the other hand, since $A = \big\{ (x,y) \in \mathbb{R}^2 \,\big|\, 0 \leqslant y \leqslant 1, 0 \leqslant x \leqslant y \big\}$, we can also evaluate the integral of f over A by

$$\int_A xy\, d\mathbb{A} = \int_0^1 \Big(\int_0^y xy\, dx \Big) dy = \int_0^1 \frac{x^2 y}{2}\Big|_{x=0}^{x=y} dy = \int_0^1 \frac{y^3}{2}\, dy = \frac{1}{8}\,.$$

Example 3.28. Let $A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1, \sqrt{x} \leqslant y \leqslant 1\}$, and $f : A \to \mathbb{R}$ be given by $f(x,y) = e^{y^3}$. Then Corollary 3.26 implies that

$$\int_A f(x,y)\, d\mathbb{A} = \int_0^1 \Big(\int_{\sqrt{x}}^1 e^{y^3} dy\Big) dx\,.$$

Since we do not know how to compute the inner integral, we look for another way of finding the integral. Observing that $A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant y \leqslant 1, 0 \leqslant x \leqslant y^2\}$, we have

$$\int_A f(x,y)\, d\mathbb{A} = \int_0^1 \Big(\int_0^{y^2} e^{y^3} dx\Big) dy = \int_0^1 y^2 e^{y^3} dy = \frac{1}{3} e^{y^3}\Big|_{y=0}^{y=1} = \frac{e-1}{3}\,.$$

Example 3.29. Let $A \subseteq \mathbb{R}^3$ be the set $\{(x_1,x_2,x_3) \in \mathbb{R}^3 \,|\, x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0, \text{and } x_1 + x_2 + x_3 \leqslant 1\}$, and $f : A \to \mathbb{R}$ be given by $f(x_1,x_2,x_3) = (x_1+x_2+x_3)^2$. Let $S = [0,1] \times [0,1] \times [0,1]$, and $\overline{f} : \mathbb{R}^3 \to \mathbb{R}$ be the extension of f by zero outside A. Then Theorem 3.23 implies that f is Riemann integrable. Write $\widehat{x}_1 = (x_2,x_3)$, $\widehat{x}_2 = (x_1,x_3)$ and $\widehat{x}_3 = (x_1,x_2)$. Theorem 3.20 implies that

$$\int_A f(x) dx = \int_S \overline{f}(x) dx\,,$$

and Theorem 3.25 implies that

$$\int_S \overline{f}(x) dx = \int_{[0,1]} \Big(\int_{[0,1]\times[0,1]} \overline{f}(\widehat{x}_3, x_3) d\widehat{x}_3\Big) dx_3\,.$$

Let $A_{x_3} = \{(x_1,x_2) \in \mathbb{R}^2 \,|\, x_1 \geqslant 0, x_2 \geqslant 0, x_1 + x_2 \leqslant 1 - x_3\}$. Then for each $x_3 \in [0,1]$,

$$\int_{[0,1]\times[0,1]} \overline{f}(\widehat{x}_3, x_3) d\widehat{x}_3 = \int_{A_{x_3}} f(\widehat{x}_3, x_3) d\widehat{x}_3 = \int_0^{1-x_3} \Big(\int_0^{1-x_3-x_2} f(x_1,x_2,x_3) dx_1\Big) dx_2\,.$$

Computing the iterated integral, we find that

$$\begin{aligned}
\int_A f(x) dx &= \int_0^1 \Big[\int_0^{1-x_3} \Big(\int_0^{1-x_3-x_2} (x_1+x_2+x_3)^2 dx_1\Big) dx_2\Big] dx_3 \\
&= \int_0^1 \Big[\int_0^{1-x_3} \frac{(x_1+x_2+x_3)^3}{3}\Big|_{x_1=0}^{x_1=1-x_3-x_2} dx_2\Big] dx_3 \\
&= \int_0^1 \Big[\int_0^{1-x_3} \Big(\frac{1}{3} - \frac{(x_2+x_3)^3}{3}\Big) dx_2\Big] dx_3 \\
&= \int_0^1 \Big(\frac{1}{4} - \frac{x_3}{3} + \frac{x_3^4}{12}\Big) dx_3 = \frac{1}{4} - \frac{1}{6} + \frac{1}{60} = \frac{15-10+1}{60} = \frac{1}{10}\,.
\end{aligned}$$

Example 3.30. In this example we compute the volume $\omega_{\rm n}$ of the n-dimensional unit ball. By the Fubini theorem,

$$\omega_{\rm n} = \int_{-1}^{1} \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}} dx_{\rm n} \cdots dx_1 \,.$$

Note that the integral $\displaystyle\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{\rm n-1}^2}} dx_{\rm n} \cdots dx_2$ is in fact $\omega_{\rm n-1}(1-x_1^2)^{\frac{\rm n-1}{2}}$, the volume of $(n-1)$-dimensional ball of radius $\sqrt{1-x_1^2}$; thus

$$\omega_{\rm n} = \int_{-1}^{1} \omega_{\rm n-1}(1-x^2)^{\frac{\rm n-1}{2}} dx = 2\,\omega_{\rm n-1} \int_0^{\frac{\pi}{2}} \cos^{\rm n}\theta\, d\theta \,. \tag{3.8}$$

Integrating by parts,

$$\begin{aligned} \int_0^{\frac{\pi}{2}} \cos^{\rm n}\theta\, d\theta &= \int_0^{\frac{\pi}{2}} \cos^{\rm n-1}\theta\, d(\sin\theta) = \cos^{\rm n-1}\theta \sin\theta\Big|_{\theta=0}^{\theta=\frac{\pi}{2}} + ({\rm n}-1)\int_0^{\frac{\pi}{2}} \cos^{\rm n-2}\theta \sin^2\theta\, d\theta \\ &= ({\rm n}-1)\int_0^{\frac{\pi}{2}} \cos^{\rm n-2}\theta (1-\cos^2\theta)\, d\theta \end{aligned}$$

which implies that

$$\int_0^{\frac{\pi}{2}} \cos^{\rm n}\theta\, d\theta = \frac{\rm n-1}{\rm n} \int_0^{\frac{\pi}{2}} \cos^{\rm n-2}\theta\, d\theta \,.$$

As a consequence,

$$\int_0^{\frac{\pi}{2}} \cos^{\rm n}\theta\, d\theta = \begin{cases} \dfrac{({\rm n}-1)({\rm n}-3)\cdots 2}{{\rm n}({\rm n}-2)\cdots 3} \displaystyle\int_0^{\frac{\pi}{2}} \cos\theta\, d\theta & \text{if n is odd}\,, \\ \dfrac{({\rm n}-1)({\rm n}-3)\cdots 1}{{\rm n}({\rm n}-2)\cdots 2} \displaystyle\int_0^{\frac{\pi}{2}} d\theta & \text{if n is even}\,, \end{cases}$$

and the recursive formula (3.8) implies that $\omega_{\rm n} = \dfrac{2\,\omega_{\rm n-2}}{n}\pi$. Further computations shows that

$$\omega_{\rm n} = \begin{cases} \dfrac{(2\pi)^{\frac{\rm n-1}{2}}}{n({\rm n}-2)\cdots 3}\omega_1 & \text{if n is odd}\,, \\ \dfrac{(2\pi)^{\frac{\rm n-2}{2}}}{n({\rm n}-2)\cdots 4}\omega_2 & \text{if n is even}\,. \end{cases}$$

Let Γ be the Gamma function defined by $\Gamma(t) = \displaystyle\int_0^\infty x^{t-1}e^{-x}\, dx$ for $t > 0$. Then $\Gamma(x+1) = x\Gamma(x)$ for all $x > 0$, $\Gamma(1) = 1$ and $\Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}$. By the fact that $\omega_1 = 2$ and $\omega_2 = \pi$, we can express $\omega_{\rm n}$ as

$$\omega_{\rm n} = \frac{\pi^{\frac{\rm n}{2}}}{\Gamma\big(\frac{\rm n+2}{2}\big)} \,.$$

3.5 The Change of Variables Formula

Fubini theorem can be used to find the integral of a (Riemann integrable) function over a rectangular domain if the iterated integrals can be evaluated. However, like the integral of a function of one variable, in many cases we need to make use of several change of variables in order to transform the integral to another integral that is easier to be evaluated. In this section, we establish the change of variables formula for the integral of functions of several variables.

Theorem 3.31 (Change of Variables Formula)**.** *Let $\mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}}$ be an open set with volume, and $\psi : \mathcal{U} \to \mathbb{R}^{\mathrm{n}}$ be an one-to-one $\mathscr{C}^1$-mapping with $\mathscr{C}^1$-inverse; that is, $\psi^{-1} : \psi(\mathcal{U}) \to \mathcal{U}$ is also continuously differentiable. Assume that the Jacobian of ψ, $\mathrm{J} = \det([D\psi])$, does not vanish in $\mathcal{U}$. If $f : \psi(\mathcal{U}) \to \mathbb{R}$ is Riemann integrable, then $(f \circ \psi)\mathrm{J}$ is Riemann integrable over $\mathcal{U}$, and*

$$\int_{\psi(\mathcal{U})} f(y)\,dy = \int_{\mathcal{U}} (f \circ \psi)(x)\big|\mathrm{J}(x)\big|\,dx = \int_{\mathcal{U}} (f \circ \psi)(x)\Big|\frac{\partial(\psi_1, \cdots, \psi_{\mathrm{n}})}{\partial(x_1, \cdots, x_{\mathrm{n}})}\Big|\,dx\,.$$

The proof of Theorem 3.31 is very lengthy and requires a bit more knowledge about the integration, so we only present the proof of a much simpler case.

Theorem 3.32. *Let $\mathrm{D} \subseteq \mathbb{R}^{\mathrm{n}}$ be an open rectangle, and $\psi : \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ be an one-to-one $\mathscr{C}^2$ mapping such that $\psi = \mathrm{Id}$ outside $B(0, r)$ for some $r > 0$; that is, $\psi(x) = x$ if $|x| \geqslant r$. Assume that the Jacobian of ψ, $\mathrm{J} = \det(\nabla\psi)$, does not vanish in $\mathbb{R}^{\mathrm{n}}$. If $f : \mathrm{D} \to \mathbb{R}$ is of class $\mathscr{C}^1$ and is compactly supported in D; that is, $\mathrm{cl}\big(\{x \in \mathrm{D} \,\big|\, f(x) \neq 0\}\big) \subseteq \mathrm{D}$, then*

$$\int_{\mathrm{D}} f(y)\,dy = \int_{\psi^{-1}(\mathrm{D})} (f \circ \psi)(x)\mathrm{J}(x)\,dx\,.$$

Proof. W.L.O.G. we can assume that $\mathrm{D} = [-R, R]^n$ is a cube and $B(0, r) \subset\subset \mathrm{D}$ (or equivalently, $0 < r < R$). Then $\psi^{-1}(\mathrm{D}) = \mathrm{D}$ since $\psi = \mathrm{Id}$ outside $B(0, R)$. Define

$$g(y_1, \cdots, y_n) = \int_{-R}^{y_1} f(z, y_2, \cdots, y_n)\,dz\,,$$

and $\mathrm{M} = \begin{bmatrix} [D(g\circ\psi)] \\ [D\psi_2] \\ [D\psi_3] \\ \vdots \\ [D\psi_n] \end{bmatrix}$. By the property of determinants and the chain rule, we find that

$$\det(\mathrm{M}) = \det\left(\begin{bmatrix} \sum\limits_{j=1}^{n}\big(\frac{\partial g}{\partial y_j}\circ\psi\big)\frac{\partial \psi_j}{\partial x_1} & \sum\limits_{j=1}^{n}\big(\frac{\partial g}{\partial y_j}\circ\psi\big)\frac{\partial \psi_j}{\partial x_2} & \cdots & \sum\limits_{j=1}^{n}\big(\frac{\partial g}{\partial y_j}\circ\psi\big)\frac{\partial \psi_j}{\partial x_n} \\ \frac{\partial \psi_2}{\partial x_1} & \frac{\partial \psi_2}{\partial x_2} & \cdots & \frac{\partial \psi_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \psi_n}{\partial x_1} & \frac{\partial \psi_n}{\partial x_2} & \cdots & \frac{\partial \psi_n}{\partial x_n} \end{bmatrix}\right)$$

$$= \det\left(\begin{bmatrix} \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\frac{\partial \psi_1}{\partial x_1} & \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\frac{\partial \psi_1}{\partial x_2} & \cdots & \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\frac{\partial \psi_1}{\partial x_n} \\ \frac{\partial \psi_2}{\partial x_1} & \frac{\partial \psi_2}{\partial x_2} & \cdots & \frac{\partial \psi_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \psi_n}{\partial x_1} & \frac{\partial \psi_n}{\partial x_2} & \cdots & \frac{\partial \psi_n}{\partial x_n} \end{bmatrix}\right)$$

$$= \big(\frac{\partial g}{\partial y_1}\circ\psi\big)\det\left(\begin{bmatrix} \frac{\partial \psi_1}{\partial x_1} & \frac{\partial \psi_1}{\partial x_2} & \cdots & \frac{\partial \psi_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \psi_n}{\partial x_1} & \frac{\partial \psi_n}{\partial x_2} & \cdots & \frac{\partial \psi_n}{\partial x_n} \end{bmatrix}\right) = (f\circ\psi)\mathrm{J}\,.$$

On the other hand, letting $\mathrm{A} = (D\psi)^{-1}$, then

$$\mathrm{Adj}(\mathrm{M})_{j1} = (-1)^{1+j}\det\big(\mathrm{M}(\widehat{1},\widehat{j})\big) = \mathrm{Adj}([D\psi])_{j1} = \mathrm{J}\mathrm{A}^j_1\,.$$

Computing the determinant by expanding along the first row, we obtain that

$$\det(\mathrm{M}) = \sum_{j=1}^{n}\mathrm{M}_{1j}\mathrm{Adj}(\mathrm{M})_{j1} = \sum_{j=1}^{n}\frac{\partial(g\circ\psi)}{\partial x_j}\mathrm{J}\mathrm{A}^j_1\,;$$

thus we conclude the identity

$$(f\circ\psi)\mathrm{J} = \sum_{j=1}^{n}\frac{\partial(g\circ\psi)}{\partial x_j}\mathrm{J}\mathrm{A}^j_1\,.$$

Therefore, with $\widehat{dx_j}$ denoting $dx_1 \cdots dx_{j-1} dx_{j+1} \cdots dx_n$, the Fubini theorem and the Piola identity imply that

$$\begin{aligned}\int_{\mathrm{D}} \big[(f \circ \psi)\mathrm{J}\big](x)\,dx &= \sum_{j=1}^{n} \int_{-R}^{R} \int_{-R}^{R} \cdots \int_{-R}^{R} \frac{\partial (g \circ \psi)}{\partial x_j} \mathrm{J}\mathrm{A}_1^j\, dx_j \widehat{dx_j} \\ &= \sum_{j=1}^{n} \int_{-R}^{R} \int_{-R}^{R} \cdots \int_{-R}^{R} \big[(g \circ \psi)\mathrm{J}\mathrm{A}_1^j\big]\Big|_{x_j=-R}^{x_j=R} \widehat{dx_j}\,.\end{aligned}$$

Since $\psi = \mathrm{Id}$ outside $B(0,r)$, we find that $\mathrm{J} = 1$ and $\mathrm{A}_1^j = \delta_{1j}$ on $\partial\mathrm{D}$; thus by the definition of g,

$$\int_{\mathrm{D}} \big[(f \circ \psi)\mathrm{J}\big](x)\,dx = \int_{-R}^{R} \int_{-R}^{R} \cdots \int_{-R}^{R} g(R, x_2, \cdots, x_n)\,\widehat{dx_1} = \int_{\mathrm{D}} f(x)\,dx\,. \qquad \square$$

Example 3.33. Suppose that $f : [0,1] \to \mathbb{R}$ is Riemann integrable and $\displaystyle\int_0^1 (1-x)f(x)\,dx = 5$. We would like to evaluate the iterated integral $\displaystyle\int_0^1 \int_0^x f(x-y)\,dydx$.

It is nature to consider the change of variables $(u,v) = (x-y, x)$ or $(u,v) = (x-y, y)$. Suppose the later case. Then $(x,y) = g(u,v) = (u+v, v)$; thus

$$\mathrm{J}_g(u,v) = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1\,.$$

Moreover, the region of integration is the triangle A with vertices $(0,0)$, $(1,0)$, $(1,1)$, and three sides $y = 0$, $x = 1$, $x = y$ correspond to $u = 0$, $u + v = 1$ and $v = 0$. Therefore, if E denotes the triangle enclosed by $u = 0$, $v = 0$ and $u + v = 1$ on the (u,v)-plane, then $g(E) = A$, and

$$\begin{aligned}\int_0^1 \int_0^x f(x-y)\,dydx &= \int_A f(x-y)\,d(x,y) = \int_{g(E)} f(x-y)\,d(x,y) \\ &= \int_E f\big(g_1(u,v) - g_2(u,v)\big)\big|\mathrm{J}_g(u,v)\big|\,d(u,v) = \int_0^1 \int_0^{1-u} f(u)\,dvdu \\ &= \int_0^1 (1-u)f(u)\,du = 5\,.\end{aligned}$$

Example 3.34. Let A be the triangular region with vertices $(0,0)$, $(4,0)$, $(4,2)$, and $f : A \to \mathbb{R}$ be given by

$$f(x,y) = y\sqrt{x-2y}\,.$$

Let $(u,v)=(x,x-2y)$. Then $(x,y)=g(u,v)=\big(u,\dfrac{u-v}{2}\big)$; thus

$$\mathrm{J}_g(u,v)=\begin{vmatrix}1 & 0\\ \frac{1}{2} & -\frac{1}{2}\end{vmatrix}=-\frac{1}{2}\,.$$

Define E as the triangle with vertices $(0,0)$, $(4,0)$, $(4,4)$. Then $A=g(E)$.

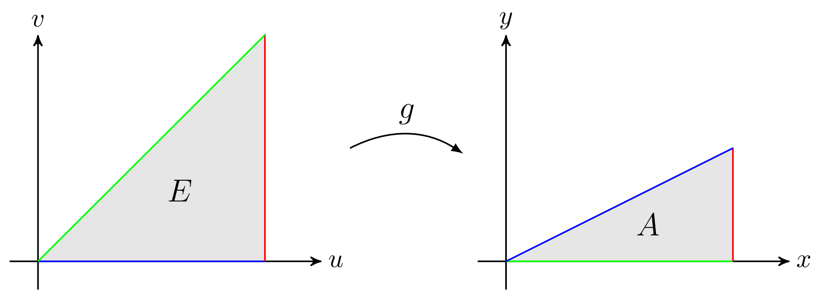


Figure 3.2: The image of E under g

Therefore,

$$\begin{aligned}\int_A f(x,y)\,d(x,y) &= \int_{g(E)} f(x,y)\,d(x,y)=\frac{1}{2}\int_E f\big(g(u,v)\big)\,d(u,v)\\ &=\frac{1}{4}\int_0^4\int_0^u (u-v)\sqrt{v}\,dvdu=\frac{1}{4}\int_0^4\Big[\frac{2}{3}uv^{\frac{3}{2}}-\frac{2}{5}v^{\frac{5}{2}}\Big]\Big|_{v=0}^{v=u}du\\ &=\frac{1}{4}\int_0^4\big(\frac{2}{3}-\frac{2}{5}\big)u^{\frac{5}{2}}\,du=\frac{1}{15}\times\frac{2}{7}u^{\frac{7}{2}}\Big|_{u=0}^{u=4}=\frac{256}{105}\,.\end{aligned}$$

Example 3.35. Let A be the region in the first quadrant of the plane bounded by the curves $xy-x+y=0$ and $x-y=1$, and $f:A\to\mathbb{R}$ be given by

$$f(x,y)=x^2y^2(x+y)e^{-(x-y)^2}\,.$$

We would like to evaluate the integral $\displaystyle\int_A f(x,y)\,d(x,y)$.

Let $(u,v)=(xy-x+y,x-y)$. Unlike the previous two examples we do not want to solve for (x,y) in terms of (u,v) but still assume that $(x,y)=g(u,v)$. By the inverse function theorem,

$$\mathrm{J}_g(u,v)\Big|_{(u,v)=g^{-1}(x,y)}=\Big|\frac{\partial(u,v)}{\partial(x,y)}\Big|^{-1}=\begin{vmatrix}y-1 & x+1\\ 1 & -1\end{vmatrix}^{-1}=\frac{1}{-y+1-x-1}=-\frac{1}{x+y}\,.$$

Moreover, the curve $xy - x + y = 0$ corresponds to $u = 0$, while the lines $x - y = 1$ and $y = 0$ correspond to $v = 1$ and $u + v = 0$, respectively; thus if E is the region enclosed by $u = 0$, $v = 1$ and $u + v = 0$, then $A = g(E)$.

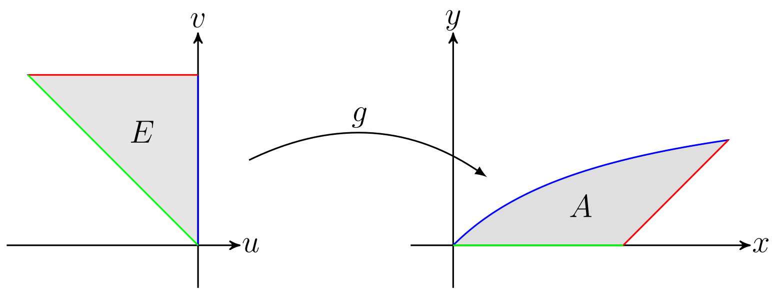


Figure 3.3: The image of E under g

Therefore,

$$\begin{aligned}\int_A f(x,y)d(x,y) &= \int_{g(E)} f(x,y)\,d(x,y) = \int_E (f\circ g)(u,v)|\mathrm{J}_g(u,v)|\,d(u,v)\\ &= \int_0^1\int_{-v}^0 (u+v)^2 e^{-v^2}\,dudv = \frac{1}{3}\int_0^1 v^3 e^{-v^2}\,dv\\ &= \frac{1}{6}\int_0^1 we^{-w}\,dw = -\frac{1}{6}(w+1)e^{-w}\Big|_{w=0}^{w=1} = -\frac{1}{6}\big(\frac{2}{e}-1\big)\,.\end{aligned}$$

Example 3.36 (Polar coordinates)**.** In $\mathbb{R}^2$, when the domain over which the integral is taken is a disk D, a particular type of change of variables is sometimes very useful for the purpose of evaluating the integral. Let $(x,y) = (x_0 + r\cos\theta, y_0 + r\sin\theta) \equiv \psi(r,\theta)$, where (x_0, y_0) is the center of D under consideration. If the radius of D is R, then D, up to removing a line segment with length R, is the image of $(0,R)\times(0,2\pi)$ under ψ. Note that the Jacobian of ψ is

$$\mathrm{J}_\psi(r,\theta) = \begin{vmatrix} \dfrac{\partial\psi_1}{\partial r} & \dfrac{\partial\psi_1}{\partial\theta} \\ \dfrac{\partial\psi_2}{\partial r} & \dfrac{\partial\psi_2}{\partial\theta}\end{vmatrix} = \begin{vmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix} = r\,.$$

Therefore, if $f : \mathrm{D}\to\mathbb{R}$ is Riemann integrable, then

$$\begin{aligned}\int_{\mathrm{D}} f(x,y)\,d(x,y) &= \int_{\psi((0,R)\times(0,2\pi))} f(x,y)\,d(x,y) = \int_{(0,R)\times(0,2\pi)} (f\circ\psi)(r,\theta)\big|\mathrm{J}_\psi(r,\theta)\big|\,d(r,\theta)\\ &= \int_{(0,R)\times(0,2\pi)} f(x_0 + r\cos\theta, y_0 + r\sin\theta)\,r\,d(r,\theta)\,.\end{aligned}$$

Example 3.37 (Cylindrical coordinates)**.** In $\mathbb{R}^3$, when the domain over which the integral is taken is a cylinder C; that is, $\mathrm{C} = \mathrm{D} \times [a,b]$ for some disk D and $-\infty < a < b < \mathbb{R}$, then the change of variables

$$\psi(r,\theta,z) = (x_0 + r\cos\theta, y_0 + r\sin\theta, z) \qquad 0 < r < R\,, 0 < \theta < 2\pi\,, a \leqslant z \leqslant b\,,$$

where (x_0, y_0) is the center of D and R is the radisu of D, is sometimes very useful for evaluating the integral. Since the Jacobian of ψ is

$$\mathrm{J}_\psi(r,\theta,z) = \begin{vmatrix} \dfrac{\partial\psi_1}{\partial r} & \dfrac{\partial\psi_1}{\partial\theta} & \dfrac{\partial\psi_1}{\partial z} \\ \dfrac{\partial\psi_2}{\partial r} & \dfrac{\partial\psi_2}{\partial\theta} & \dfrac{\partial\psi_2}{\partial z} \\ \dfrac{\partial\psi_3}{\partial r} & \dfrac{\partial\psi_3}{\partial\theta} & \dfrac{\partial\psi_3}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r\,,$$

we must have

$$\begin{aligned} \int_{\mathrm{C}} f(x,y,z)\,d(x,y,z) &= \int_{\psi((0,R)\times(0,2\pi)\times[a,b])} f(x,y,z)\,d(x,y,z) \\ &= \int_{(0,R)\times(0,2\pi)\times[a,b]} (f\circ\psi)(r,\theta,z)\big|\mathrm{J}_\psi(r,\theta,z)\big|\,d(r,\theta,z) \\ &= \int_{(0,R)\times(0,2\pi)\times[a,b]} f(x_0 + r\cos\theta, y_0 + r\sin\theta, z)\,r\,d(r,\theta,z)\,. \end{aligned}$$

Example 3.38 (Spherical coordinates)**.** In $\mathbb{R}^3$, when the domain over which the integral is taken is a ball B, the change of variables

$$\psi(\rho,\theta,\phi) = (x_0 + \rho\cos\theta\sin\phi, y_0 + \rho\sin\theta\sin\phi, z_0 + \rho\cos\phi) \quad 0 < \rho < R, 0 < \theta < 2\pi, 0 < \phi < \pi,$$

where (x_0, y_0, z_0) is the center of B and R is the radius of B, is often used to evaluate the integral a function over B. Since the Jacobian of ψ is

$$\begin{aligned} \mathrm{J}_\psi(\rho,\theta,\phi) &= \begin{vmatrix} \dfrac{\partial\psi_1}{\partial\rho} & \dfrac{\partial\psi_1}{\partial\theta} & \dfrac{\partial\psi_1}{\partial\phi} \\ \dfrac{\partial\psi_2}{\partial\rho} & \dfrac{\partial\psi_2}{\partial\theta} & \dfrac{\partial\psi_2}{\partial\phi} \\ \dfrac{\partial\psi_3}{\partial\rho} & \dfrac{\partial\psi_3}{\partial\theta} & \dfrac{\partial\psi_3}{\partial\phi} \end{vmatrix} = \begin{vmatrix} \cos\theta\sin\phi & -\rho\sin\theta\sin\phi & \rho\cos\theta\cos\phi \\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \\ \cos\phi & 0 & -\rho\sin\phi \end{vmatrix} \\ &= -\rho^2\cos^2\theta\sin^3\phi - \rho^2\sin^2\theta\sin\phi\cos^2\phi - \rho^2\cos^2\theta\sin\phi\cos^2\phi - \rho^2\sin^2\theta\sin^3\phi \\ &= -\rho^2\sin^3\phi - \rho^2\sin\phi\cos^2\phi = -\rho^2\sin\phi\,, \end{aligned}$$

if the radius of B is R, we must have

$$\begin{aligned}\int_{\mathrm{B}} f(x,y,z)\,d(x,y,z) &= \int_{\psi((0,R)\times(0,2\pi)\times(0,\pi))} f(x,y,z)\,d(x,y,z)\\ &= \int_{(0,R)\times(0,2\pi)\times(0,\pi)} (f\circ\psi)(\rho,\theta,\phi)\big|\mathrm{J}_\psi(\rho,\theta,\phi)\big|\,d(\rho,\theta,\phi)\\ &= \int_{(0,R)\times(0,2\pi)\times(0,\pi)} f(x_0+\rho\cos\theta\sin\phi, y_0+\rho\sin\theta\sin\phi, z_0+\rho\cos\phi)\,\rho^2\sin\phi\,d(r,\theta,z)\,.\end{aligned}$$

Chapter 4

Vector Calculus

4.1 The Line Integrals

4.1.1 Curves

Definition 4.1. A subset $C \subseteq \mathbb{R}^n$ is called a ***curve*** if C is the image of an interval $I \subseteq \mathbb{R}$ under the continuous map $\gamma : I \to \mathbb{R}^n$ (that is, $C = \gamma(I)$). The continuous map $\gamma : I \to \mathbb{R}^n$ is called a ***parametrization*** of the curve. A curve C is called ***simple*** if it has an injective parametrization; that is, there exists $\gamma : I \to \mathbb{R}^n$ such that $\gamma(I) = C$ and $\gamma(x) = \gamma(y)$ implies that $x = y$. A curve C with parametrization $\gamma : I \to \mathbb{R}^n$ is called ***closed*** if $I = [a, b]$ for some closed interval $[a, b] \subseteq \mathbb{R}$ and $\gamma(a) = \gamma(b)$. A ***simple closed*** curve C is a closed curve with parametrization $\gamma : [a, b] \to \mathbb{R}^n$ such that γ is one-to-one on (a, b).

Example 4.2. A line segment joining two points $P_0, P_1 \in \mathbb{R}^n$ is a curve. It can be parameterized by $\gamma : [0, 1] \to \mathbb{R}^n$ defined by $\gamma(t) = tP_1 + (1 - t)P_0$.

Example 4.3. A circle on the plane is a simple closed curve. In fact, a circle centered at the (x_0, y_0) with radius r has the following parametrization: $\gamma : [0, 2\pi] \to \mathbb{R}^2$ defined by $\gamma(\theta) = (x_0 + r\cos\theta, y_0 + r\sin\theta)$.

Example 4.4. Figure eight is the zero level set of $F(x, y) = x^4 - a^2(x^2 - y^2)$ for some $a \neq 0$. It can also be parameterized by $\gamma : [0, 4\pi] \to \mathbb{R}^2$ defined by $\gamma(\theta) = \big(a\cos\frac{\theta}{2}, \frac{a}{2}\sin\theta\big)$.

Definition 4.5 (Length of Curves)**.** The length of curve $C \subseteq \mathbb{R}^n$ parameterized by $\gamma : [a, b] \to \mathbb{R}^n$ is defined as the number

$$\ell(C) \equiv \sup\left\{ \sum_{i=1}^{k} \big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^n} \,\middle|\, k \in \mathbb{N} \text{ and } a = t_0 < t_1 < \cdots < t_k = b \right\}.$$

Definition 4.6 (Rectifiable curves)**.** A curve $C \subseteq \mathbb{R}^n$ with parametrization $\gamma : I \to \mathbb{R}^n$ is called ***rectifiable*** if there is an homeomorphism $\varphi : \widetilde{I} \to I$, where $\widetilde{I}$ is again an interval, such that the map $\gamma \circ \varphi : \widetilde{I} \to \mathbb{R}^n$ is Lipschitz.

Remark 4.7. 1. By an homeomorphism it means a continuous bijection whose inverse is also continuous.

2. We can think of a curve as an equivalence class of continuous maps $\gamma : I \to \mathbb{R}^n$, where two parametrization $\gamma : I \to \mathbb{R}^n$ and $\widetilde{\gamma} : \widetilde{I} \to \mathbb{R}^n$ are equivalent if and only if there is an homeomorphism $\varphi : \widetilde{I} \to I$ such that $\widetilde{\gamma} = \gamma \circ \varphi$. Each element of the equivalence class is a parametrization of the curve and thus a rectifiable curve is a curve which has a Lipschitz continuous parametrization.

3. The length of a rectifiable curve parameterized by $\gamma : [a, b] \to \mathbb{R}^n$ is finite since by choosing a Lipschitz parametrization $\widetilde{\gamma} : [c, d] \to \mathbb{R}^n$, the number

$$\Big\{ \sum_{i=1}^{k} \big\|\widetilde{\gamma}(t_i) - \widetilde{\gamma}(t_{i-1})\big\|_{\mathbb{R}^n} \,\Big|\, k \in \mathbb{N} \text{ and } c = t_0 < t_1 < \cdots < t_k = d \Big\}$$

is bounded from above by $M(d - c)$, where M is the Lipschitz constant of $\widetilde{\gamma}$.

Example 4.8 (Non-rectifiable curves)**.** Let $C \subseteq \mathbb{R}^2$ be a curve parameterized by

$$\gamma(t) = \begin{cases} \big(t, t\sin\dfrac{\pi}{t}\big) & \text{if } t \in (0, 1]\,, \\ (0, 0) & \text{if } t = 0\,. \end{cases}$$

Since

$$\ell\big(\gamma([\frac{1}{n+1}, \frac{1}{n}])\big) \geqslant \big\|\gamma(\frac{1}{n+1}) - \gamma(\frac{1}{n+1/2})\big\|_{\mathbb{R}^2} + \big\|\gamma(\frac{1}{n+1/2}) - \gamma(\frac{1}{n})\big\|_{\mathbb{R}^2} \geqslant \frac{2}{n+1/2}$$

and $\sum\limits_{n=1}^{\infty} \dfrac{2}{n+1/2} = \infty$, by the remark above we conclude that $\gamma([0, 1])$ is not a rectifiable curve.

Definition 4.9. A curve $C \subseteq \mathbb{R}^n$ is said to be of class $\mathscr{C}^k$ or a $\mathscr{C}^k$-curve if there exists a parametrization $\gamma : I \to \mathbb{R}^n$ such that γ is k-times continuously differentiable. Such a parametrization is called a $\mathscr{C}^k$-parametrization of the curve. If there exists a parametrization $\gamma : I \to \mathbb{R}$ which is of class $\mathscr{C}^k$ for all $k \in \mathbb{N}$, then the curve is said to be ***smooth***. A curve $C \subseteq \mathbb{R}^n$ is said to be ***regular*** if there exists a $\mathscr{C}^1$-parametrization $\gamma : I \to \mathbb{R}^n$ such that $\gamma'(t) \neq \mathbf{0}$ for all $t \in I$.

Theorem 4.10. *Let $C \subseteq \mathbb{R}^n$ be a curve with $\mathscr{C}^1$-parametrization $\gamma : [a,b] \to \mathbb{R}^n$. Then*

$$\ell(C) = \int_a^b \|\gamma'(t)\|_{\mathbb{R}^n}\, dt\,.$$

Proof. Let $\varepsilon > 0$ be given. Since $\gamma : [a,b] \to \mathbb{R}^n$ is $\mathscr{C}^1$, there exists $\delta > 0$ such that

$$\big\|\gamma'(t) - \gamma'(s)\big\|_{\mathbb{R}^n} < \frac{\varepsilon}{4\sqrt{n}(b-a)} \quad \text{whenever} \quad s,t \in [a,b], |s-t| < \delta\,.$$

By the definition of the length of curves, there exists a partition $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_k = b\}$ of $[a,b]$ such that

$$\ell(C) - \frac{\varepsilon}{4} < \sum_{i=1}^k \big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^n} \leqslant \ell(C)\,.$$

W.L.O.G., we can assume that $\|\mathcal{P}\| < \delta$. For each component γ_j of γ, the mean value theorem implies that for some $\xi_i \in [t_{i-1}, t_i]$,

$$\gamma_j(t_i) - \gamma_j(t_{i-1}) = \gamma_j'(\xi_i)(t_i - t_{i-1})\,;$$

thus for each $i \in \{1, \cdots, k\}$ and $s_i \in [t_{i-1}, t_i]$,

$$\big|\gamma_j(t_i) - \gamma_j(t_{i-1}) - \gamma_j'(s_i)(t_i - t_{i-1})\big| \leqslant \big|\gamma_j'(\xi_i) - \gamma_j'(s_i)\big||t_i - t_{i-1}| < \frac{\varepsilon}{4\sqrt{n}(b-a)}|t_i - t_{i-1}|\,.$$

As a consequence, for each $i \in \{1, \cdots, k\}$ and $s_i \in [t_{i-1}, t_i]$,

$$\begin{aligned}
&\Big|\big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^n} - \big\|\gamma'(s_i)\big\|_{\mathbb{R}^n}|t_i - t_{i-1}|\Big| < \Big|\big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^n} - \big\|\gamma'(s_i)(t_i - t_{i-1})\big\|_{\mathbb{R}^n}\Big| \\
&\quad \leqslant \big\|\gamma(t_i) - \gamma(t_{i-1}) - \gamma'(s_i)(t_i - t_{i-1})\big\|_{\mathbb{R}^n} \leqslant \Big[\sum_{j=1}^n \Big(\frac{\varepsilon}{4\sqrt{n}(b-a)}|t_i - t_{i-1}|\Big)^2\Big]^{\frac{1}{2}} \\
&\quad < \frac{\varepsilon}{4(b-a)}|t_i - t_{i-1}|
\end{aligned}$$

which further implies that

$$\Big|\sum_{i=1}^k \big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^n} - \sum_{i=1}^k \big\|\gamma'(s_i)\big\|_{\mathbb{R}^n}|t_i - t_{i-1}|\Big| < \frac{\varepsilon}{4}\,.$$

Therefore, for $a = t_0 \leqslant s_0 \leqslant t_1 \leqslant s_1 \cdots \leqslant s_k \leqslant t_k = b$,

$$\ell(C) - \frac{\varepsilon}{2} < \sum_{i=1}^k \big\|\gamma'(s_i)\big\|_{\mathbb{R}^n}|t_i - t_{i-1}| < \ell(C) + \frac{\varepsilon}{2}\,.$$

Since $\|\gamma'\|$ is Riemann integrable over $[a,b]$, we must have

$$\ell(C) - \varepsilon < L(\|\gamma'\|_{\mathbb{R}^n}, \mathcal{P}) \leqslant \int_a^b \big\|\gamma'(t)\big\|_{\mathbb{R}^n} dt \leqslant U(\|\gamma'\|_{\mathbb{R}^n}, \mathcal{P}) < \ell(C) + \varepsilon\,,$$

and the theorem is concluded because $\varepsilon > 0$ is given arbitrarily. $\square$

Example 4.11. The length of the elliptic helix C parameterized by

$$\gamma(t) = (a\cos t, b\sin t, ct) \qquad t \in \big[0, \frac{\pi}{2}\big]$$

can be computed by

$$\ell(C) = \int_0^{\frac{\pi}{2}} \|\gamma'(t)\|_{\mathbb{R}^3} dt = \int_0^{\frac{\pi}{2}} \sqrt{a^2\sin^2 t + b^2\cos^2 t + c^2}\, dt\,.$$

1. When $a < b$, letting $k = \sqrt{\dfrac{b^2 - a^2}{b^2 + c^2}}$, then

$$\ell(C) = \sqrt{b^2 + c^2} \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2\sin^2 t}\, dt\,.$$

2. When $a > b$, letting $k = \sqrt{\dfrac{a^2 - b^2}{a^2 + c^2}}$, then

$$\ell(C) = \sqrt{a^2 + c^2} \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2\cos^2 t}\, dt = \sqrt{a^2 + c^2} \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2\sin^2 t}\, dt\,.$$

The integral $E(k,\phi) \equiv \displaystyle\int_0^{\phi} \sqrt{1 - k^2\sin^2 t}\, dt$, where $0 < k^2 < 1$, is called the ***elliptic integral function of the second kind***, and $\mathrm{E}(k) \equiv E\big(k, \frac{\pi}{2}\big)$ is called the ***complete elliptic integral of the second kind***.

Definition 4.12. Let $C \subseteq \mathbb{R}^{\mathrm{n}}$ be a curve with finite length. An ***arc-length parametrization*** of C is an injective parametrization $\gamma : [a,b] \to \mathbb{R}^{\mathrm{n}}$ such that the length of the curve $\gamma([a,s])$ is exactly $s - a$; that is,

$$\ell\big(\gamma([a,s])\big) = s - a \qquad \forall\, s \in [a,b]\,.$$

Example 4.13. Let C be the circle centered at the origin with radius R. Then the parametrization

$$\gamma(s) = \big(R\cos\frac{s}{R}, R\sin\frac{s}{R}\big) \qquad s \in [0, 2\pi R]\,,$$

is an arc-length parametrization of C. To see this, we note that

$$\ell\big(\gamma([0,s])\big) = \int_0^s \big\|\gamma'(t)\big\|_{\mathbb{R}^2} dt = \int_0^s \big\|\big(-\sin\frac{s}{R}, \cos\frac{s}{R}\big)\big\|_{\mathbb{R}^2} dt = \int_0^s dt = s \qquad \forall\, s \in [0, 2\pi R]\,.$$

In general, the arc-length parametrization of a rectifiable curve exists, and we have the following

Theorem 4.14. *Let $C \subseteq \mathbb{R}^n$ be a rectifiable simple curve. Then there exists a arc-length parametrization of C.*

Proof. We only prove the case that C has a regular $\mathscr{C}^1$-parametrization $\gamma : [a, b] \to \mathbb{R}^n$.

Let $s(t) = \displaystyle\int_a^t \|\gamma'(t')\|_{\mathbb{R}^n}\,dt'$. Note that the $s : [a, b] \to \mathbb{R}$ is strictly increasing since the fundamental theorem of Calculus implies that $s'(t) = \|\gamma'(t)\|_{\mathbb{R}^n} > 0$ for all $t \in [a, b]$. The Inverse Function Theorem (Theorem A.10) then guarantees the existence of a $\mathscr{C}^1$-inverse $u : [0, \ell(C)] \to [a, b]$ and we have $u'(t) = \dfrac{1}{s'(u(t))}$. Define $\widetilde{\gamma} = \gamma \circ u$. Then the chain rule implies that $\widetilde{\gamma} : [0, \ell(C)] \to \mathbb{R}^n$ is a $\mathscr{C}^1$-parametrization of C, and Theorem 4.10 implies that

$$\begin{aligned}\ell\big(\widetilde{\gamma}([0, s])\big) &= \int_0^s \|\widetilde{\gamma}'(t)\|_{\mathbb{R}^n}\,dt = \int_0^s \|\gamma'(u(t))u'(t)\|_{\mathbb{R}^n}\,dt = \int_0^s \|\gamma'(u(t))\|_{\mathbb{R}^n}\big|u'(t)\big|\,dt \\ &= \int_0^s s'(u(t))\frac{1}{\big|s'(u(t))\big|}\,dt = \int_0^s dt = s\end{aligned}$$

which implies that $\widetilde{\gamma} : [0, \ell(C)]$ is an arc-length parametrization of C. □

Theorem 4.15. *Let $C \subseteq \mathbb{R}^n$ be a $\mathscr{C}^1$-curve with an arc-length parametrization $\gamma : I \to \mathbb{R}^n$. Then $\|\gamma'(s)\|_{\mathbb{R}^n} = 1$ for all $s \in I$.*

Proof. Suppose that $I = [a, b]$. Since $\gamma : I \to \mathbb{R}^n$ is an arc-length parametrization of C, we must have

$$s - a = \int_a^s \|\gamma'(t)\|_{\mathbb{R}^n}\,dt \qquad \forall\, t \in I\,.$$

Differentiating both sides of the equality above in t, the fundamental theorem of Calculus implies that $1 = \|\gamma'(s)\|_{\mathbb{R}^n}$ for all $s \in I$. □

4.1.2 The line element and line integrals

Line elements

Definition 4.16. A curve $C \subseteq \mathbb{R}^n$ is said to be piecewise $\mathscr{C}^k$ (smooth, regular) if there exists a parametrization $\gamma : [a, b] \to \mathbb{R}^n$ and a finite set of points $\{a = t_0 < t_1 < \cdots < t_N = b\}$ such that $\gamma : [t_i, t_{i+1}] \to \mathbb{R}^n$ is $\mathscr{C}^k$ (smooth, regular) for all $i \in \{0, 1, \cdots, N-1\}$.

Definition 4.17. Let $\mathscr{R}_{\mathcal{C}}$ be the collection of all piecewise regular curves. The line element is a set function $s : \mathscr{R}_{\mathcal{C}} \to \mathbb{R}$ that satisfies the following properties:

1. $s(C) > 0$ for all $C \in \mathscr{R}_{\mathcal{C}}$.

2. If $C \in \mathscr{R}_{\mathcal{C}}$ is the union of finitely many regular curves $C_1, \cdots, C_k$ that do not overlap except at their end-points, then
$$s(C) = s(C_1) + \cdots + s(C_k)\,.$$

3. The value of s agrees with the length on straight line segments; that is,
$$s(L) = \ell(L) \qquad \text{for all line segaments } L\,.$$

Line integrals of scalar functions

Definition 4.18. Let $C \subseteq \mathbb{R}^{\mathrm{n}}$ be a simple rectifiable curve with an injective Lipschitz parametrization $\gamma : [a,b] \to \mathbb{R}^{\mathrm{n}}$, and $f : C \to \mathbb{R}$ be a real-valued function. The ***line integral*** of f along C, denoted by $\displaystyle\int_C f\,ds$, is the number

$$\sup\Big\{\sum_{i=1}^{k}\big(\inf_{\xi\in\gamma([t_{i-1},t_i])} f(\xi)\big)\ell\big(\gamma([t_{i-1},t_i])\big)\,\Big|\,k\in\mathbb{N}, a=t_0<t_1<\cdots<t_k=b\Big\}$$

provided that it is identical to

$$\inf\Big\{\sum_{i=1}^{k}\big(\sup_{\xi\in\gamma([t_{i-1},t_i])} f(\xi)\big)\ell\big(\gamma([t_{i-1},t_i])\big)\,\Big|\,k\in\mathbb{N}, a=t_0<t_1<\cdots<t_k=b\Big\}\,.$$

When C is a closed curve, we also use $\displaystyle\oint_C f\,ds$ to denote the line integral of f along C to emphasize that the curve C is a closed loop.

Remark 4.19. Since the parametrization γ is required to be injective, the line integral of f along C is independent of the choice of the parametrization.

Remark 4.20. In particular, if $f \equiv 1$, then $\ell(C) = \displaystyle\int_C 1\,ds \equiv \int_C ds$.

Remark 4.21. If the curve C is a line segment $\big\{(x,0)\,\big|\,a \leqslant x \leqslant b\big\}$, then the line integral of f along C is simply the Riemann integral of f over $[a,b]$ (by treating f as a function of x).

Remark 4.22 (The interpretation of the line integrals)**.** Let C be a piecewise smooth curve, and $f(x)$ denote the density of the curve C at position x. Suppose that f is continuous on C and $x = \gamma(t)$. Then $f(x)$ is computed by

$$f(x) = f(\gamma(t)) = \lim_{\Delta t \to 0} \frac{\mathrm{m}(\gamma([t, t+\Delta t]))}{\ell(\gamma([t, t+\Delta t]))},$$

where $\mathrm{m}(\cdot)$ denotes the mass. Let $\varepsilon > 0$ be given. Then by the continuity of $f \circ \gamma$ and the definition of limit, there exists $\delta > 0$ such that

$$\big|(f \circ \gamma)(t) - (f \circ \gamma)(s)\big| < \frac{\varepsilon}{4\ell(C)} \quad \text{if} \quad t, s \in [a, b],\ |t - s| < \delta$$

and

$$\big|f(\gamma(t))\ell(\gamma([t, t+\Delta t])) - \mathrm{m}(\gamma([t, t+\Delta t]))\big| \leqslant \ell(\gamma([t, t+\Delta t]))\frac{\varepsilon}{4\ell(C)} \quad \text{if} \quad |\Delta t| < \delta\,;$$

thus if $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_k = b\}$ is a partition of $[a, b]$ with $\|\mathcal{P}\| < \delta$, the total mass of the curve $\mathrm{m}(C)$, given by $\mathrm{m}(C) = \sum\limits_{i=1}^{k} \mathrm{m}(\gamma([t_{i-1}, t_i]))$, validates the following estimate:

$$\Big|\mathrm{m}(C) - \sum_{i=1}^{k} f(\gamma(s_{i-1}))\ell(\gamma([t_{i-1}, t_i]))\Big| \leqslant \frac{\varepsilon}{2}\,.$$

As a consequence,

$$\mathrm{m}(C) - \varepsilon < \sum_{i=1}^{k} \inf_{\xi \in \gamma([t_{i-1}, t_i])} f(\xi)\ell(\gamma([t_{i-1}, t_i])) \leqslant \sum_{i=1}^{k} \sup_{\xi \in \gamma([t_{i-1}, t_i])} f(\xi)\ell(\gamma([t_{i-1}, t_i])) < \mathrm{m}(C) + \varepsilon$$

which implies that the line integral of f along C is exactly the mass of the curve.

Theorem 4.23. *Let $C \subseteq \mathbb{R}^{\mathrm{n}}$ be a simple curve with $\mathscr{C}^1$-parametrization $\gamma : [a, b] \to \mathbb{R}^{\mathrm{n}}$, and $f : C \to \mathbb{R}$ be a real-valued continuous function. Then*

$$\int_C f\, ds = \int_a^b f\big(\gamma(t)\big)\|\gamma'(t)\|_{\mathbb{R}^{\mathrm{n}}}\, dt\,. \tag{4.1}$$

Proof. Let $\varepsilon > 0$ be given. Since $f \circ \gamma$ and γ' are continuous on $[a, b]$, $|f \circ \gamma| + \|\gamma'\|_{\mathbb{R}^{\mathrm{n}}} \leqslant M$ on $[a, b]$ for some $M > 0$, and there exists $\delta > 0$ such that

$$\big|(f \circ \gamma)(s) - (f \circ \gamma)(t)\big| < \frac{\varepsilon}{8(M+1)(b-a)} \quad \text{whenever} \quad s, t \in [a, b],\ |s - t| < \delta$$

and

$$\big\|\gamma'(s)-\gamma'(t)\big\|_{\mathbb{R}^n} < \frac{\varepsilon}{8(M+1)(b-a)} \quad \text{whenever} \quad s,t\in[a,b],\ |s-t|<\delta\,.$$

Moreover, since $f\circ\gamma$ and γ' are both continuous on $[a,b]$, the integral $\displaystyle\int_a^b f\big(\gamma(t)\big)\|\gamma'(t)\|_{\mathbb{R}^n}\,dt$ exists; thus there exists a partition $\mathcal{P}=\{a=t_0<t_1<\cdots<t_k=b\}$ of $[a,b]$ with $\|\mathcal{P}\|<\delta$ such that

$$\sum_{i=1}^k\Big(\sup_{s\in[t_{i-1},t_i]}\big(f(\gamma(s))\|\gamma'(s)\|_{\mathbb{R}^n}\big)-\inf_{s\in[t_{i-1},t_i]}\big(f(\gamma(s))\|\gamma'(s)\|_{\mathbb{R}^n}\big)\Big)|t_i-t_{i-1}|<\frac{\varepsilon}{2}\,. \tag{4.2}$$

Let $s_i,r_i\in[t_{i-1},t_i]$ be such that

$$\sup_{t\in[t_{i-1},t_i]}\big(f(\gamma(t))\|\gamma'(t)\|_{\mathbb{R}^n}\big)=f(\gamma(s_i))\big\|\gamma'(s_i)\big\|_{\mathbb{R}^n} \quad \text{and} \quad \sup_{\xi\in\gamma([t_{i-1},t_i])}f(\xi)=f(\gamma(r_i))\,.$$

Moreover, by Theorem 4.10 and the mean value theorem for integrals, there exists $q_i\in[t_{i-1},t_i]$ such that

$$\ell(\gamma([t_{i-1},t_i]))=\int_{t_{i-1}}^{t_i}\big\|\gamma'(s)\big\|_{\mathbb{R}^n}\,ds=\big\|\gamma'(q_i)\big\|_{\mathbb{R}^n}|t_i-t_{i-1}|\,;$$

thus

$$\Big|\ell(\gamma([t_{i-1},t_i]))-\big\|\gamma'(s_i)\big\|_{\mathbb{R}^n}|t_i-t_{i-1}|\Big|\leqslant\frac{\varepsilon}{8(M+1)(b-a)}|t_i-t_{i-1}|\,.$$

Therefore, by the fact that $s_i,r_i,q_i\in[t_{i-1},t_i]$ and $|t_i-t_{i-1}|<\delta$,

$$\begin{aligned}
&\Big|\sup_{s\in[t_{i-1},t_i]}\big(f(\gamma(s))\|\gamma'(s)\|_{\mathbb{R}^n}\big)|t_i-t_{i-1}|-\sup_{\xi\in\gamma([t_{i-1},t_i])}f(\xi)\ell(\gamma([t_{i-1},t_i]))\Big|\\
&\quad=\Big|f(\gamma(s_i))\big\|\gamma'(s_i)\big\|_{\mathbb{R}^n}-f(\gamma(r_i))\big\|\gamma'(q_i)\big\|_{\mathbb{R}^n}\Big||t_i-t_{i-1}|\\
&\quad\leqslant\big|f(\gamma(s_i))-f(\gamma(r_i)\big|\big\|\gamma'(s_i)\big\|_{\mathbb{R}^n}|t_i-t_{i-1}|+\big|f(\gamma(r_i))\big|\big\|\gamma'(s_i)-\gamma'(q_i)\big\|_{\mathbb{R}^n}|t_i-t_{i-1}|\\
&\quad<\frac{\varepsilon}{4(b-a)}|t_i-t_{i-1}|\,,
\end{aligned}$$

and summing the inequality above over i we obtain that

$$\Big|\sum_{i=1}^k\sup_{s\in[t_{i-1},t_i]}\big(f(\gamma(s))\|\gamma'(s)\|_{\mathbb{R}^n}\big)|t_i-t_{i-1}|-\sum_{i=1}^k\sup_{\xi\in\gamma([t_{i-1},t_i])}f(\xi)\ell(\gamma([t_{i-1},t_i]))\Big|<\frac{\varepsilon}{4}\,.$$

Similarly,

$$\Big|\sum_{i=1}^k\inf_{s\in[t_{i-1},t_i]}\big(f(\gamma(s))\|\gamma'(s)\|_{\mathbb{R}^n}\big)|t_i-t_{i-1}|-\sum_{i=1}^k\inf_{\xi\in\gamma([t_{i-1},t_i])}f(\xi)\ell(\gamma([t_{i-1},t_i]))\Big|<\frac{\varepsilon}{4}\,;$$

thus using (4.2) we find that

$$\begin{aligned}\int_a^b (f\circ\gamma)(t)\|\gamma'(t)\|_{\mathbb{R}^n}\,dt-\varepsilon &< L\big((f\circ\gamma)\|\gamma'\|_{\mathbb{R}^n},\mathcal{P}\big)-\frac{\varepsilon}{4}\\ &\leqslant \sum_{i=1}^k \inf_{\xi\in\gamma([t_{i-1},t_i])} f(\xi)\ell(\gamma([t_{i-1},t_i]))\leqslant \sum_{i=1}^k \sup_{\xi\in\gamma([t_{i-1},t_i])} f(\xi)\ell(\gamma([t_{i-1},t_i]))\\ &\leqslant U\big((f\circ\gamma)\|\gamma'\|_{\mathbb{R}^n},\mathcal{P}\big)+\frac{\varepsilon}{4}<\int_a^b (f\circ\gamma)(t)\|\gamma'(t)\|_{\mathbb{R}^n}\,dt+\varepsilon\,.\end{aligned}$$

Since $\varepsilon>0$ is chosen arbitrary, we conclude (4.1). ▫

Example 4.24. Let C be the upper half part of the circle centered at the origin with radius $R>0$ in the xy-plane. Evaluate the line integral $\displaystyle\int_C y\,ds$.

First, we parameterize C by

$$\gamma(t)=(R\cos t,R\sin t)\qquad t\in[0,\pi]\,.$$

Then

$$\int_C y\,ds=\int_0^\pi R\sin t\big\|(-R\sin t,R\cos t)\big\|_{\mathbb{R}^2}dt=\int_0^\pi R^2\sin t\,dt=2R^2\,.$$

Example 4.25. Find the mass of a wire lying along the first octant part of the curve of intersection of the elliptic paraboloid $z=2-x^2-2y^2$ and the parabolic cylinder $z=x^2$ between $(0,1,0)$ and $(1,0,1)$ if the density of the wire at position (x,y,z) is $\varrho(x,y,z)=xy$.

Note that we can parameterize the curve C by

$$\gamma(t)=(t,\sqrt{1-t^2},t^2)\qquad t\in[0,1]\,.$$

Therefore, the mass of the curve can be computed by

$$\begin{aligned}\int_C\varrho\,ds&=\int_0^1 t\sqrt{1-t^2}\,\Big\|(1,\frac{-t}{\sqrt{1-t^2}},2t)\Big\|_{\mathbb{R}^3}dt=\int_0^1 t\sqrt{1-t^2}\,\frac{\sqrt{1-t^2+t^2+4t^2(1-t^2)}}{\sqrt{1-t^2}}\,dt\\ &=\int_0^1 t\sqrt{2-(1-2t^2)^2}\,dt=\frac14\int_{-1}^1\sqrt{2-u^2}du=\frac12\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}2\cos^2\theta\,d\theta=\frac{\pi}{8}(\pi+2)\,.\end{aligned}$$

Line integrals of vector fields

We recall that a vector field is a vector-valued function whose domain and co-domain are subsets of identical Euclidean space $\mathbb{R}^n$.

Let C be a simple regular curve parameterized by $\gamma : I \to \mathbb{R}^n$, and $\boldsymbol{F} : C \to \mathbb{R}^n$ be a vector field. The ***line integral of F along** C **in the direction of*** γ (or the ***oriented line integral of F along*** C) is defined as the line integral of the scalar function $\boldsymbol{F} \cdot \mathbf{T}$ along C, where $\mathbf{T}$ is the unit tangent of C given by

$$\mathbf{T} = \frac{\gamma'}{\|\gamma'\|_{\mathbb{R}^n}} \circ \gamma^{-1} \quad \text{on} \quad C\,. \tag{4.3}$$

Given another parametrization $\phi : \widetilde{I} \to \mathbb{R}^n$ of C such that $(\phi' \circ \phi^{-1}) \cdot (\gamma' \circ \gamma^{-1}) > 0$ (that is, the orientation of C given by ϕ and γ are the same), using the chain rule we obtain that

$$\gamma' = \frac{d}{dt}(\phi \circ \phi^{-1} \circ \gamma)(t) = (\phi' \circ \phi^{-1} \circ \gamma)(t)(\phi^{-1} \circ \gamma)'(t)\,. \tag{4.4}$$

Since $\phi^{-1} \circ \gamma : I \to \widetilde{I}$, $(\phi^{-1} \circ \gamma)'$ is a scalar function; thus (4.4) and the fact that $(\phi' \circ \phi^{-1}) \cdot (\gamma' \circ \gamma^{-1}) > 0$ implies that $\gamma' \circ \gamma^{-1} = c(\phi' \circ \phi^{-1})$ for some positive scalar function $c : C \to \mathbb{R}$. Therefore,

$$\frac{\phi'}{\|\phi'\|_{\mathbb{R}^n}} \circ \phi^{-1} = \frac{\gamma'}{\|\gamma'\|_{\mathbb{R}^n}} \circ \gamma^{-1} \quad \text{on} \quad C\,. \tag{4.5}$$

In other words, the tangent vector $\mathbf{T}$ is well-defined on C; thus the line integral of $\boldsymbol{F}$ along C in the direction of the parametrization γ is a well-defined quantity.

Suppose that $I = [a, b]$. Using (4.1), we find that

$$\int_C \boldsymbol{F} \cdot \mathbf{T}\, ds = \int_a^b (\boldsymbol{F} \circ \gamma)(t) \cdot \frac{\gamma'(t)}{\|\gamma'(t)\|_{\mathbb{R}^n}} \|\gamma'(t)\|_{\mathbb{R}^n}\, dt = \int_a^b (\boldsymbol{F} \circ \gamma)(t) \cdot \gamma'(t)\, dt\,.$$

Let $\boldsymbol{r} : \widetilde{I} \to \mathbb{R}^n$ be an arc-length parametrization of C such that $(\boldsymbol{r}' \circ \boldsymbol{r}^{-1}) \cdot (\gamma' \circ \gamma^{-1}) > 0$ on C. Then (4.5) implies that $\mathbf{T} = \dfrac{d\boldsymbol{r}}{ds}$. In terms of notation, we also write $\mathbf{T}\, ds$ as $d\boldsymbol{r}$; thus

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int_C \boldsymbol{F} \cdot \mathbf{T}\, ds = \int_a^b (\boldsymbol{F} \circ \gamma)(t) \cdot \gamma'(t)\, dt\,.$$

Remark 4.26 (The interpretation of line integrals of vector fields)**.** Consider the work done by moving an object along a smooth curve C parameterized by $\gamma : I \to \mathbb{R}^n$ with a continuous variable force $\boldsymbol{F} : C \to \mathbb{R}^n$ from $\gamma(a)$ to $\gamma(b)$ (that is, in the direction of the parametrization of γ). Since the work done by a constant force is the inner product of the displacement and the force, we find the the work done by the force $\boldsymbol{F}$ along the small portion $\gamma([t_i, t_{i+1}])$, from $\gamma(t_i)$ to $\gamma(t_{i+1})$, of the curve, where $|t_i - t_{i+1}| \ll 1$, is approximately

$$(\boldsymbol{F} \cdot \mathbf{T})(\gamma(t_i))\ell\big(\gamma([t_i, t_{i+1}])\big) \equiv \boldsymbol{F}(\gamma(t_i)) \cdot \mathbf{T}(\gamma(t_i))\ell\big(\gamma([t_i, t_{i+1}])\big)\,.$$

Summing over all the portions, we conclude that the work done by the force $\boldsymbol{F}$ along the curve C, in the direction of the parametrization γ, is approximately

$$\sum_{i=0}^{k-1}(\boldsymbol{F}\cdot\mathbf{T})(\gamma(t_i))\ell\big(\gamma([t_i,t_{i+1}])\big)$$

which converges to the line integral $\displaystyle\int_C (\boldsymbol{F}\cdot\mathbf{T})\,ds$. Therefore, the line integral of vector fields $\boldsymbol{F}$ along C in the direction of the parametrization γ is simply the work done by the force $\boldsymbol{F}$ in moving an object along the curve C from the starting point to the end point.

Example 4.27. Let $\boldsymbol{F}(x,y)=(y^2,2xy)$. Evaluate the line integral $\displaystyle\int_C \boldsymbol{F}\cdot d\boldsymbol{r}$ from $(0,0)$ to $(1,1)$ along

1. the straight line $y=x$,
2. the curve $y=x^2$, and
3. the piecewise smooth path consisting of the straight line segments from $(0,0)$ to $(0,1)$ and from $(0,1)$ to $(1,1)$.

For the straight line case, we parameterize the path by $\gamma(t)=(t,t)$ for $t\in[0,1]$. Then

$$\int_C \boldsymbol{F}\cdot d\boldsymbol{r}=\int_0^1(t^2,2t^2)\cdot(1,1)dt=\int_0^1 3t^2dt=1\,.$$

For the case of parabola, we parameterize the path by $\gamma(t)=(t,t^2)$ for $t\in[0,1]$. Then

$$\int_C \boldsymbol{F}\cdot d\boldsymbol{r}=\int_0^1(t^4,2t^3)\cdot(1,2t)dt=\int_0^1 5t^4dt=1\,.$$

For the piecewise linear case, we let C_1 denote the line segment joining $(0,0)$ and $(0,1)$, and let C_2 denote the line segment joining $(0,1)$ and $(1,1)$. Note that we can parameterize C_1 and C_2 by

$$\gamma_1(t)=(0,t)\quad t\in[0,1]\quad\text{and}\quad\gamma_2(t)=(t,1)\quad t\in[0,1]\,,$$

respectively. Therefore,

$$\int_C \boldsymbol{F}\cdot d\boldsymbol{r}=\int_{C_1}\boldsymbol{F}\cdot d\boldsymbol{r}+\int_{C_2}\boldsymbol{F}\cdot d\boldsymbol{r}=\int_0^1(t^2,0)\cdot(0,1)\,dt+\int_0^1(1,2t)\cdot(1,0)\,dt=1\,.$$

We note that in this example the line integrals of $\boldsymbol{F}$ over three different paths joining $(0,0)$ and $(1,1)$ are identical.

Example 4.28. Let $\boldsymbol{F}(x,y) = (y,-x)$. Evaluate the line integral $\displaystyle\int_C \boldsymbol{F}\cdot d\boldsymbol{r}$ from $(1,0)$ to $(0,-1)$ along

1. the straight line segment joining these points, and
2. three-quarters of the circle of unit radius centered at the origin and traversed counter-clockwise.

For the first case, we parameterize the path by $\gamma(t) = (1-t,-t)$ for $t \in [0,1]$. Then

$$\int_C \boldsymbol{F}\cdot d\boldsymbol{r} = \int_0^1 (-t, t-1)\cdot(-1,-1)\,dt = \int_0^1 1\,dt = 1\,.$$

For the second case, we parameterize the path by $\gamma(t) = (\cos t, \sin t)$ for $t \in \left[0, \frac{3\pi}{2}\right]$. Then

$$\int_C \boldsymbol{F}\cdot d\boldsymbol{r} = \int_0^{\frac{3\pi}{2}} (\sin t, -\cos t)\cdot(-\sin t, \cos t)\,dt = \int_0^{\frac{3\pi}{2}} (-1)\,dt = -\frac{3\pi}{2}\,.$$

We note that in this example the line integrals of $\boldsymbol{F}$ over different paths joining $(1,0)$ and $(0,-1)$ might be different.

4.2 Conservative Vector Fields

In the previous section, we define the line integral of a force along a curve in a given orientation. In Example 4.27, we see that the line integrals along three different paths connecting two given points are the same, while in Example 4.28 the line integrals along two different paths (connecting two given points) are different. In this section, we are interested in the rule of judging whether the line integral is path independent or not.

Definition 4.29 (Conservative Fields)**.** A vector field $\mathbf{F} : \mathcal{D} \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ is said to be ***conservative*** if $\mathbf{F} = \nabla\phi$ for some scalar function $\varphi : \mathcal{D} \to \mathbb{R}$. Such a ϕ is called a (scalar) potential for $\mathbf{F}$ on $\mathcal{D}$.

Theorem 4.30. *Let $\mathcal{D}$ be an open, connected domain in $\mathbb{R}^{\mathrm{n}}$, and let $\mathbf{F}$ be a smooth vector field defined on $\mathcal{D}$. Then the following three statements are equivalent:*

(1) $\mathbf{F}$ *is conservative in $\mathcal{D}$.*

(2) $\displaystyle\oint_C \mathbf{F}\cdot d\mathbf{r} = 0$ *for every piecewise smooth, closed curve C in $\mathcal{D}$.*

(3) *Given and two point $P_0, P_1 \in \mathcal{D}$, $\displaystyle\int_C \mathbf{F}\cdot d\mathbf{r}$ has the same value for all piecewise smooth curves in $\mathcal{D}$ starting at P_0 and ending at P_1.*

Proof. (1) $\Rightarrow$ (2): Suppose that $\boldsymbol{F} = \nabla\phi$ in $\mathcal{D}$ for some scalar function $\phi : \mathcal{D} \to \mathbb{R}$. Let $C \subseteq \mathbb{R}^{\rm n}$ be a piecewise smooth closed curve parameterized by $\gamma : [a,b] \to \mathbb{R}^{\rm n}$ such that $\gamma : [t_{i-1}, t_i] \to \mathbb{R}^{\rm n}$ is smooth for all $1 \leqslant i \leqslant N$, where $a = t_0 < t_1 < \cdots < t_N = b$. Let $C_i = \gamma([t_{i-1}, t_i])$. Then the chain rule implies that

$$\begin{aligned}\oint_C \boldsymbol{F}\cdot d\boldsymbol{r} &= \sum_{i=1}^N \int_{C_i} \nabla\phi\cdot d\boldsymbol{r} = \sum_{i=1}^N \int_{t_{i-1}}^{t_i} (\nabla\phi\circ\gamma)(t)\cdot\gamma'(t)\,dt \\ &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \frac{d}{dt}(\phi\circ\gamma)(t)\,dt = \sum_{i=1}^N (\phi\circ\gamma)(t)\Big|_{t=t_{i-1}}^{t=t_i} = \phi(\gamma(b)) - \phi(\gamma(a)) = 0\,.\end{aligned}$$

(2) $\Rightarrow$ (3): Let C_1 and C_2 be two piecewise smooth curves in $\mathcal{D}$ starting at P_0 and ending at P_1 parameterized by $\gamma_1 : [a,b] \to \mathbb{R}^{\rm n}$ and $\gamma_2 : [c,d] \to \mathbb{R}^{\rm n}$, respectively. Define $\gamma : [a, b+d-c] \to \mathbb{R}^{\rm n}$ by

$$\gamma(t) = \begin{cases} \gamma_1(t) & \text{if } t \in [a,b]\,, \\ \gamma_2(b+d-t) & \text{if } t \in [b, b+d-c]\,. \end{cases}$$

Then $C = \gamma([a, b+d-c])$ is a piecewise smooth closed curve; thus

$$\begin{aligned} 0 = \oint_C \boldsymbol{F}\cdot d\boldsymbol{r} &= \int_a^b (\boldsymbol{F}\circ\gamma_1)(t)\cdot\gamma_1'(t)\,dt - \int_b^{b+d-c} (\boldsymbol{F}\circ\gamma_2)(b+d-t)\gamma_2'(b+d-t)\,dt \\ &= \int_{C_1} \boldsymbol{F}\cdot d\boldsymbol{r} - \int_c^d (\boldsymbol{F}\circ\gamma_2)(t)\gamma_2'(t)dt = \int_{C_1} \boldsymbol{F}\cdot d\boldsymbol{r} - \int_{C_2} \boldsymbol{F}\cdot d\boldsymbol{r}\,. \end{aligned}$$

(3) $\Rightarrow$ (1): Let $P_0 \in \mathcal{D}$. For $x \in \mathcal{D}$, define $\phi(x) = \displaystyle\int_C \boldsymbol{F}\cdot d\boldsymbol{r}$, where C is any piecewise smooth curve starting at P_0 and ending at x. Note that by assumption, $\phi : \mathcal{D} \to \mathbb{R}$ is well-defined.

Choose $\delta > 0$ such that $B(x,\delta) \subseteq \mathcal{D}$. Let C be a piecewise smooth curve joining P_0, and L be the line segment joining x and $x + h\mathrm{e}_j$, where $0 < h < \delta$ and $\mathrm{e}_j = (0, \cdots, 0, 1, 0, \cdots, 0)$ is the unit vector whose j-th component is 1. Then with the parametrization of L: $\gamma(t) = x + t\mathrm{e}_j$ for $t \in [0,h]$, we have

$$\frac{\phi(x+h\mathrm{e}_j) - \phi(x)}{h} = \frac{1}{h}\int_L \boldsymbol{F}\cdot d\boldsymbol{r} = \frac{1}{h}\int_0^h \boldsymbol{F}(x+t\mathrm{e}_j)\cdot\mathrm{e}_j\,dt\,;$$

thus passing to the limit as $h \to 0$, we find that

$$\frac{\partial \phi}{\partial x_j}(x) = \boldsymbol{F}(x) \cdot \mathrm{e}_j \,.$$

As a consequence, $\boldsymbol{F}(x) = (\nabla\phi)(x)$ which implies that $\boldsymbol{F}$ is conservative. $\square$

Let $\mathcal{D} \subseteq \mathbb{R}^2$, and $\boldsymbol{F} = (M, N) : \mathcal{D} \to \mathbb{R}^2$. If $\boldsymbol{F}$ is conservative, then $M = \phi_x$ and $N = \phi_y$ for some scalar function $\phi : \mathcal{D} \to \mathbb{R}$; thus if ϕ is of class $\mathscr{C}^2$, we must have $M_y = N_x$. In other words, if $\boldsymbol{F} : \mathcal{D} \to \mathbb{R}^2$ is a smooth vector field, then it is necessary that $M_y = N_x$. The converse statement is not true in general, and we have the following counter-example.

Example 4.31. Let $\mathcal{D} \subseteq \mathbb{R}^2$ be the annular region $\mathcal{D} = \big\{(x, y) \,\big|\, 1 < x^2 + y^2 < 4\big\}$, and consider the vector field $\boldsymbol{F}(x, y) = \big(\frac{y}{x^2+y^2}, \frac{-x}{x^2+y^2}\big)$. Then

$$\frac{\partial}{\partial y}\frac{y}{x^2+y^2} = \frac{x^2-y^2}{(x^2+y^2)^2} = \frac{\partial}{\partial x}\frac{-x}{x^2+y^2}\,;$$

however, if $\boldsymbol{F} = \nabla\phi$ for some differentiable scalar function $\phi : \mathcal{D} \to \mathbb{R}$, we must have

$$\phi_x(x, y) = \frac{y}{x^2+y^2}$$

which further implies that

$$\phi(x, y) = \arctan\frac{x}{y} + f(y)\,.$$

Using that $\phi_y(x, y) = \frac{y}{x^2+y^2}$, we conclude that f is a constant function; thus

$$\phi(x, y) = \arctan\frac{x}{y} + C\,.$$

Since ϕ is not differentiable on the positive x-axis, $\boldsymbol{F} \neq \nabla\phi$.

Definition 4.32. A connected domain $\mathcal{D}$ is said to be ***simply connected*** if every simple closed curve can be continuously shrunk to a point in $\mathcal{D}$ without any part ever passing out of $\mathcal{D}$.

Theorem 4.33. *Let $\mathcal{D} \subseteq \mathbb{R}^2$ be simply connected, and $\boldsymbol{F} = (M, N) : \mathcal{D} \to \mathbb{R}^2$ be of class $\mathscr{C}^1$. If $M_y = N_x$, then $\boldsymbol{F}$ is conservative.*

The theorem above can be proved using Theorem 4.30 and Green's theorem (Theorem 4.90), and is left till Section 4.8 (where Green's theorem is introduced).

4.3 The Surface Integrals

4.3.1 Surfaces

Definition 4.34. A subset $\Sigma \subseteq \mathbb{R}^3$ is called a surface if for each $p \in \Sigma$, there exist an open neighborhood $\mathcal{U} \subseteq \Sigma$ of p, an open set $\mathcal{V} \subseteq \mathbb{R}^2$, and a continuous map $\varphi : \mathcal{U} \to \mathcal{V}$ such that $\varphi : \mathcal{U} \to \mathcal{V}$ is one-to-one, onto, and its inverse $\psi = \varphi^{-1}$ is also continuous. Such a pair $\{\mathcal{U}, \varphi\}$ is called a coordinate chart (or simply chart) at p, and $\{\mathcal{V}, \psi\}$ is called a (local) parametrization at p.

Remark 4.35. In some literatures the surface is defined in the following equivalent but reversed way: A subset $\Sigma \subseteq \mathbb{R}^3$ is a surface if for each $p \in \Sigma$, there exists a neighborhood $\mathcal{U} \subseteq \mathbb{R}^3$ of p and a map $\psi : \mathcal{V} \to \mathcal{U} \cap \Sigma$ of an open set $\mathcal{V} \subseteq \mathbb{R}^2$ onto $\mathcal{U} \cap \Sigma \subseteq \mathbb{R}^3$ such that ψ is a homeomorphism; that is, ψ has an inverse $\varphi = \psi^{-1} : \mathcal{U} \cap \Sigma \to \mathcal{V}$ which is continuous. The mapping ψ is called a parametrization or a system of (local) coordinates in (a neighborhood of) p.

Definition 4.36 (Regular surfaces)**.** A surface $\Sigma \subseteq \mathbb{R}^3$ is said to be regular if for each $p \in \Sigma$, there exists a differentiable local parametrization $\{\mathcal{V}, \psi\}$ of Σ at p such that $D\psi(q)$, the derivative of ψ at q, has full rank for all $q \in \mathcal{V}$; that is, $D\psi(q) : \mathbb{R}^2 \to \mathbb{R}^3$ is one-to-one for all $q \in \mathcal{V}$. The range of the map $D\psi\big(\psi^{-1}(p)\big)$ is called the ***tangent plane*** of Σ at p, and is denoted by $\mathbf{T}_p\Sigma$.

In the following, we always assume that $D\psi(q)$ has full rank for all $q \in \mathcal{V}$ if $\{\mathcal{V}, \psi\}$ is a local parametrization of a regular surface $\Sigma \subseteq \mathbb{R}^3$.

Remark 4.37. Write $\psi : \mathcal{V} \to \Sigma$ as

$$\psi(u, v) = \big(x(u, v), y(u, v), z(u, v)\big) .$$

Then if $q = (u_0, v_0)$,

$$\big[(D\psi)(q)\big] = \begin{bmatrix} x_u(u_0, v_0) & x_v(u_0, v_0) \\ y_u(u_0, v_0) & y_v(u_0, v_0) \\ z_u(u_0, v_0) & z_v(u_0, v_0) \end{bmatrix} = \big[[\psi_{,1}]\vdots[\psi_{,2}]\big] .$$

The injectivity of $D\psi(q)$ is then translated to that the two vectors

$$\psi_{,1}(u_0, v_0) \equiv \psi_u(u_0, v_0) = \big(x_u(u_0, v_0), y_u(u_0, v_0), z_u(u_0, v_0)\big)$$
$$\psi_{,2}(u_0, v_0) \equiv \psi_v(u_0, v_0) = \big(x_v(u_0, v_0), y_v(u_0, v_0), z_v(u_0, v_0)\big)$$

are linearly independent. Therefore, the range of $D\psi(q)$ is the span of the two vectors $\psi,_1(q)$ and $\psi,_2(q)$ and is indeed a plane for all $q \in \mathcal{V}$.

Let $p \in \Sigma$ and $q = \psi^{-1}(p)$. Since $D\psi(q)$ is injective, each $\boldsymbol{v} \in \mathbf{T}_p\Sigma$ corresponds a unique vector $(a, b) \in \mathbb{R}^2$ such that $\boldsymbol{v} = a\psi,_1(q) + b\psi,_2(q)$. This vector $(a, b) \in \mathbb{R}^2$ satisfies $[\boldsymbol{v}] = \big[D\psi(q)\big][a, b]^{\mathrm{T}}$, and can be computed by

$$\begin{bmatrix} a \\ b \end{bmatrix} = \Big(\big[D\psi(q)\big]^{\mathrm{T}}\big[D\psi(q)\big]\Big)^{-1}\big[D\psi(q)\big]^{\mathrm{T}}[\boldsymbol{v}]\,.$$

Example 4.38. Let $\mathbb{S}^2 = \big\{(x, y, z) \in \mathbb{R}^3 \,\big|\, x^2 + y^2 + z^2 = 1\big\}$ be the unit sphere in $\mathbb{R}^3$. If $p = (x_0, y_0, z_0) \in \mathbb{S}^2$, then either x_0, y_0 or z_0 is non-zero. Suppose that $z_0 \neq 0$. Choose $r > 0$ such that $(x - x_0)^2 + (y - y_0)^2 < 1$. Define

$$\psi(x, y) = \left\{ \begin{array}{cl} (x, y, \sqrt{1 - x^2 - y^2}) & \text{if } z_0 > 0\,, \\ (x, y, -\sqrt{1 - x^2 - y^2}) & \text{if } z_0 < 0\,, \end{array} \right.$$

$\mathcal{V} = B\big((x_0, y_0), r\big)$, and $\mathcal{U} = \psi(\mathcal{V})$. Then $\psi : \mathcal{V} \to \mathcal{U}$ is a bijection. Let $\varphi = \psi^{-1}$. Then $\{\mathcal{U}, \varphi\}$ is a coordinate chart at p; thus $\mathbb{S}^2$ is a surface.

There exists another coordinate chart. Let $\mathcal{U}_1 = \mathbb{S}^2\backslash(0, 0, -1)$ and $\mathcal{U}_2 = \mathbb{S}^2\backslash(0, 0, 1)$. Define the map $\varphi_1 : \mathcal{U}_1 \to \mathbb{R}^2$ by that $\varphi_1(p)$ is the unique point on $\mathbb{R}^2$ such that $(0, 0, -1)$, $\varphi_1(p)$ and $(x, y, 0)$ are on the same straight line. Similarly, define $\varphi_2 : \mathcal{U}_2 \to \mathbb{R}^2$ by that $\varphi_2(p)$ is the unique point on $\mathbb{R}^2$ such that $(0, 0, 1)$, $\varphi_2(p)$ and $(x, y, 0)$ are on the same straight line. It is easy to check that if $p \in \mathbb{S}^2$, then either $\{\mathcal{U}_1, \varphi_1\}$ or $\{\mathcal{U}_2, \varphi_2\}$ is a coordinate chart at p.

A third kind of coordinate chart is given as follows. Let $\mathcal{U} = (0, 2\pi) \times (0, \pi)$, and define

$$\psi(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi).$$

Then $\psi : \mathcal{U} \to \mathbb{S}^2\backslash\{(x, 0, z) \,|\, 0 \leqslant x \leqslant 1, x^2 + z^2 = 1\}$ is a continuous bijection with a continuous inverse. We note that for any $\mathcal{U} = (\theta_0, \theta_0 + 2\pi) \times (\phi_0, \phi_0 + \pi)$, ψ is a homeomorphism between $\mathcal{U}$ and an open subset of $\mathbb{S}^2$.

Next, we would like to define the derivative of f when $f : \Sigma \to \mathbb{R}^{\mathrm{n}}$ is a vector-valued function. We first talk about what the directional derivative is. Let $\Sigma \subseteq \mathbb{R}^3$ be a regular surface, $p \in \Sigma$, and $\boldsymbol{v} \in \mathbf{T}_p\Sigma$. It is intuitive to define the directional derivative of f at p in the direction $\boldsymbol{v}$ by

$$\frac{d}{dt}\Big|_{t=0}(f \circ \boldsymbol{x})(t)\,, \tag{4.6}$$

if the derivative exists, where $\boldsymbol{x} : (-\delta, \delta) \to \Sigma$ is a $\mathscr{C}^1$-parametrization of a curve on Σ such that $\boldsymbol{x}(0) = p$ and $\boldsymbol{x}'(0) = \boldsymbol{v}$. The first question arising naturally is that if the derivative in (4.6) depends on the choices of $\boldsymbol{x}$. Suppose that $\boldsymbol{y} : (-\delta, \delta) \to \Sigma$ is a $\mathscr{C}^1$-parametrization of another curve on Σ such that $\boldsymbol{y}(0) = p$ and $\boldsymbol{y}'(0) = \boldsymbol{v}$ (note that the curve $\boldsymbol{x}((-\delta, \delta))$ and $\boldsymbol{y}((-\delta, \delta))$ in general are different). Let $\{\mathcal{V}, \psi\}$ be a parametrization of Σ at p, and $q = \psi^{-1}(p)$. Then the chain rule (Theorem 2.49) implies that

$$\boldsymbol{v} = \boldsymbol{x}'(0) = \frac{d}{dt}\Big|_{t=0}(\psi \circ \psi^{-1} \circ \boldsymbol{x})(t) = (D\psi)(q)\Big(\frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{x})(t)\Big)$$

and similarly, $\boldsymbol{v} = (D\psi)(q)\Big(\frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{y})(t)\Big)$. Therefore,

$$(D\psi)(q)\Big(\frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{x})(t)\Big) = (D\psi)(q)\Big(\frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{y})(t)\Big).$$

The injectivity of $(D\psi)(\psi^{-1}(p))$ then shows that

$$\frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{x})(t) = \frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{y})(t).$$

Using the chain rule again,

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0}(f \circ \boldsymbol{x})(t) &= \frac{d}{dt}\Big|_{t=0}(f \circ \psi \circ \psi^{-1} \circ \boldsymbol{x})(t) = D(f \circ \psi)(\psi^{-1}(p))\Big(\frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{x})(t)\Big)\\ &= D(f \circ \psi)(\psi^{-1}(p))\Big(\frac{d}{dt}\Big|_{t=0}(\psi^{-1} \circ \boldsymbol{y})(t)\Big) = \frac{d}{dt}\Big|_{t=0}(f \circ \boldsymbol{y})(t).\end{aligned}$$

In other words, the derivative in (4.6) is independent of the choice of $\boldsymbol{x}$ as long as $\boldsymbol{x}(0) = p$ and $\boldsymbol{x}'(0) = \boldsymbol{v}$. This observation implies the following

Theorem 4.39. *Let $\Sigma \subseteq \mathbb{R}^3$ be a regular surface, $\{\mathcal{V}_1, \psi_1\}$ and $\{\mathcal{V}_2, \psi_2\}$ be two local $\mathscr{C}^1$-parametrizations of Σ at a point $p \in \Sigma$, and $\mathcal{U} = \psi_1(\mathcal{V}_1) \cap \psi_2(\mathcal{V}_2) \subseteq \Sigma$. Then for $(i, j) = (1, 2)$ and $(2, 1)$, the transition function $\psi_j^{-1} \circ \psi_i : \psi_i^{-1}(\mathcal{U}) \to \psi_j^{-1}(\mathcal{U})$ is of class $\mathscr{C}^1$.*

Proof. We first note that $\psi_j^{-1} \circ \psi_i$ is continuous on $\psi_i^{-1}(\mathcal{U})$. Moreover, by the chain rule we find that $\dfrac{\partial(\psi_j^{-1} \circ \psi_i)}{\partial u}$ is the unique 2-vector satisfying

$$\Big[\frac{\partial \psi_i}{\partial u}(u, v)\Big] = \Big[\frac{\partial}{\partial u}(\psi_j \circ \psi_j^{-1} \circ \psi_i)(u, v)\Big] = \Big[(D\psi_j)(\psi_j^{-1} \circ \psi_i)(u, v)\Big]\Big[\frac{\partial(\psi_j^{-1} \circ \psi_i)}{\partial u}(u_i, v_i)\Big].$$

Similarly, $\dfrac{\partial(\psi_j^{-1} \circ \psi_i)}{\partial v}$ is the unique 2-vector satisfying

$$\Big[\frac{\partial \psi_i}{\partial v}(u, v)\Big] = \Big[\frac{\partial}{\partial v}(\psi_j \circ \psi_j^{-1} \circ \psi_i)(u, v)\Big] = \Big[(D\psi_j)(\psi_j^{-1} \circ \psi_i)(u, v)\Big]\Big[\frac{\partial(\psi_j^{-1} \circ \psi_i)}{\partial v}(u_i, v_i)\Big].$$

Therefore, we obtain that

$$\big[D\psi_i\big] = \big[(D\psi_j)\circ(\psi_j^{-1}\circ\psi_i)\big]\Big[\big[\frac{\partial(\psi_j^{-1}\circ\psi_i)}{\partial u}\big]\vdots\big[\frac{\partial(\psi_j^{-1}\circ\psi_i)}{\partial v}\big]\Big]. \tag{4.7}$$

Since $\big[D\psi_j\big]$ has full rank, $\big[D\psi_j\big]^{\mathrm{T}}\big[D\psi_j\big]$ is an invertible 2×2 matrix (for if $A^{\mathrm{T}}Ax = 0$ then $\|Ax\|^2_{\mathbb{R}^{\mathrm{n}}} = x^{\mathrm{T}}A^{\mathrm{T}}Ax = 0$ which implies $x = 0$ since A has full rank); thus (4.7) implies that

$$\Big[\big[\frac{\partial(\psi_j^{-1}\circ\psi_i)}{\partial u}\big]\vdots\big[\frac{\partial(\psi_j^{-1}\circ\psi_i)}{\partial v}\big]\Big] = \Big(\big(\big[D\psi_j\big]^{\mathrm{T}}\big[D\psi_j\big]\big)\circ(\psi_j^{-1}\circ\psi_i)\Big)^{-1}\big[(D\psi_j)\circ(\psi_j^{-1}\circ\psi_i)\big]^{\mathrm{T}}\big[D\psi_i\big];$$

thus the partial derivatives of $\psi_j^{-1}\circ\psi_i$ exist and are continuous. Theorem 2.30 then implies that $\psi_j^{-1}\circ\psi_i$ is of class $\mathscr{C}^1$. □

Similar to how the directional derivative is defined, we intend to define the differentiability of f through the differentiability of the function $f\circ\psi : \mathcal{V}\to\mathbb{R}^{\mathrm{n}}$, where $\{\mathcal{V},\psi\}$ is a local parametrization of Σ (at some point). Again, we need to talk about if this definition depends on the choice of local parametrizations. Nevertheless, if $\{\mathcal{V}_1,\psi_1\}$ and $\{\mathcal{V}_2,\psi_2\}$ are two $\mathscr{C}^1$-local parametrization of Σ at p, and $f\circ\psi_1$ is differentiable at $\psi_1^{-1}(p)$, then the chain rule and Theorem 4.39 imply that $f\circ\psi_2$ is also differentiable at $\psi_2^{-1}(p)$ since $f\circ\psi_2 = (f\circ\psi_1)\circ(\psi_1^{-1}\circ\psi_2)$. This induces the following

Definition 4.40. Let $\Sigma\subseteq\mathbb{R}^3$ be a $\mathscr{C}^1$-regular surface. A scalar function $f:\Sigma\to\mathbb{R}$ is said to be differentiable at $p\in\Sigma$ if for every parametrization $\{\mathcal{V},\psi\}$ of Σ at p, the function $f\circ\psi:\mathcal{V}\to\mathbb{R}^{\mathrm{n}}$ is differentiable at $\psi^{-1}(p)$. The derivative of f at p, denoted by df_p, is a linear map on $T_p\Sigma$ satisfying

$$(df_p)(\boldsymbol{v}) = \frac{d}{dt}\Big|_{t=0}(f\circ\boldsymbol{x})(t),$$

where $\boldsymbol{x}:(-\delta,\delta)\to\Sigma$ is a $\mathscr{C}^1$-parametrization of a curve on Σ such that $\boldsymbol{x}(0) = p$ and $\boldsymbol{x}'(0) = \boldsymbol{v}$. A scalar function $f:\Sigma\to\mathbb{R}$ is said to be of class $\mathscr{C}^1$ if $f\circ\psi$ is of class $\mathscr{C}^1$ for all local parametrization $\{\mathcal{V},\psi\}$.

4.3.2 The metric tensor and the first fundamental form

Definition 4.41 (Metric)**.** Let $\Sigma\subseteq\mathbb{R}^3$ be a regular surface. The metric tensor associated with the local parametrization $\{\mathcal{V},\psi\}$ (at $p\in\Sigma$) is the matrix $g = [g_{\alpha\beta}]_{2\times 2}$ given by

$$g_{\alpha\beta} = \psi,_\alpha\cdot\psi,_\beta = \sum_{i=1}^{3}\frac{\partial\psi^i}{\partial y_\alpha}\frac{\partial\psi^i}{\partial y_\beta} \qquad \text{in} \quad \mathcal{V}.$$

Proposition 4.42. *Let $\Sigma \subseteq \mathbb{R}^3$ be a regular surface, and $g = [g_{\alpha\beta}]_{2\times 2}$ be the metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$). Then the metric tensor g is positive definite; that is,*

$$\sum_{\alpha,\beta=1}^{2} g_{\alpha\beta} v^\alpha v^\beta > 0 \qquad \forall\, \boldsymbol{v} = \sum_{\gamma=1}^{2} v^\gamma \frac{\partial \psi}{\partial y^\gamma} \neq \mathbf{0}\,.$$

Proof. Since $D\psi$ has full rank on $\mathcal{V}$, every tangent vector $\boldsymbol{v}$ can be expressed as the linear combination of $\left\{\frac{\partial \psi}{\partial y_1}, \frac{\partial \psi}{\partial y_2}\right\}$. Write $\boldsymbol{v} = \sum\limits_{\gamma=1}^{2} v^\gamma \frac{\partial \psi}{\partial y^\gamma}$. Then if $\boldsymbol{v} \neq \mathbf{0}$,

$$0 < \|\boldsymbol{v}\|_{\mathbb{R}^3}^2 = \sum_{i=1}^{3} \sum_{\alpha,\beta=1}^{2} v^\alpha \frac{\partial \psi^i}{\partial y_\alpha} v^\beta \frac{\partial \psi^i}{\partial \psi_\beta} = \sum_{\alpha,\beta=1}^{2} g_{\alpha\beta} v^\alpha v^\beta\,. \qquad \square$$

Definition 4.43 (The first fundamental form)**.** Let $\Sigma \subseteq \mathbb{R}^3$ be a regular surface, and $g = [g_{\alpha\beta}]_{2\times 2}$ be the metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$). The first fundamental form associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$) is the scalar function $\mathrm{g} = \det(g)$.

Theorem 4.44. *Let $\Sigma \subseteq \mathbb{R}^3$ be a regular surface, and $\{\mathcal{V}, \psi\}$ be a local parametrization at $p \in \Sigma$. Then*

$$\sqrt{\mathrm{g}} = \|\psi,_1 \times \psi,_2\|_{\mathbb{R}^3}\,. \tag{4.8}$$

Proof. Using the permutation symbol and Kronecker's delta, we have

$$\begin{aligned}
\|\psi,_1 \times \psi,_2\|_{\mathbb{R}^3}^2 &= \sum_{i=1}^{3}\Big(\sum_{j,k=1}^{3} \varepsilon_{ijk}\psi^j,_1 \psi^k,_2\Big)\Big(\sum_{r,s=1}^{3} \varepsilon_{irs}\psi^r,_1 \psi^s,_2\Big) \\
&= \sum_{j,k,r,s=1}^{3}\Big[\Big(\sum_{i=1}^{3}\varepsilon_{ijk}\varepsilon_{irs}\Big)\psi^j,_1 \psi^k,_2 \psi^r,_1 \psi^s,_2\Big] \\
&= \sum_{j,k,r,s=1}^{3}\big(\delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}\big)\psi^j,_1 \psi^k,_2 \psi^r,_1 \psi^s,_2\ ,
\end{aligned}$$

where we use the identity

$$\sum_{i=1}^{3}\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr} \tag{4.9}$$

to conclude the last equality. Therefore,

$$\begin{aligned}
\|\psi,_1 \times \psi,_2\|_{\mathbb{R}^3}^2 &= \sum_{j,k=1}^{3}\big(\psi^j,_1 \psi^k,_2 \psi^j,_1 \psi^k,_2 - \psi^j,_1 \psi^k,_2 \psi^j,_2 \psi^k,_1\big) \\
&= g_{11}g_{22} - g_{12}g_{21} = \det(g) = \mathrm{g}\,.
\end{aligned}$$

Finally, (4.8) is concluded from the fact that g is positive definite. $\square$

Remark 4.45. Let $L \in \mathscr{B}(\mathbb{R}^2; \mathbf{T}_p\Sigma)$ be given by

$$L(a\mathrm{e}_1 + b\mathrm{e}_2) = a\psi,_1 + b\psi,_2 \ ,$$

where $\mathcal{B}_2 = \{\mathrm{e}_1, \mathrm{e}_2\}$ is the standard basis of $\mathbb{R}^2$. Let $\mathcal{B}' = \{e_1, e_2\}$ be an orthonormal basis of $\mathbf{T}_p\Sigma$, and $\mathcal{B}_3 = \{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$ be the standard basis of $\mathbb{R}^3$. Then

$$[L]_{\mathcal{B}_2,\mathcal{B}'} = \begin{bmatrix} \psi,_1 \cdot e_1 & \psi,_2 \cdot e_1 \\ \psi,_1 \cdot e_2 & \psi,_2 \cdot e_2 \end{bmatrix} = \begin{bmatrix} [e_1]^{\mathrm{T}}_{\mathcal{B}_3} \\ [e_2]^{\mathrm{T}}_{\mathcal{B}_3} \end{bmatrix} \Big[[\psi,_1]_{\mathcal{B}_3} \vdots [\psi,_2]_{\mathcal{B}_3} \Big] .$$

By the fact that $\{e_1, e_2\}$ is an orthonormal basis,

$$\begin{aligned} [L]^{\mathrm{T}}_{\mathcal{B}_2,\mathcal{B}'}[L]_{\mathcal{B}_2,\mathcal{B}'} &= \begin{bmatrix} [\psi,_1]^{\mathrm{T}}_{\mathcal{B}_3} \\ [\psi,_2]^{\mathrm{T}}_{\mathcal{B}_3} \end{bmatrix} \Big[[e_1]_{\mathcal{B}_3} \vdots [e_2]_{\mathcal{B}_3} \Big] \begin{bmatrix} [e_1]^{\mathrm{T}}_{\mathcal{B}_3} \\ [e,_2]^{\mathrm{T}}_{\mathcal{B}_3} \end{bmatrix} \Big[[\psi,_1]_{\mathcal{B}_3} \vdots [\psi,_2]_{\mathcal{B}_3} \Big] \\ &= \begin{bmatrix} [\psi,_1]^{\mathrm{T}}_{\mathcal{B}_3} \\ [\psi,_2]^{\mathrm{T}}_{\mathcal{B}_3} \end{bmatrix} \Big[[\psi,_1]_{\mathcal{B}_3} \vdots [\psi,_2]_{\mathcal{B}_3} \Big] = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} , \end{aligned}$$

where $[g_{\alpha\beta}]_{2\times 2}$ is the metric tensor associated with the parametrization $\{\mathcal{V}, \psi\}$. Therefore, $\det([L]_{\mathcal{B}_2,\mathcal{B}'}) = \sqrt{\mathrm{g}}$ as long as $\mathcal{B}'$ is an orthonormal basis of $\mathbf{T}_p\Sigma$.

Since a natural way to write $L\boldsymbol{v}$, where $\boldsymbol{v} = a\mathrm{e}_1 + b\mathrm{e}_2 \in \mathbb{R}^2$, is

$$L\boldsymbol{v} = \Big[[\psi,_1] \vdots [\psi,_2] \Big] \begin{bmatrix} a \\ b \end{bmatrix} = \big[\nabla\psi \big] \begin{bmatrix} a \\ b \end{bmatrix} ,$$

sometimes we also use $\nabla\psi$ to denote L, and then write $\sqrt{\mathrm{g}}$ as $\det(\nabla\psi)$ (even though $[\nabla\psi]$ is a 3×2 matrix) and call $\sqrt{\mathrm{g}}$ the Jacobian of the map ψ.

Example 4.46. Let Σ be the sphere centered at the origin with radius R. Consider the local parametrization $\psi(\theta, \phi) = (R\cos\theta\sin\phi, R\sin\theta\sin\phi, R\cos\phi)$ with $(\theta, \phi) \in \mathcal{V} \equiv (0, 2\pi) \times (0, \pi)$. Then

$$\begin{aligned} \psi,_1(\theta, \phi) &= \psi_\theta(\theta, \phi) = (-R\sin\theta\sin\phi, R\cos\theta\sin\phi, 0) \, , \\ \psi,_2(\theta, \phi) &= \psi_\phi(\theta, \phi) = (R\cos\theta\cos\phi, R\sin\theta\cos\phi, -R\sin\phi) \, ; \end{aligned}$$

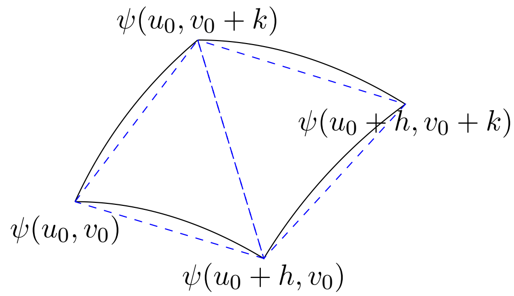
thus the metric tensor and the first fundamental form associated with the parametrization $\{\mathcal{V}, \psi\}$ are

$$g(\theta, \phi) = \big[D\psi \big]^{\mathrm{T}} \big[D\psi \big](\theta, \phi) = \begin{bmatrix} R^2\sin^2\phi & 0 \\ 0 & R^2 \end{bmatrix}$$

and $\mathrm{g} = \det(g) = R^4 \sin^2\phi$.

What does the first fundamental form do for us?

Let $p = \psi(u_0, v_0)$ be a point in Σ. Then the surface area of the region $\psi\big([u_0, u_0 + h] \times [v_0, v_0 + k]\big)$, where h, k are very small, can be approximated by the sum of the area of two triangles, one with vertices $\psi(u_0, v_0)$, $\psi(u_0 + h, v_0)$, $\psi(u_0, v_0 + k)$ and the other with vertices $\psi(u_0 + h, v_0)$, $\psi(u_0, v_0 + k)$, $\psi(u_0 + h, v_0 + k)$.



Here we remark that the approximation of the surface area of a regular $\mathscr{C}^1$-surface obeys

$$\lim_{(h,k)\to(0,0)} \frac{\text{the surface area of } \psi\big([u_0, u_0 + h] \times [v_0, v_0 + k]\big)}{\text{the sum of area of the two triangles given in the context}} = 1\,. \tag{4.10}$$

The area of the triangle with vertices $\psi(u_0, v_0)$, $\psi(u_0 + h, v_0)$, $\psi(u_0, v_0 + k)$ is

$$A_1 = \frac{1}{2}\Big\|\big(\psi(u_0 + h, v_0) - \psi(u_0, v_0)\big) \times \big(\psi(u_0, v_0 + k) - \psi(u_0, v_0)\big)\Big\|_{\mathbb{R}^3}\,.$$

By the mean value theorem, for each component $j \in \{1, 2, 3\}$, we have

$$\begin{aligned}
\psi^j(u_0 + h, v_0) - \psi^j(u_0, v_0) &= \psi,_1(u_0 + \theta_1^j h, v_0)h\,,\\
\psi^j(u_0, v_0 + k) - \psi^j(u_0, v_0) &= \psi,_2(u_0, v_0 + \theta_2^j k)k
\end{aligned}$$

for some $\theta_i^j \in (0, 1)$; thus if ψ is of class $\mathscr{C}^1$,

$$\begin{aligned}
\psi(u_0 + h, v_0) - \psi(u_0, v_0) &= \psi,_1(u_0, v_0)h + \boldsymbol{E}_1(u_0, v_0; h)h\,,\\
\psi(u_0, v_0 + k) - \psi(u_0, v_0) &= \psi,_2(u_0, v_0)k + \boldsymbol{E}_2(u_0, v_0; k)k\,,
\end{aligned}$$

where $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are bounded vector-valued functions satisfying that $\lim\limits_{h\to 0} \boldsymbol{E}_1(u_0, v_0; h) = 0$ and $\lim\limits_{k\to 0} \boldsymbol{E}_2(u_0, v_0; k) = 0$. Therefore,

$$\lim_{(h,k)\to(0,0)} \frac{\big(\psi(u_0+h, v_0) - \psi(u_0, v_0)\big) \times \big(\psi(u_0, v_0+k) - \psi(u_0, v_0)\big)}{hk} - \psi,_1(u_0, v_0) \times \psi,_2(u_0, v_0) = \mathbf{0}\,.$$

Since $\sqrt{\mathrm{g}} = \|\psi_{,1} \times \psi_{,2}\|_{\mathbb{R}^3}$, we have

$$A_1 = \frac{1}{2}\sqrt{\mathrm{g}(u_0, v_0)}hk + f_1(u_0, v_0; h, k)hk$$

for some function f_1 which converges to 0 as $(h, k) \to (0, 0)$ and is bounded since $\nabla\psi$ is bounded. Similarly, the area of the triangle with vertices $\psi(u_0 + h, v_0)$, $\psi(u_0, v_0 + k)$, $\psi(u_0 + h, v_0 + k)$ is

$$A_2 = \frac{1}{2}\sqrt{\mathrm{g}(u_0, v_0)}hk + f_2(u_0, v_0; h, k)hk\,.$$

Taking (4.10) into account, we find that

$$\text{the surface area of } \psi\big([u_0, u_0 + h] \times [v_0, v_0 + k]\big) = \sqrt{\mathrm{g}(u_0, v_0)}hk + f(u_0, v_0; h, k)hk \tag{4.11}$$

for some bounded function $f(\cdot, \cdot; \cdot, \cdot)$ which converges to 0 as the last two variables h, k approach 0.

Now consider the surface area of $\psi([a, a + L] \times [b, b + W])$. Let $\varepsilon > 0$ be given. Choose $N > 0$ such that

$$\big|f(u, v; h, k)\big| < \frac{\varepsilon}{2LW} \quad \forall\, 0 < h < \frac{L}{N}, 0 < k < \frac{W}{N} \text{ and } (u, v) \in [a, a + L] \times [b, b + W]\,,$$

and

$$\Big|\sum_{j=1}^{m}\sum_{i=1}^{n}\sqrt{\mathrm{g}\big(a + \frac{i-1}{n}L, b + \frac{j-1}{m}M\big)}\frac{L}{n}\frac{W}{m} - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big| < \frac{\varepsilon}{2} \quad \text{if } n, m \geqslant N\,.$$

Then for $n, m \geqslant N$, with (h, k) denoting $\big(\frac{L}{n}, \frac{W}{m}\big)$ (4.11) implies that

$$\begin{aligned}
&\Big|\text{the surface area of } \psi([a, a + L] \times [b, b + W]) - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big| \\
&\quad = \Big|\sum_{j=1}^{m}\sum_{i=1}^{n}\text{the surface area of } \psi([a + (i-1)h, a + ih] \times [b + (j-1)k, b + jk]) \\
&\qquad - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big| \\
&\quad \leqslant \Big|\sum_{j=1}^{m}\sum_{i=1}^{n}\sqrt{\mathrm{g}\big(a + (i-1)h, b + (j-1)k\big)}hk - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big| \\
&\qquad + \Big|\sum_{j=1}^{m}\sum_{i=1}^{n} f(a + (i-1)h, b + (j-1)k; h, k)hk\Big| \\
&\quad < \frac{\varepsilon}{2} + \frac{\varepsilon}{2LW}\sum_{j=1}^{m}\sum_{i=1}^{n} hk = \varepsilon\,.
\end{aligned}$$

The discussion above verifies the following

Theorem 4.47. *Let $\Sigma \subseteq \mathbb{R}^3$ be a regular $\mathscr{C}^1$-surface, $\{\mathcal{V}, \psi\}$ be a local $\mathscr{C}^1$-parametrization of Σ at p, and g be the first fundamental form associated with $\{\mathcal{V}, \psi\}$. Then*

$$\textit{the surface area of } \psi(\mathcal{V}) = \int_{\mathcal{V}} \sqrt{\mathrm{g}}\, d\mathbb{A}\,.$$

Example 4.48. Recall from Example 4.46 that the first fundamental form g of the parametrization $\{\mathcal{V}, \psi\}$ of the 2-sphere centered at the origin with radius R, where

$$\psi(\theta, \phi) = (R\cos\theta\sin\phi, R\sin\theta\sin\phi, R\cos\phi)$$

and $\mathcal{V} = (0, 2\pi) \times (0, \pi)$, is given by $\mathrm{g}(\theta, \phi) = R^4 \sin^2\phi$. Therefore,

$$\begin{aligned}\text{the surface area of } \psi\big((0, 2\pi) \times (0, \pi)\big) &= \int_{(0,2\pi)\times(0,\pi)} R^2 \sin\phi\, d(\theta, \phi) \\ &= R^2 \int_0^{2\pi} \int_0^{\pi} \sin\phi\, d\phi d\theta = 4\pi R^2\,.\end{aligned}$$

Since the difference of the 2-sphere and $\psi\big((0, 2\pi) \times (0, \pi)\big)$ has zero area, we find that the surface area of the 2-sphere with radius R is $4\pi R^2$.

4.3.3 The surface element and the surface integral

Let $\Sigma \subseteq \mathbb{R}^3$ be a regular surface, and $\{\mathcal{V}, \psi\}$ be a parametrization of Σ such that $\psi(\mathcal{V}) = \Sigma$. If $f : \Sigma \to \mathbb{R}$ is a bounded continuous function, the surface integral of f over Σ, denoted by $\displaystyle\int_\Sigma f\, dS$, is defined by

$$\int_\Sigma f\, dS = \int_{\mathcal{V}} (f \circ \psi)\sqrt{\mathrm{g}}\, d\mathbb{A}\,. \tag{4.12}$$

In particular, if $f \equiv 1$, the number $\displaystyle\int_\Sigma dS \equiv \int_\Sigma 1\, dS$ is the surface area of Σ.

Since the surface integrals defined by (4.12) seems to depend on a given parametrization, before proceeding we show that the surface integral is indeed independent of the choice of the parametrizations. Suppose that $\{\mathcal{V}_1, \psi_1\}$ and $\{\mathcal{V}_2, \psi_2\}$ are two local $\mathscr{C}^1$-parametrizations of a regular surface Σ at p, g_1, g_2 denote the metric tensors associated with the parametrizations $\{\mathcal{V}_1, \psi_1\}$, $\{\mathcal{V}_2, \psi_2\}$, respectively, and $\mathrm{g}_1 = \det(g_1)$, $\mathrm{g}_2 = \det(g_2)$ are corresponding first fundamental forms. Let $\Psi = \psi_2^{-1} \circ \psi_1$. Then the change of variables formula (Theorem 3.31) implies that

$$\int_{\mathcal{V}_2} (f \circ \psi_2)\sqrt{\mathrm{g}_2}\, d\mathbb{A} = \int_{\mathcal{V}_1} (f \circ \psi_2 \circ \Psi)\big(\sqrt{\mathrm{g}_2} \circ \Psi\big)|J_\Psi|\, d\mathbb{A} = \int_{\mathcal{V}_1} (f \circ \psi_1)\big(\sqrt{\mathrm{g}_2} \circ \Psi\big)|J_\Psi|\, d\mathbb{A}\,,$$

where J_Ψ is the Jacobian of the map Ψ. Using (4.7), we find that

$$\big[D\Psi\big]^{\mathrm{T}}\big[(D\psi_2)\circ\Psi\big]^{\mathrm{T}}\big[(D\psi_2)\circ\Psi\big]\big[D\Psi\big] = \big[D\psi_1\big]^{\mathrm{T}}\big[D\psi_1\big]\,;$$

thus by the fact that $\mathrm{g}_1 = \det\big(\big[D\psi_1\big]^{\mathrm{T}}\big[D\psi_1\big]\big)$ and $\mathrm{g}_2 = \det\big(\big[D\psi_2\big]^{\mathrm{T}}\big[D\psi_2\big]\big)$, we obtain that

$$\det\big(\big[D\Psi\big]\big)^2(\mathrm{g}_2\circ\Psi) = \mathrm{g}_1\,.$$

Since $J_\Psi = \det\big(\big[D\Psi\big]\big)$, the identity above implies that $|J_\Psi|(\sqrt{\mathrm{g}_2}\circ\Psi) = \sqrt{\mathrm{g}_1}$, so we conclude that

$$\int_{\mathcal{V}_1}(f\circ\psi_1)\sqrt{\mathrm{g}_1}\,d\mathbb{A} = \int_{\mathcal{V}_2}(f\circ\psi_2)\sqrt{\mathrm{g}_2}\,d\mathbb{A}\,. \tag{4.13}$$

Therefore, the surface integral of f over Σ is independent of the choice of parametrizations of Σ. In particular, the surface area of a regular $\mathscr{C}^1$-surface which can be parameterized by a global parametrization is also independent of the choice of parametrizations.

Example 4.49. Let $\Sigma \subseteq \mathbb{R}^3$ be the upper half sphere; that is, $\Sigma = \big\{(x,y,z)\in\mathbb{R}^3 \,\big|\, x^2+y^2+z^2=R^2, z>0\big\}$, and $\{\mathcal{V},\psi\}$ be a global parametrization of Σ given by

$$\psi(u,v) = (u,v,\sqrt{R^2-u^2-v^2})\,,\quad (u,v)\in\mathcal{V} = \big\{(u,v)\in\mathbb{R}^2 \,\big|\, u^2+v^2 \leqslant R^2\big\}\,.$$

To find the surface area using this parametrization, we first compute $\{\psi,_1, \psi,_2\}$ as follows:

$$\psi,_1(u,v) = \Big(1,0,\frac{-u}{\sqrt{R^2-u^2-v^2}}\Big) \quad\text{and}\quad \psi,_2(u,v) = \Big(0,1,\frac{-v}{\sqrt{R^2-u^2-v^2}}\Big)\,,$$

thus the first fundamental form associated with the parametrization $\{\mathcal{V},\psi\}$ is

$$\begin{aligned}\mathrm{g}(u,v) &= \|\psi,_1(u,v)\times\psi,_2(u,v)\|^2_{\mathbb{R}^3} = \Big\|\Big(\frac{u}{\sqrt{R^2-u^2-v^2}},\frac{v}{\sqrt{R^2-u^2-v^2}},1\Big)\Big\|^2_{\mathbb{R}^3}\\ &= \frac{R^2}{R^2-u^2-v^2}\,.\end{aligned}$$

Therefore, the surface area of Σ is

$$\begin{aligned}\int_\Sigma dS &= \int_{\mathcal{V}}\frac{R}{\sqrt{R^2-u^2-v^2}}\,d\mathbb{A} = \int_{-R}^{R}\int_{-\sqrt{R^2-u^2}}^{\sqrt{R^2-u^2}}\frac{R}{\sqrt{R^2-u^2-v^2}}\,dvdu\\ &= R\int_{-R}^{R}\arcsin\frac{v}{\sqrt{R^2-u^2}}\Big|_{v=-\sqrt{R^2-u^2}}^{v=\sqrt{R^2-u^2}}\,du = R\int_{-R}^{R}\pi\,du = 2\pi R^2\,.\end{aligned}$$

Note the the computation above also shows that the surface area of the sphere in $\mathbb{R}^3$ with radius R is $4\pi R^2$ which is the same as what we have conclude in Example 4.48.

Remark 4.50. The example above provides one specific way of evaluating the surface integrals: if the surface Σ is in fact a subset of the graph of a function $\varphi : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$; that is, $\Sigma \subseteq \big\{x, y, \varphi(x, y)) \,\big|\, (x, y) \in \mathcal{D}\big\}$, then Σ has a global parametrization

$$\psi(x, y) = \big(x, y, \varphi(x, y)\big)\,, \qquad (x, y) \in \mathcal{V}\,,$$

where $\mathcal{V}$ is the projection of Σ onto the xy-plane along the z-direction. Then the first fundamental form associated to this parametrization is

$$\mathrm{g}(x, y) = \|\psi,_1(x, y) \times \psi,_2(x, y)\|^2_{\mathbb{R}^3} = 1 + \Big|\frac{\partial \varphi}{\partial x}(x, y)\Big|^2 + \Big|\frac{\partial \varphi}{\partial y}(x, y)\Big|^2\,;$$

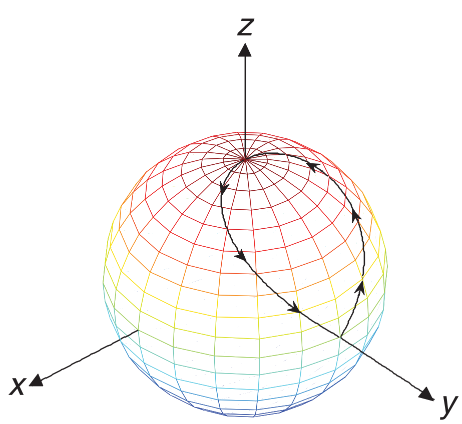
thus the surface integral of f over Σ is

$$\int_\Sigma f\, dS = \int_{\mathcal{V}} f\big(x, y, \varphi(x, y)\big)\sqrt{1 + \Big|\frac{\partial \varphi}{\partial x}(x, y)\Big|^2 + \Big|\frac{\partial \varphi}{\partial y}(x, y)\Big|^2}\, d(x, y)\,.$$

Example 4.51. Let C be a smooth curve parameterized by

$$\boldsymbol{r}(t) = (\cos t \sin t, \sin t \sin t, \cos t)\,, \qquad t \in \Big[-\frac{\pi}{2}, \frac{\pi}{2}\Big]\,.$$

The clearly C is on the unit sphere $\mathbb{S}^2$ since $\|\boldsymbol{r}(t)\|_{\mathbb{R}^3} = 1$ for all $t \in \Big[-\frac{\pi}{2}, \frac{\pi}{2}\Big]$. Since C is a closed curve, C divides $\mathbb{S}^2$ into two parts. Let Σ denote the part with smaller area (see the following figure), and we are interested in finding the surface area of Σ.



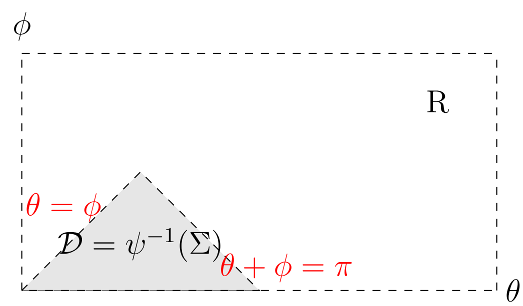
To compute the surface area of Σ, we need to find a way to parameterize Σ. Naturally we try to parameterize Σ using the spherical coordinate. In other words, let $\mathrm{R} = (0, 2\pi) \times (0, \pi)$ and $\psi : \mathrm{R} \to \mathbb{R}^3$ be defined by

$$\psi(\theta, \phi) = (\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi)\,,$$

and we would like to find a region $\mathcal{D} \subseteq \mathrm{R}$ such that $\psi(\mathcal{D}) = \Sigma$.

Suppose that $\gamma(t) = \big(\theta(t), \varphi(t)\big)$, $t \in \Big[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\Big]$, is a curve in R such that $(\psi \circ \gamma)(t) = \boldsymbol{r}(t)$. Then for $t \in \big[0, \frac{\pi}{2}\big]$, the identity $\cos t = \cos\phi(t)$ implies that $\phi(t) = t$; thus the identities $\cos t \sin t = \cos\theta(t)\cos\phi(t)$ and $\sin t \sin t = \sin\theta(t)\cos\phi(t)$ further imply that $\theta(t) = t$.

On the other hand, for $t \in \big[-\frac{\pi}{2}, 0\big]$, the identity $\cos t = \cos\phi(t)$, where $\phi(t) \in (0, \pi)$, implies that $\phi(t) = -t$; thus the identities $\cos t \sin t = \cos\theta(t)\sin\phi(t)$ and $\sin t \sin t = \sin\theta(t)\sin\phi(t)$ further imply that $\theta(t) = \pi + t$.



Since the first fundamental form associate with $\{\mathrm{R}, \psi\}$ is the first fundamental form associated with $\{\mathrm{R}, \psi\}$ is

$$\begin{aligned}
\mathrm{g}(u, v) &= \big\|(\psi_\theta \times \psi_\phi)(u, v)\big\|^2_{\mathbb{R}^3} \\
&= \big\|(-\sin\theta\sin\phi, \cos\theta\sin\phi, 0) \times (\cos\theta\cos\phi, \sin\theta\cos\phi, -\sin\phi)\big\|^2_{\mathbb{R}^3} \\
&= \big\|(-\cos\theta\sin^2\phi, -\sin\theta\sin^2\phi, -(\sin^2\theta + \cos^2\theta)\sin\phi\cos\phi)\big\|^2_{\mathbb{R}^3} \\
&= (\cos^2\theta + \sin^2\theta)\sin^4\phi + \sin^2\phi\cos^2\phi = \sin^2\phi\,,
\end{aligned}$$

the area of the desired surface can be computed by

$$\begin{aligned}
\int_\Sigma dS &= \int_{\psi^{-1}(\Sigma)} \sqrt{\mathrm{g}}\, d\mathbb{A} = \int_0^{\frac{\pi}{2}} \int_\phi^{\pi - \phi} \sin\phi\, d\theta d\phi = \int_0^{\frac{\pi}{2}} (\pi - 2\phi)\sin\phi\, d\phi \\
&= \big(-\pi\cos\phi + 2\phi\cos\phi - 2\sin\phi\big)\Big|_{\phi=0}^{\phi=\frac{\pi}{2}} = \pi - 2\,.
\end{aligned}$$

Another way to parameterize Σ is to view Σ as the graph of function $z = \sqrt{1 - x^2 - y^2}$ over $\mathcal{D}$, where $\mathcal{D}$ is the projection of Σ along z-axis onto xy-plane. We note that the boundary of $\mathcal{D}$ can be parameterized by

$$\widetilde{\boldsymbol{r}}(t) = (\cos t \sin t, \sin t \sin t)\,, \qquad t \in \Big[-\frac{\pi}{2}, \frac{\pi}{2}\Big]\,.$$

Let $(x,y) \in \partial\mathcal{D}$. Then $x^2+y^2=y$; thus Σ can also be parameterized by $\psi:\mathcal{D}\to\mathbb{R}^3$, where

$$\psi(x,y)=\big(x,y,\sqrt{1-x^2-y^2}\big) \quad \text{and} \quad \mathcal{D}=\big\{(x,y)\,\big|\,x^2+y^2\leqslant y\big\}\,.$$

Therefore, with f denoting the function $f(x,y)=\sqrt{1-x^2-y^2}$, Remark 4.50 implies that the surface area of Σ can be computed by

$$\begin{aligned}\int_{\mathcal{D}}\sqrt{1+f_x^2+f_y^2}\,d\mathbb{A} &= \int_0^1\int_{-\sqrt{y-y^2}}^{\sqrt{y-y^2}}\frac{1}{\sqrt{1-x^2-y^2}}\,dxdy\\ &= \int_0^1 \arcsin\frac{x}{\sqrt{1-y^2}}\Big|_{x=-\sqrt{y-y^2}}^{x=\sqrt{y-y^2}}dy = 2\int_0^1\arcsin\frac{\sqrt{y}}{\sqrt{1+y}}\,dy\,;\end{aligned}$$

thus making a change of variable $y=\tan^2\theta$ we conclude that

$$\begin{aligned}\text{the surface area of }\Sigma &= 2\int_0^{\frac{\pi}{4}}\arcsin\frac{\tan\theta}{\sec\theta}\,d(\tan^2\theta) = 2\int_0^{\frac{\pi}{4}}\theta\,d\big(\tan^2\theta\big)\\ &= 2\Big[\theta\tan^2\theta\Big|_{\theta=0}^{\theta=\frac{\pi}{4}} - \int_0^{\frac{\pi}{4}}\tan^2\theta d\theta\Big]\\ &= 2\Big[\frac{\pi}{4}-\int_0^{\frac{\pi}{4}}(\sec^2\theta-1)\,d\theta\Big] = 2\Big[\frac{\pi}{4}-(\tan\theta-\theta)\Big|_{\theta=0}^{\theta=\frac{\pi}{4}}\Big]\\ &= 2\Big[\frac{\pi}{4}-\Big(1-\frac{\pi}{4}\Big)\Big] = \pi-2\,.\end{aligned}$$

As noticed in Remark 4.45, the first fundamental form $\sqrt{\mathrm{g}}$ associated with the parametrization $\{\mathcal{V},\psi\}$ can be viewed as the Jacobian of the map ψ. Therefore, we arrive at the conclusion that dS " = "$\sqrt{\mathrm{g}}\,d\mathbb{A}$. dS is called the ***surface element***. Moreover, similar to the reason provided in Remark 4.22, the surface integral of a positive continuous function f over Σ, where f is considered as the mass density of the surface given by

$$f(x)=\lim_{\substack{\mathrm{diam}(\Delta)\to 0\\ \psi^{-1}(x)\in\Delta}}\frac{\text{the mass of }\psi(\Delta)}{\text{the surface area of }\psi(\Delta)}$$

is the total mass of the surface.

Next, we study the surface area of general regular surfaces that cannot be parameterized using a single pair $\{\mathcal{V},\psi\}$. Let $\Sigma\subseteq\mathbb{R}^3$ be a regular surface, and $\{\mathcal{V}_i,\psi_i\}_{i\in\mathcal{I}}$ be a collection of local parametrizations satisfying that for each $p\in\Sigma$ there exists $i\in\mathcal{I}$ such that $\{\mathcal{V}_i,\psi_i\}$ is a local parametrization of Σ at p. If there exists a countable collection of non-negative functions $\{\zeta_j\}_{j\in\mathcal{J}}$ defined on Σ such that

1. For each $j \in \mathcal{J}$, $\operatorname{spt}(\zeta_j) \equiv$ the closure of $\{x \in \Sigma \,|\, \zeta_j(x) \neq 0\} \subseteq \mathcal{V}_i$ for some $i \in \mathcal{I}$;

2. $\sum_{j\in\mathcal{J}} \zeta_j(x) = 1$ for all $x \in \Sigma$,

then intuitively we can compute the surface area by

$$\int_\Sigma dS = \sum_{j\in\mathcal{J}} \int_\Sigma \zeta_j \, dS \,, \tag{4.14}$$

where the surface integral of ζ_j over Σ is defined by (4.12) since $\operatorname{spt}(\zeta_j) \subseteq \psi(\mathcal{V}_i)$ and $\zeta_j = 0$ outside $\operatorname{spt}(\zeta_j)$. In other words, each term on the right-hand side of (4.14) can be evaluated by

$$\int_\Sigma \zeta_j \, dS = \int_{\mathcal{V}_i} (\zeta_j \circ \psi_i)\sqrt{\mathrm{g}_i} \, dS \,.$$

if $\operatorname{spt}(\zeta_j) \subseteq \psi_i(\mathcal{V}_i)$. Similarly, for a bounded continuous function f defined on Σ, the surface integral of f over Σ can be defined by

$$\int_\Sigma f \, dS = \sum_{j\in\mathcal{J}} \int_\Sigma (\zeta_j f) \, dS = \sum_{j\in\mathcal{J}} \sum_{\substack{\text{choose one } i \text{ such that} \\ \operatorname{spt}(\zeta_j) \subseteq \psi_i(\mathcal{V}_i)}} \int_{\mathcal{V}_i} (\zeta_j f) \circ \psi_i \sqrt{\mathrm{g}_i} \, dS \,. \tag{4.15}$$

Remark 4.52. Defining the surface integrals of a function as above, a question arises naturally: is the surface integral given by (4.15) independent of the choice of the parametrization and the partition-of-unity? In other words, if a regular $\mathscr{C}^k$-surface Σ admits two collections of local parametrization $\{\mathcal{U}_i, \varphi_i\}_{i\in\mathcal{I}}$ and $\{\mathcal{V}_j, \psi_j\}_{j\in\mathcal{J}}$, and $\{\zeta_i\}_{i\in\mathcal{I}}$ and $\{\lambda_j\}_{j\in\mathcal{J}}$ are $\mathscr{C}^k$-partition-of-unity subordinate to $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$ and $\{\mathcal{V}_j\}_{j\in\mathcal{J}}$, respectively. Is it true that

$$\sum_{i\in\mathcal{I}} \sum_{\substack{\text{choose one } i \text{ such that} \\ \operatorname{spt}(\zeta_j) \subseteq \varphi_i(\mathcal{U}_i)}} \int_{\mathcal{U}_i} (\zeta_i f) \circ \varphi_i \sqrt{\mathrm{g}_i} \, dS = \sum_{j\in\mathcal{J}} \sum_{\substack{\text{choose one } j \text{ such that} \\ \operatorname{spt}(\lambda_k) \subseteq \psi_j(\mathcal{V}_j)}} \int_{\mathcal{V}_j} (\lambda_j f) \circ \psi_j \sqrt{g_j} \, dS \,,$$

where g_i and g_j are the first fundamental form associated with the parametrization $\{\mathcal{U}_i, \varphi_i\}$ and $\{\mathcal{V}_j, \psi_j\}$, respectively.

The answer to the question above is affirmative, and the surface integral given by (4.15) is indeed independent of the choice of parametrization of the surface and the partition-of-unity; however, we will not prove this and only treat this as a known fact.

Now we focus on the existence of a collection of functions $\{\zeta_j\}_{j\in\mathcal{J}}$ discussed above.

Definition 4.53. A collection of subsets of $\mathbb{R}^{\mathrm{n}}$ is said to be ***locally finite*** if for every point $x \in \mathbb{R}^{\mathrm{n}}$ there exists $r > 0$ such that $B(x, r)$, the ball centered at x with radius r, intersects at most finitely many sets in this collection.

Definition 4.54 (Partition of Unity)**.** Let $A \subseteq \mathbb{R}^n$ be a subset. A collection of functions $\{\zeta_j\}_{j\in\mathcal{J}}$ is said to be a ***partition-of-unity*** of A if

1. $0 \leqslant \zeta_j \leqslant 1$ for all $j \in \mathcal{J}$.
2. The collection of sets $\big\{\mathrm{spt}(\zeta_j)\big\}_{j\in\mathcal{J}}$ is locally finite.
3. $\sum\limits_{j\in\mathcal{J}} \zeta_j(x) = 1$ for all $x \in A$.

Let $\{\mathcal{U}_j\}_{j\in\mathcal{J}}$ be an open cover of A; that is, $\mathcal{U}_j$ is open for all $j \in \mathcal{J}$ and $A \subseteq \bigcup_{j\in\mathcal{J}} \mathcal{U}_j$. A partition-of-unity $\{\zeta_j\}_{j\in\mathcal{J}}$ of A is said to be ***subordinate*** to $\{\mathcal{U}_j\}_{j\in\mathcal{J}}$ (or $\{\mathcal{U}_j\}_{j\in\mathcal{J}}$ has a subordinate partition-of-unity of A) if $\mathrm{spt}(\zeta_j) \subseteq \mathcal{U}_j$ for all $j \in \mathcal{J}$.

We note the if $\{\zeta_j\}_{j\in\mathcal{J}}$ is a partition-of-unity of A, then the property of local finiteness of $\{\mathrm{spt}(\zeta_j)\}_{j\in\mathcal{J}}$ ensures that for each point $x \in A$ has a neighborhood on which all but finitely many λ_j's are zero.

Lemma 4.55. *Let $A \subseteq \mathbb{R}^n$ be a subset, $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$ be an open cover of A, and $\{\mathcal{V}_j\}_{j\in\mathcal{J}}$ be a collection of open sets such that each $\mathcal{V}_j$ is a subset of some $\mathcal{U}_i$; that is, for each $j \in \mathcal{J}$, $\mathcal{V}_j \subseteq \mathcal{U}_i$ for some $i \in \mathcal{I}$. If $\{\mathcal{V}_j\}_{j\in\mathcal{J}}$ has a subordinate $\mathscr{C}^k$-partition-of-unity of A, so has $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$.*

Proof. Let $\{\zeta_j\}_{j\in\mathcal{J}}$ be a partition-of-unity of A subordinate to $\{\mathcal{V}_j\}_{j\in\mathcal{J}}$, and $f : \mathcal{J} \to \mathcal{I}$ be a map such that $\mathcal{V}_j \subseteq \mathcal{U}_{f(j)}$ (we note that such f in general is not unique). Define $\chi_i : \mathbb{R}^n \to [0,1]$ by

$$\chi_i(x) = \sum_{j\in f^{-1}(i)} \zeta_j(x)\,. \tag{4.16}$$

Then clearly $\mathrm{spt}(\chi_i) \subseteq \mathcal{U}_i$ and $\sum\limits_{i\in\mathcal{I}} \chi_i(x) = 1$ for all $x \in A$. Moreover, since the sum (4.16) is a finite sum, χ_i is of class $\mathscr{C}^k$ for all $i \in \mathcal{I}$ since ζ_j if of class $\mathscr{C}^k$ for all $j \in \mathcal{J}$. Now we show that $\big\{\mathrm{spt}(\chi_i)\big\}_{i\in\mathcal{I}}$ is locally finite. Let $x \in \mathbb{R}^n$ be given. By the local finiteness of $\big\{\mathrm{spt}(\zeta_j)\big\}_{j\in\mathcal{J}}$ there exists $r > 0$ such that $\#\big\{j \in \mathcal{J} \,\big|\, B(x,r) \cap \mathrm{spt}(\zeta_j) \neq \varnothing\big\} < \infty$. By the fact that $f^{-1}(i_1) \cap f^{-1}(i_2) = \varnothing$ if $i_1 \neq i_2$ (that is, each $j \in \mathcal{J}$ belongs to $f^{-1}(i)$ for exactly one $i \in \mathcal{I}$) and that

$$y \in B(x,r) \cap \mathrm{spt}(\chi_i) \quad \Leftrightarrow \quad y \in B(x,r) \cap \mathrm{spt}(\zeta_j) \text{ for some } j \in f^{-1}(i)\,,$$

we must have

$$\#\big\{i \in \mathcal{I} \,\big|\, B(x,r) \cap \mathrm{spt}(\chi_i) \neq \varnothing\big\} \leqslant \#\big\{j \in \mathcal{J} \,\big|\, B(x,r) \cap \mathrm{spt}(\zeta_j) \neq \varnothing\big\} < \infty\,. \qquad \square$$

Theorem 4.56. *Let $\Sigma \subseteq \mathbb{R}^3$ be a regular $\mathscr{C}^k$-surface. Then every open cover of Σ has a subordinate $\mathscr{C}^k$-partition-of-unity of Σ.*

Proof. Let $\{\mathcal{O}_i\}_{i\in\mathcal{I}}$ be a given open cover of Σ. Let $\{\mathcal{U}_j, \varphi_j\}_{j\in\mathcal{J}}$ be a collection of $\mathscr{C}^k$-charts of Σ such that $\{\mathcal{U}_j\}_{j\in\mathcal{J}}$ is a locally finite open cover of Σ and for each $j \in \mathcal{J}$, $\overline{\mathcal{U}_j} \subseteq \mathcal{O}_i$ for some $i \in \mathcal{I}$. By Lemma 4.55, it suffices to find a $\mathscr{C}^k$-partition-of-unity of Σ subordinate to $\{\mathcal{U}_j\}_{j\in\mathcal{J}}$.

W.L.O.G., we can assume that $\mathcal{U}_j$ and $\mathcal{V}_j \equiv \varphi(\mathcal{U}_j)$ is bounded for all $j \in \mathcal{J}$. Define $\psi_j = \varphi_j^{-1}$. Then $\{\mathcal{V}_j, \psi_j\}_{j\in\mathcal{J}}$ is a collection of local parametrization of Σ. Choose a collection of open sets $\{\mathcal{W}_j\}_{j\in\mathcal{J}}$ such that $\overline{\mathcal{W}_j} \subseteq \mathcal{V}_j$ for all $j \in \mathcal{J}$ and $\big\{\psi_j(\mathcal{W}_j)\big\}_{j\in\mathcal{J}}$ is still an open cover of Σ. For each $j \in \mathcal{J}$, let $\big\{B_k^{(j)}\big\}_{k=1}^{N_j}$ be a collection of open balls satisfying $\overline{\mathcal{W}_j} \subseteq \bigcup\limits_{k=1}^{N_j} B_k^{(j)}$ and $\mathrm{cl}(B_k^{(j)}) \subseteq \mathcal{V}_j$ for all $k \in \{1, \cdots, N_j\}$. For $j \in \mathcal{J}$ and $k \in \{1, \cdots, N_j\}$, with $c_{j,k}$ and $r_{j,k}$ denoting the center and the radius of $B_k^{(j)}$, respectively, let

$$\mu_{(j,k)}(x) = \begin{cases} \exp\Big(\dfrac{1}{\|x - c_{j,k}\|_{\mathbb{R}^2}^2 - r_{j,k}^2}\Big) & \text{if } x \in B_k^{(j)}\,, \\ 0 & \text{if } x \notin B_k^{(j)}\,, \end{cases}$$

and then define $\chi_j : \mathbb{R}^2 \to \mathbb{R}$ by $\chi_j(x) = \sum\limits_{k=1}^{N_j} \mu_{(j,k)}(x)$. Then $\chi_j > 0$ in $\overline{\mathcal{W}_j}$, and $\chi_j = 0$ outside $\bigcup\limits_{k=1}^{N_j} B_k^{(j)}$. Further define

$$\lambda_j(x) = \begin{cases} (\chi_j \circ \varphi_j)(x) & \text{if } x \in \mathcal{U}_j\,, \\ 0 & \text{if } x \in \mathcal{U}_j^\complement\,. \end{cases}$$

Then $\lambda_j > 0$ on $\psi_j(W_j)$ which implies that $\sum\limits_{j\in\mathcal{J}} \lambda_j > 0$. Finally, we define $\zeta_j = \dfrac{\lambda_j}{\sum_{j\in\mathcal{J}} \lambda_j}$. Then $\{\zeta_j\}_{j\in\mathcal{J}}$ is a $\mathscr{C}^k$-partition-of-unity subordinate to the open cover $\{\mathcal{U}_j\}_{j\in\mathcal{J}}$. □

Definition 4.57 (Piecewise Regular Surface)**.** A surface $\Sigma \subseteq \mathbb{R}^3$ is said to be piecewise regular if there are finite many curves $C_1, \cdots, C_k$ such that $\Sigma \backslash \bigcup\limits_{i=1}^{k} C_i$ is a disjoint union of regular surfaces.

Definition 4.58. Let $\Sigma \subseteq \mathbb{R}^3$ be a piecewise regular surface such that Σ is the disjoint union of regular surfaces Σ_i, where $i \in \mathcal{I}$ for some finite index set $\mathcal{I}$. For a continuous function $f : \Sigma \to \mathbb{R}$, the surface integral of f over Σ, still denoted by $\displaystyle\int_\Sigma f\,dS$, is defined by

$$\int_\Sigma f\,dS = \sum_{i\in\mathcal{I}} \int_{\Sigma_i} f\,dS\,.$$

Definition 4.59. Let $\mathscr{R}_\Sigma$ be the collection of piecewise regular surfaces in $\mathbb{R}^3$. The surface element is a set function $\mathscr{S} : \mathscr{R}_\Sigma \to \mathbb{R}$ that satisfies the following properties:

1. $\mathscr{S}(\Sigma) > 0$ for all $\Sigma \in \mathscr{R}_\Sigma$.

2. If Σ is the union of finitely many regular surfaces $\Sigma_1, \cdots, \Sigma_k$ that do not overlap except at their boundaries, then
$$\mathscr{S}(\Sigma) = \mathscr{S}(\Sigma_1) + \cdots + \mathscr{S}(\Sigma_k)\,.$$

3. The value of $\mathscr{S}$ agrees with the area on planar surfaces; that is,
$$\mathscr{S}(\mathcal{P}) = \mathbb{A}(\mathcal{P}) \qquad \text{for all planar surfaces } \mathcal{P}.$$

4.4 Oriented Surfaces

In the study of surfaces, orientability is a property that measures whether it is possible to make a consistent choice of surface normal vector at every point. A choice of surface normal allows one to use the right-hand rule to define a "counter-clockwise" direction of loops in the surface that is required in the presentation of the Stokes theorem (Theorem 4.86), a main result in vector calculus which will be introduced later in Section 4.7.2.

Definition 4.60. A regular surface $\Sigma \subseteq \mathbb{R}^3$ is said to be ***oriented*** if there exists a continuous vector-valued function $\mathbf{N} : \Sigma \to \mathbb{R}^3$ such that $\|\mathbf{N}\|_{\mathbb{R}^3} = 1$ and for all $p \in \Sigma$, $\mathbf{N} \cdot \boldsymbol{v} = 0$ for all $\boldsymbol{v} \in \mathbf{T}_p\Sigma$. Such a vector-field $\mathbf{N}$ is called a ***unit normal*** of Σ.

Suppose that $\Sigma \subseteq \mathbb{R}^3$ is a connected regular surface. Since at each $p \in \Sigma$ the tangent plane $\mathbf{T}_p\Sigma$ of Σ at p has two normal directions, Σ has at most two continuous unit normal vector fields. If in addition that Σ is oriented, there are exactly two continuous unit normal vector fields of Σ, and one is the opposite of the other. The two unit normal vector fields define two ***sides*** of the surface.

Suppose further that this oriented surface Σ is the boundary of an open set $\Omega \subseteq \mathbb{R}^3$ (for example, a sphere is the boundary of a ball), then one of the unit normal vector fields $\mathbf{N} : \partial\Omega \to \mathbb{R}^3$ has the property that $p + t\mathbf{N}(p) \notin \Omega$ for all small but positive t. Such a normal is called the ***outward-pointing unit normal*** of $\partial\Omega$, and the opposite of the outward-pointing unit normal of $\partial\Omega$ is called the ***inward-pointing unit normal*** of $\partial\Omega$.

Example 4.61. Consider the unit sphere $\mathbb{S}^2 = \left\{(x,y,z) \in \mathbb{R}^3 \,\middle|\, x^2+y^2+z^2=1\right\}$. Then $\mathbf{N} : \mathbb{S}^2 \to \mathbb{R}^3$ defined by $\mathbf{N}(p) = p$, where the right-hand side is treated as the vector $p - 0$, is a continuous unit normal vector field on Σ; thus $\mathbb{S}^2$ is an oriented surface. Let $B(0,1) = \left\{(x,y,z) \in \mathbb{R}^3 \,\middle|\, x^2+y^2+z^2<1\right\}$ be the unit ball in $\mathbb{R}^3$. Then $\mathbf{N}$ is the outward-pointing unit normal of $\partial B(0,1)$.

Let $\Sigma \subseteq \mathbb{R}^3$ be a regular surface, $p \in \Sigma$, and $\{\mathcal{V}, \psi\}$ be a local parametrization of Σ at p. Since $\psi_{,1}$ and $\psi_{,2}$ are linearly independent, $\psi_{,1} \times \psi_{,2} \neq \mathbf{0}$; thus the vector $\boldsymbol{n}$ given by

$$\boldsymbol{n} = \frac{\psi_{,1} \times \psi_{,2}}{\|\psi_{,1} \times \psi_{,2}\|_{\mathbb{R}^3}} \circ \psi^{-1}$$

is a unit normal vector field on $\psi(\mathcal{V})$. As a consequence, a regular $\mathscr{C}^1$-surface that can be parameterized by one single parametrization $\{\mathcal{V}, \psi\}$; that is, $\Sigma = \psi(\mathcal{V})$, is always oriented. Such a normal vector fields is said to be compatible with the parametrization $\{\mathcal{V}, \psi\}$. To be more precise, we have the following

Definition 4.62. Let $\Sigma \subseteq \mathbb{R}^3$ be an oriented $\mathscr{C}^1$-surface, and $\mathbf{N} : \Sigma \to \mathbb{R}^3$ be a continuous unit normal vector field of Σ. For each $p \in \mathcal{V}$, $\mathbf{N}$ is said to be compatible with a local parametrization $\{\mathcal{V}, \psi\}$ of Σ at p if $\det\big([\psi, 1 \,\vdots\, \psi_{,2} \,\vdots\, \mathbf{N} \circ \psi]\big) > 0$.

The following example provides a famous regular surface which is not oriented.

Example 4.63. A Möbius strip/band is a surface obtained, conceptually, by half-twisting a paper strip and then joining the ends of the strip together to form a loop (see the following figure for the idea).

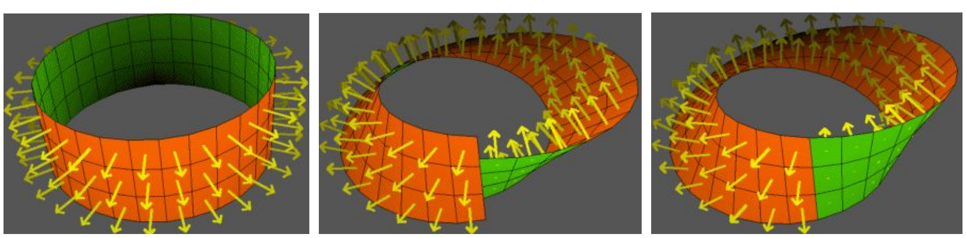


Figure 4.1: Normal vector fields on a Möbius strip

As one can see from Figure 4.1, a Möbius strip is not oriented. To see this mathematically, consider the following Möbius strip

$$\mathcal{M} = \left\{\left(-\,(2 + v\cos\frac{u}{2})\sin u, (2 + v\cos\frac{u}{2})\cos u, v\sin\frac{u}{2}\right) \,\middle|\, (u,v) \in [0,2\pi] \times (-1,1)\right\}$$

and choose a local parametrization $\psi : \mathcal{V} \to \mathbb{R}^3$ given by

$$\psi(u,v) = \Big(-(2+v\cos\frac{u}{2})\sin u, (2+v\cos\frac{u}{2})\cos u, v\sin\frac{u}{2}\Big)\,,$$

where $(u,v) \in \mathcal{V} \equiv (0,2\pi)\times(-1,1)$.

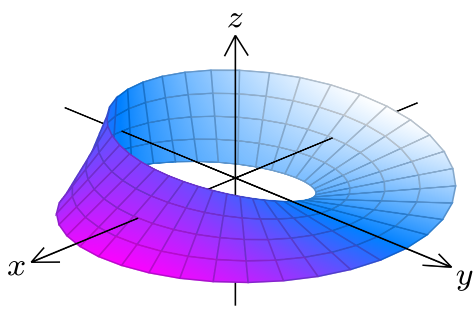


Figure 4.2: The Möbius strip/band $\psi([0,2\pi]\times[-1,1])$

Then the unit normal vector field on $\psi(\mathcal{V})$ compatible with the parametrization $\{\mathcal{V},\psi\}$ is

$$\begin{aligned}(\mathbf{N}\circ\psi)(u,v) &= \frac{\psi,_1\times\psi,_2}{\|\psi,_1\times\psi,_2\|_{\mathbb{R}^3}} = \frac{2}{\sqrt{v^2+(4+2v\cos(u/2))^2}}\times\\ &\quad\times\Big(\frac{v}{2}\cos u + (2+v\sin\frac{u}{2})\sin\frac{u}{2}\sin u,\\ &\qquad -\frac{v}{2}\sin u + (2+v\cos\frac{u}{2})\sin\frac{u}{2}\cos u, -(2+v\cos\frac{u}{2})\cos\frac{u}{2}\Big)\,,\end{aligned}$$

but $\mathbf{N}$ does not have a continuous extension on $\mathcal{M}$ since if $\widetilde{\mathbf{N}}$ is a continuous extension of $\mathbf{N}$; that is, $\widetilde{\mathbf{N}}$ is a unit normal vector field on $\mathcal{M}$ and $\mathbf{N} = \widetilde{\mathbf{N}}$ on $\psi(\mathcal{V})$, then

$$(0,0,-1) = \lim_{u\to 0^+}(\mathbf{N}\circ\psi)(u,0) = \widetilde{\mathbf{N}}(2,0,0) = \lim_{u\to 2\pi^-}(\mathbf{N}\circ\psi)(u,0) = (0,0,1)$$

which is a contradiction.

Another way of seeing that $\mathcal{M}$ is not oriented is the following. Let $r(t) = G(t,0) = (2\cos t, 2\sin t, 0)$, and $C = r([0,2\pi]) \subseteq \mathcal{M}$ be a closed curve on $\mathcal{M}$. If there is a continuous unit normal vector field $\widetilde{\mathbf{N}}$ on $\mathcal{M}$, then $\widetilde{\mathbf{N}}$ is also continuous on C. However, $\widetilde{\mathbf{N}}$ is never continuous on C since by moving $\mathbf{N}$ continuously along C, starting from $r(0)$ and moving along C in the direction r' and back to $r(0) = r(2\pi)$, we obtain a different vector which implies that $\widetilde{\mathbf{N}}\circ r$ is not continuous at $r(0) = r(2\pi) = (2,0,0)$.

Definition 4.64. An open set $\Omega \subseteq \mathbb{R}^3$ is said to be of class $\mathscr{C}^k$ if the boundary $\partial\Omega$ is a regular $\mathscr{C}^k$-surface.

Theorem 4.65. *Let $\Omega \subseteq \mathbb{R}^3$ be a bounded open set of class $\mathscr{C}^1$. Then $\partial\Omega$ is oriented.*

4.5 Manifolds, Charts, Atlas and Differentiable Structure

In the following, we introduce a more abstract concept, the so-called manifolds, which is a generalization of regular surfaces.

Definition 4.66. A topological space $\mathcal{M}$ is called an n-dimensional ***manifold*** if it is locally homeomorphic to $\mathbb{R}^n$; that is, there is an open cover $\mathscr{U} = \{\mathcal{U}_i\}_{i\in\mathcal{I}}$ of $\mathcal{M}$ such that for each $i \in \mathcal{I}$ there is a map $\varphi_i : \mathcal{U}_i \to \mathbb{R}^n$ which maps $\mathcal{U}_i$ homeomorphically onto an open subset of $\mathbb{R}^n$. The pair $\{\mathcal{U}_i, \varphi_i\}$ is called a ***chart*** (or coordinate system) with domain $\mathcal{U}_i$, and $\{\varphi_i(\mathcal{U}_i), \varphi_i^{-1}\}$ is called a local parametrization of $\mathcal{M}$. The collection of charts $\Phi = \{\mathcal{U}_i, \varphi_i\}_{i\in\mathcal{I}}$ is called an ***atlas***.

Two charts $\{\mathcal{U}_i, \varphi_i\}$ and $\{\mathcal{U}_j, \varphi_j\}$ are said to be $\mathscr{C}^r$***-compatible*** or ***have*** $\mathscr{C}^r$***-overlap*** if the coordinate change

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(\mathcal{U}_i \cap \mathcal{U}_j) \to \varphi_j(\mathcal{U}_i \cap \mathcal{U}_j)$$

is of class $\mathscr{C}^r$. An atlas Φ on $\mathcal{M}$ is called $\mathscr{C}^r$ if every pair of its charts is $\mathscr{C}^r$-compatible. A maximal $\mathscr{C}^r$-atlas α on $\mathcal{M}$ is called a ***differentiable structure***, and the pair $\{M, \alpha\}$ is called a manifold of class $\mathscr{C}^r$.

A function $f : \mathcal{M} \to \mathbb{R}$ is said to be of class $\mathscr{C}^r$ if $f \circ \varphi_i^{-1} : \mathcal{U}_i \to \mathbb{R}$ is of class $\mathscr{C}^r$ for all charts $\{\mathcal{U}_i, \varphi_i\}$.

In particular, a regular $\mathscr{C}^1$-curve $C \subseteq \mathbb{R}^3$ is a one-dimensional $\mathscr{C}^1$-manifold, and a regular $\mathscr{C}^1$-surface $\Sigma \subseteq \mathbb{R}^3$ is a two-dimensional $\mathscr{C}^1$-manifold.

Definition 4.67 (Metric)**.** Let $\Sigma \subseteq \mathbb{R}^n$ be a $(n-1)$-dimensional manifold. The metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$) is the matrix $g = [g_{\alpha\beta}]_{(n-1)\times(n-1)}$ given by

$$g_{\alpha\beta} = \psi,_\alpha \cdot \psi,_\beta = \sum_{i=1}^{n} \frac{\partial \psi^i}{\partial y_\alpha} \frac{\partial \psi^i}{\partial y_\beta} \qquad \text{in} \quad \mathcal{V}\,.$$

Proposition 4.68. *Let $\Sigma \subseteq \mathbb{R}^n$ be a $(n-1)$-dimensional manifold, and $g = [g_{\alpha\beta}]_{(n-1)\times(n-1)}$ be the metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$). Then the metric tensor g is positive definite; that is,*

$$\sum_{\alpha,\beta=1}^{n-1} g_{\alpha\beta} v^\alpha v^\beta > 0 \qquad \forall\, \boldsymbol{v} = \sum_{\gamma=1}^{n-1} v^\gamma \frac{\partial \psi}{\partial y^\gamma} \neq \mathbf{0}\,.$$

Definition 4.69 (The first fundamental form)**.** Let $\Sigma \subseteq \mathbb{R}^{\rm n}$ be a $(n-1)$-dimensional manifold, and $g = [g_{\alpha\beta}]_{({\rm n}-1)\times({\rm n}-1)}$ be the metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$). The first fundamental form associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$) is the scalar function ${\rm g} = \det(g)$.

Definition 4.70 (Surface integrals)**.** Let $\mathcal{M}$ be an $({\rm n}-1)$-dimensional $\mathscr{C}^1$-manifold, $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$ be a collection of charts of $\mathcal{M}$ and $\{\zeta_i\}_{i\in\mathcal{I}}$ is a partition-of-unity of $\mathcal{M}$ subordinate to $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$. The "surface integral" (or simply integral) of a scalar function $f : \mathcal{M} \to \mathbb{R}$ over $\mathcal{M}$, denoted by $\displaystyle\int_{\mathcal{M}} f\,dS$, is defined by

$$\int_{\mathcal{M}} f\,dS = \sum_{i\in I} \int_{\varphi_i(\mathcal{U}_i)} \big[(\zeta_i f)\circ\varphi^{-1}\big]\sqrt{{\rm g}_i}\,dx\,,$$

where ${\rm g}_i$ is the first fundamental form associated with the parametrization $\big\{\varphi_i(\mathcal{U}_i), \varphi^{-1}\big\}$.

Remark 4.71. Let $C \subseteq \mathbb{R}^3$ be a regular $\mathscr{C}^1$-curve. The line integral of a scalar function $f : C \to \mathbb{R}$ over C is the "surface integral" of f over C defined in (4.70). In other words, $dS = ds$ in the case that $\mathcal{M}$ is a one-dimensional manifold.

4.5.1 Some useful identities

Let $\Sigma \subseteq \mathbb{R}^{\rm n}$ be the boundary of an open set Ω (thus an oriented surface), $\{\mathcal{V}, \psi\}$ be a local parametrization of Σ, and $\mathbf{N} : \Sigma \to \mathbb{R}^{\rm n}$ be the normal vector on Σ which is compatible with the parametrization ψ; that is,

$$\det\big(\big[\,\psi_{,1} \,\vdots\, \psi_{,2} \,\vdots\, \cdots \,\vdots\, \psi_{,{\rm n}-1} \,\vdots\, \mathbf{N}\circ\psi\,\big]\big) > 0\,.$$

Define $\Psi(y', y_n) = \psi(y') + y_n(\mathbf{N}\circ\psi)(y')$. Then $\Psi : \mathcal{V} \times (-\varepsilon, \varepsilon) \to \mathcal{T}$ for some tubular neighborhood $\mathcal{T}$ of Σ.

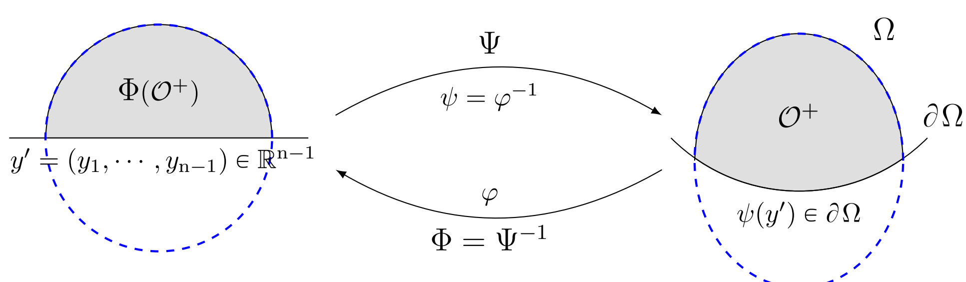


Figure 4.3: The map Ψ constructed from the local parametrization $\{\mathcal{V}, \psi\}$

Since $(\nabla\Psi)\big|_{\{y_n=0\}} = \big[\,\psi_{,1} \,\vdots\, \psi_{,2} \,\vdots\, \cdots \,\vdots\, \psi_{,\mathrm{n}-1} \,\vdots\, \mathbf{N}\circ\psi\,\big]$, Corollary 1.65 and 1.66 implies that

$$\det(\nabla\Psi)^2\big|_{\{y_\mathrm{n}=0\}} = \big[\det\big((\nabla\Psi)^\mathrm{T}\big)\det(\nabla\Psi)\big]\Big|_{\{y_\mathrm{n}=0\}} = \det\big((\nabla\Psi)^\mathrm{T}\nabla\Psi\big)\Big|_{\{y_\mathrm{n}=0\}}$$
$$= \det\left(\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{(\mathrm{n}-1)1} & 0 \\ g_{21} & g_{22} & \cdots & g_{(\mathrm{n}-1)2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{(\mathrm{n}-1)1} & g_{(\mathrm{n}-1)2} & \cdots & g_{(\mathrm{n}-1)(\mathrm{n}-1)} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}\right) = \mathrm{g}\,.$$

Defining J as the Jacobian of the map Ψ; that is, $\mathrm{J} = \det(\nabla\Psi)$, then the identity above implies that

$$\mathrm{J} = \sqrt{\mathrm{g}} \qquad \text{on} \quad \{y_\mathrm{n} = 0\}\,.$$

Moreover, letting A denote the inverse of the Jacobian matrix of Ψ; that is, $\mathrm{A} = (\nabla\Psi)^{-1}$, and letting $\big[g^{\alpha\beta}\big]_{(\mathrm{n}-1)\times(\mathrm{n}-1)}$ be the inverse matrix of $\big[g_{\alpha\beta}\big]_{(\mathrm{n}-1)\times(\mathrm{n}-1)}$, we find that

$$\mathrm{A}\big|_{\{y_n=0\}} = \left[\,\sum_{\alpha=1}^{\mathrm{n}-1} g^{1\alpha}\psi_{,\alpha} \,\vdots\, \cdots\cdots\cdots\cdots \,\vdots\, \sum_{\alpha=1}^{\mathrm{n}-1} g^{(\mathrm{n}-1)\alpha}\psi_{,\alpha} \,\vdots\, \mathbf{N}\circ\psi\,\right]^\mathrm{T}.$$

As a consequence,

$$(\mathrm{J}\mathrm{A}^\mathrm{T}\mathrm{e}_\mathrm{n})\big|_{\{y_\mathrm{n}=0\}} = \sqrt{\mathrm{g}}\,(\mathbf{N}\circ\psi)\,. \tag{4.17}$$

4.6 The Divergence Theorem

Two differential operators play important roles in vector calculus. The first one is called the **divergence operator** which measures the flux of a vector field, and the second one is called the **curl operator** which measures the circulation (the speed of rotation) of a vector field. We will study this two operators in the following two sections.

4.6.1 Flux integrals

Let $\Sigma \subseteq \mathbb{R}^3$ be an oriented surface with a fixed unit normal vector field $\mathbf{N} : \Sigma \to \mathbb{R}^3$, and $\boldsymbol{u} : \Sigma \to \mathbb{R}^3$ be a vector-valued function. The flux integral of $\boldsymbol{u}$ over Σ with given orientation $\mathbf{N}$ is the surface integral of $\boldsymbol{u}\cdot\mathbf{N}$ over Σ.

Physical interpretation

Let $\Omega \subseteq \mathbb{R}^3$ be an open set which stands for a fluid container and fully contains some liquid such as water, and $\boldsymbol{u} : \Omega \to \mathbb{R}^3$ be a vector-field which stands for the fluid velocity; that is, $\boldsymbol{u}(x)$ is the fluid velocity at point $x \in \Omega$. Furthermore, let $\Sigma \subseteq \Omega$ be a surface immersed in the fluid with given orientation $\mathbf{N}$, and $c : \Omega \to \mathbb{R}$ be the concentration of certain material dissolving in the liquid. Then the amount of the material carried across the surface in the direction $\mathbf{N}$ by the fluid in a time period of Δt is

$$\Delta t \cdot \int_\Sigma c\boldsymbol{u} \cdot \mathbf{N}\, dS\,.$$

Therefore, $\displaystyle\int_\Sigma c\boldsymbol{u}\cdot\mathbf{N}\, dS$ is the instantaneous amount of the material carried across the surface in the direction $\mathbf{N}$ by the fluid.

Example 4.72. Find the flux integral of the vector field $\boldsymbol{F}(x, y, z) = (x, y^2, z)$ upward through the first octant part Σ of the cylindrical surface $x^2 + z^2 = a^2$, $0 < y < b$.

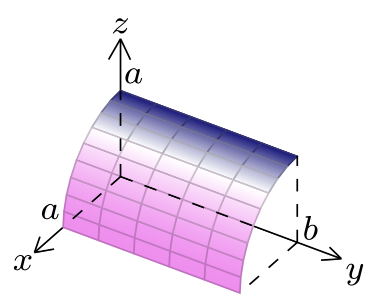


Figure 4.4: The surface Σ

Fist, we parameterize Σ by

$$\psi(u, v) = (u, v, \sqrt{a^2 - u^2}), \quad (u, v) \in \mathcal{V} = (0, a) \times (0, b)\,.$$

Since the first fundamental form g associated with $\{\mathcal{V}, \psi\}$ is $\mathrm{g} = \|\psi_{,1} \times \psi_{,2}\|^2_{\mathbb{R}^3} = \dfrac{a^2}{a^2 - u^2}$, and the upward-pointing unit normal is $\mathbf{N}(x, y, z) = (\dfrac{x}{a}, 0, \dfrac{z}{a})$, we have

$$\begin{aligned}\int_\Sigma \boldsymbol{F} \cdot \mathbf{N}\, dS &= \int_{\mathcal{V}} \frac{1}{a}(u^2 + a^2 - u^2)\frac{a}{\sqrt{a^2 - u^2}}\, d(u, v) = a^2 \int_{\mathcal{V}} \frac{1}{\sqrt{a^2 - u^2}}\, d(u, v) \\ &= a^2 \int_0^b \int_0^a \frac{1}{\sqrt{a^2 - u^2}}\, du dv = a^2 b \arcsin \frac{u}{a}\Big|_{u=0}^{u=a} = \frac{\pi a^2 b}{2}\,.\end{aligned}$$

4.6.2 Measurements of the flux - the divergence operator

Let $\Omega \subseteq \mathbb{R}^3$ be an open set, and $\boldsymbol{u} : \Omega \to \mathbb{R}^3$ be a $\mathscr{C}^1$ vector field. Suppose that $\mathcal{O}$ is a bounded open set of class $\mathscr{C}^1$ such that $\overline{\mathcal{O}} \subseteq \Omega$ with outward-point unit normal vector field $\mathbf{N}$. Then the flux integral of $\boldsymbol{u}$ over $\partial\mathcal{O}$ in the direction $\mathbf{N}$ is

$$\int_{\partial\mathcal{O}} \boldsymbol{u} \cdot \mathbf{N}\, dS\,.$$

Consider a special case that $\mathcal{O} = B(a, r)$ for some ball in $\mathbb{R}^3$ centered at a with radius r. We first compute $\displaystyle\int_{\partial B(a,r)} \boldsymbol{u}^3 \mathbf{N}_3\, dS$. Consider

$$\begin{aligned}
\psi_+(x_1, x_2) &= \big(x_1, x_2, a_3 + \sqrt{r^2 - (x_1 - a)^2 - (x_2 - a_2)^2}\big)\,, \qquad (x_1, x_2) \in D(a, r)\,,\\
\psi_-(x_2, x_2) &= \big(x_1, x_2, a_3 - \sqrt{r^2 - (x_1 - a)^2 - (x_2 - a_2)^2}\big)\,, \qquad (x_1, x_2) \in D(a, r)\,,
\end{aligned}$$

where $D(a, r)$ is the disk in $\mathbb{R}^2$ given by $\big\{(x_1, x_2) \in \mathbb{R}^2 \,\big|\, (x_1 - a_1)^2 + (x_2 - a_2)^2 \leqslant r^2\big\}$. Since $\partial B(a, r) \backslash \big(\psi_+(D(a, r)) \cup \psi_-(D(a, r))$ is the equator of the sphere $\partial B(a, r)$ which has zero area, we must have

$$\int_{\partial B(a,r)} \boldsymbol{u}^3 \mathbf{N}_3\, dS = \int_{\psi_+(D(a,r))} \boldsymbol{u}^3 \mathbf{N}_3\, dS + \int_{\psi_-(D(a,r))} \boldsymbol{u}^3 \mathbf{N}_3\, dS\,.$$

Note that $(\mathbf{N} \circ \psi_\pm)(x_1, x_2) = \dfrac{1}{r}\big(\psi_\pm(x_1, x_2) - a\big)$. In view of Example 4.49, we have

$$\begin{aligned}
&\int_{\psi_+(D(a,r))} \boldsymbol{u}^3 \mathbf{N}_3\, dS \\
&\quad = \int_{D(a,r)} \boldsymbol{u}^3(\psi_+(x_1, x_2)) \frac{\sqrt{r^2 - (x_1 - a_1)^2 - (x_2 - a_2)^2}}{r} \frac{r}{\sqrt{r^2 - (x_1 - a_1)^2 - (x_2 - a_2)^2}}\, d\mathbb{A} \\
&\quad = \int_{D(a,r)} \boldsymbol{u}^3(\psi_+(x_1, x_2))\, d\mathbb{A}\,.
\end{aligned}$$

and similarly,

$$\int_{\psi_+(D(a,r))} \boldsymbol{u}^3 \mathbf{N}_3\, dS = -\int_{D(a,r)} \boldsymbol{u}^3(\psi_-(x_1, x_2))\, d\mathbb{A}\,.$$

Therefore,

$$\begin{aligned}
\int_{\partial B(a,r)} \boldsymbol{u}^3 \mathbf{N}_3\, dS &= \int_{D(a,r)} \big[\boldsymbol{u}^3(\psi_+(x_1, x_2)) - \boldsymbol{u}^3(\psi_-(x_1, x_2))\big]\, d\mathbb{A} \\
&= \int_{D(a,r)} \Big(\int_{a_3 - \sqrt{r^2 - (x_1 - a_1)^2 - (x_2 - a_2)^2}}^{a_3 + \sqrt{r^2 - (x_1 - a_1)^2 - (x_2 - a_2)^2}} \frac{\partial \boldsymbol{u}^3}{\partial x_3}(x_1, x_2, x_3)\, dx_3\Big) d\,\mathbb{A} \\
&= \int_{B(a,r)} \frac{\partial \boldsymbol{u}^3}{\partial x_3}\, dx\,.
\end{aligned}$$

Similarly,

$$\int_{\partial B(a,r)} \boldsymbol{u}^1 \mathbf{N}_1 \, dS = \int_{B(a,r)} \frac{\partial u^1}{\partial x_1} \, dx \quad \text{and} \quad \int_{\partial B(a,r)} \boldsymbol{u}^2 \mathbf{N}_2 \, dS = \int_{B(a,r)} \frac{\partial u^2}{\partial x_2} \, dx \,;$$

thus we conclude that

$$\int_{\partial B(a,r)} \boldsymbol{u} \cdot \mathbf{N} \, dS = \int_{B(a,r)} \sum_{i=1}^{3} \frac{\partial \boldsymbol{u}^i}{\partial x_i} \, dx \,.$$

The computation above motivates the following

Definition 4.73 (The divergence operator)**.** Let $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ be a vector field. The divergence of $\boldsymbol{u}$ is a scalar function defined by

$$\operatorname{div}\boldsymbol{u} = \sum_{i=1}^{\mathrm{n}} \frac{\partial \boldsymbol{u}^i}{\partial x_i} \,.$$

Definition 4.74. A vector field $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ is called *solenoidal* or *divergence-free* if $\operatorname{div}\boldsymbol{u} = 0$ in Ω.

4.6.3 The divergence theorem

Theorem 4.75 (The divergence theorem)**.** *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded Lipschitz domain, and $\boldsymbol{v} \in \mathscr{C}^1(\Omega) \cap \mathscr{C}(\overline{\Omega})$. Then*

$$\int_{\Omega} \operatorname{div}\boldsymbol{v} \, dx = \int_{\partial\Omega} \boldsymbol{v} \cdot \mathbf{N} \, dS \,,$$

where $\mathbf{N}$ is the outward-pointing unit normal of Ω.

Proof. To embrace the beauty of geometry (and the context that we have introduced), we prove the case that Ω is a bounded open set of class $\mathscr{C}^3$.

Let $\{\mathcal{U}_m\}_{m=1}^{K}$ be an open cover of $\partial\Omega$ such that for each $m \in \{1, \cdots, K\}$ there exists a $\mathscr{C}^3$-parametrization $\psi_m : \mathcal{V}_m \subseteq \mathbb{R}^{\mathrm{n}-1} \to \mathcal{U}_m$ which is compatible with the orientation $\mathbf{N}$; that is,

$$\det\big([\psi_{m,1} \,\vdots\, \cdots \,\vdots\, \psi_{m,\mathrm{n}-1} \,\vdots\, \mathbf{N} \circ \psi_m]\big) > 0 \qquad \text{on} \quad \mathcal{V}_m \,.$$

Define $\vartheta_m(y', y_{\mathrm{n}}) = \psi_m(y') + y_{\mathrm{n}}(\mathbf{N} \circ \psi_m)(y')$ as in Section 4.5.1. Then there exists $\varepsilon_m > 0$ such that $\vartheta_m : \mathcal{V}_m \times (-\varepsilon_m, \varepsilon_m) \to \mathcal{W}_m$ is a $\mathscr{C}^2$-diffeomorphism for some open set in $\mathbb{R}^{\mathrm{n}}$ such that $\vartheta_m : \mathcal{V}_m \times (-\varepsilon_m, 0) \to \Omega \cap \mathcal{W}_m$ while $\vartheta_m : \mathcal{V}_m \times (0, \varepsilon_m) \to \operatorname{int}(\Omega^{\complement}) \cap \mathcal{W}_m$.

Choose an open set $\mathcal{W}_0 \subseteq \mathbb{R}^{\mathrm{n}}$ such that $\overline{\mathcal{W}_0} \subseteq \Omega$ and $\overline{\Omega} \subseteq \bigcup\limits_{m=0}^{K} \mathcal{W}_m$, and define ϑ_0 as the identity map. Let $0 \leqslant \zeta_m \leqslant 1$ in $\mathscr{C}_c^\infty(\mathcal{U}_m)$ denote a partition-of-unity of $\overline{\Omega}$ subordinate to the open covering $\{\mathcal{W}_m\}_{m=0}^K$; that is,

$$\sum_{m=0}^{K} \zeta_m = 1 \quad \text{and} \quad \mathrm{spt}(\zeta_m) \subseteq \mathcal{U}_m \quad \forall\, m\,.$$

Let $\mathrm{J}_m = \det(\nabla \vartheta_m)$, $\mathrm{A}_m = (\nabla \vartheta_m)^{-1}$, and g_m denote the first fundamental form associated with $\{\mathcal{V}_m, \psi_m\}$. Using (4.17), $\sqrt{\mathrm{g}_m}(\mathbf{N} \circ \vartheta_m) = \mathrm{J}_m(\mathrm{A}_m)^{\mathrm{T}}\mathrm{e}_{\mathrm{n}}$ on $\mathcal{V}_m \times \{0\}$ for $m \in \{1, \cdots, K\}$. Therefore, making change of variable $x = \vartheta_m(y)$ in each $\mathcal{W}_m$ we find that

$$\begin{aligned}
\int_{\partial\Omega} \boldsymbol{v} \cdot \mathbf{N}\, dS &= \sum_{m=1}^{K} \int_{\partial\Omega \cap \mathcal{W}_m} \zeta_m(\boldsymbol{v} \cdot \mathbf{N})\, dS \\
&= \sum_{m=1}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times \{y_{\mathrm{n}}=0\}} (\zeta_m \circ \vartheta_m)(\boldsymbol{v}^i \circ \vartheta_m)(\mathbf{N}^i \circ \vartheta_m)\sqrt{\mathrm{g}_m}\, dy' \\
&= \sum_{m=1}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times \{y_{\mathrm{n}}=0\}} (\zeta_m \circ \vartheta_m)(\boldsymbol{v}^i \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^{\mathrm{n}}\, dy' \\
&= \sum_{m=1}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times (-\varepsilon_m, 0)} \frac{\partial}{\partial y_{\mathrm{n}}}\big[(\zeta_m \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^{\mathrm{n}}(\boldsymbol{v}^i \circ \vartheta_m)\big]\, dy\,.
\end{aligned}$$

On the other hand, for $\alpha \in \{1, \cdots, \mathrm{n}-1\}$ and $i \in \{1, \cdots, \mathrm{n}\}$,

$$\int_{\mathcal{V}_m \times (-\varepsilon_m, 0)} \frac{\partial}{\partial y_\alpha}\big[(\zeta_m \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^{\alpha}(\boldsymbol{v}^i \circ \vartheta_m)\big]\, dy = 0\,;$$

thus the Piola identity (2.6) implies that

$$\begin{aligned}
\int_{\partial\Omega} \boldsymbol{v} \cdot \mathbf{N}\, dS &= \sum_{m=1}^{K} \sum_{i,j=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times (-\varepsilon_m, 0)} \frac{\partial}{\partial y_j}\big[(\zeta_m \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^{j}(\boldsymbol{v}^i \circ \vartheta_m)\big]\, dy \\
&= \sum_{m=1}^{K} \sum_{i,j=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times (-\varepsilon_m, 0)} \mathrm{J}_m(\mathrm{A}_m)_i^{j}(\zeta_m \circ \vartheta_m)_{,j}\,(\boldsymbol{v}^i \circ \vartheta_m)\, dy \\
&\quad + \sum_{m=1}^{K} \sum_{i,j=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times (-\varepsilon_m, 0)} (\zeta_m \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^{j}(\boldsymbol{v}^i \circ \vartheta_m)_{,j}\, dy\,.
\end{aligned}$$

Making change of variable $y = \vartheta_m^{-1}(x)$ in each $\mathcal{V}_m \times (-\varepsilon_m, 0)$ again, by the fact that

$$\sum_{i,j=1}^{\mathrm{n}} (\mathrm{A}_m)_i^j(\boldsymbol{v}^i \circ \theta_m)_{,j} = (\mathrm{div}\boldsymbol{v}) \circ \theta_m \qquad \text{and} \qquad \int_{\mathcal{W}_0} \mathrm{div}\big(\zeta_0 \boldsymbol{v}\big)\, dx = 0\,,$$

we conclude that

$$\begin{aligned}\int_{\partial\Omega} \boldsymbol{v}\cdot\mathbf{N}\,dS &= \int_{\mathcal{W}_0} \operatorname{div}\big(\zeta_0\boldsymbol{v}\big)\,dx + \sum_{m=1}^{K}\int_{\mathcal{W}_m} (\boldsymbol{v}\cdot\nabla_x)\zeta_m\,dx + \sum_{m=1}^{K}\int_{\mathcal{W}_m} \zeta_m \operatorname{div}\boldsymbol{v}\,dx \\ &= \sum_{m=0}^{K}\int_{\mathcal{W}_m} (\boldsymbol{v}\cdot\nabla_x)\zeta_m\,dx + \sum_{m=0}^{K}\int_{\mathcal{W}_m} \zeta_m \operatorname{div}\boldsymbol{v}\,dx \\ &= \int_\Omega (\boldsymbol{v}\cdot\nabla_x)1\,dx + \int_\Omega \operatorname{div}\boldsymbol{v}\,dx = \int_\Omega \operatorname{div}\boldsymbol{v}\,dx\,. \end{aligned}$$ □

Letting $\boldsymbol{v} = (0,\cdots,0,f,0,\cdots,0) = f\mathrm{e}_i$, we obtain the following

Corollary 4.76. *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded Lipschitz domain, and $f \in \mathscr{C}^1(\Omega)\cap\mathscr{C}(\overline{\Omega})$. Then*

$$\int_\Omega \frac{\partial f}{\partial x_i}\,dx = \int_{\partial\Omega} f\,\mathbf{N}_i\,dS\,,$$

where $\mathbf{N}_i$ is the i-th component of the outward-pointing unit normal $\mathbf{N}$ of Ω.

Letting $\boldsymbol{v}$ be the product of a scalar function and a vector-valued function in Theorem 4.75, we conclude the following

Corollary 4.77. *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded Lipschitz domain, and $\boldsymbol{v} \in \mathscr{C}^1(\Omega;\mathbb{R}^{\mathrm{n}})\cap\mathscr{C}(\overline{\Omega};\mathbb{R}^{\mathrm{n}})$ be a vector-valued function and $\varphi \in \mathscr{C}^1(\Omega)\cap\mathscr{C}(\overline{\Omega})$ be a scalar function. Then*

$$\int_\Omega \varphi\operatorname{div}\boldsymbol{v}\,dx = \int_{\partial\Omega} (\boldsymbol{v}\cdot\mathbf{N})\varphi\,dS - \int_\Omega \boldsymbol{v}\cdot\nabla\varphi\,dx\,, \tag{4.18}$$

where $\mathbf{N}$ is the outward-pointing unit normal on $\partial\Omega$.

Example 4.78. Let Ω be the the first octant part bounded by the cylindrical surface $x^2+z^2=a^2$ and the plane $y=b$, and $\boldsymbol{F}:\Omega\to\mathbb{R}^3$ be a vector-valued function defined by $\boldsymbol{F}(x,y,z) = (x,y^2,z)$.

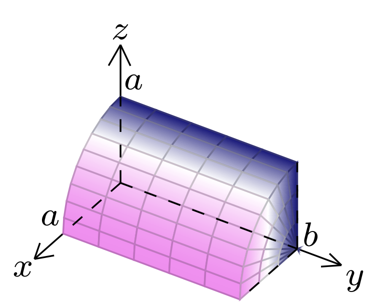


Figure 4.5: The domain Ω and its five pieces of boundaries

With $\mathbf{N}$ denoting the outward-pointing unit normal of $\partial\Omega$,

$$\begin{aligned}\int_\Omega \operatorname{div}\boldsymbol{F}\,d(x,y,z) &= \int_0^a\int_0^b\int_0^{\sqrt{a^2-x^2}}(2+2y)\,dzdydx = (b^2+2b)\int_0^a\int_0^{\sqrt{a^2-x^2}}dzdx\\ &= \frac{\pi a^2(b^2+2b)}{4}\,.\end{aligned}$$

On the other hand, we note that the boundary of Ω has five parts: Σ as given in Example 4.72, two rectangles $\mathrm{R}_1 = \{x=0\}\times[0,b]\times[0,a]$, $\mathrm{R}_2 = [0,a]\times[0,b]\times\{z=0\}$, and two quarter disc $\mathrm{D}_1 = \big\{(x,0,z)\in\mathbb{R}^3\,\big|\,x^2+z^2\leqslant a^2, x,z\geqslant 0\big\}$ and $\mathrm{D}_2 = \big\{(x,b,z)\in\mathbb{R}^3\,\big|\,x^2+z^2\leqslant a^2, x,z\geqslant 0\big\}$. Therefore,

$$\begin{aligned}\int_{\mathrm{R}_1}\boldsymbol{F}\cdot\mathbf{N}\,dS &= \int_0^a\int_0^b(0,y^2,z)\cdot(-1,0,0)\,dydz = 0\,,\\ \int_{\mathrm{R}_2}\boldsymbol{F}\cdot\mathbf{N}\,dS &= \int_0^a\int_0^b(x,y^2,0)\cdot(0,0,-1)\,dydx = 0\,,\\ \int_{\mathrm{D}_1}\boldsymbol{F}\cdot\mathbf{N}\,dS &= \int_0^a\int_0^{\sqrt{a^2-x^2}}(x,0,z)\cdot(0,-1,0)\,dzdx = 0\,,\end{aligned}$$

and

$$\int_{\mathrm{D}_1}\boldsymbol{F}\cdot\mathbf{N}\,dS = \int_0^a\int_0^{\sqrt{a^2-x^2}}(x,b^2,z)\cdot(0,1,0)\,dzdx = b^2\int_0^a\int_0^{\sqrt{a^2-x^2}}dzdx = \frac{\pi a^2b^2}{4}\,.$$

Together with the result in Example 4.72, we find that

$$\begin{aligned}\int_{\partial\Omega}\boldsymbol{F}\cdot\mathbf{N}\,dS &= \Big(\int_\Sigma+\int_{\mathrm{R}_1}+\int_{\mathrm{R}_2}+\int_{\mathrm{D}_1}+\int_{\mathrm{D}_2}\Big)\boldsymbol{F}\cdot\mathbf{N}\,dS = \frac{\pi a^2b^2}{4}+\frac{\pi a^2 b}{2} = \frac{\pi a^2(b^2+2b)}{4}\\ &= \int_\Omega\operatorname{div}\boldsymbol{F}\,d(x,y,z)\,.\end{aligned}$$

4.6.4 The divergence theorem on surfaces with boundary

This section is devoted to the divergence theorem on surfaces in $\mathbb{R}^3$ instead of domains of $\mathbb{R}^\mathrm{n}$. To do so, we need to define what the divergence operator on a surface is, and this requires that we first define the vector fields on which the surface divergence operator acts.

Definition 4.79. Let $\Sigma\subseteq\mathbb{R}^3$ be an open $\mathscr{C}^1$-surface; that is, Σ is of class $\mathscr{C}^1$ and $\Sigma\cap\partial\Sigma = \varnothing$. A vector field $\boldsymbol{u}$ defined on Σ is called a tangent vector field on Σ, denoted by $\boldsymbol{u}\in\mathbf{T}\Sigma$, if $\boldsymbol{u}\cdot\mathbf{N}=0$ on Σ, where $\mathbf{N}:\Sigma\to\mathbb{S}^2$ is a unit normal vector field on Σ.

Having established (4.18), we find that the divergence operator div is the formal adjoint of the operator $-\nabla$. The following definition is motivated by this observation.

Definition 4.80 (The surface gradient and the surface divergence)**.** Let $\Sigma \subseteq \mathbb{R}^{\rm n}$ be a regular $\mathscr{C}^1$-surface. The surface gradient of a function $f : \Sigma \to \mathbb{R}$, denoted by $\nabla_\Sigma f$, is a vector-valued function from Σ to $\mathbf{T}_p\Sigma$ given, in a local parametrization $\{\mathcal{V}, \psi\}$, by

$$(\nabla_\Sigma f) \circ \psi = \sum_{\alpha,\beta=1}^{\rm n-1} g^{\alpha\beta} \frac{\partial(f \circ \psi)}{\partial y_\alpha} \frac{\partial \psi}{\partial y_\beta} ,$$

where $[g^{\alpha\beta}]$ is the inverse matrix of the metric tensor $[g_{\alpha\beta}]$ associated with $\{\mathcal{V}, \psi\}$, and $\Big\{\dfrac{\partial \psi}{\partial y_\beta}\Big\}_{\beta=1}^{2}$ are tangent vectors to Σ.

The surface divergence operator div_Σ is defined as the formal adjoint of $-\nabla_\Sigma$; that is, if $\boldsymbol{u} \in \mathbf{T}\Sigma$, then

$$-\int_\Sigma \boldsymbol{u} \cdot \nabla_\Sigma f \, dS = \int_\Sigma f \, \mathrm{div}_\Sigma \boldsymbol{u} \, dS \qquad \forall\, f \in \mathscr{C}_c^1(\Sigma; \mathbb{R}) .$$

In a local parametrization $(\mathcal{V}, \psi)$,

$$(\mathrm{div}_\Sigma \boldsymbol{u}) \circ \psi = \frac{1}{\sqrt{\rm g}} \sum_{\alpha,\beta=1}^{\rm n-1} \frac{\partial}{\partial y_\alpha} \Big[\sqrt{\rm g} g^{\alpha\beta} \Big((\boldsymbol{u} \circ \psi) \cdot \frac{\partial \psi}{\partial y_\beta} \Big) \Big] ,$$

where ${\rm g} = \det(g)$ is the first fundamental form associated with $\{\mathcal{V}, \psi\}$.

Remark 4.81. Suppose that $f : \mathcal{O} \subseteq \mathbb{R}^3 \to \mathbb{R}$ for some open set containing Σ. Then the surface gradient of f at $p \in \Sigma$ is the projection of the gradient vector $(\nabla f)(p)$ onto the tangent plane $T_p\Sigma$. In other words, let $\mathbf{N} : \Sigma \to \mathbb{R}^3$ be a continuous unit normal vector field on Σ, then

$$(\nabla_\Sigma f)(p) = (\nabla f)(p) - \big[(\nabla f)(p) \cdot \mathbf{N}(p)\big] \mathbf{N}(p) \quad \text{(or simply } \nabla_\Sigma f = \nabla f - (\nabla f \cdot \mathbf{N})\mathbf{N}) .$$

Definition 4.82 (Surfaces with Boundary)**.** An oriented $\mathscr{C}^k$-surface $\Sigma \subseteq \mathbb{R}^3$ is said to have $\mathscr{C}^\ell$-boundary $\partial\Sigma$ if there exists a collection of pairs $\{\mathcal{V}_m, \psi_m\}_{m=1}^K$, called a collection of local parametrization of $\overline{\Sigma}$, if

1. $\mathcal{V}_m \subseteq \mathbb{R}^2$ is open and $\psi_m : \mathcal{V}_m \to \mathbb{R}^3$ is one-to-one map of class $\mathscr{C}^k$ for all $m \in \{1, \cdots, K\}$;

2. $\psi_m(\mathcal{V}_m) \cap \Sigma \neq \varnothing$ for all $m \in \{1, \cdots, K\}$ and $\overline{\Sigma} \subseteq \bigcup_{m=1}^K \psi_m(\mathcal{V}_m)$;

3. $\psi_m : \mathcal{V}_m \to \psi_m(\mathcal{V}_m)$ is a $\mathscr{C}^k$-diffeomorphism if $\psi_m(\mathcal{V}_m) \subseteq \Sigma$;

4. $\psi_m : \mathcal{V}_m^+ \equiv \mathcal{V}_m \cap \{y_2 > 0\} \to \psi_m(\mathcal{V}_m) \cap \Sigma$ is a $\mathscr{C}^k$-diffeomorphism if $\mathcal{U}_m \cap \partial\Sigma \neq \varnothing$;

5. $\psi_m : \mathcal{V}_m \cap \{y_2 = 0\} \to \mathcal{U}_m \cap \partial\Sigma$ is of class $\mathscr{C}^\ell$ if $\mathcal{U}_m \cap \partial\Sigma \neq \varnothing$.

Now we are in the position of stating the divergence theorem on surfaces with boundary.

Theorem 4.83. *Let $\Sigma \subseteq \mathbb{R}^3$ be an oriented $\mathscr{C}^1$-surface with $\mathscr{C}^1$-boundary $\partial\Sigma$, $\mathbf{N} : \Sigma \to \mathbb{S}^2$ be a continuous unit normal vector field on Σ, and $\mathbf{T} : \partial\Sigma \to \mathbb{S}^2$ be tangent vector on $\partial\Sigma$ such that $\mathbf{T}$ is compatible with $\mathbf{N}$ (which means $\mathbf{T} \times \mathbf{N}$ points away from Σ). Then*

$$\int_{\partial\Sigma} \boldsymbol{u} \cdot (\mathbf{T} \times \mathbf{N})\, ds = \int_\Sigma \mathrm{div}_\Sigma \boldsymbol{u}\, dS \qquad \forall\, \boldsymbol{u} \in \mathbf{T}\Sigma \cap \mathscr{C}^1(\Sigma; \mathbb{R}^3) \cap \mathscr{C}(\overline{\Sigma}; \mathbb{R}^3)\,,$$

where div_Σ *is the surface divergence operator.*

Proof. Let $\{\mathcal{V}_m, \psi_m\}_{m=1}^K$ denote a collection of local parametrization of $\overline{\Sigma}$ such that $\psi_m(\mathcal{V}_m) \cap \partial\Sigma = \varnothing$ for $1 \leqslant m \leqslant J$, and $\psi_m(\mathcal{V}_m) \cap \partial\Sigma$ is non-empty and connected for $J+1 \leqslant m \leqslant K$. W.L.O.G., we can assume that $\mathcal{V}_m = B_m \equiv B(0, r_m)$ for some $r_m > 0$. Write $\mathcal{U}_m = \psi_m(\mathcal{V}_m)$, and let $\{g_m\}_{m=1}^K$ be the associated metric tensor, as well as the associated first fundamental form $\mathrm{g}_m = \det(g_m)$. Let $\{\zeta_m\}_{m=1}^K$ be a partition-of-unity of $\overline{\Sigma}$ subordinate to $\{\mathcal{U}_m\}_{m=1}^K$. Then

$$\begin{aligned}
\int_\Sigma \mathrm{div}_\Sigma \boldsymbol{u}\, dS &= \sum_{m=1}^K \int_{\mathcal{U}_m \cap \Sigma} \zeta_m \mathrm{div}_\Sigma \boldsymbol{u}\, dS \\
&= \sum_{m=1}^J \sum_{\alpha,\beta=1}^2 \int_{B_m} (\zeta_m \circ \psi_m) \frac{\partial}{\partial y_\alpha} \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \Big((\boldsymbol{u} \circ \psi_m) \cdot \frac{\partial \psi_m}{\partial y_\beta} \Big) \Big] dy \\
&\quad + \sum_{m=J+1}^K \sum_{\alpha,\beta=1}^2 \int_{B_m^+} (\zeta_m \circ \psi_m) \frac{\partial}{\partial y_\alpha} \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \Big((\boldsymbol{u} \circ \psi_m) \cdot \frac{\partial \psi_m}{\partial y_\beta} \Big) \Big] dy\,.
\end{aligned}$$

Let $\boldsymbol{n}$ denote the outward-pointing unit normal on either ∂B_m for $1 \leqslant m \leqslant J$ or ∂B_m^+ for $J+1 \leqslant m \leqslant K$. Since $\zeta_m \circ \vartheta_m = 0$ on $\partial B(0, r_m)$ for $1 \leqslant m \leqslant J$, and $\zeta_m \circ \vartheta_m = 0$ on

$\{y_2 > 0\} \cap \partial B(0, r_m)$ for $J+1 \leqslant m \leqslant K$, the divergence theorem (on $\mathbb{R}^2$) implies that

$$
\begin{aligned}
\int_\Sigma \operatorname{div}_\Sigma \boldsymbol{u}\, dS &= -\sum_{m=1}^{K} \sum_{\alpha,\beta=1}^{2} \int_{\psi_m^{-1}(\mathcal{U}_m\cap\Sigma)} \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\psi_m)\cdot \frac{\partial \psi_m}{\partial y_\beta}\big)\Big] \frac{\partial}{\partial y_\alpha}(\zeta_m\circ\psi_m)\, dy \\
&\qquad + \sum_{m=J+1}^{K} \sum_{\alpha,\beta=1}^{2} \int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)\boldsymbol{n}_\alpha \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\psi_m)\cdot \frac{\partial \psi_m}{\partial y_\beta}\big)\Big]\, dy_1 \\
&= -\sum_{m=1}^{K} \int_{\psi_m^{-1}(\mathcal{U}_m\cap\Sigma)} (\boldsymbol{u}\cdot\nabla_\Sigma\zeta_m)\circ\psi_m\sqrt{\mathrm{g}_m}\, dy \\
&\qquad + \sum_{m=J+1}^{K} \int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)(\boldsymbol{u}\circ\psi_m)\cdot \Big[\sum_{\alpha,\beta=1}^{2} \boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\frac{\partial \psi_m}{\partial y_\beta}\Big]\, dy_1\,.
\end{aligned}
$$

Since

$$
\sum_{m=1}^{K} \int_{\psi_m^{-1}(\mathcal{U}_m\cap\Sigma)} (\boldsymbol{u}\cdot\nabla_\Sigma\zeta_m)\circ\psi_m\sqrt{\mathrm{g}_m}\, dy = \sum_{m=1}^{K}\int_{\mathcal{U}_m\cap\Sigma} (\boldsymbol{u}\cdot\nabla_\Sigma\zeta_m)\, dS = \int_\Sigma \big(\boldsymbol{u}\cdot\nabla\sum_{m=1}^{K}\zeta_m\big)\, dS = 0\,,
$$

we conclude that

$$
\int_\Sigma \operatorname{div}_\Sigma \boldsymbol{u}\, dS = \sum_{m=J+1}^{K} \int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)(\boldsymbol{u}\circ\psi_m)\cdot \Big[\sum_{\alpha,\beta=1}^{2} \boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\frac{\partial \psi_m}{\partial y_\beta}\Big]\, dy_1\,.
$$

On the other hand,

$$
\begin{aligned}
\int_{\partial\Sigma} \boldsymbol{u}\cdot(\mathbf{T}\times\mathbf{N})\, ds &= \sum_{m=J+1}^{K}\int_{\partial\Sigma\cap\mathcal{U}_m} \zeta_m \boldsymbol{u}\cdot(\mathbf{T}\times\mathbf{N})\, ds \\
&= \sum_{m=J+1}^{K}\int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)(\boldsymbol{u}\circ\psi_m)\cdot\Big[(\mathbf{T}\times\mathbf{N})\circ\psi_m\Big|\frac{\partial\psi_m}{\partial y_1}\Big|\Big]\, dy_1\,.
\end{aligned}
$$

Therefore, the theorem can be concluded as long as we can show that

$$
\sum_{\alpha,\beta=1}^{2} \boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m}\, g_m^{\alpha\beta}\,\frac{\partial \psi_m}{\partial y_\beta} = (\mathbf{T}\times\mathbf{N})\circ\psi_m\,\Big|\frac{\partial\psi_m}{\partial y_1}\Big| \qquad \text{on} \quad B_m\cap\{y_2=0\}\,. \tag{4.19}
$$

Let $\boldsymbol{\tau}_m = \sum\limits_{\alpha,\beta=1}^{2} \boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\dfrac{\partial \psi_m}{\partial y_\beta}$ on $B_m\cap\{y_2=0\}$. Since $\boldsymbol{n}_\alpha = -\delta_{2\alpha}$, we find that $\boldsymbol{\tau}_m\cdot\dfrac{\partial\psi_m}{\partial y_1} = 0$ on $B_m\cap\{y_2=0\}$; thus

$$
\boldsymbol{\tau}_m\cdot(\mathbf{T}\circ\psi_m) = 0 \qquad \text{on} \quad B_m\cap\{y_2=0\}\,.
$$

Moreover, noting that $\boldsymbol{\tau}_m$ is a linear combination of tangent vectors $\dfrac{\partial \psi_m}{\partial y_\beta}$, we must have

$$\boldsymbol{\tau}_m \cdot (\mathbf{N} \circ \psi_m) = 0 \qquad \text{on} \quad B_m \cap \{y_2 = 0\}\,.$$

As a consequence,

$$\boldsymbol{\tau}_m \,/\!/\, (\mathbf{T} \times \mathbf{N}) \circ \psi_m \qquad \text{on} \quad B_m \cap \{y_2 = 0\}\,.$$

Since $(\mathbf{T} \times \mathbf{N})$ points away from Σ, while $\dfrac{\partial \psi_m}{\partial y_2} \circ \psi_m^{-1}\Big|_{\partial\Sigma}$ points toward Σ, by the fact that

$$\boldsymbol{\tau}_m \cdot \frac{\partial \psi_m}{\partial y_2} = \sum_{\alpha,\beta=1}^{2} \boldsymbol{n}_\alpha \sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \frac{\partial \psi_m}{\partial y_\beta} \cdot \frac{\partial \psi_m}{\partial y_2} = -\sqrt{\mathrm{g}_m} g_m^{22} < 0\,,$$

we must have $\boldsymbol{\tau}_m \cdot (\mathbf{T} \times \mathbf{N}) \circ \psi_m > 0$ on $B_m \cap \{y_2 = 0\}$. In other words,

$$\boldsymbol{\tau}_m = |\boldsymbol{\tau}_m| (\mathbf{T} \times \mathbf{N}) \circ \psi_m \qquad \text{on} \quad B_m \cap \{y_2 = 0\}\,.$$

Finally, since

$$\boldsymbol{\tau}_m \cdot \boldsymbol{\tau}_m = \sum_{\alpha,\beta,\gamma,\delta=1}^{2} \mathrm{g}_m \, \boldsymbol{n}_\alpha \, \boldsymbol{n}_\gamma \, g_m^{\alpha\beta} g_m^{\gamma\delta} \, \frac{\partial \psi_m}{\partial y_\beta} \cdot \frac{\partial \psi_m}{\partial y_\delta} = \mathrm{g}_m g_m^{22} = g_{m11} = \Big| \frac{\partial \psi_m}{\partial y_1} \Big|^2,$$

we conclude that $\boldsymbol{\tau}_m = \Big| \dfrac{\partial \psi_m}{\partial y_1} \Big| (\mathbf{T} \times \mathbf{N}) \circ \psi_m$ on $\{y_2 = 0\}$; thus (4.19) is established. □

Remark 4.84. On $\partial\Sigma$, the vector $\mathbf{T} \times \mathbf{N}$ is “tangent” to Σ and points away from Σ. In other words, $\mathbf{T} \times \mathbf{N}$ can be treated as the “outward-pointing” unit “normal” of $\partial\Sigma$ which makes the divergence theorem on surfaces more intuitive.

4.7 The Stokes Theorem

4.7.1 Measurements of the circulation - the curl operator

We consider the circulation or the speed of rotation of a vector field u about an axis in the direction $\mathbf{N}$. Let P be a plane passing thorough a point a and having normal $\mathbf{N}$, and C_r be a circle on the plane P centered at a with radius r. Pick the orientation of the unit tangent vector $\mathbf{T}$ which is compatible with the unit normal $\mathbf{N}$ (see Figure 4.6 for reference).

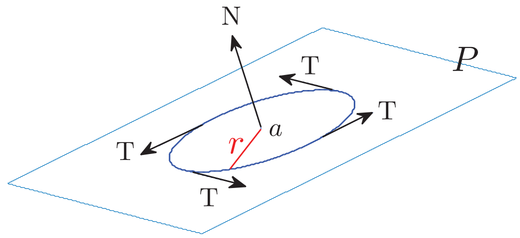


Figure 4.6: the circulation about an axis in direction $\mathbf{N}$

Since the instantaneous angular velocity of a vector field u along the circle C_r is measured by $\dfrac{\boldsymbol{u}\cdot\mathbf{T}}{r}$, it is quite reasonable to measure the circulation of u along C_r by averaging the angular velocity; that is, we consider the quantity

$$\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds \tag{4.20}$$

as a (constant multiple of) measurement of the speed of rotation. The limit of the quantity above, as $r\to 0$, is then a good measurement of the rotation speed of $\boldsymbol{u}$ at the point a about the axis in the direction $\mathbf{N}$.

Since we expect that this measurement does not depend on the choice of coordinate systems, we start from letting P be the x_1x_2-plane, and $\mathbf{N}=(0,0,1)$, $\mathbf{T}=(-\sin\theta,\cos\theta,0)$. By the change of variable $ds=rd\theta$ and the L'Hôspital rule,

$$\begin{aligned}
\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds &= \lim_{r\to 0}\int_0^{2\pi}\frac{\boldsymbol{u}^2\big(a+(r\cos\theta,r\sin\theta,0)\big)\cos\theta-\boldsymbol{u}^1\big(a+(r\cos\theta,r\sin\theta,0)\big)\sin\theta}{2\pi r}\,d\theta\\
&=\frac{1}{2\pi}\int_0^{2\pi}\Big[\boldsymbol{u}^2_{,1}(a)\cos^2\theta+\boldsymbol{u}^2_{,2}(a)\cos\theta\sin\theta-\boldsymbol{u}^1_{,1}(a)\cos\theta\sin\theta-\boldsymbol{u}^1_{,2}(a)\sin^2\theta\Big]d\theta\\
&=\frac{1}{2}\big[\boldsymbol{u}^2_{,1}(a)-\boldsymbol{u}^1_{,2}(a)\big]=\frac{1}{2}\sum_{i,j=1}^{2}\varepsilon_{3ij}\boldsymbol{u}^j_{,i}(a).
\end{aligned}\tag{4.21}$$

Now suppose the general case that $\mathbf{N}\neq e_3$. There is an orthonormal matrix $\mathrm{O}=[\widehat{\mathrm{e}}_1|\widehat{\mathrm{e}}_2|\widehat{\mathrm{e}}_3]$ so that $\mathrm{O}\mathrm{e}_3=\mathbf{N}$. As a consequence, $\widehat{\mathrm{e}}_3=\mathbf{N}$, $\widehat{\mathrm{e}}_j=\mathrm{O}\mathrm{e}_j$ for $j=1,2$, $\mathbf{T}=\mathrm{O}\boldsymbol{\tau}$ with $\boldsymbol{\tau}=(-\sin\theta,\cos\theta,0)$, and the limit of the quantity in (4.20) is given by

$$\begin{aligned}
&\lim_{r\to 0}\frac{1}{2\pi r}\int_0^{2\pi}\boldsymbol{u}(a+r\cos\theta\widehat{\mathrm{e}}_1+r\sin\theta\widehat{\mathrm{e}}_2)\cdot(\mathrm{O}\boldsymbol{\tau})\,d\theta\\
&\quad=\lim_{r\to 0}\frac{1}{2\pi r}\int_0^{2\pi}(\mathrm{O}\boldsymbol{v})(a+r\cos\theta\widehat{\mathrm{e}}_1+r\sin\theta\widehat{\mathrm{e}}_2)\cdot(\mathrm{O}\boldsymbol{\tau})\,d\theta\\
&\quad=\lim_{r\to 0}\frac{1}{2\pi r}\int_0^{2\pi}\boldsymbol{v}(a+r\cos\theta\widehat{\mathrm{e}}_1+r\sin\theta\widehat{\mathrm{e}}_2)\cdot(-\sin\theta,\cos\theta,0)\,d\theta,
\end{aligned}$$

where $\boldsymbol{v} = \mathrm{O}^{\mathbf{T}}\boldsymbol{u}$, and the identity that $(\mathrm{O}\boldsymbol{v})\cdot(\mathrm{O}\boldsymbol{\tau}) = \boldsymbol{v}\cdot\boldsymbol{\tau}$ is used to deduce the last equality. By the L'Hôspital rule again,

$$\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds = \frac{1}{2}\sum_{j=1}^{3}\big[\boldsymbol{v}^2_{,j}(a)\,\widehat{\mathrm{e}}^j_1 - \boldsymbol{v}^1_{,j}(a)\,\widehat{\mathrm{e}}^j_2\big].$$

In fact, we expect this to hold since if using $x' = \mathrm{O}^{\mathbf{T}}x$ as the new coordinate, by (4.21) and the chain rule we obtain that

$$\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds = \frac{1}{2}\Big[\frac{\partial\boldsymbol{v}^2}{\partial x'_1} - \frac{\partial\boldsymbol{v}^1}{\partial x'_2}\Big](a') = \frac{1}{2}\sum_{j=1}^{3}\Big[\frac{\partial\boldsymbol{v}^2}{\partial x_j}(a)\,\widehat{\mathrm{e}}^j_1 - \frac{\partial\boldsymbol{v}^1}{\partial x_j}(a)\,\widehat{\mathrm{e}}^j_2\Big].$$

Finally, we note that $\boldsymbol{v}^\ell_{,j} = \sum\limits_{k=1}^{3}\boldsymbol{u}^k_{,j}\,\mathrm{e}_k\cdot\widehat{\mathrm{e}}_\ell = \sum\limits_{k=1}^{3}\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}^k_\ell$ for $\ell = 1,2,3$; thus

$$\begin{aligned}\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds &= \frac{1}{2}\sum_{j=1}^{3}\big[\boldsymbol{v}^2_{,j}\,\widehat{\mathrm{e}}^j_1 - \boldsymbol{v}^1_{,j}\,\widehat{\mathrm{e}}^j_2\big] = \frac{1}{2}\sum_{j,k=1}^{3}\boldsymbol{u}^k_{,j}\big[\widehat{\mathrm{e}}^k_2\,\widehat{\mathrm{e}}^j_1 - \widehat{\mathrm{e}}^k_1\,\widehat{\mathrm{e}}^j_2\big]\\ &= \frac{1}{2}\sum_{j,k,r,s=1}^{3}(\delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr})\,\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}^r_1\,\widehat{\mathrm{e}}^s_2\,,\end{aligned}$$

where $\delta_{\cdot\cdot}$'s are the Kronecker deltas. Due to the following useful identity

$$\sum_{i=1}^{3}\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}\,, \tag{4.22}$$

we conclude that

$$\begin{aligned}\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds &= \frac{1}{2}\sum_{i,j,k,r,s=1}^{3}\varepsilon_{ijk}\varepsilon_{irs}\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}^r_1\,\widehat{\mathrm{e}}^s_2 = \frac{1}{2}\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\boldsymbol{u}^k_{,j}(\widehat{\mathrm{e}}_1\times\widehat{\mathrm{e}}_2)^i\\ &= \frac{1}{2}\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}^i_3 = \frac{1}{2}\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\boldsymbol{u}^k_{,j}\mathbf{N}_i.\end{aligned}$$

This motivates the following

Definition 4.85 (The curl operator)**.** Let $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$, $\mathrm{n} = 2$ or $\mathrm{n} = 3$, be a vector field.

1. For $\mathrm{n} = 2$, the curl of $\boldsymbol{u}$ is a scalar function defined by

$$\mathrm{curl}\boldsymbol{u} = \sum_{i,j=1}^{2}\varepsilon_{3ij}\,\boldsymbol{u}^j_{,i}\,.$$

2. For n = 3, the curl of $\boldsymbol{u}$ is a vector-valued function defined by

$$(\mathrm{curl}\boldsymbol{u})^i = \sum_{j,k=1}^{3} \varepsilon_{ijk} \boldsymbol{u}^k_{,j} \,.$$

The function $\mathrm{curl}\boldsymbol{u}$ is also called the *vorticity* of $\boldsymbol{u}$, and is usually denoted by one single Greek letter ω.

Having the curl operator defined, for the three-dimensional case the circulation of a vector field $\boldsymbol{u}$ on the plane with normal $\mathbf{N}$ is given by $\dfrac{\mathrm{curl}\boldsymbol{u}\cdot\mathbf{N}}{2}$.

4.7.2 The Stokes theorem

The path we choose to circle around the point a does not have to be a circle. However, in such a case the average of the angular velocity no longer makes sense (since $\boldsymbol{u}\cdot\mathbf{T}$ might not contribute to the motion in the angular direction), and we instead consider the limit of the following quantity

$$\lim_{\mathrm{A}\to 0} \frac{1}{\mathrm{A}} \oint_C \boldsymbol{u}\cdot\mathbf{T}\, ds,$$

where A is the area enclosed by C. This limit is always $\mathrm{curl}\boldsymbol{u}\cdot\mathbf{N}$ because of the famous Stokes' theorem.

Theorem 4.86 (The Stokes theorem)**.** *Let $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ be a smooth vector field, and Σ be a $\mathscr{C}^1$-surface with $\mathscr{C}^1$-boundary $\partial\Sigma$ in Ω. Then*

$$\int_{\partial\Sigma} \boldsymbol{u}\cdot\mathbf{T}\, ds = \int_{\Sigma} \mathrm{curl}\boldsymbol{u}\cdot\mathbf{N}\, dS\,,$$

where $\mathbf{N}$ and $\mathbf{T}$ are compatible normal and tangent vector fields.

To prove the Stokes theorem, we first establish the following

Lemma 4.87. *Let $\Omega \subseteq \mathbb{R}^3$ be a bounded Lipschitz domain, and $\boldsymbol{w} : \Omega \to \mathbb{R}^{\mathrm{n}}$ be a mooth vector-valued function. If $\Sigma \subseteq \Omega$ is an oriented $\mathscr{C}^1$-surface with normal $\mathbf{N}$, then*

$$\mathrm{curl}\boldsymbol{w}\cdot\mathbf{N} = \mathrm{div}_\Sigma(\boldsymbol{w}\times\mathbf{N}) \qquad \textit{on} \quad \Sigma\,. \tag{4.23}$$

Proof. Let $\mathcal{O} \subseteq \Omega$ be a $\mathscr{C}^1$-domain such that $\Sigma \subseteq \partial\mathcal{O}$ and $\mathbf{N}$ is the outward-pointing unit normal on $\partial\mathcal{O}$. In other words, Σ is part of the boundary of $\mathcal{O}$. Since

$$(\nabla\varphi)^i = \frac{\partial\varphi}{\partial\mathbf{N}}\mathbf{N}^i + (\nabla_{\partial\mathcal{O}}\varphi)^i \qquad \text{on} \quad \partial\mathcal{O}\,,$$

by the divergence theorem we conclude that for all $\varphi \in \mathscr{C}^1(\overline{\mathcal{O}})$,

$$\begin{aligned}\int_{\partial\mathcal{O}} (\text{curl}\boldsymbol{w} \cdot \mathbf{N})\varphi\, dS &= \int_{\mathcal{O}} \text{curl}\boldsymbol{w} \cdot \nabla\varphi\, dx = \int_{\partial\mathcal{O}} (\mathbf{N} \times \boldsymbol{w}) \cdot \nabla\varphi\, dS \\ &= \int_{\partial\mathcal{O}} (\mathbf{N} \times \boldsymbol{w}) \cdot \nabla_{\partial\mathcal{O}}\varphi\, dS = \int_{\partial\mathcal{O}} \text{div}_{\partial\mathcal{O}}(\boldsymbol{w} \times \mathbf{N})\varphi\, dS\,.\end{aligned}$$

Identity (4.23) is concluded since φ can be chosen arbitrarily on Σ. □

Proof of the Stokes theorem. Using (4.23) and then applying the divergence theorem on surfaces with boundary (Theorem 4.83), we find that

$$\int_\Sigma \text{curl}\boldsymbol{u} \cdot \mathbf{N}\, dS = \int_\Sigma \text{div}_\Sigma(\boldsymbol{u} \times \mathbf{N})\, dS = \int_{\partial\Sigma} (\boldsymbol{u} \times \mathbf{N}) \cdot (\mathbf{T} \times \mathbf{N})\, ds = \int_{\partial\Sigma} (\boldsymbol{u} \cdot \mathbf{T})\, ds$$

in which the identity $(\boldsymbol{u} \times \mathbf{N}) \cdot (\mathbf{T} \times \mathbf{N}) = \boldsymbol{u} \cdot \mathbf{T}$ is used. □

Example 4.88. Let Σ be the surface given in Example 4.51, and $\boldsymbol{F} : \mathbb{R}^3 \to \mathbb{R}^3$ be a vector-valued function given by $\boldsymbol{F}(x, y, z) = (y, -x, 0)$. Then by the definition of line integral,

$$\begin{aligned}\oint_C \boldsymbol{F} \cdot dr &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\sin^2 t, -\cos t \sin t, 0) \cdot \big(\cos^2 t - \sin^2 t, 2\sin t \cos t, -\sin t\big) dt \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \big(\sin^2 t \cos^2 t - \sin^4 t - 2\sin^2 t \cos^2 t\big) dt \\ &= -\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2 t dt = -\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1 - \cos 2t}{2} dt = -\Big(\frac{t}{2} - \frac{\sin 2t}{4}\Big)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = -\frac{\pi}{2}\,.\end{aligned}$$

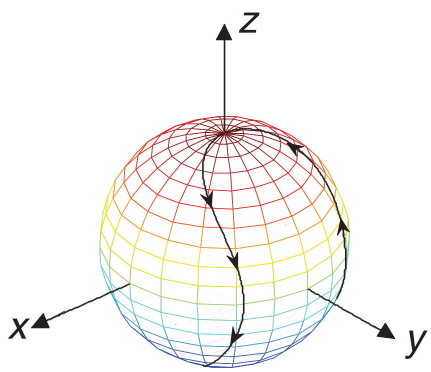
while by the fact that $\text{curl}\boldsymbol{F} = (0, 0, -2)$, the Stokes theorem implies that

$$\begin{aligned}\oint_C \boldsymbol{F} \cdot d\boldsymbol{r} &= \int_\Sigma (0, 0, -2) \cdot \mathbf{N}\, dS = \int_{\psi^{-1}(\Sigma)} -2\cos\phi \sin\phi\, d(\theta, \phi) = -2\int_0^{\frac{\pi}{2}} \int_\phi^{\pi-\phi} \sin\phi \cos\phi\, d\theta d\phi \\ &= -\int_0^{\frac{\pi}{2}} (\pi - 2\phi) \sin 2\phi\, d\phi = \big(\frac{\pi}{2} \cos 2\phi - \phi \cos 2\phi + \frac{1}{2} \sin 2\phi\big)\Big|_{\phi=0}^{\phi=\frac{\pi}{2}} \\ &= -\frac{\pi}{2} - \frac{\pi}{2} + \frac{\pi}{2} = -\frac{\pi}{2}\,.\end{aligned}$$

Example 4.89. Let C be a smooth curve parameterized by

$$\boldsymbol{r}(t) = \big(\cos(\sin t) \sin t, \sin(\sin t) \sin t, \cos t\big)\,, \qquad t \in [0, 2\pi]\,.$$

Then the curve C is a closed curve on $\mathbb{S}^2$, and divide $\mathbb{S}^2$ into two parts. Let Σ denote the part with smaller area.



As in Example 4.51 and Example 4.88, we would like to find the area of Σ, and verify the Stokes theorem for the special case that $\boldsymbol{F} : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$\boldsymbol{F}(x, y, z) = (y, -x, 0)\,.$$

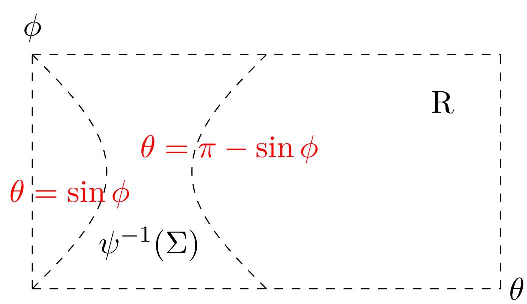
To find the surface area of Σ, we need to parameterize Σ. As in Example 4.51, we look for $\gamma(t) = \big(\theta(t), \phi(t)\big)$, $t \in [0, 2\pi]$, such that $\psi\big(\gamma(t)\big) = \boldsymbol{r}(t)$, where $\psi : \mathrm{R} \equiv (0, 2\pi) \times (0, \pi)$ is given by $\psi(\theta, \phi) = (\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi)\,.$

For $t \in (0, \pi)$, since $\cos t = \cos\phi(t)$ and $\phi(t) \in (0, \pi)$, we must have $\phi(t) = t$; thus the two identities $\cos(\sin t) \sin t = \cos\theta(t) \sin\phi(t)$ and $\sin(\sin t) \sin t = \sin\theta(t) \sin\phi(t)$ further imply that $\theta(t) = \sin t$. Therefore, the curve $\boldsymbol{r}\big((0, \pi)\big)$ corresponds to $\theta = \sin\phi$, $\phi \in (0, \pi)$, on R.

On the other hand, for $t \in (\pi, 2\pi)$, the identity $\cos\phi(t) = \cos t$ implies that $\phi(t) = 2\pi - t$. The two identities $\cos(\sin t) \sin t = \cos\theta(t) \sin\phi(t)$ and $\sin(\sin t) \sin t = \sin\theta(t) \sin\phi(t)$ further imply that

$$\cos(\sin t) = -\cos\theta(t) \quad \text{and} \quad \sin(\sin t) = -\sin\theta(t) \qquad t \in (\pi, 2\pi)\,.$$

Therefore, $\theta(t) = \pi + \sin t$ which implies that the curve $\boldsymbol{r}\big((\pi, 2\pi)\big)$ corresponds to $\theta = \pi - \sin\phi$, $\phi \in (0, \pi)$, on R.



Therefore, the surface area of Σ is

$$\int_0^\pi \int_{\sin\phi}^{\pi-\sin\phi} \sin\phi\, d\theta d\phi = \int_0^\pi (\pi - 2\sin\phi)\sin\phi\, d\phi = -\Big(\pi\cos\phi + \phi - \frac{\sin(2\phi)}{2}\Big)\Big|_{\phi=0}^{\phi=\pi} = \pi\,.$$

Next, we compute the line integral $\displaystyle\oint_C \boldsymbol{F}\cdot d\boldsymbol{r}$. First, we note that

$$\boldsymbol{r}'(t) = (-\sin(\sin t)\sin t\cos t + \cos(\sin t)\cos t, \cos(\sin t)\sin t\cos t + \sin(\sin t)\cos t, -\sin t)\,;$$

thus

$$\begin{aligned}(\boldsymbol{F}\circ\boldsymbol{r})(t)\cdot\boldsymbol{r}'(t) &= -\sin^2(\sin t)\sin^2 t\cos t + \sin(\sin t)\cos(\sin t)\sin t\cos t\\ &\qquad - \cos^2(\sin t)\sin^2 t\cos t - \sin(\sin t)\cos(\sin t)\sin t\cos t\\ &= -\sin^2 t\cos t\,.\end{aligned}$$

As a consequence,

$$\oint_C \boldsymbol{F}\cdot d\boldsymbol{r} = -\int_0^{2\pi} \sin^2 t\cos t\, dt = -\frac{1}{3}\sin^3 t\Big|_{t=0}^{t=2\pi} = 0\,.$$

On the other hand,

$$\begin{aligned}\int_\Sigma \mathrm{curl}\boldsymbol{F}\cdot\mathbf{N}\, dS &= \int_0^\pi\int_{\sin\phi}^{\pi-\sin\phi} (0,0,-2)\cdot(\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi)\sin\phi\, d\theta d\phi\\ &= -2\int_0^\pi \sin\phi\cos\phi(\pi - 2\sin\phi)\, d\phi\\ &= \Big(\frac{\pi}{2}\cos 2\phi + \frac{4}{3}\sin^3\phi\Big)\Big|_{\phi=0}^{\phi=\pi} = 0\,.\end{aligned}$$

4.8 Green's Theorem

In most of materials Green's theorem is introduced prior to the divergence theorem and the Stokes theorem; however, we treat Green's theorem as a corollary of the divergence theorem (Theorem 4.75), the Stokes theorem (Theorem 4.86) and Theorem 4.83.

Theorem 4.90 (Green's theorem)**.** *Let $\mathcal{D}$ be a bounded domain whose boundary $\partial\mathcal{D}$ is piecewise smooth, and $M, N : \mathcal{D}\to\mathbb{R}$ be of class $\mathscr{C}^1$. Then*

$$\oint_{\partial\mathcal{D}} (M,N)\cdot d\boldsymbol{r} = \int_{\mathcal{D}} (N_x - M_y)\, d\mathbb{A}\,,$$

where the line integral (*on the left-hand side of the identity above*) *is taken so that the curve is counter-clockwise oriented.*

Proof 1. Let $\boldsymbol{u}(x,y) = \big(N(x,y), -M(x,y)\big)$ be a vector-valued function defined on the 2-dimensional domain $\mathcal{D}$. Suppose that $\partial\mathcal{D}$ is parameterized by $\boldsymbol{r}(t) = \big(x(t), y(t)\big)$ for $t \in [a,b]$, where $\boldsymbol{r}'$ points in the counter-clockwise direction. Then with $\mathbf{N}$ denoting the outward-pointing unit normal of $\partial\mathcal{D}$, the divergence theorem implies that

$$\oint_{\partial\mathcal{D}} (M,N)\cdot d\boldsymbol{r} = \oint_{\partial\mathcal{D}} \boldsymbol{u}\cdot\mathbf{N}\,ds = \int_{\mathcal{D}} \operatorname{div}\boldsymbol{u}\,d\mathbb{A} = \int_{\mathcal{D}} (N_x - M_y)\,d\mathbb{A}\,. \qquad \square$$

Proof 2. Let $\boldsymbol{F}(x,y,z) = \big(M(x,y), N(x,y), 0\big)$ be a vector-valued function defined in a subset of $\mathbb{R}^3$. Then

$$\operatorname{curl}\boldsymbol{F} = (0,0,N_x - M_y)\,;$$

thus the Stokes theorem implies that

$$\begin{aligned}\oint_{\partial\mathcal{D}} (M,N)\cdot d\boldsymbol{r} &= \int_{\partial\mathcal{D}} \boldsymbol{F}\cdot\mathbf{T}\,ds = \int_{\mathcal{D}} \operatorname{curl}\boldsymbol{F}\cdot\mathbf{N}\,dS = \int_{\mathcal{D}} (0,0,N_x - M_y)\cdot(0,0,1)\,d\mathbb{A}\\ &= \int_{\mathcal{D}} (N_x - M_y)\,d\mathbb{A}\,.\end{aligned} \qquad \square$$

Proof 3. Let $\Sigma = \mathcal{D}\times\{z=0\}$. Then Σ is a surface with boundary and the upward-pointing unit normal $\mathbf{N} = (0,0,1)$. Let $\boldsymbol{F}:\Sigma\to\mathbb{R}^3$ and $\boldsymbol{u}:\mathcal{D}\to\mathbb{R}^2$ be vector-valued functions defined by $\boldsymbol{F}(x,y,z) = \big(N(x,y), -M(x,y), 0)$ and $\boldsymbol{u}(x,y) = \big(N(x,y), -M(x,y)\big)$, respectively. We note that if $\partial\mathcal{D}$ is parameterized by $\boldsymbol{r}(t) = (x(t), y(t), 0)$, then

$$\mathbf{T}\times\mathbf{N} = \frac{1}{\|\boldsymbol{r}'(t)\|_{\mathbb{R}^3}}\big(x'(t), y'(t), 0\big)\times(0,0,1) = \frac{1}{\|\boldsymbol{r}'(t)\|_{\mathbb{R}^3}}\big(y'(t), -x'(t), 0\big)\,;$$

thus by the fact that the surface divergence operator $\operatorname{div}_\Sigma$ is the same as the 2-d divergence operator (since Σ is flat), Theorem 4.83 implies that

$$\oint_{\partial\mathcal{D}} (M,N)\cdot d\boldsymbol{r} = \oint_{\partial\mathcal{D}} \boldsymbol{F}\cdot(\mathbf{T}\times\mathbf{N})\,ds = \int_{\Sigma} \operatorname{div}_\Sigma\boldsymbol{F}\,dS = \int_{\mathcal{D}} \operatorname{div}\boldsymbol{u}\,d\mathbb{A} = \int_{\mathcal{D}} (N_x - M_y)\,d\mathbb{A}\,. \qquad \square$$

Corollary 4.91. *Let* $\mathrm{R}\subseteq\mathbb{R}^2$ *be a domain enclosed by a simple closed curve* C *which is parameterized by* $\boldsymbol{r}(t) = \big(x(t), y(t)\big)$ *for* $t\in[a,b]$. *Suppose* $\boldsymbol{r}'$ *points in the counter-clockwise direction. Then*

$$\textit{the area of } \mathrm{R} = \frac{1}{2}\int_a^b \big(x(t)y'(t) - y(t)x'(t)\big)dt\,.$$

Proof. The corollary is concluded by applying Green's theorem to the special case: $M(x,y) = -y$ and $N(x,y) = x$. $\square$

Example 4.92. Compute the area enclosed by the Cardioid which has a polar representation $r = (1 - \sin\theta)$ with $\theta \in [0, 2\pi]$.

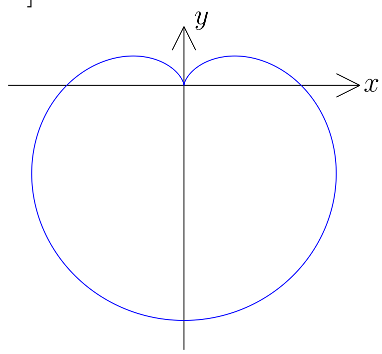


Figure 4.7: The Cardioid

Given the polar representation $r = (1 - \sin\theta)$, a parametrization of the Cardioid is

$$\boldsymbol{r}(t) = \big(x(t), y(t)\big) = \big((1-\sin t)\cos t, (1-\sin t)\sin t\big) \qquad t \in [0, 2\pi]\,.$$

Then Corollary 4.91 implies that the area enclosed by the Cardioid is

$$\begin{aligned}
&\frac{1}{2}\int_0^{2\pi} \big[(1-\sin t)\cos t\big(-\cos t\sin t + (1-\sin t)\cos t\big) \\
&\qquad - (1-\sin t)\sin t\big(-\cos^2 t - (1-\sin t)\sin t\big)\big]\,dt \\
&\quad = \frac{1}{2}\int_0^{2\pi}(1-\sin t)\big[\cos^2 t - 2\sin t\cos^2 t + \sin t\cos^2 t + \sin^2 t - \sin^3 t\big]\,dt \\
&\quad = \frac{1}{2}\int_0^{2\pi}(1-\sin t)(1-\sin t\cos^2 t - \sin^3 t)dt = \frac{1}{2}\int_0^{2\pi}(1-\sin t)^2 dt = \frac{3\pi}{2}\,.
\end{aligned}$$

Before finishing this chapter, we would like to establish an unproven theorem: Theorem 4.33. We recall Theorem 4.33 as follows.

Theorem 4.33. Let $\mathcal{D} \subseteq \mathbb{R}^2$ be simply connected, and $\boldsymbol{F} = (M, N) : \mathcal{D} \to \mathbb{R}^2$ be of class $\mathscr{C}^1$. If $M_y = N_x$, then $\boldsymbol{F}$ is conservative.

Proof of Theorem 4.33. By Theorem 4.30, it suffices to show that $\oint_C \boldsymbol{F}\cdot d\boldsymbol{r} = 0$ for all piecewise smooth closed curve $C \in \mathcal{D}$. Nevertheless, if C is a piecewise closed curve and R is the region enclosed by C, by the fact that $\mathcal{D}$ is simply connected, we must have $\partial\mathrm{R} = C$. Therefore, Green's theorem implies that

$$\oint_C (M, N)\cdot d\boldsymbol{r} = \int_{\mathrm{R}} (N_x - M_y)\, d\mathbb{A} = 0\,. \qquad \square$$

Chapter 5

Additional Topics

5.1 Reynolds' Transport Theorem

Let Ω_1 and Ω_2 be two Lipschitz domains of $\mathbb{R}^{\rm n}$ with outward-pointing unit normal $\mathbf{N}$ and n, respectively, and the map $\psi : \begin{cases} \Omega_1 \;\to\; \Omega_2 \\ \partial\Omega_1 \to \partial\Omega_2 \\ \;\; y \;\;\mapsto x = \psi(y) \end{cases}$ be a diffeomorphism; that is, ψ is one-to-one and onto, and has smooth inverse. Let $f \in \mathscr{C}^1(\Omega_2) \cap \mathscr{C}(\overline{\Omega}_2)$, and $F = f \circ \psi$ which in turns belongs to $\mathscr{C}^1(\Omega_1) \cap \mathscr{C}(\overline{\Omega}_1)$. By the divergence theorem,

$$\int_{\Omega_2} \frac{\partial f}{\partial x_i}(x)\, dx = \int_{\partial\Omega_2} (f \boldsymbol{n}_i)(x)\, dS_x \,.$$

On the other hand, by the chain rule we have that

$$\frac{\partial F}{\partial y_i} = \frac{\partial (f \circ \psi)}{\partial y_i} = \sum_{j=1}^{\rm n} \Big[\frac{\partial f}{\partial x_j} \circ \psi\Big] \frac{\partial \psi^j}{\partial y_i}\,;$$

thus if $\mathrm{A} = (\nabla\psi)^{-1}$,

$$\frac{\partial f}{\partial x_i} \circ \psi = \sum_{j=1}^{\rm n} \mathrm{A}^j_i \frac{\partial F}{\partial y_j}\,. \tag{5.1}$$

Letting $\mathrm{J} = \det(\nabla\psi)$ be the Jacobian of ψ, by the change of variable $y = \psi(y)$ and the Piola identity,

$$\int_{\Omega_2} \frac{\partial f}{\partial x_i}(x)\, dx = \int_{\Omega_1} \frac{\partial f}{\partial x_i}\big(\psi(y)\big) \det(\nabla\psi)(y) dy = \sum_{j=1}^{\rm n} \int_{\Omega_1} \frac{\partial}{\partial y_j}(\mathrm{J}\mathrm{A}^j_i F)\, dy.$$

The divergence theorem again implies that

$$\int_{\Omega_2} \frac{\partial f}{\partial x_i}(x)\, dx = \sum_{j=1}^{\rm n} \int_{\Omega_1} \mathrm{J}\mathrm{A}^j_i F \mathbf{N}_j dS_y$$

which further implies that

$$\int_{\partial\Omega_2} (f\boldsymbol{n})(x)\, dS_x = \int_{\partial\Omega_1} F \frac{\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}}{|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|} |\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}| dS_y\,. \tag{5.2}$$

Let $\psi^*(dS_x)$ denote the pull-back of the surface element dS_x having the property that for any function h defined on $\partial\Omega_2 = \psi(\partial\Omega_1)$,

$$\int_{\psi(\partial\Omega_1)} h(x)\, dS_x = \int_{\partial\Omega_1} (h \circ \psi)(y) \psi^*(dS_x)\,;$$

in other words, $\psi^*(dS_x) = \sqrt{\mathrm{g}(y)}\, dS_y$ for some "Jacobian" $\sqrt{\mathrm{g}}$ of the map $\psi : \partial\Omega_1 \to \partial\Omega_2$. Therefore, (5.2) suggests that

$$\int_{\partial\Omega_2} f\boldsymbol{n} dS = \int_{\partial\Omega_1} \big[(f\boldsymbol{n}) \circ \psi\big](y)\psi^*(dS_x) = \int_{\partial\Omega_1} (f \circ \psi) \frac{\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}}{|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|} |\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}| dS_y\,.$$

Since f can be chosen arbitrarily, the equality above suggests that

$$\boldsymbol{n} \circ \psi = \frac{\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}}{|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|} = \frac{\mathrm{A}^{\mathrm{T}}\mathbf{N}}{|\mathrm{A}^{\mathrm{T}}\mathbf{N}|} \tag{5.3}$$

and

$$\psi^*(dS_x) = |\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}| dS_y\,. \tag{5.4}$$

We finish this section by the following

Theorem 5.1 (Reynolds' transport theorem)**.** *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a smooth domain, $\psi : \Omega \times [0, T] \to \mathbb{R}^{\mathrm{n}}$ be a diffeomorphism, $\Omega(t) = \psi(\Omega, t)$ and $f(x, t)$ be a function defined on $\Omega(t)$. Then*

$$\frac{d}{dt} \int_{\Omega(t)} f(x, t)\, dx = \int_{\Omega(t)} f_t(x, t)\, dx + \int_{\partial\Omega(t)} (\sigma f)(x, t)\, dS_x\,, \tag{5.5}$$

where σ is the speed of the boundary in the direction of outward pointing normal of $\partial\Omega(t)$; that is, with $\boldsymbol{n}$ denoting the outward-pointing unit normal of $\Omega(t)$,

$$\sigma = (\psi_t \circ \psi^{-1}) \cdot \boldsymbol{n}\,.$$

Proof. By the change of variable formula,

$$\int_{\Omega(t)} f(x, t)\, dx = \int_{\Omega} f(\psi(y, t), t) \det(\nabla\psi)(y, t)\, dy.$$

Let $f(\psi(y,t),t) = F(y,t)$, $\mathrm{A} = (\nabla\psi)^{-1}$, and $\mathrm{J} = \det(\nabla\psi)$. By (1.3) and (5.1), we find that

$$\begin{aligned}\frac{d}{dt}\int_{\Omega(t)} f(x,t)\,dx &= \int_\Omega \Big[f_t(\psi(y,t),t) + \psi_t(y,t)\cdot(\nabla_x f)(\psi(y,t),t)\Big]\mathrm{J}(y,t)dy \\ &\qquad + \sum_{i,j=1}^{\mathrm{n}} \int_\Omega F(y,t)(\mathrm{J}\mathrm{A}^j_i\psi^i_{t,j})(y,t)\,dy \\ &= \int_\Omega f_t(\psi(y,t),t)dy + \sum_{i,j=1}^{\mathrm{n}} \int_\Omega \Big[\psi^i_t\mathrm{A}^j_i F_{,j}\,\mathrm{J} + F\mathrm{J}\mathrm{A}^j_i\psi^i_{t,j}\Big](y,t)\,dy \\ &= \int_\Omega (f_t\circ\psi)\mathrm{J}dy + \sum_{i,j=1}^{\mathrm{n}} \int_\Omega \big(\mathrm{J}\mathrm{A}^j_i\psi^i_t F\big)_{,j}\,dy,\end{aligned}$$

where the Piola identity (2.6) is used to conclude the last equality. The divergence theorem then implies that

$$\frac{d}{dt}\int_{\Omega(t)} f(x,t)\,dx = \int_\Omega (f_t\circ\psi)\mathrm{J}\,dy + \sum_{i,j=1}^{\mathrm{n}} \int_{\partial\Omega} \mathrm{J}\mathrm{A}^j_i\mathbf{N}_j\psi^i_t F dS_y\,.$$

As a consequence, changing back to the variable x on the right-hand side, by (5.3) and (5.4) we conclude that

$$\frac{d}{dt}\int_{\Omega(t)} f(x,t)\,dx = \int_{\Omega(t)} f_t(x,t)\,dx + \sum_{i,j=1}^{\mathrm{n}} \int_{\partial\Omega(t)} (\sigma f)(x,t)\,dS_x\,. \qquad \square$$

5.2 Eulerian and Lagrangian Coordinates

We have seen that the diffeomorphism $\psi : \Omega \to \Omega(t)$ plays an important role in the Reynolds transport theorem Theorem 5.1. In fluid dynamics, if the fluid domain is carried by the fluid velocity; that is, the boundary of the fluid domain moves along with the fluid velocity, then there is a natural map with domain Ω and range $\Omega(t)$, and we focus a little bit on this map in this sub-section.

Let $\Omega(t) \subseteq \mathbb{R}^{\mathrm{n}}$ be a (time dependent) domain, and $u(\cdot,t) : \Omega(t) \to \mathbb{R}^{\mathrm{n}}$ be a smooth vector field. We say that $\partial\Omega(t)$ moves along with u if any smooth curve $\big\{x(t) \in \mathbb{R}^{\mathrm{n}} \,\big|\, t \in [0,T]\big\}$ satisfying $x(t) \in \partial\Omega(t)$ also satisfies that

$$x'(t)\cdot n\big(x(t),t\big) = u\big(x(t),t\big)\cdot n\big(x(t),t\big), \tag{5.6}$$

where n again denotes the outward-pointing unit normal of $\Omega(t)$. We remark that using the notation $\Omega(t)$, we include the possibility that the fluid domain may vary in time, while in a lot of applications, $\Omega \equiv \Omega(0) = \Omega(t)$ for all t. Now suppose that $\partial\Omega(t)$ moves along with u. Let $\eta : \Omega \to \mathbb{R}^n$ be the unique solution to the ODE

$$\eta_t(\alpha, t) = u\big(\eta(\alpha, t), t\big) \qquad \forall\, \alpha \in \Omega\,, t \in (0, T)\,, \tag{5.7a}$$

$$\eta(\alpha, 0) = \alpha \qquad \forall\, \alpha \in \Omega, \tag{5.7b}$$

here we assume that the solution exists up to time T. The value of $x = \eta(\alpha, t)$ is the location of the fluid particle at time t which is initially positioned at $\alpha \in \Omega$. By (5.6), we must have $\eta(\Omega, t) = \Omega(t)$.

A time independent coordinate system used in the co-domain of η is called the *Eulerian coordinate*. We note that since in general $\Omega(t)$ varies continuously in time, the Eulerian coordinate is usually defined on a subset of $\mathbb{R}^n$ larger than $\Omega(t)$. In fact, the Cartesian coordinate is one of the most important Eulerian coordinate system since $\Omega(t) \subseteq \mathbb{R}^n$ for all $t > 0$. On the contrary, the coordinate used in the domain of η is called the *Lagrangian coordinate*. Since the Lagrangian coordinate is used to identify the initial position of fluid particles, it is often called the *material coordinate* as well. In short, the Eulerian coordinate is used to describe the (larger) background space (so each x corresponds to a point in space which might not a point in the fluid), while the Lagrangian coordinate is used to describe the particle in the fluid (so each α corresponds to a particle in the fluid).

Let us explain what these two coordinate systems are doing. Suppose that a kind of censor (whose volume and mass are both zero in the mathematical setting so that it does not affect any physics) is designed to measure certain physical quantity. The censor can be fixed at a point x in space so that the readings indicate the value of that physical quantity at x for various time. On the other hand, we may set the censor to flow with the fluid (the fluid will carry the censor). If the censor initially is position at a given point α, then the readings of the censor indicate the value of the quantity at the particle which initially locates at position α. In other words, a function with variables in Lagrangian coordinate is a function defined on material particles inside the fluids, while a function with variables in Eulerian coordinate is a function defined on space.

Theorem 5.2. *Let $u : \Omega(t) \times (0, T) \to \mathbb{R}^n$ be a smooth vector field, and the flow map $\eta(\cdot, t) :$ $\left\{\begin{array}{rcl} \Omega & \to & \Omega(t) \\ \alpha & \mapsto & \eta(\alpha, t) \end{array}\right.$ be defined by* (5.7)*. Then u is divergence-free if and only if* $\det(\nabla\eta) \equiv 1$ *for all $t > 0$.*

Proof. Let $\mathrm{J} = \det(\nabla\eta)$, and $\mathrm{A} = (\nabla\eta)^{-1}$. By (1.2) and (1.3),

$$\mathrm{J}_t = \sum_{i,j=1}^{\mathrm{n}} \mathrm{JA}_i^j \frac{\partial \eta_t^i}{\partial \alpha_j} = \sum_{i,j=1}^{\mathrm{n}} \mathrm{JA}_i^j \frac{\partial (u^i \circ \eta)}{\partial \alpha_j} = \sum_{i,j,k=1}^{\mathrm{n}} \mathrm{JA}_i^j \Big(\frac{\partial u^i}{\partial x_k} \circ \eta\Big) \frac{\partial \eta^k}{\partial \alpha_j}.$$

Since $\mathrm{A} = (\nabla\eta)^{-1}$, $\sum\limits_{j=1}^{\mathrm{n}} \mathrm{A}_i^j \eta_{,j}^k = \delta_{ik}$; thus

$$\mathrm{J}_t = \mathrm{J}(\mathrm{div} u) \circ \eta. \tag{5.8}$$

The theorem is then concluded by the fact that $\mathrm{J}|_{t=0} = 1$ since η is the identity map at $t = 0$. $\square$

Corollary 5.3. *Let $u(\cdot, t) : \Omega(t) \to \mathbb{R}^{\mathrm{n}}$ be a smooth divergence-free vector field, and η be the corresponding flow map (which is assumed to exist up to time T as well). If $\mathcal{U} \subseteq \Omega \equiv \Omega(0)$ is a smooth domain and*

$$\mathcal{U}(t) = \Big\{ x \in \mathbb{R}^{\mathrm{n}} \,\Big|\, x = \eta(\alpha, t) \text{ for some } \alpha \in \mathcal{U} \Big\};$$

that is, $\mathcal{U}(t)$ is the image of $\mathcal{U}$ under the map η at time t, then

$$\text{the volume of } \mathcal{U} = \text{the volume of } \mathcal{U}(t) \qquad \forall\, t \in (0, T).$$

Proof. Let $|\mathcal{O}|$ denote the Lebesgue measure of set $\mathcal{O}$. Then

$$|\mathcal{U}(t)| = \int_{\mathcal{U}(t)} dx = \int_{\mathcal{U}} \det(\nabla\eta)(\alpha) d\alpha = \int_{\mathcal{U}} d\alpha = |\mathcal{U}|. \qquad \square$$

Remark 5.4. If the fluid velocity is divergence-free, then the corollary above says that the volume of a region carried by the fluid is constant in time. For this reason we sometimes also called solenoidal vector fields *incompressible*.

5.2.1 The material derivative

In continuum mechanics, the material derivative describes the time rate of change of some physical quantity (like heat or momentum) for a material element subjected to a space-and-time-dependent velocity field. To be more precise, the material derivative, sometimes called the *substantial derivative*, denoted by $\frac{D}{Dt}$, is defined by

$$\frac{DF}{Dt} = \frac{\partial F}{\partial t} + \sum_{i=1}^{\mathrm{n}} u^i \frac{\partial F}{\partial x_i} = F_t + (u \cdot \nabla) F, \tag{5.9}$$

where F is a physical quantity in Eulerian variable, and u is the fluid velocity field. Let η be the flow map associated to u, and define $f = F \circ \eta$; that is, $f(\alpha, t) = F\big(\eta(\alpha, t), t\big)$, then

$$\frac{\partial}{\partial t} f(\alpha, t) = \big[F_t + (u \cdot \nabla)F\big] \circ \eta = \frac{DF}{Dt} \circ \eta\,.$$

Therefore, the composition of the material derivative of a function and the flow map is the time rate of change of the composition of that function and the flow map.'

5.3 The particle trajectory and streamlines

Not yet completed!!!

5.4 Exercises

In this set of exercise, the Einstein summation convention is used.

Problem 1. Complete the following.

1. Let $\delta_{\cdot\cdot}$'s denote the Kronecker deltas. Prove (4.9); that is, show that
$$\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}\,. \tag{4.9}$$

2. Let $\mathrm{O} \subseteq \mathbb{R}^3$ be an open domain, and $\boldsymbol{u} : \mathrm{O} \to \mathbb{R}^3$ be a smooth vector field. Denote twice the anti-symmetric part of $\nabla \boldsymbol{u}$ as Ω; that is, $\Omega_{ij} = \boldsymbol{u}^i{}_{,j} - \boldsymbol{u}^j{}_{,i}$. Show that
$$\Omega_{kj} = \varepsilon_{ijk}\boldsymbol{\omega}^i\,, \tag{5.10}$$
where $\boldsymbol{\omega} = \mathrm{curl}\boldsymbol{u}$ is the vorticity of $\boldsymbol{u}$.

3. Use (4.9) to show the following identities:
   (a) $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u} \cdot \boldsymbol{v})\boldsymbol{w}$ if $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are three 3-vectors.
   (b) $\mathrm{curl}\mathrm{curl}\boldsymbol{u} = -\Delta\boldsymbol{u} + \nabla\mathrm{div}\boldsymbol{u}$ if $\boldsymbol{u} : \mathrm{O} \to \mathbb{R}^3$ is smooth.
   (c) $\boldsymbol{u} \times \mathrm{curl}\boldsymbol{u} = \dfrac{1}{2}\nabla(|\boldsymbol{u}|^2) - (\boldsymbol{u} \cdot \nabla)\boldsymbol{u}$ if $\boldsymbol{u} : \mathrm{O} \to \mathbb{R}^3$ is smooth.

4. Use (5.10) to show that
$$\mathrm{curl}\big[(\boldsymbol{u} \cdot \nabla)\boldsymbol{u}\big] = (\boldsymbol{u} \cdot \nabla)\boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{u} + (\mathrm{div}\boldsymbol{u})\boldsymbol{\omega}$$
if $\boldsymbol{u} : \mathrm{O} \to \mathbb{R}^3$ is smooth.

Problem 2. Let $\psi(\cdot,t):\Omega\to\Omega(t)$ be a diffeomorphism as defined in Theorem 5.1, and $\mathrm{J}=\det(\nabla\psi)$ and $\mathrm{A}=(\nabla\psi)^{-1}$. Complete the proof of the Piola identity, identities (2.7), (5.3) and (5.4) by the following argument:

1. Let $\boldsymbol{u}(\cdot,t):\Omega(t)\to\mathbb{R}^{\mathrm{n}}$ be a smooth vector field. Show that

$$\int_{\Omega(t)}\operatorname{div}\boldsymbol{u}\,dx=\int_{\Omega}\mathrm{JA}_i^j(\boldsymbol{u}\circ\psi)^i_{,j}\,dy\,;$$

thus by the divergence theorem,

$$\int_{\partial\Omega(t)}\boldsymbol{u}\cdot\boldsymbol{n}\,dS_x=\int_{\partial\Omega}\mathrm{JA}_i^j(\boldsymbol{u}\circ\psi)^i\mathbf{N}_j\,dS_y-\int_{\Omega}(\mathrm{JA}_i^j)_{,j}\,(\boldsymbol{u}\circ\psi)^i\,dy\,. \tag{5.11}$$

2. Using (5.11),

$$\int_{\Omega}(\mathrm{JA}_i^j)_{,j}\,(\boldsymbol{u}\circ\psi)^i\,dy=0\qquad\forall\,\boldsymbol{u}(\cdot,t):\Omega(t)\to\mathbb{R}^{\mathrm{n}}\text{ vanishing on }\partial\Omega(t).$$

As a consequence, the Piola identity is valid.

3. By the Piola identity, (5.11) implies that

$$\int_{\partial\Omega(t)}\boldsymbol{u}\cdot\boldsymbol{n}\,dS_x=\int_{\partial\Omega}\mathrm{JA}_i^j(\boldsymbol{u}\circ\psi)^i\mathbf{N}_j\,dS_y\qquad\forall\,\boldsymbol{u}(\cdot,t):\Omega(t)\to\mathbb{R}^{\mathrm{n}}\text{ smooth.}$$

Therefore, identities (5.3) and (5.4) are also valid.

4. Using identity (2.7) (which is obtained independent of the Piola identity) to show that

$$\mathrm{J}_{,k}=\mathrm{JA}_i^j\psi^i_{,jk}\,.$$

Appendix A

Appendix

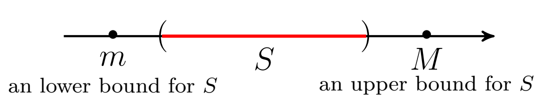
A.1 Properties of Real Numbers

Definition A.1. Let $\varnothing \neq S \subseteq \mathbb{R}$. A number $M \in \mathbb{R}$ is called an ***upper bound*** (上界) for S if $x \leqslant M$ for all $x \in S$, and a number $m \in \mathbb{R}$ is called a ***lower bound*** (下界) for S if $x \geqslant m$ for all $x \in S$. If there is an upper bound for S, then S is said to be ***bounded from above***, while if there is a lower bound for S, then S is said to be ***bounded from below***. A number $b \in \mathbb{R}$ is called a ***least upper bound*** (最小上界) if

1. b is an upper bound for S, and
2. if M is an upper bound for S, then $M \geqslant b$.

A number a is called a ***greatest lower bound*** (最大下界) if

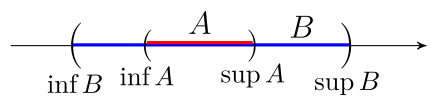
1. a is a lower bound for S, and
2. if m is a lower bound for S, then $m \leqslant a$.



If S is not bounded above, the least upper bound of S is set to be ∞, while if S is not bounded below, the greatest lower bound of S is set to be $-\infty$. The least upper bound of S is also called the ***supremum*** of S and is usually denoted by lub$S$ or $\sup S$, and "the" greatest lower bound of S is also called the ***infimum*** of S, and is usually denoted by glb$S$ or $\inf S$. If $S = \varnothing$, then $\sup S = -\infty$, $\inf S = \infty$.

Remark A.2. The least upper bound and the greatest lower bound of S need not be a member of S.

Remark A.3. The reason for defining $\sup\varnothing = -\infty$ and $\inf\varnothing = \infty$ is as follows: if $\varnothing \neq A \subseteq B$, then $\sup A \leqslant \sup B$ and $\inf A \geqslant \inf B$.



Since $\varnothing$ is a subset of any other sets, we shall have $\sup\varnothing$ is smaller then any real number, and $\inf\varnothing$ is greater than any real number. However, this "definition" would destroy the property that $\inf A \leqslant \sup A$.

The "definition" of $\sup\varnothing$ and $\inf\varnothing$ is purely artificial. One can also define $\sup\varnothing = \infty$ and $\inf\varnothing = -\infty$.

Definition A.4. An ***open interval*** in $\mathbb{R}$ is of the form (a, b) which consists of all $x \in \mathbb{R} \ni a < x < b$. A ***closed interval*** in $\mathbb{R}$ is of the form $[a, b]$ which consists of all $x \in \mathbb{R} \ni a \leqslant x \leqslant b$.

Proposition A.5. *Let $S \subseteq \mathbb{R}$ be non-empty. Then*

1. $b = \sup S \in \mathbb{R}$ *if and only if*
   (a) *b is an upper bound of S.*
   (b) $\forall\, \varepsilon > 0$, $\exists\, x \in S \ni x > b - \varepsilon$.
2. $a = \inf S \in \mathbb{R}$ *if and only if*
   (a) *a is a lower bound of S.*
   (b) $\forall\, \varepsilon > 0$, $\exists\, x \in S \ni x < a + \varepsilon$.

Proof. "$\Rightarrow$" (a) is part of the definition of being a least upper bound.

(b) If M is an upper bound of S, then we must have $M \geqslant b$; thus $b - \varepsilon$ is not an upper bound of S. Therefore, $\exists\, x \in S \ni x > b - \varepsilon$.

"$\Leftarrow$" We only need to show that if M is an upper bound of S, then $M \geqslant b$. Assume the contrary. Then $\exists\, M$ such that M is an upper bound of S but $M < b$. Let $\varepsilon = b - M$, then there is no $x \in S \ni x > b - \varepsilon$. $\rightarrow\leftarrow$ $\square$

The Completeness Axiom（實數完備性公設）

Every subset of $\mathbb{R}$ which is bounded from above has a least upper bound.

Definition A.6 (Cauchy sequence)**.** A sequence $\{x_k\}_{k=1}^{\infty}$ in $\mathbb{R}$ is said to be ***Cauchy*** if for every $\varepsilon > 0$, there exists $N > 0$ such that $|x_k - x_\ell| < \varepsilon$ whenever $k, \ell \geqslant N$.

Theorem A.7. *Every Cauchy sequence in $\mathbb{R}$ converges.*

A.2 Properties of Continuous Functions

Theorem A.8 (Uniform Continuity)**.**

Theorem A.9 (Mean Value Theorem)**.**

Theorem A.10 (Inverse Function Theorem)**.** *Let $f : (a, b) \to \mathbb{R}$ be differentiable, and f' is sign-definite; that is, $f'(x) > 0$ for all $x \in (a, b)$ or $f'(x) < 0$ for all $x \in (a, b)$. Then $f : (a, b) \to f((a, b))$ is a bijection, and f^{-1}, the inverse function of f, is differentiable on $f((a, b))$, and*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)} \qquad \forall\, x \in (a, b)\,. \tag{A.1}$$

Proof. W.L.O.G. we assume that $f'(x) > 0$ for all $x \in (a, b)$. Then f is strictly increasing; thus f^{-1} exists.
Claim: $f^{-1} : f((a, b)) \to (a, b)$ is continuous.
Proof of claim: Let $y_0 = f(x_0) \in f((a, b))$, and $\varepsilon > 0$ be given. Then $f((x_0 - \varepsilon, x_0 + \varepsilon)) = \big(f(x_0 - \varepsilon), f(x_0 + \varepsilon)\big)$ since f is continuous on (a, b) and $(x_0 - \varepsilon, x_0 + \varepsilon)$ is connected. Let $\delta = \min\{f(x_0) - f(x_0 - \varepsilon), f(x_0 + \varepsilon) - f(x_0)\big\}$. Then $\delta > 0$, and

$$(y_0 - \delta, y_0 + \delta) = \big(f(x_0) - \delta, f(x_0) + \delta\big) \subseteq f((x_0 - \varepsilon, x_0 + \varepsilon))\,;$$

thus by the injectivity of f,

$$f^{-1}((y_0 - \delta, y_0 + \delta)) \subseteq f^{-1}(f((x_0 - \varepsilon, x_0 + \varepsilon))) = (x_0 - \varepsilon, x_0 + \varepsilon) = (f^{-1}(y_0) - \varepsilon, f^{-1}(y_0) + \varepsilon)\,.$$

The inclusion above implies that f^{-1} is continuous at y_0.

Writing $y = f(x)$ and $x = f^{-1}(y)$. Then if $y_0 = f(x_0) \in f((a, b))$,

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)}\,.$$

Since f^{-1} is continuous on $f((a,b))$, $x \to x_0$ as $y \to y_0$; thus

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{x \to x_0} \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{f'(x_0)}$$

which implies that f^{-1} is differentiable at y_0. $\square$