An introduction to the Euclidean space
Lecture notes, Spring 2025
(Version: February 18, 2025)

Pu-Zhao Kow

DEPARTMENT OF MATHEMATICAL SCIENCES, NATIONAL CHENGCHI UNIVERSITY
Email address: pzkow@g.nccu.edu.tw

Preface

This lecture note is prepared for the course *Geometry* during Spring Semester 2025 (113-2), which explains the points, lines, surfaces, as well as other objects in Euclidean spaces, based on some selected materials in [**Apo74, BN10, ConNA, dC76**]. This lecture note is prepared for beginners, and we will focus on concepts ("big picture") rather than go through all details of the rigorous proofs. One also can see e.g. the monographs [**Cha06, Lee13**] for more advance topics. The lecture note may updated during the course. If we have time, probably we also can discuss the d-dimensional spherical harmonics [**EF14**] as well as the vector spherical harmonics [**BEG85**].

Title. Geometry (Spring Semester 2025, 3 credits)

Lectures. Tuesday (13:10 ~ 16:00), February 17, 2025 ~ June 20, 2025

Language. Chinese and English. Materials will be prepared in English.

Instructor. Pu-Zhao Kow (Email: `pzkow@g.nccu.edu.tw`)

Office hour. Tuesday (16:10 ~ 17:00)

Completion. The course can be taken for credit by attending the lectures, returning written solutions (60%) and taking taking midterm and final exams (each 20%).

Contents

CHAPTER 1

Points in Euclidean space

The main theme of this chapter is to discuss the points and sets in Euclidean space.

1.1. Definition of Euclidean space

We first introduce the Euclidean spaces suggested by the title of this lecture note, based on [**Apo74**, Chapter 3], [**BN10**, Section 1.5] as well as [**BV04**, Chapter 2].

DEFINITION 1.1.1. Let $n \geq 1$ be an integer. An ordered set of n real numbers $(x_1, x_2, \cdots, x_n)$ is called an *n-dimensional point* or a *vector with n components*. Throughout this lecture note, *points* or *vectors* will usually denoted by bold letters, for example:

$$\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \quad \text{or} \quad \boldsymbol{y} = (y_1, y_2, \cdots, y_n).$$

The number x_k is called the k^{th} coordinate of $\boldsymbol{x}$. The set of all n-dimensional points is called an n-dimensional Euclidean space, usually denoted as $\mathbb{R}^n$. In other words,

$$\mathbb{R}^n := \{\boldsymbol{x} = (x_1, \cdots, x_n) : x_k \in \mathbb{R} \text{ for } k = 1, \cdots, n\}.$$

We introduce some standard operations on n-dimensional points:

DEFINITION 1.1.2. For each $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ and $\boldsymbol{y} = (y_1, \cdots, y_n) \in \mathbb{R}^n$, we define:

(a) **Equality.** $\boldsymbol{x} = \boldsymbol{y}$ if and only if $x_k = y_k$ for all $k = 1, \cdots, n$.
(b) **Linear combinations.** For each $a, b \in \mathbb{R}$, we define $a\boldsymbol{x} + b\boldsymbol{y} := (ax_1 + by_1, \cdots, ax_n + by_n)$, with *zero vector* (also called *origin*) $\mathbf{0} = (0, \cdots, 0)$.
(c) **Inner (dot) product.** $\boldsymbol{x} \cdot \boldsymbol{y} := \sum_{k=1}^n x_k y_k \equiv x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$.
(d) **Norm (length).** $|\boldsymbol{x}| := (\boldsymbol{x} \cdot \boldsymbol{x})^{1/2} = (\sum_{k=1}^n x_k^2)^{1/2}$. The number $|\boldsymbol{x} - \boldsymbol{y}|$ is called the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$.

REMARK 1.1.3. When $n = 1$, one sees that $|x| = \sqrt{x^2}$. In other words, the 1-dimensional Euclidean norm is identical to the absolute value. Therefore we will use the notation $|\cdot|$ rather than $\|\cdot\|$ for Euclidean norm. We usually reserve the notation $\|\cdot\|$ for function spaces, which we will not going discuss in this lecture note.

It is easy to see that:

(i) **Positive definite.** $|\boldsymbol{x}| \geq 0$ and the equality holds if and only if $\boldsymbol{x} = \mathbf{0}$.
(ii) **Absolute homogeneous.** $|a\boldsymbol{x}| = |a||\boldsymbol{x}|$ for all $a \in \mathbb{R}$.

(iii) **Symmetry.** $|\boldsymbol{x}-\boldsymbol{y}| = |\boldsymbol{y}-\boldsymbol{x}|$.

LEMMA 1.1.4. *For each $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$, one has the* Cauchy-Schwartz inequality $|\boldsymbol{x}\cdot\boldsymbol{y}| \leq |\boldsymbol{x}||\boldsymbol{y}|$ *and the* triangle inequality $|\boldsymbol{x}\pm\boldsymbol{y}| \leq |\boldsymbol{x}|+|\boldsymbol{y}|$.

PROOF. If either $\boldsymbol{x}=\boldsymbol{0}$ or $\boldsymbol{y}=\boldsymbol{0}$, then we we have nothing to proof. We now consider the case when both $\boldsymbol{x}\neq\boldsymbol{0}$ and $\boldsymbol{y}\neq\boldsymbol{0}$. In this case, we have $|\boldsymbol{x}|\neq 0$ and $|\boldsymbol{y}|\neq 0$ and we can define

$$\hat{\boldsymbol{x}} := \frac{\boldsymbol{x}}{|\boldsymbol{x}|} \quad \text{and} \quad \hat{\boldsymbol{y}} := \frac{\boldsymbol{y}}{|\boldsymbol{y}|}.$$

Since

$$\begin{aligned} 0 \leq |\hat{\boldsymbol{x}}\pm\hat{\boldsymbol{y}}|^2 &= \sum_{k=1}^n (\hat{x}_k \pm \hat{y}_k)^2 = \sum_{k=1}^n \hat{x}_k^2 + \sum_{k=1}^n \hat{y}_k^2 \pm 2\sum_{k=1}^n \hat{x}_k\hat{y}_k \\ &= |\hat{\boldsymbol{x}}|^2 + |\hat{\boldsymbol{y}}|^2 \pm 2\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{y}} = 2 \pm 2\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{y}}. \end{aligned}$$

Now we have $\mp\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{y}} \leq 1$ and this is equivalent to $|\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{y}}| \leq 1$, which immediately implies the Cauchy-Schwartz inequality $|\boldsymbol{x}\cdot\boldsymbol{y}| \leq |\boldsymbol{x}||\boldsymbol{y}|$.

With the Cauchy-Schwartz inequality at hand, then the rest is easy: Using similar computations, one sees that

$$|\boldsymbol{x}\pm\boldsymbol{y}|^2 = |\boldsymbol{x}|^2 + |\boldsymbol{y}|^2 \pm 2\boldsymbol{x}\cdot\boldsymbol{y}.$$

Now we use the Cauchy-Schwartz inequality we see that

$$|\boldsymbol{x}\pm\boldsymbol{y}|^2 \leq |\boldsymbol{x}|^2 + |\boldsymbol{y}|^2 + 2|\boldsymbol{x}||\boldsymbol{y}| = (|\boldsymbol{x}|+|\boldsymbol{y}|)^2,$$

which concludes the triangle inequality $|\boldsymbol{x}\pm\boldsymbol{y}| \leq |\boldsymbol{x}|+|\boldsymbol{y}|$. □

COROLLARY 1.1.5. *For each $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$, one has $|\boldsymbol{x}-\boldsymbol{y}| \geq ||\boldsymbol{x}|-|\boldsymbol{y}||$.*

PROOF. By observing that $\boldsymbol{x} = (\boldsymbol{x}-\boldsymbol{y})+\boldsymbol{y}$, the triangle inequality implies

$$|\boldsymbol{x}| \leq |\boldsymbol{x}-\boldsymbol{y}| + |\boldsymbol{y}|,$$

which gives $|\boldsymbol{x}|-|\boldsymbol{y}| \leq |\boldsymbol{x}-\boldsymbol{y}|$. Interchanging the role between $\boldsymbol{x}$ and $\boldsymbol{y}$ yields $|\boldsymbol{y}|-|\boldsymbol{x}| \leq |\boldsymbol{y}-\boldsymbol{x}| = |\boldsymbol{x}-\boldsymbol{y}|$. Putting these two inequalities together, we see that

$$\pm(|\boldsymbol{x}|-|\boldsymbol{y}|) \leq |\boldsymbol{x}-\boldsymbol{y}|,$$

which is equivalent to our desired inequality. □

It is also convenient to introduce the notion of matrix.

DEFINITION 1.1.6. An $n\times m$ (real) matrix is a rectangular array with n rows and m columns. Elements of the array are called the *entries* of the matrix. We usually denote the matrix by indexed letters, for example,

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix}$$

with entries $A_{jk} \in \mathbb{R}$. The set of all $n \times m$ matrices is usually denoted as $\mathbb{R}^{n\times m}$.

REMARK 1.1.7. Sometimes we simply abuse the notation by writing $A = (A_{jk})$.

We also can define some operations similar to Definition 1.1.2:

DEFINITION 1.1.8. For each $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{n\times m}$, we define:

(a) **Equality.** $A = B$ if and only if $A_{jk} = B_{jk}$ for all $j = 1,\cdots,n$ and $k = 1,\cdots,m$.
(b) **Linear combinations.** For each $a,b \in \mathbb{R}$, we define $(aA+bB)_{jk} = aA_{jk} + bB_{jk}$ for all $j = 1,\cdots,n$ and $k = 1,\cdots,m$.
(c) **Inner product.** $A : B := \sum_{j=1}^{n} \sum_{k=1}^{m} A_{jk}B_{jk}$.
(d) **Frobenius norm.** $|A| := (A : A)^{1/2} = (\sum_{j=1}^{n} \sum_{k=1}^{m} A_{jk}^2)^{1/2}$.

We usually *do not identify* the set $\mathbb{R}^{n\times m}$ given in Definition 1.1.6 with $\mathbb{R}^{nm}$ due to the following operations:

DEFINITION 1.1.9. For each $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times p}$:

(a) the **transpose** $A^\intercal \in \mathbb{R}^{m\times n}$ is defined by $(A^\intercal)_{kj} = A_{jk}$ for all $j = 1,\cdots,n$ and $k = 1,\cdots,m$.
(b) the **matrix multiplication** $AB \in \mathbb{R}^{n\times p}$ is defined by $(AB)_{j\ell} = \sum_{k=1}^{m} A_{jk}B_{k\ell}$ for all $j = 1,\cdots,n$ and $\ell = 1,\cdots,p$.

In view of the matrix multiplication operation above, we define the identity matrix $\mathrm{Id} \in \mathbb{R}^{m\times m}$ by

$$\mathrm{Id}_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

which satisfies $A\mathrm{Id} = A$ and $\mathrm{Id}B = B$ for all matrices A and B with suitable dimensions.

However, we usually identify the elements Euclidean space $\mathbb{R}^n$ given in Definition 1.1.1 with the elements in $\mathbb{R}^{n\times 1}$ (called the *column vectors*) given in Definition 1.1.6 since we can expressed the inner product in terms of transpose and matrix multiplication:

$$\boldsymbol{x}\cdot\boldsymbol{y} = \boldsymbol{x}^\intercal\boldsymbol{y} = \boldsymbol{y}^\intercal\boldsymbol{x} \quad \text{for all } \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n \cong \mathbb{R}^{n\times 1}.$$

We also denote $\boldsymbol{e}_j$ be the j^{th} column of the identity matrix I. In terms of ordered set, it can be expressed as

$$\boldsymbol{e}_1 = (1,0,0,\cdots,0)$$
$$\boldsymbol{e}_2 = (0,1,0,\cdots,0)$$

and so on. It is easy to see that $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ is *orthonormal* (with respect to the inner product in Definition 1.1.2):

$$\boldsymbol{e}_j\cdot\boldsymbol{e}_k=\begin{cases}1 & \text{whenever } j=k,\\ 0 & \text{whenever } j\neq k.\end{cases}$$

It is also important to see that $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ forms an basis of $\mathbb{R}^n$: given any $\boldsymbol{x}\in\mathbb{R}^n$, there exists a unique n-tuple $(x_1,\cdots,x_n)$ with $x_k\in\mathbb{R}$ such that

$$\boldsymbol{x}=\sum_{k=1}^n x_k\boldsymbol{e}_k\equiv x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_k.$$

REMARK 1.1.10. When $n=2$, some authors denote

$$\boldsymbol{i}=\boldsymbol{e}_1=(1,0),\quad \boldsymbol{j}=\boldsymbol{e}_2=(0,1).$$

REMARK 1.1.11. When $n=3$, some authors denote

$$\boldsymbol{i}=\boldsymbol{e}_1=(1,0,0),\quad \boldsymbol{j}=\boldsymbol{e}_2=(0,1,0),\quad \boldsymbol{k}=\boldsymbol{e}_e=(0,0,1).$$

1.2. A geometry description on dot product

Before giving a geometry interpretation for inner product mentioned in Definition 1.1.2, lets us recall the trigonometry functions.

Let B be the unit ball in $\mathbb{R}^2$ with radius 1 with boundary ∂B, that is,

$$B=\left\{\boldsymbol{x}\in\mathbb{R}^2:|\boldsymbol{x}|<1\right\},\quad \partial B=\left\{\boldsymbol{x}\in\mathbb{R}^2:|\boldsymbol{x}|=1\right\}.$$

Let π be the area of B, which is approximated by $\pi\approx 3.14159\cdots$. The length of ∂B (also known as *circumference* or the *perimeter* of the unit circle B) is 2π. Let L_1 and L_2 be two straight lines both connecting the origin and the "infinity", and let

$$P_1:=L_1\cap\partial B,\quad P_2:=L_2\cap\partial B,$$

and we see that the circle ∂B is partitioned into two parts, says Γ_1 and Γ_2, by the points P_1 or P_2. Intuitively, it is natural to define the angle between the lines L_1 and L_2 by the length of Γ_1 or Γ_2, but however this may cause some trouble in mathematics, since this is not a function, since both choices Γ_1 and Γ_2 correspond to the same geometry.

In order to make the definition rigorous, we define angle with direction (or orientation). Starting from the point P_1, which corresponds to line L_1, we rotate it towards the *counter-clockwise* direction and stop at P_2 (not necessarily stop at the first meeting), which corresponds to the line L_2. Now we define:

(1.2.1) the *angle* θ (in radian) from L_1 to L_2 by
the displacement of the locus mentioned above.

The angle defined by (1.2.1) is oriented, one sees that the angle from L_2 to L_1 is $-\theta$. In some occasion, we sometimes refer the non-negative number $|\theta|$ the undirected angle *between* the lines L_1 and L_2.

REMARK 1.2.1. Since π is transcendental (roughly speaking, it cannot be expressed in terms of countable (see the following remark) many binary operators $+,-,\times,\div$ as well as roots $\sqrt[m]{\cdot}$ with $m = 2,3,\cdots$), it is not so convenient in some applications (e.g. aviation). We usually normalize the angle as follows: the angle $\tilde{\theta}$ (in degree) is defined by $\tilde{\theta} := \frac{360}{2\pi}\theta$, where $\theta \in [0,2\pi)$ is the angle in radian. The reason we choose 360 is it is dividable by many integers, including $2,3,4,5,6,8,9,10,\cdots$.

REMARK 1.2.2. It is a bit technical to explain the term "countable" is a rigorous way. Roughly speaking, a set S which consists of infinitely many elements is said to be *countable* if we can arrange its elements and labeled them using natural number:

$$S = \{s_1, s_2, \cdots\}.$$

The set of natural number $\mathbb{N}$, the set of integers $\mathbb{Z}$ and the set of rational numbers $\mathbb{Q}$ are all countable. Without causing any ambiguity, those sets only consist finitely many elements are also said to be countable. However, the set of real numbers $\mathbb{R}$ is not countable, that is, it is not possible to write $\mathbb{R}$ in the form of $\{r_1, r_2, \cdots\}$. The open interval $(a,b) := \{x \in \mathbb{R} : a < x < b\}$ is also not countable. It is not surprisingly that the following facts hold true:

- If A is countable and $B \subset A$, then B is also countable.
- If B is uncountable and $B \subset A$, then A is also uncountable.

With the oriented angle at hand, we now can define the trigonometric functions, as in Figure 1.2.1 below:

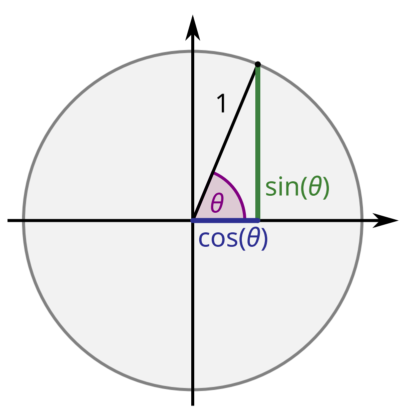


FIGURE 1.2.1. Definition of sine function and cosine function: Stephan Kulla (User:Stephan Kulla), CC0, via Wikimedia Commons

Since the angle is oriented, thus the trigonometric also has sign, for example, $\sin\theta < 0$ when $\frac{\pi}{2} < \theta < \frac{3}{2}\pi$. According to the above definition, we also see that

$$\sin : \mathbb{R} \to [-1,1] \quad \text{and} \quad \cos : \mathbb{R} \to [-1,1]$$

are both surjective functions, but not injective. The definition of sine and cosine function immediately gives

$$(\cos\theta)^2 + (\sin\theta)^2 = 1 \quad \text{for all } \theta \in \mathbb{R}.$$

Some special values are showed in Figure 1.2.2 below:

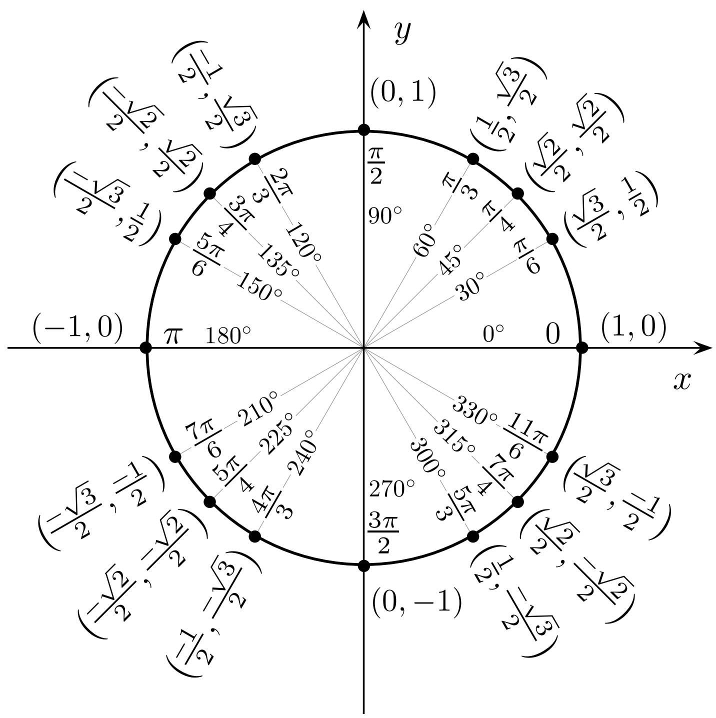


FIGURE 1.2.2. Some special values of $(\cos\theta, \sin\theta)$: Gustavb (talk · contribs), Public domain, via Wikimedia Commons

It is remarkable to mention that

$$\sin(-\theta) = -\sin\theta \quad \text{for all } \theta \in \mathbb{R},$$
$$\cos(-\theta) = \cos\theta \quad \text{for all } \theta \in \mathbb{R},$$

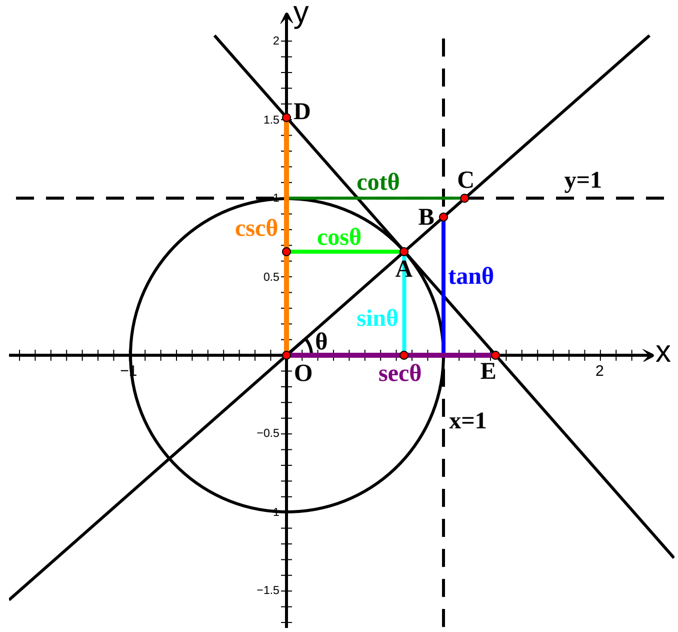


FIGURE 1.2.3. Six trigonometric functions: Onmaditque, CC BY-SA 4.0, via Wikimedia Commons

that is, the sine function is odd while the cosine function is even. The other trigonometric functions are defined as follows (see Figure 1.2.3):

$$\begin{aligned}
\tan\theta &:= \frac{\sin\theta}{\cos\theta} \quad \text{for all } \theta \in \mathbb{R}\setminus\left(\pi\mathbb{Z}+\frac{\pi}{2}\right),\\
\sec\theta &:= \frac{1}{\cos\theta} \quad \text{for all } \theta \in \mathbb{R}\setminus\left(\pi\mathbb{Z}+\frac{\pi}{2}\right),\\
\cot\theta &:= \frac{\cos\theta}{\sin\theta} \quad \text{for all } \theta \in \mathbb{R}\setminus\pi\mathbb{Z},\\
\mathrm{cosec}\theta \equiv \csc\theta &:= \frac{1}{\sin\theta} \quad \text{for all } \theta \in \mathbb{R}\setminus\pi\mathbb{Z}.
\end{aligned}$$

One see that

$$\cot\theta = \frac{1}{\tan\theta} \quad \text{for all } \theta \in \mathbb{R}\setminus\left(\pi\mathbb{Z}\cup\left(\pi\mathbb{Z}+\frac{\pi}{2}\right)\right) = \mathbb{R}\setminus\frac{\pi}{2}\mathbb{Z},$$

which means that the identity only holds true in restricted domain.

In order to define inverse functions, we usually (unless stated) consider

$$\begin{aligned}
\sin &: [-\pi/2, \pi/2] \to [-1,1] \\
\cos &: [0,\pi] \to [-1,1], \\
\tan &: (-\pi/2, \pi/2) \to \mathbb{R}, \\
\cot &: (0,\pi) \to \mathbb{R}, \\
\sec &: [0,\pi] \setminus \{\pi/2\} \to \mathbb{R}_{\leq -1} \cup \mathbb{R}_{\geq 1}, \\
\operatorname{cosec} \equiv \csc &: [-\pi/2, \pi/2] \setminus \{0\} \to \mathbb{R}_{\leq -1} \cup \mathbb{R}_{\geq 1},
\end{aligned}$$

which are bijective, and hence the corresponding inverse functions, called the *inverse trigonometric functions*, are defined as :

$$\begin{aligned}
\arcsin \equiv \sin^{-1} &: [-1,1] \to [-\pi/2, \pi/2], \\
\arccos \equiv \cos^{-1} &: [-1,1] \to [0,\pi], \\
\arctan \equiv \tan^{-1} &: \mathbb{R} \to (-\pi/2, \pi/2), \\
\cot^{-1} &: \mathbb{R} \to (0,\pi), \\
\sec^{-1} &: \mathbb{R}_{\leq -1} \cup \mathbb{R}_{\geq 1} \to [0,\pi] \setminus \{\pi/2\}, \\
\csc^{-1} &: \mathbb{R}_{\leq -1} \cup \mathbb{R}_{\geq 1} \to [-\pi/2, \pi/2] \setminus \{0\}.
\end{aligned}$$

Now we are ready to give a geometry interpretation for the inner product given in Definition 1.1.2:

THEOREM 1.2.3. *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ are nonzero vectors in $\mathbb{R}^n$, then the not oriented angle $\theta \in [0,\pi]$ between $\boldsymbol{x}$ and $\boldsymbol{y}$ is given by*

$$\theta := \cos^{-1}\left(\frac{\boldsymbol{x}}{|\boldsymbol{x}|} \cdot \frac{\boldsymbol{y}}{|\boldsymbol{y}|} \right).$$

SKETCH OF PROOF. By considering the rotation and translation, it is suffice to consider $n = 2$ and $\boldsymbol{x} = \boldsymbol{e}_1$. The proof of the theorem is sketched in Figure 1.2.3 above. ☐

EXERCISE 1.2.4. Sketch the function $\sin^{-1} \circ \sin : \mathbb{R} \to \mathbb{R}$ and $\cos^{-1} \circ \cos : \mathbb{R} \to \mathbb{R}$.

REMARK 1.2.5 (Euler formula and trigonometric functions). Here we also explain a simple way to derive trigonometric identities. We *formally* write the *imaginary number* $\mathrm{i} := \sqrt{-1}$, one can see e.g. my other lecture note [**Kow23**], which is much more advance, for a precise definition. The Euler formula reads:

$$e^{\mathrm{i}\theta} := \cos\theta + \mathrm{i}\sin\theta \quad \text{for all } \theta \in \mathbb{R}.$$

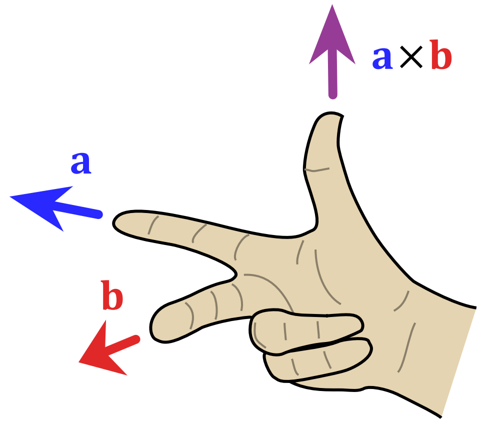


FIGURE 1.2.4. Right-hand rule for cross product: Acdx, CC BY-SA 3.0, via Wikimedia Commons

Performing some *formal* computations lead (more precisely, the *de Moivre theorem*)

$$\begin{aligned}
&\cos(\theta_1+\theta_2)+\mathrm{i}\sin(\theta_1+\theta_2)=e^{\mathrm{i}(\theta_1+\theta_2)}\\
&\quad=e^{\mathrm{i}\theta_1}e^{\mathrm{i}\theta_2}=(\cos\theta_1+\mathrm{i}\sin\theta_1)(\cos\theta_2+\mathrm{i}\sin\theta_2)\\
&\quad=\cos\theta_1\cos\theta_2+\mathrm{i}^2\sin\theta_1\sin\theta_2+\mathrm{i}(\cos\theta_1\sin\theta_2+\sin\theta_1\cos\theta_2)\\
&\quad=(\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2)+\mathrm{i}(\cos\theta_1\sin\theta_2+\sin\theta_1\cos\theta_2).
\end{aligned}$$

Comparing the real and imaginary parts lead us to the *sum-to-product formula*:

$$\begin{aligned}
\cos(\theta_1+\theta_2)&=\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2,\\
\sin(\theta_1+\theta_2)&=\cos\theta_1\sin\theta_2+\sin\theta_1\cos\theta_2.
\end{aligned}$$

Choosing $\theta_1=\theta_2=\theta$ leads to the *multiple angle formula*. It is easy to obtain further generalization by consider $e^{\mathrm{i}(\theta_1+\theta_2+\theta_3)}=e^{\mathrm{i}\theta_1}e^{\mathrm{i}\theta_2}e^{\mathrm{i}\theta_3}$ and so on. From this, it is easy to derive the *product-to-sum formula*, for example,

$$\begin{aligned}
&\cos(\theta_1+\theta_2)+\cos(\theta_1-\theta_2)\\
&\quad=\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2+\cos\theta_1\cos(-\theta_2)-\sin\theta_1\sin(-\theta_2)\\
&\quad=\cos\theta_1\cos\theta_2-\cancel{\sin\theta_1\sin\theta_2}+\cos\theta_1\cos\theta_2+\cancel{\sin\theta_1\sin\theta_2}\\
&\quad=2\cos\theta_1\cos\theta_2.
\end{aligned}$$

The other three product-to-sum formula can be easily obtained by considering

$$\cos(\theta_1+\theta_2)-\cos(\theta_1-\theta_2),\quad\sin(\theta_1+\theta_2)+\sin(\theta_1-\theta_2),\quad\sin(\theta_1+\theta_2)-\sin(\theta_1-\theta_2),$$

here we left the details for readers as an exercise.

REMARK 1.2.6. When $n = 3$, the *cross product* $\boldsymbol{x} \times \boldsymbol{y}$ is defined by the formula

$$\boldsymbol{x} \times \boldsymbol{y} := (|\boldsymbol{x}||\boldsymbol{y}| \sin\theta)\boldsymbol{n}, \tag{1.2.2}$$

where $\theta \in [0,\pi]$ is the undirected angle between $\boldsymbol{x}$ and $\boldsymbol{y}$ given by Theorem 1.2.3 and $\boldsymbol{n}$is a unit vector perpendicular to the plane containing $\boldsymbol{x}$ and $\boldsymbol{y}$, with direction as indicated in Figure 1.2.4 above. In fact, the cross product can be expressed as

$$\boldsymbol{x} \times \boldsymbol{y} = (x_2y_3 - x_3y_2)\boldsymbol{e}_1 + (x_3y_1 - x_1y_3)\boldsymbol{e}_2 + (x_1y_2 - x_2y_1)\boldsymbol{e}_3,$$

or sometimes we *abuse the notation* by writing

$$\boldsymbol{x} \times \boldsymbol{y} = \det \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

in terms of *formal* determinant. It is important to point out that $|\boldsymbol{x} \times \boldsymbol{y}|$ is the area of the parallelogram generated by the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$.

1.3. The Euclidean topology

1.3.1. Open sets. Before we further proceed, we first classify the sets in the Euclidean space. In view of the norm given in Definition 1.1.2, it is natural to consider the following object:

DEFINITION 1.3.1. The (open) *ball* in $\mathbb{R}^n$ (also known as *n-dimensional ball* or simply *n-ball*) with radius $r > 0$ and centered at $\boldsymbol{x} \in \mathbb{R}^n$ is defined by

$$B_R(\boldsymbol{x}) := \{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| < R\}.$$

In order to simplify the notations, we also denote $B_R = B_R(\boldsymbol{0})$.

By using the above definition, we can classify the points in any set in $\mathbb{R}^n$:

DEFINITION 1.3.2. Given any set $S \subset \mathbb{R}^n$, we define the *interior* of S by

$$\mathrm{int}\,(S) := \{\boldsymbol{x} \in S : \text{there exists } \varepsilon = \varepsilon(\boldsymbol{x}) > 0 \text{ such that } B_\varepsilon(\boldsymbol{x}) \subset S\}.$$

The elements in $\mathrm{int}\,(S)$ are called the *interior points*.

DEFINITION 1.3.3. A set $\Omega \subset \mathbb{R}^n$ is said to be *open* if and only if $\Omega = \mathrm{int}\,(\Omega)$.

REMARK 1.3.4. The empty set $\emptyset$ is also an open set. An open set Ω contained z sometimes called the *neighborhood* of z.

DEFINITION 1.3.5. The *Euclidean topology on* $\mathbb{R}^n$ is the collection $\mathscr{T}$ consisting of all open sets given in Definition 1.3.3, in other words,

$$\mathscr{T} = \{\Omega \subset \mathbb{R}^n : \Omega = \mathrm{int}\,(\Omega)\}.$$

Here we consider the Zermelo–Fraenkel set theory. Roughly speaking, we view the subsets in $\mathbb{R}^n$ as "level-1" objects. The points in subsets in $\mathbb{R}^n$ are viewed as "level-0" objects. The set for which consists of other sets as a "level-2" objects are usually referred the *collection* – rather than the term "set" – in order to distinguish them (for example the collection $\mathscr{T}$ mentioned in Definition 1.3.5). We also remark that the set consists of collections (i.e. "level-2" objects) as a *superset*, which is natural to be labeled as "level-3" object. We distinguish the notations "$\in$" and "$\subset$" as follows:

- We write $\boldsymbol{x} \in S$ for "level-0" object $\boldsymbol{x}$ (point) and for "level-1" object S (set); we write $\Omega \in \mathscr{T}$ for "level-1" object Ω (set) and for "level-2" object $\mathscr{T}$ (collection); and so on.
- We write $X \subset Y$ for two "level-1" objects X and Y (sets); we write $\mathscr{T} \subset \mathscr{U}$ for two "level-2" objects $\mathscr{T}$ and $\mathscr{U}$ (collections); and so on.

If we identify the notations "$\in$" and "$\subset$" without introducing the concept of "levels", the Russell's paradox exhibits the "set of all sets that do not contain themselves":

$$\{X : X \text{ is a set and } X \notin X\},$$

which is not allow in the modern/standard mathematical logic.

EXERCISE 1.3.6. Let $\mathscr{T}$ is the Euclidean topology on $\mathbb{R}^n$. Show that:

(a) $\emptyset, \mathbb{R}^n \in \mathscr{T}$.
(b) Any arbitrary union of members of $\mathscr{T}$ belongs to $\mathscr{T}$.
(c) Any finite intersection of members of $\mathscr{T}$ belongs to $\mathscr{T}$.

In plain words, Exercise 1.3.6(b) means that: If $\{\Omega_\alpha\}_{\alpha\in\Lambda}$ is any collection of open sets then $\bigcup_{\alpha\in\Lambda}\Omega_\alpha$ is also an open set. Here the index set Λ can be arbitrary abstract set, even may uncountable. In plain words, Exercise 1.3.6(c) means that: If m is a given positive integer and $\Omega_1, \cdots, \Omega_m$ are open set, then $\bigcap_{k=1}^m \Omega_m$ is also open. In fact, one can define other topology on $\mathbb{R}^n$, but – for simplicity – here we only consider Euclidean topology throughout this lecture note.

EXERCISE 1.3.7. We define the closed ball $\overline{B_1} := \{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| \leq 1\}$. Show that $\overline{B_1}$ is not an open set in $\mathbb{R}^n$.

EXAMPLE 1.3.8. Now we consider $\Omega_k = B_{1+\frac{1}{k}}$ for $k \in \mathbb{N}$. Note that $\{\Omega_k\}_{k\in\mathbb{N}}$ is a collection of open sets, but however we see that

$$\bigcap_{k\in\mathbb{N}} \Omega_k = \overline{B_1},$$

which is not open by Exercise 1.3.7. This example suggests that the result in Exercise 1.3.6(c) is already optimal.

EXERCISE 1.3.9. Let $-\infty < a_k < b_k < +\infty$ for $k = 1, \cdots, n$, and we define the *open interval* (or *open rectangle*)

$$\prod_{k=1}^{n}(a_k, b_k) := \{\boldsymbol{x} = (x_1, \cdots, x_n) : a_k < x_k < b_k \text{ for all } k = 1, \cdots, n\}. \tag{1.3.1}$$

Show that $\prod_{k=1}^{n}(a_k, b_k)$ is an open set in $\mathbb{R}^n$.

Let Ω be a nonempty open set in $\mathbb{R}^n$. For each $\boldsymbol{x} \in \mathbb{R}^n$, there exists $\varepsilon = \varepsilon(\boldsymbol{x}) > 0$ such that $B_{\varepsilon(\boldsymbol{x})} \subset \Omega$, thus we see that

$$\Omega = \bigcup_{\boldsymbol{x} \in \Omega} B_{\varepsilon(\boldsymbol{x})},$$

which shows that each open set can be expressed as a union of open balls. In other words, we can regard the collection of open balls as a *topological basis*[1] of the Euclidean topology.

EXERCISE 1.3.10. Let Ω be a nonempty set in $\mathbb{R}^n$. Show that the following are equivalent:

(a) Ω is open;
(b) For each $\boldsymbol{x} \in \Omega$, there exists an open interval $I(\boldsymbol{x}) \ni \boldsymbol{x}$ such that $I(\boldsymbol{x}) \subset \Omega$.
(c) For each $\boldsymbol{x} \in \Omega$, there exists an open interval $I(\boldsymbol{x}) \ni \boldsymbol{x}$ such that $\Omega = \bigcup_{\boldsymbol{x} \in \Omega} I(\boldsymbol{x})$.

The above exercise means that the collection of open interval is also a topological basis of the same Euclidean topology.

1.3.2. Topological closed sets and closure of sets.

DEFINITION 1.3.11. A set F in $\mathbb{R}^n$ is said to be *topological closed* if its complement $S^{\complement} := \mathbb{R}^n \setminus S \equiv \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \notin S\}$ is open.

REMARK 1.3.12. In this lecture note, we will keeping emphasizing the term "topological closed" rather than simply say "closed" to prevent ambiguity with the below mentioned "closed curved".

EXERCISE 1.3.13. Let $-\infty < a_k < b_k < +\infty$ for $k = 1, \cdots, n$, and we define the *closed interval* (or *closed rectangle*)

$$\prod_{k=1}^{n}[a_k, b_k] := \{\boldsymbol{x} = (x_1, \cdots, x_n) : a_k \leq x_k \leq b_k \text{ for all } k = 1, \cdots, n\}. \tag{1.3.2}$$

Show that $\prod_{k=1}^{n}[a_k, b_k]$ is a topological closed set in $\mathbb{R}^n$, but it is not an open set in $\mathbb{R}^n$.

From Exercise 1.3.6, we immediately see the following corollary.

COROLLARY 1.3.14.

(a) $\emptyset, \mathbb{R}^n$ *are topological closed sets in* $\mathbb{R}^n$.

[1] Not to be confused with the vector basis basis.

(b) *Any arbitrary intersection of topological closed sets is also topological closed.*
(c) *Any finite union of topological closed sets is also topological closed.*

DEFINITION 1.3.15. A point $\boldsymbol{x} \in \mathbb{R}^n$ is called an *adherent point* of a set $S \subset \mathbb{R}^n$ if

$$S \cap B_\varepsilon(\boldsymbol{x}) \neq \emptyset \quad \text{for all } \varepsilon > 0.$$

A point $\boldsymbol{x} \in \mathbb{R}^n$ is called a *limit point* (or *accumulation point*) of a set $S \subset \mathbb{R}^n$ if

$$S \cap (B_\varepsilon(\boldsymbol{x}) \setminus \{\boldsymbol{x}\}) \neq \emptyset \quad \text{for all } \varepsilon > 0.$$

EXERCISE 1.3.16. Let $\boldsymbol{x} \in \mathbb{R}^n$ be a limit point of $S \subset \mathbb{R}^n$. Show that for each $\varepsilon > 0$ the set $S \cap B_\varepsilon(\boldsymbol{x})$ contains infinitely many points of S.

The point $\boldsymbol{x} \in \mathbb{R}^n$ mentioned in Definition 1.3.15 is not necessarily contained is S, therefore it is meaningful to introduce the following definition:

DEFINITION 1.3.17. The *closure* of S is defined by

$$\overline{S} := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \text{ is an adherent point of } S\}.$$

The boundary ∂S of a set S is defined by

$$\partial S := \left\{\boldsymbol{x} \in \mathbb{R}^n : B_\varepsilon(\boldsymbol{x}) \cap S \neq \emptyset \text{ and } B_\varepsilon(\boldsymbol{x}) \cap S^\complement \neq \emptyset \text{ for all } \varepsilon > 0\right\}.$$

The set of limit points of S is defined by (also known as *derived set*)

$$S' := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \text{ is a limit point of } S\}.$$

EXERCISE 1.3.18. Show that $\partial S = \overline{S} \setminus \mathrm{int}(S)$ for all set $S \subset \mathbb{R}^n$.

It is easy to see that $S' \subset \overline{S}$ and $\overline{S} = S \cup S'$, therefore it is interesting to classify the points in terms of the following definition:

DEFINITION 1.3.19. If $\boldsymbol{x} \in S$ but it is not a limit point of S, then $\boldsymbol{x}$ is called an *isolated point* of S.

The following exercise demonstrates how Definition 1.3.15 generalizes the notion of "limit of sequence/functions".

EXERCISE 1.3.20. Determine all the limit points of the set $\left\{(x, f(x)) : x = \frac{1}{n} \text{ for some } n \in \mathbb{N}\right\}$ if:

$$\text{(a) } f(x) = x \sin\frac{1}{x}, \quad \text{(b) } f(x) = \sin\frac{1}{x}, \quad \text{(c) } f(x) = \frac{1}{x}\sin\frac{1}{x}.$$

We now show the following important characterization for closed set in $\mathbb{R}^n$.

THEOREM 1.3.21. *A set $S \subset \mathbb{R}^n$ is topological closed if and only if $S = \overline{S}$.*

PROOF. We first assume that $S \subset \mathbb{R}^n$ is closed. By the definition of $\overline{S}$, it holds that $S \subset \overline{S}$, therefore it is suffice to show that $\overline{S} \setminus S = \emptyset$. Suppose the contrary that there exists $\boldsymbol{x} \in \overline{S} \setminus S$. Since $S^{\complement}$ is open, then there exists $\varepsilon = \varepsilon(\boldsymbol{x}) > 0$ such that $B_\varepsilon(\boldsymbol{x}) \subset S^{\complement}$, in other words,

$$B_\varepsilon(\boldsymbol{x}) \cap S = \emptyset,$$

which contradicts with the definition of $\boldsymbol{x} \in \overline{S}$. Therefore we conclude $\overline{S} \setminus S = \emptyset$ by using a contradiction argument.

Conversely, we now assume that $S = \overline{S}$. Let $\boldsymbol{x} \in S^{\complement} = \overline{S}^{\complement}$ (i.e. $\boldsymbol{x}$ is not an adherent point of S), which means that there exists $\varepsilon > 0$ such that

$$S \cap B_\varepsilon(\boldsymbol{x}) = \emptyset, \quad \text{in other words,} \quad B_\varepsilon(\boldsymbol{x}) \subset S^{\complement}.$$

By arbitrariness of $\boldsymbol{x} \in S^{\complement}$, we conclude that $S^{\complement}$ is open, that is, S is closed. □

COROLLARY 1.3.22. *For each $S \subset \mathbb{R}^n$, its closure $\overline{S}$ is topological closed.*

COROLLARY 1.3.23. *A set $S \subset \mathbb{R}^n$ is topological closed if and only if $S \supset S'$.*

1.3.3. Precompact sets.

DEFINITION 1.3.24. A set $S \subset \mathbb{R}^n$ is said to be bounded if $S \subset B_R$ for some $R > 0$.

In fact, all bounded sets are precompact, more precisely:

THEOREM 1.3.25 (Bolzano-Weierstrass, see e.g. [**Apo74**, Theorem 3.24]). *If a bounded set $S \subset \mathbb{R}^n$ contains infinitely many points, then there exists a limit point of S.*

Since the proof of the Bolzano-Weierstrass theorem is quite technical, here we skip the details there. Here we remind the readers that *Bolzano-Weierstrass theorem holds true since $\mathbb{R}^n$ is a finite dimensional vector space* and it is not possible to extend the result for infinite dimensional space[2].

1.3.4. Compact sets.

DEFINITION 1.3.26. Let Λ be arbitrary abstract set. A collection of open sets $\{\Omega_\lambda\}_{\lambda \in \Lambda}$ is called an *open cover* of a set $S \subset \mathbb{R}^n$ if

$$S \subset \bigcup_{\lambda \in \Lambda} \Omega_\lambda.$$

An open cover $\{\Omega_\lambda\}_{\lambda \in \Lambda}$ of S can be viewed as an approximation of S. One may try to find $\Lambda_0 \subset \Lambda$ such that $\{\Omega_\lambda\}_{\lambda \in \Lambda_0}$ is still an open cover of S, especially when Λ is infinitely uncountable. This leads the following definition.

[2]The set $S := \{(0,\cdots,0,\overbrace{1}^{k^{\text{th}}\text{-tuple}},0,0,\cdots)\}_{k \in \mathbb{N}}$ is bounded in the space ℓ^2 ("infinite dimension Euclidean space"), but all points in S are isolated, thus S does not have a limit point in ℓ^2.

DEFINITION 1.3.27. A set $K \subset \mathbb{R}^n$ is said to be *compact* if and only if the following properties hold true: If $\{\Omega_\lambda\}_{\lambda\in\Lambda}$ is an open cover of K, then there exists a finite set $\Lambda_0 \subset \Lambda$ such that $\{\Omega_\lambda\}_{\lambda\in\Lambda_0}$ is still an open cover of K.

We begin with the following trivial example.

EXAMPLE 1.3.28. Let $S = \{s_1, \cdots, s_m\} \subset \mathbb{R}^n$ be a finite set. Let $\{\Omega_\lambda\}_{\lambda\in\Lambda}$ be any open cover of S. For each $k = 1, \cdots, m$, there exists $\lambda_k \in \Lambda$ such that

$$s_k \in \Omega_{\lambda_k},$$

therefore we see that $\{\Omega_{\lambda_k}\}_{k=1}^m$ is still an open cover of S. Hence we conclude that such set S is compact.

EXAMPLE 1.3.29. The open interval $(0,1) \subset \mathbb{R}^1$ is not compact: By considering the open cover $\{(1/n,1)\}_{n\in\mathbb{N}}$, we see that $\{(1/n,1)\}_{n\in\Lambda_0}$ must not be an open cover of $(0,1)$ for any finite subset $\Lambda_0 \subset \mathbb{N}$.

However, it is not easy to verify whether a given set is compact or not using Definition 1.3.27 directly. By using the advantage that $\mathbb{R}^n$ is a finite dimensional vector space, one has the following property:

THEOREM 1.3.30 (Heine-Borel, see e.g. [**Apo74**, Theorem 3.31]). *Let S be a subset of $\mathbb{R}^n$. The following are equivalent:*

(a) *S is compact.*
(b) *S is closed and bounded.*
(c) *Every infinite subset of S has a limit point and such limit point is an element in S.*

The proof of the Heine-Borel theorem is quite technical, and therefore we do not exhibit the details here. Again we remind the readers that the result cannot be extended for infinite dimensional space, see e.g. the Arzela-Ascoli theorem [**Rud87**, Theorem 11.28] for continuous functions. In general, the closed unit ball $\{x \in X : \|x\| \leq 1\}$ in infinite dimensional space X equipped with norm $\|\cdot\|$ need not to be compact (despite it is closed and bounded), see e.g. the Kakutani theorem for reflexive spaces (see e.g. [**Bre11**, Theorem 3.17]). See also the Banach-Alaoglu theorem (see e.g. [**Bre11**, Theorem 3.16]) regarding the compactness of the closed unit ball $\{x \in X : \|x\| \leq 1\}$.

EXERCISE 1.3.31. Let $\{\mathscr{K}^{(k)}\}_{k\in\mathbb{N}}$ be a sequence of nonempty compact sets in $\mathbb{R}^n$ such that $\mathscr{K}^{(1)} \supset \mathscr{K}^{(2)} \supset \mathscr{K}^{(3)} \supset \cdots$. Show that $\bigcap_{k\in\mathbb{N}} \mathscr{K}^{(k)} \neq \emptyset$. [Hint: consider the complement of $\mathscr{K}^{(k)}$. Here we remind that $\mathbb{R}^n$ is not compact.]

EXERCISE 1.3.32. Let $\{\mathscr{K}^{(t)}\}_{t\in(0,1)}$ be a collection of nonempty compact sets in $\mathbb{R}^n$ such that $\mathscr{K}^{(t_1)} \subset \mathscr{K}^{(t_2)}$ for all $0 < t_1 < t_2 < 1$. Show that $\bigcap_{t\in(0,1)} \mathscr{K}^{(t)} \neq \emptyset$.

1.3.5. Connected sets.

DEFINITION 1.3.33. Let S be any set in $\mathbb{R}^n$. A subset $S_0 \subset S$ is said to be *relative open* in S if there exists an open set $\Omega \subset \mathbb{R}^n$ such that $S_0 = S \cap \Omega$. Similarly, a subset $S_1 \subset S$ is said to be relative topological closed in S if there exists a topological closed set $F \subset \mathbb{R}^n$ such that $S_1 = S \cap F$. A set S is said to be *connected* if the following holds:

$$\begin{aligned} &\text{if } S_0 \subset S \text{ is both relative open and relative topological closed in } S \\ &\text{then either } S_0 = \emptyset \text{ or } S_0 = S. \end{aligned} \tag{1.3.3}$$

REMARK 1.3.34 (Relative open sets in open sets). If S is an open set (resp. topological closed set) in $\mathbb{R}^n$ and $S_0 \subset S$, then S_0 is open (resp. topological closed) in $\mathbb{R}^n$ if and only if S_0 is relative open (resp. relative topological closed) in S. This can be easily see by the trivial set inclusion $S_0 = S \cap S_0$.

It is make sense to say that a set S is said to be disconnected if (1.3.3) does not hold. This means that there exists $\emptyset \neq S_0 \subsetneq S$

$$\begin{aligned} &\text{there exists } \emptyset \neq S_0 \subsetneq S \text{ such that} \\ &S_0 \text{ is both relative open and relative topological closed in } S. \end{aligned}$$

In this case, if we define $S_1 := S \setminus S_0$, it is easy to see that $\emptyset \neq S_1 \subsetneq S_0$ is also both relative open and relative topological closed in S. Therefore one see that S_0 and S_1 are both disjoint (open) components of S.

DEFINITION 1.3.35. We denote $[\boldsymbol{x},\boldsymbol{y}]$ the line segment with endpoints $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^n$, that is,

$$[\boldsymbol{x},\boldsymbol{y}] := \{\alpha\boldsymbol{x} + \beta\boldsymbol{y} : \alpha \geq 0, \beta \geq 0, \alpha + \beta = 1\}.$$

A polygonal line is a finite union of line segments of the form $[\boldsymbol{x}_0,\boldsymbol{x}_1] \cup [\boldsymbol{x}_1,\boldsymbol{x}_2] \cup \cdots \cup [\boldsymbol{x}_{n-1},\boldsymbol{x}_n]$.

REMARK 1.3.36. When $n = 1$, we see that Definition 1.3.35 is simply a closed interval in the real line $\mathbb{R}$.

LEMMA 1.3.37. *Let Ω be an open set in $\mathbb{R}^n$. Then Ω is connected if and only if for any $a, b \in \Omega$ there exists a polygonal line in Ω connecting a and b. We also called an open connected set a* region *(when $n = 2$) or* domain.

PROOF OF THE IMPLICATION "$\Rightarrow$" IN LEMMA 1.3.37. Let $\boldsymbol{a} \in \Omega$ and let

$$A := \{\boldsymbol{x} \in \Omega : \text{there exists a polygonal line connecting } \boldsymbol{a} \text{ and } \boldsymbol{x}\}.$$

It is clear that $\boldsymbol{a} \in A$, which shows that $A \neq \emptyset$. Given any $\boldsymbol{x} \in A \subset \Omega$, since Ω is open, then there exists $\varepsilon = \varepsilon(\boldsymbol{x}) > 0$ such that $B_\varepsilon(\boldsymbol{x}) \subset \Omega$. Clearly any point in $B_\varepsilon(\boldsymbol{x})$ can be connected to $\boldsymbol{x}$ by using a straight line, then any point in $B_\varepsilon(\boldsymbol{x})$ can be connected to $\boldsymbol{a}$ by a polygonal line. In other words, $B_\varepsilon(\boldsymbol{x}) \subset A$. By arbitrariness of $\boldsymbol{x} \in A$, we conclude that A is open in $\mathbb{R}^n$, and hence also relative open in Ω. Similar argument shows that $\Omega \setminus A$ is also relative open in Ω. This shows that A is relative topological closed in Ω. Since $A \neq \emptyset$, then $A = \Omega$. $\square$

Despite the implication "$\Leftarrow$" sounds natural, but its proof is surprisingly technical (and require some advance knowledge on connected sets). We will skip this part here. One can refer, for example, my lecture note for complex analysis [Kow23] for a detailed proof.

REMARK 1.3.38. The above exhibits a standard argument when dealing with open connected set:

(1) First show that the target set A (i.e. the set of the property which we wish to show) is nonempty.
(2) Show that A is relative open.
(3) Show that $\Omega \setminus A$ is relative open.

To show an open set is connected, one of course can try to construct a continuous path

In my opinion, even though Lemma 1.3.37 sounds natural, but in practical it is not convenient to manipulate. The abstract definition in Definition 1.3.33 is much more convenient to use. The following exercise exhibits a counter-intuitive example of connected set.

EXERCISE 1.3.39. Let $S = S_1 \cup S_2$ where

$$S_1 = \left\{(x,y) \in \mathbb{R}^2 : x = 0\right\}, \quad S_2 = \left\{(x,y) \in \mathbb{R}^2 : x > 0, y = \sin\frac{1}{x}\right\}.$$

Show that S is topologically connected (despite $S_1 \cap S_2 = \emptyset$). [Hint: Note that both S_1 and S_2 are topologically connected. In order to show that S is topological connected, we need to show that S_1 (and S_2) cannot be both relative open and relative closed in S. Note that S_1 is closed in $\mathbb{R}^2$, and thus it is relative closed in S. Therefore, one only need to show that S_1 is not relative open in S.]

1.3.6. Convex sets. We begin with the following definition.

DEFINITION 1.3.40. A set $C \subset \mathbb{R}^n$ is *convex* if the line segment (Definition 1.3.35) between any two points in C lies in C, that is,

$$[\boldsymbol{x}_1, \boldsymbol{x}_2] \subset C \quad \text{for all } \boldsymbol{x}_1, \boldsymbol{x}_2 \in C.$$

EXERCISE 1.3.41. Let Λ be an abstract index set (not necessarily countable) and let $\{S_\alpha\}_{\alpha\in\Lambda}$ be a collection of convex sets in $\mathbb{R}^n$. Show that the intersection $\bigcap_{\alpha\in\Lambda} S_\alpha$ is convex.

EXERCISE 1.3.42. Let C be a convex set in $\mathbb{R}^n$, show that its closure $\overline{C}$ and its interior $\operatorname{int}(C)$ are also convex.

EXERCISE 1.3.43. Let C_1 and C_2 be convex sets in $\mathbb{R}^n$. Show that $C_1 + C_2 := \{\boldsymbol{x} + \boldsymbol{y} : \boldsymbol{x}_1 \in C_1, \boldsymbol{x}_2 \in C_2\}$ is convex.

EXERCISE 1.3.44. Let $C_1 \subset \mathbb{R}^n$ and $C_2 \subset \mathbb{R}^m$ be convex. Show that $C_1 \times C_2 := \{(\boldsymbol{x},\boldsymbol{y}) : \boldsymbol{x}_1 \in C_1, \boldsymbol{x}_2 \in C_2\}$ is a convex set in $\mathbb{R}^{n+m}$.

EXERCISE 1.3.45. Let C be a convex set in $\mathbb{R}^n$ and let $A \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Show that the set $AC + b := \{A\boldsymbol{x} + \boldsymbol{b} : \boldsymbol{x} \in C\}$ is a convex set in $\mathbb{R}^m$.

DEFINITION 1.3.46. A point of the form $\theta_1\boldsymbol{x}_1 + \cdots + \theta_k\boldsymbol{x}_k \in \mathbb{R}^n$ with $\theta_1 + \cdots + \theta_k = 1$ and $\theta_i \geq 0$ for all $i = 1,\cdots,k$ is called a convex combination of the points $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k \in \mathbb{R}^n$. The convex hull of a set C, denoted as $\operatorname{conv}(C)$, is the set of all convex combinations of points in C.

A convex combination of points can be interpreted as a *mixture* or *weighted average* of the points (with weights θ_i). It is easy to see that

$$C \text{ is convex if and only if } C = \operatorname{conv}(C).$$

EXERCISE 1.3.47. Let S be any set in $\mathbb{R}^n$. Show that its convex hull $\operatorname{conv}(S)$ is identical to the intersection of all convex sets containing S.

DEFINITION 1.3.48. A *hyperplane* is a set of the form $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}\cdot\boldsymbol{x} = b\}$ for some *normal vector* $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and the *offset parameter* $b \in \mathbb{R}$.

A hyperplane divides $\mathbb{R}^n$ into two half spaces:

DEFINITION 1.3.49. A (topological closed) *half space* is a set of the form $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}\cdot\boldsymbol{x} \leq b\}$ for some $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and $b \in \mathbb{R}$.

We now state a theorem which is fundamental in convex optimization:

THEOREM 1.3.50 (Separating hyperplane theorem [**BV04**, Section 2.5.1]). *Suppose that C and D are nonempty convex sets in $\mathbb{R}^n$ with $C \cap D = \emptyset$, then there exist $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and $b \in \mathbb{R}$ such that*

$$C \subset \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}\cdot\boldsymbol{x} \leq b\} \quad \textit{and} \quad D \subset \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}\cdot\boldsymbol{x} \geq b\}. \tag{1.3.4}$$

The proof sounds easy, but the proof is quite technical. It is important to mention that, by assuming the axiom of choice, one can prove an analogue result for abstract spaces ("infinite dimension space"), which is called the *first geometry form Hahn-Banach theorem* [**Bre11**, Theorem 1.6]. Despite this is a geometric result, but this has an analytic form [**Bre11**, Theorem 1.1], which is fundamental in functional analysis.

However, we see that the result in Theorem 1.3.50 is not good enough. For example, we choose convex sets $C = D = \{\boldsymbol{0}\}$ and we see that C and D are separated in the sense of (1.3.4). This arises the following question:

QUESTION 1.3.51. *Given any convex sets C and D in $\mathbb{R}^n$, is it always possible to* strictly separate *C and D by a hyperplane or not? More precisely, is it always possible to find $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $b \in \mathbb{R}$ such that*

$$C \subset \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a} \cdot \boldsymbol{x} < b\} \quad \text{and} \quad D \subset \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a} \cdot \boldsymbol{x} > b\} \tag{1.3.5}$$

or not?

However, the answer of Question 1.3.51 even when both C and D are closed, see the following exercise:

EXERCISE 1.3.52. Given an example of two topological closed convex sets C and D in $\mathbb{R}^2$ with $C \cap D = \emptyset$ which cannot be strictly separated in the sense of (1.3.5). Justify your answers.

THEOREM 1.3.53 (Separating hyperplane theorem). *Suppose that C and D are nonempty convex sets in $\mathbb{R}^n$ with $C \cap D = \emptyset$. If C is topological closed and D is compact, then there exist $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $b \in \mathbb{R}$ such that*

$$C \subset \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a} \cdot \boldsymbol{x} < b\} \quad \text{and} \quad D \subset \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a} \cdot \boldsymbol{x} > b\}.$$

It is important to mention that, by assuming the axiom of choice, one can prove an analogue result for abstract spaces ("infinite dimensional space"), which is called the *second geometry form Hahn-Banach theorem* [**Bre11**, Theorem 1.7]. This geometric result also leads analytic a result [**Bre11**, Corollary 1.8] which is fundamental to verify bases of abstract sets (for example the Fourier series, see e.g. my lecture note [**Kow22**]).

DEFINITION 1.3.54. Let S be a set in $\mathbb{R}^n$ and let $\boldsymbol{x}_0 \in \partial S$. If there exists $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that $\boldsymbol{a} \cdot \boldsymbol{x} \leq \boldsymbol{a} \cdot \boldsymbol{x}_0$ for all $\boldsymbol{x} \in S$, then the hyperplane $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a} \cdot \boldsymbol{x} = \boldsymbol{a} \cdot \boldsymbol{x}_0\}$ is called a *supporting hyperplane* to S at the point $\boldsymbol{x}_0$.

We now state a basic result regarding the supporting hyperplane, which is readily proved from the separating hyperplane theorem (Theorem 1.3.50):

THEOREM 1.3.55 (Supporting hyperplane theorem [**BV04**, Section 2.5.2]). *If C is a nonempty convex set in $\mathbb{R}^n$, then for each $\boldsymbol{x}_0 \in \partial C$ there exists a supporting hyperplane to C at $\boldsymbol{x}_0$.*

Combining the above theorem with Corollary 1.3.14 and Exercise 1.3.41, we immediately conclude the following corollary.

COROLLARY 1.3.56. *A closed convex set in $\mathbb{R}^n$ is the intersection of all (topological closed) half spaces that contain it.*

It is also interesting to mention a partial converse:

THEOREM 1.3.57 (A partial converse of supporting hyperplane theorem [**BV04**, Exercise 2.27]). *Let C be a topological closed set in $\mathbb{R}^n$ such that* $\mathrm{int}\,(C) \neq \emptyset$. *If for each point in its boundary ∂C has a supporting hyperplane, then C is convex.*

1.4. Limits and continuous functions

Let us consider the functions

$$f(x) = \frac{x^2}{x} \text{ for all } x \in \mathbb{R} \setminus \{0\}, \quad g(x) = x \text{ for all } x \in \mathbb{R}.$$

We see that $f(x) = g(x)$ for all $x \in \mathbb{R} \setminus \{0\}$ but $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are different functions, since f is not well-defined at $x = 0$. Intuitively, they should be almost the same: When x is "sufficiently close" to 0, then $f(x)$ is an "approximation" of $g(0) = 0$. But, what is the precise meaning of "sufficiently close" and "approximation"? We now quantify this intuition by introducing the following concept.

DEFINITION 1.4.1. Let S be any subset in $\mathbb{R}^n$ (not necessarily open), let $\boldsymbol{x}_0 \in S$ and let $f : S \setminus \{\boldsymbol{x}_0\} \to \mathbb{R}$ be a function.

(a) We say that $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = L$ for some $L \in \mathbb{R}$ if the following holds: Given any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\boldsymbol{x} \in S \cap B_\delta(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\} \quad \text{implies} \quad |f(\boldsymbol{x}) - L| < \varepsilon. \tag{1.4.1}$$

(b) We say that $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = +\infty$ if the following holds: Given any $M > 0$, there exists $\delta = \delta(M) > 0$ such that

$$\boldsymbol{x} \in S \cap B_\delta(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\} \quad \text{implies} \quad f(\boldsymbol{x}) > M.$$

(c) We say that $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = -\infty$ if the following holds: Given any $M > 0$, there exists $\delta = \delta(M) > 0$ such that

$$\boldsymbol{x} \in S \cap B_\delta(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\} \quad \text{implies} \quad f(\boldsymbol{x}) < -M.$$

We say that the *limit* $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists (sometimes we emphasize the limit exists in $[-\infty, +\infty]$) if either one of the (a), (b) or (c) above holds. We say that the *limit* $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists in $\mathbb{R}$ if the (a) above holds. If S is an open set, we simply denote

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) := \lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}).$$

The idea is: Lets find a third-party judge, which is absolutely fair, give a tolerance level $\varepsilon > 0$ (ε is the Greek letter corresponding to English character "e", which represents the "error"), we then decide a reasonable distance $\delta > 0$ (δ is the Greek letter corresponding to English character "d", which represents the "distance") depends on the tolerance error ε given by the judge (and we usually emphasize the dependence by writing $\delta = \delta(\varepsilon)$). Similar philosophy appeared in the cases (b) and (c) as well. We see that (1.4.1) can be rephrased as

$$\boldsymbol{x} \in S \cap B_\delta(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\} \quad \text{implies} \quad f(\boldsymbol{x}) \in B_\varepsilon(L),$$

that is,

$$f(S \cap B_\delta(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\}) := \{f(\boldsymbol{x}) \in \mathbb{R} : S \cap B_\delta(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\}\} \subset B_\varepsilon(L).$$

LEMMA 1.4.2 (Basic properties of limits). *Let S be any subset in $\mathbb{R}^n$, let $\boldsymbol{x}_0 \in S$ and let $f, g : S \setminus \{\boldsymbol{x}_0\} \to \mathbb{R}$ be a function. If both limits $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})$ and $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x})$ exist in $\mathbb{R}$, then the following holds true:*

(a) *For each $c_1 \in \mathbb{R}$ and $c_2 \in \mathbb{R}$, the limit $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0}(c_1 f(\boldsymbol{x}) + c_2 g(\boldsymbol{x}))$ exists in $\mathbb{R}$ and satisfies*

$$\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0}(c_1 f(\boldsymbol{x}) + c_2 g(\boldsymbol{x})) = c_1 \lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) + c_2 \lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x}).$$

(b) *If $f(\boldsymbol{x}) \leq g(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}_0\}$, then*

$$\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) \leq \lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x}).$$

(c) *The limit $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) g(\boldsymbol{x})$ exists in $\mathbb{R}$ and satisfies*

$$\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) g(\boldsymbol{x}) = \left(\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})\right)\left(\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x})\right).$$

(d) *If we additionally assume that $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x}) \neq 0$, then the limit $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})}$ exists in $\mathbb{R}$ and satisfies*

$$\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = \frac{\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})}{\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x})}.$$

REMARK 1.4.3. Some results above can be extended for limits in $[-\infty, +\infty]$, here we will not going to point out all these details here. The main point here is one has to avoid the "indefinite form" such as "$\infty - \infty$" or "$\frac{\infty}{\infty}$".

Unfortunately, for the case when $n \geq 2$ there is no result which is analogue to the L' Hôpital's rule which holds true when $n = 1$. In addition, one cannot define the notion "left" and "right". In practical, it is not easy to check that whether the limit exists or not. Here suggests another perspective which may helpful.

DEFINITION 1.4.4. Let S be any subset in $\mathbb{R}^n$, let $\boldsymbol{x}_0 \in S$ and let $f, g : S \setminus \{\boldsymbol{x}_0\} \to \mathbb{R}$ be functions. We define the *limit superior* or *upper limit* by

$$\limsup_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x}) := \lim_{r \to 0_+}\left(\sup_{S \cap B_r(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\}} g\right), \tag{1.4.2}$$

where $\lim_{r \to 0_+}$ is the right limit on $\mathbb{R}$ and

$$\sup_{S \cap B_r(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\}} g = \inf\{M \in [-\infty, +\infty] : M > g(\boldsymbol{x}) \text{ for all } \boldsymbol{x} \in S \cap B_r(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\}\}$$

where the infimum can be understood in terms of Lebesgue measure. Similarly, we define the *limit inferior* or lower *limit* by

$$\liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) := \lim_{r\to 0_+} \left(\inf_{S\cap B_r(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\}} g \right), \tag{1.4.3}$$

where $\lim_{r\to 0_+}$ is the right limit on $\mathbb{R}$ and

$$\inf_{S\cap B_r(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\}} g = \sup \{ M\in [-\infty,+\infty] : M < g(\boldsymbol{x}) \text{ for all } \boldsymbol{x}\in S\cap B_r(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\} \}.$$

If S is an open set in $\mathbb{R}^n$, we simply denote

$$\limsup_{\boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) := \limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) \quad \text{and} \quad \liminf_{\boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) := \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}).$$

One sees that the mapping

$$r\mapsto \sup_{S\cap B_r(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\}} g$$

is monotone non-increasing, therefore the right limit (1.4.2) always exists in $[-\infty,+\infty]$. Similar arguments shows that the right limit (1.4.3) always exists in $[-\infty,+\infty]$ as well. It is worth to mention the following result:

THEOREM 1.4.5. *Let S be any subset in $\mathbb{R}^n$, let $\boldsymbol{x}_0\in S$ and let $f : S\setminus\{\boldsymbol{x}_0\}\to\mathbb{R}$ be a function.*

(a) *If $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists in $[-\infty,+\infty]$, then*

$$\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) = \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) = \lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}). \tag{1.4.4}$$

(b) *If $\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) = \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x})$, then $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists in $[-\infty,+\infty]$ and (1.4.4) holds.*

EXERCISE 1.4.6. Prove Theorem 1.4.5.

In order to simplify our statements, here we emphasize the difference between the term "near" and "at".

DEFINITION 1.4.7. Let S be any subset in $\mathbb{R}^n$, let $\boldsymbol{x}_0\in S$ and let $f : S\setminus\{\boldsymbol{x}_0\}\to\mathbb{R}$ be a function. We say that a certain property of f holds for all $\boldsymbol{x}\in S\setminus\{\boldsymbol{x}_0\}$ near $\boldsymbol{x}_0$ if there exists $\delta>0$ such that the property of f holds true for all $\boldsymbol{x}\in S\cap B_\delta(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\}$.

Here we emphasize some basic properties of limits (Lemma 1.4.2) cannot directly extend for limit superior/inferior. We exhibit basic properties for limit superior/inferior in the following exercise.

EXERCISE 1.4.8. Let S be any subset in $\mathbb{R}^n$, let $\boldsymbol{x}_0\in S$ and let $f,g : S\setminus\{\boldsymbol{x}_0\}\to\mathbb{R}$ be functions.

(a) Show that

$$\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} (f(\boldsymbol{x})+g(\boldsymbol{x})) \le \limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}) + \limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}).$$

(b) Show that

$$\liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} (f(\boldsymbol{x})+g(\boldsymbol{x})) \geq \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}) + \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}).$$

(c) If $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x})$ exists in $\mathbb{R}$, show that

$$\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} (f(\boldsymbol{x})+g(\boldsymbol{x})) = \limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}) + \lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}),$$
$$\liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} (f(\boldsymbol{x})+g(\boldsymbol{x})) = \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}) + \lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}).$$

(d) If $f(\boldsymbol{x}) \leq g(\boldsymbol{x})$ for all $\boldsymbol{x} \in S\setminus\{\boldsymbol{x}_0\}$ *near*[3] $\boldsymbol{x}_0$, show that

$$\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}) \leq \limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) \quad \text{and} \quad \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}) \leq \liminf_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}).$$

(e) If $f(\boldsymbol{x}) \geq 0$ and $g(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in S\setminus\{\boldsymbol{x}_0\}$ *near*[4] $\boldsymbol{x}_0$, show that

$$\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x})g(\boldsymbol{x}) \leq \left(\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x})\right)\left(\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x})\right).$$

(f) If $|f(\boldsymbol{x})| \leq M$ for all $\boldsymbol{x} \in S\setminus\{\boldsymbol{x}_0\}$ *near*[5] $\boldsymbol{x}_0$ and $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x})$ exists in $\mathbb{R}$ with $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x}) \geq 0$, show that

$$\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x})g(\boldsymbol{x}) = \left(\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x})\right)\left(\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g(\boldsymbol{x})\right).$$

In the particular case when $g(\boldsymbol{x}) = c \geq 0$ for all for all $\boldsymbol{x} \in S\setminus\{\boldsymbol{x}_0\}$ near $\boldsymbol{x}_0$, we reach

$$\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} cf(\boldsymbol{x}) = c\limsup_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}).$$

EXERCISE 1.4.9 (Squeeze theorem). Let S be any subset in $\mathbb{R}^n$, let $\boldsymbol{x}_0 \in S$ and let $f, g_1, g_2 : S\setminus\{\boldsymbol{x}_0\} \to \mathbb{R}$ be functions. If

$$g_1(\boldsymbol{x}) \leq f(\boldsymbol{x}) \leq g_2(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x} \in S\setminus\{\boldsymbol{x}_0\} \text{ near } \boldsymbol{x}_0,$$

and both $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g_1(\boldsymbol{x})$ and $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g_2(\boldsymbol{x})$ exist with

$$\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g_1(\boldsymbol{x}) = \lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g_2(\boldsymbol{x}),$$

show that $\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists and

$$\lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} f(\boldsymbol{x}) = \lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g_1(\boldsymbol{x}) = \lim_{S\ni \boldsymbol{x}\to \boldsymbol{x}_0} g_2(\boldsymbol{x}).$$

---

[3] There exists $\delta > 0$ such that $f(\boldsymbol{x}) \leq g(\boldsymbol{x})$ for all $\boldsymbol{x} \in S\cap B_\delta(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\}$.

[4] There exists $\delta > 0$ such that $f(\boldsymbol{x}) \geq 0$ and $g(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in S\cap B_\delta(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\}$.

[5] There exists $\delta > 0$ such that $|f(\boldsymbol{x})| \leq M$ for all $\boldsymbol{x} \in S\cap B_\delta(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\}$.

It is helpful to see that $\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = L \in \mathbb{R}$ if and only if

$$\limsup_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} |f(\boldsymbol{x}) - L| = 0,$$

and we can estimate $|f(\boldsymbol{x}) - L|$ using triangle inequality, which is much convenient in practical.

DEFINITION 1.4.10. Let S be any subset in $\mathbb{R}^n$ (not necessarily open) and let $f : S \to \mathbb{R}$ be a function. We say that the function f is continuous at $\boldsymbol{x}_0 \in S$ if

$$\lim_{S \ni \boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = f(\boldsymbol{x}_0). \tag{1.4.5}$$

(a) We say that the function f is continuous *near* $\boldsymbol{x}_0 \in S$ if there exists $\delta > 0$ such that the function f is continuous at all points in $S \cap B_\delta(\boldsymbol{x}_0)$.
(b) If f is continuous at all points in S, then we simply say that the function $f : S \to \mathbb{R}$ is continuous, and we denote $f \in C(S)$.

EXAMPLE 1.4.11. Let $n = 1$ and let $f : [0,1] \to \mathbb{R}$ be a continuous function. At the endpoint $x = 0$, we see that

$$f(0) = \lim_{x \to 0_+} f(x),$$

while at the endpoint $x = 1$, we see that

$$f(1) = \lim_{x \to 1_-} f(x),$$

where $\lim_{x \to 1_-}$ denotes the left limit.

Similar to Definition 1.4.7, we again emphasize the difference between the terms "near" and "at":

DEFINITION 1.4.12. Let S be any subset in $\mathbb{R}^n$, let $\boldsymbol{x}_0 \in S$ and let $f : S \to \mathbb{R}$ be a function. We say that a certain property of f holds for all $\boldsymbol{x} \in S$ near $\boldsymbol{x}_0$ if there exists $\delta > 0$ such that the property of f holds true for all $\boldsymbol{x} \in S \cap B_\delta(\boldsymbol{x}_0)$.

It is important to observe the following:

EXERCISE 1.4.13. Let S be any subset in $\mathbb{R}^n$ (not necessarily open) and let $f : S \to \mathbb{R}$ be a function. Show that the following are equivalent:

(a) the function $f : S \to \mathbb{R}$ is continuous;
(b) for each open set I in $\mathbb{R}$, the preimage $f^{-1}(I) := \{\boldsymbol{x} \in S : f(\boldsymbol{x}) \in I\}$ is relative open in S.
(c) for each topological closed set I in $\mathbb{R}$, the preimage $f^{-1}(I) := \{\boldsymbol{x} \in S : f(\boldsymbol{x}) \in I\}$ is relative topological closed in S.

EXERCISE 1.4.14. Let S be a connected subset in $\mathbb{R}^n$ and let $f \in C(S)$. Show that $f(S)$ is a connected subset in $\mathbb{R}^n$.

EXERCISE 1.4.15. Let K be a compact subset in $\mathbb{R}^n$ and let $f \in C(K)$. Show that

$$f(K) := \{f(\boldsymbol{x}) \in \mathbb{R} : \boldsymbol{x} \in K\}$$

is a compact subset in $\mathbb{R}^n$ and there exist $\boldsymbol{x}_{\max} \in K$ and $\boldsymbol{x}_{\min} \in K$ such that

$$f(\boldsymbol{x}_{\max}) \geq f(\boldsymbol{x}) \geq f(\boldsymbol{x}_{\min}) \quad \text{for all } \boldsymbol{x} \in K.$$

(**Hint.** The first result can be proved by using the definition of compact sets, while the second result can be proved using Exercise 1.3.32.)

EXERCISE 1.4.16. Construct a topological closed set F in $\mathbb{R}$ and a function $f : F \to \mathbb{R}$ such that $f(F)$ is not closed in $\mathbb{R}$.

EXERCISE 1.4.17. Let S be any set in $\mathbb{R}^n$, and we define the distance function $\operatorname{dist}(\cdot, S) : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ by

$$\operatorname{dist}(\boldsymbol{x}, S) := \inf_{\boldsymbol{y} \in S} |\boldsymbol{x} - \boldsymbol{y}| \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^n.$$

Show that $\operatorname{dist}(\cdot, S) : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is 1-Lipschitz continuous, that is,

$$|\operatorname{dist}(\boldsymbol{x}_1, S) - \operatorname{dist}(\boldsymbol{x}_2, S)| \leq |\boldsymbol{x}_1 - \boldsymbol{x}_2| \quad \text{for all } \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n.$$

EXERCISE 1.4.18. Let K be a compact subset in $\mathbb{R}^n$ and let F be a topological closed subset in $\mathbb{R}^n$. If $K \cap F = \emptyset$, show that

$$\operatorname{dist}(K, F) := \inf_{\boldsymbol{x} \in K, \boldsymbol{y} \in F} |\boldsymbol{x} - \boldsymbol{y}| > 0.$$

EXERCISE 1.4.19. Construct two topological closed subset F_1 and F_2 in $\mathbb{R}^n$ such that $F_1 \cap F_2 = \emptyset$ but $\operatorname{dist}(F_1, F_2) = 0$.

It is also interesting to mention the following theorem, which is a consequence of Urysohn's lemma:

THEOREM 1.4.20 ([**Rud87**, Tietze's extension theorem]). *Let K be a compact subset of $\mathbb{R}^n$ and let $f \in C(K)$. Then there exists an $F \in C(\mathbb{R}^n)$ such that*

$$\operatorname{supp}(F) := \overline{\{\boldsymbol{x} \in \mathbb{R}^n : F(\boldsymbol{x}) \neq 0\}} \text{ is a compact set in } \mathbb{R}^n$$

and $F(\boldsymbol{x}) = f(\boldsymbol{x})$ for all $\boldsymbol{x} \in K$.

We finally end this section by exhibit an example to illustrate that, even in the one-dimensional space, the function can have a complex geometry.

EXAMPLE 1.4.21. We consider the function (it is not possible to plot the function)

$$f : (0,1) \to \mathbb{R}, f(x) = \begin{cases} \frac{1}{q} & ,\text{if } x = \frac{p}{q} \in (0,1) \cap \mathbb{Q}, q > 0, \gcd(p,q) = 1, \\ 0 & ,\text{if } x \in (0,1) \setminus \mathbb{Q}. \end{cases} \tag{1.4.6}$$

In view of Theorem 1.4.5, it is easy to show that f is not continuous at all $x_1 \in (0,1)\cap\mathbb{Q}$, since

$$\liminf_{x\to x_1} f(x) = 0 < f(x_1).$$

We now show that f is continuous at all $x_0 \in (0,1)\setminus\mathbb{Q}$. Since $f(x)\geq 0$ for all $x\in(0,1)$ and $f(x_0)=0$, it is suffice to show $\limsup_{x\to x_0} f(x) = 0$. For each integer $q\in\mathbb{N}$, we define the set of rational number with denominator at most q, that is,

$$\mathbb{Q}_q := \mathbb{Z}\cup\frac{1}{2}\mathbb{Z}\cup\frac{1}{3}\mathbb{Z}\cup\cdots\cup\frac{1}{q}\mathbb{Z}.$$

One sees that $(0,1)\cap\mathbb{Q}_q$ is a finite set, i.e. there are only finitely many points in that set. Since $x_0\in(0,1)\setminus\mathbb{Q}$, then

$$\operatorname{dist}(x_0,(0,1)\cap\mathbb{Q}_q) = \min_{x\in(0,1)\cap\mathbb{Q}_q}|x-x_0| > 0.$$

This means that, if we define

$$r_q := \frac{1}{2}\operatorname{dist}(x_0,(0,1)\cap\mathbb{Q}_q),$$

we see that the set $B_{r_q}(x_0)\setminus\{x_0\}$ only consists of rational number with denominator $\geq q+1$, therefore

$$\sup_{B_{r_q}(x_0)\setminus\{x_0\}} f \leq \frac{1}{q+1}.$$

Hence, we see that

$$\limsup_{x\to x_0} f(x) = \lim_{r\to 0+}\left(\sup_{B_r(x_0)\setminus\{x_0\}} f\right) \leq \sup_{B_{r_q}(x_0)\setminus\{x_0\}} f \leq \frac{1}{q+1} \quad \text{for all } q\in\mathbb{N}.$$

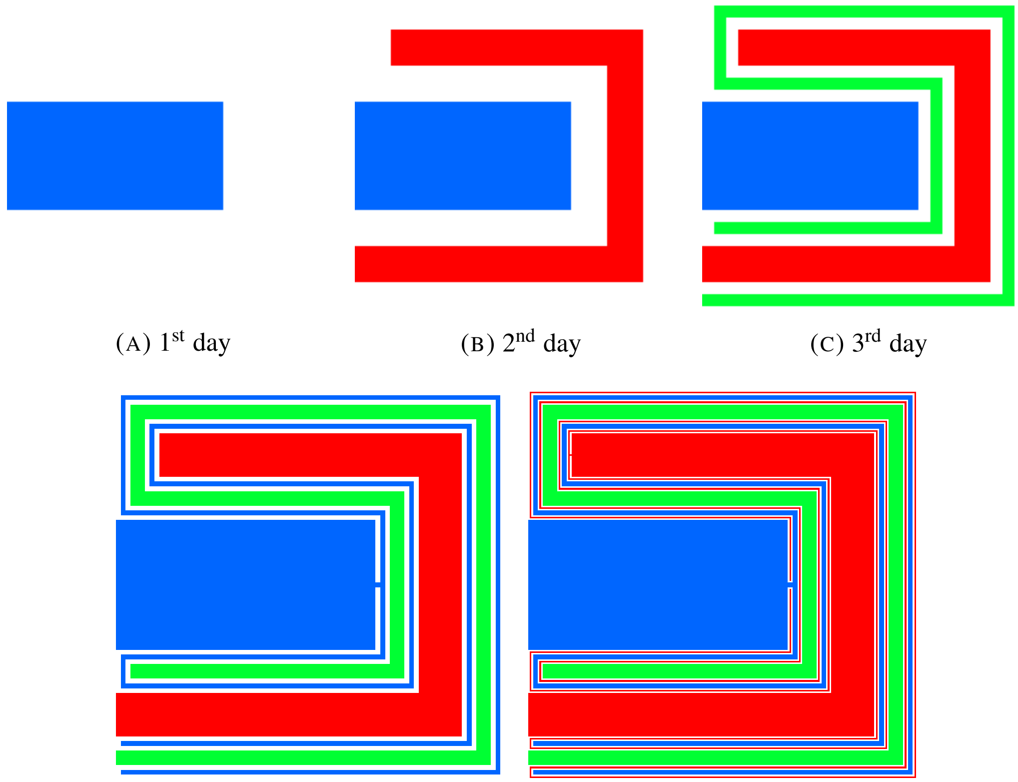
Since the left hand side is independent of q, by arbitrariness of $q\in\mathbb{N}$, we now conclude that $\limsup_{x\to x_0} f(x)=0$, and hence f is continuous at $x_0\in(0,1)\setminus\mathbb{Q}$. We have showed that the function f given in (1.4.6) is continuous at each irrational point, but discontinuous at each rational point.

1.5. Paradoxical spaces: Wada's lakes

The main theme of this section is to exibit a counter-intuitive example in $\mathbb{R}^2$, called the Wada's lakes. These examples were first introduced by Yoneyama [**Yon17**], which is similar to the construction by Brower [**Bro10**]. Here we will follow the approach in the Wikipedia page (`https://en.wikipedia.org/wiki/Lakes_of_Wada`).

The simplest version of the Wada's lakes are three disjoint connected open sets in the topological closed unit square $S_0 := [0,1]\times[0,1]\subset\mathbb{R}^2$ with the counter-intuitive property that they all have the same boundary. Begin with a closed unit square S_0 of dry land, we dig 3 lakes according to the following rule:

- On k^{th} day ($k\in\mathbb{N}$), we extend the m^{th} lake (where $m\equiv k \mod 3$) so that it is open and connected and passes within a distance $1/k$ of all remaining dry land.



(A) 1^{st} day (B) 2^{nd} day (C) 3^{rd} day

(D) 4^{th} day (E) 5^{th} day

FIGURE 1.5.1. First 5 days of the construction of Wada's lakes: 本日晴天, CC0, via Wikimedia Commons

- This should be done so that the remaining dry land S_k remains homeomorphic to a closed unit square $[0,1]\times[0,1]\subset\mathbb{R}^2$ (i.e. there exists a continuous bijective function $\phi_k : S_k \to S_0$ such that its inverse $\phi_k^{-1} : S_0 \to S_k$ is also continuous).

REMARK 1.5.1. By using Exercise 1.4.14 and Exercise 1.4.15, since S_0 is connected and compact, then the remaining dry land S_k at the k^{th} day also must be connected and compact. Roughly speaking, the remaining dry land S_k at the k^{th} day can be understood as a corridor of width $\leq 1/k$.

From the construction, it is easy to see that $S_{k+1}\subset S_k$. After an infinite number of days, the remaining dry land is given by

$$S_\infty := \bigcap_{k\in\mathbb{N}} S_k,$$

which is nonempty by Exercise 1.3.31, and the remaining dry land S_∞ is exactly the boundary of each of the 3 lakes.

EXAMPLE 1.5.2. The first 5 days might be (see Figure 1.5.1):

(1) On 1st day ($k = 1$), we dig the 1st lake ($m = 1$, in blue) of width 3^{-1} passing within $3^{-1}\sqrt{2}$ of all dry land, and the remaining dry land is S_1.
(2) On 2nd day ($k = 2$), we dig the 2nd lake ($m = 2$, in red) of width 3^{-2} passing within $3^{-2}\sqrt{2}$ the dry land S_1, and the remaining dry land is S_2.
(3) On 3rd day ($k = 3$), we dig the 3rd lake ($m = 3$, in green) of width 3^{-3} passing within $3^{-3}\sqrt{2}$ the dry land S_2, and the remaining dry land is S_3.
(4) On 4th day ($k = 4$), we extend the 1st lake ($m = 1$, in blue) by a channel of width 3^{-4} passing within $3^{-4}\sqrt{2}$ the dry land S_3, and the remaining dry land is S_4.
(5) On 5th day ($k = 5$), we extend the 2nd lake ($m = 2$, in red) by a channel of width 3^{-5} passing within $3^{-5}\sqrt{2}$ the dry land S_4, and the remaining dry land is S_5.

On kth day, we extend the mth lake (where $m \equiv k \mod 3$) by a channel of width 3^{-k} passing within $3^{-k}\sqrt{2}$ the dry land S_{k-1}, and the remaining dry land is S_k.

REMARK 1.5.3. A variation of this construction can produce a countable infinite number of connected lakes with the same boundary: Instead of extending the lakes in the order $1,2,3,1,2,3,1,2,3,\cdots$ as described above, extend them in the order

$$1,1,2,1,2,3,1,2,3,4,1,2,3,4,5,\cdots.$$

1.6. A paradoxical set: the Cantor ternary set

The Cantor ternary set $C \subset [0,1]$ is created by iteratively deleting the open middle third from a set of line segments. First of all, we delete the open middle third $(\frac{1}{3}, \frac{2}{3})$ from the interval $C_0 = [0,1]$, leaving two line segments

$$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Next, the open middle third of each of these remaining segments is deleted, leaving four line segments

$$C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right],$$

and so on. This construction is equivalent by the following recurrence relation:

$$C_0 := [0,1], \quad C_k := \frac{1}{3}C_{k-1} \cup \left(\frac{2}{3} + \frac{C_{k-1}}{3}\right) \equiv \left\{\frac{1}{3}x \cup \left(\frac{2}{3} + \frac{x}{3}\right) : x \in C_{k-1}\right\}.$$

One sees that C_k are nonempty compact sets with $C_k \subset C_{k-1}$ for all $k \in \mathbb{N}$. Therefore we can use Exercise 1.3.31 to guarantee the Cantor ternary set

$$C := \bigcap_{k \in \mathbb{N}} C_k$$

is a nonempty compact set. Since the Cantor set is defined as the set of points not excluded, the proportion (more precisely, the Lebesgue measure) of the unit interval remaining can be found by

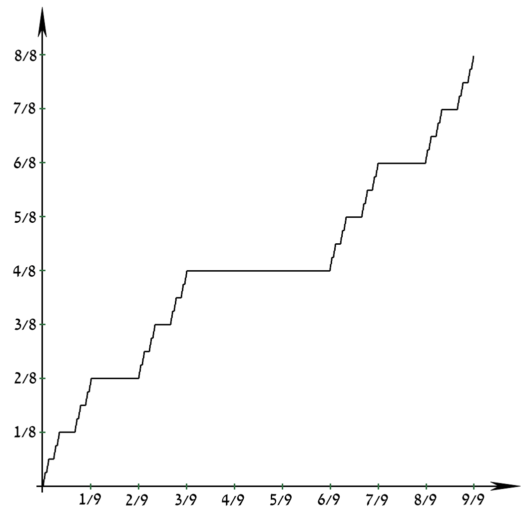


FIGURE 1.6.1. The Cantor function $c : [0,1] \to [0,1]$: CantorEscalier.svg: Theonderivative work: Amirki, CC BY-SA 3.0, via Wikimedia Commons

total length removed. This total is the geometric progression

$$\sum_{k=0}^{\infty} \frac{2^k}{3^{k+1}} = \frac{1}{3}\left(\frac{1}{1-\frac{2}{3}}\right) = 1,$$

so that the Lebesgue measure ("total length") of C is zero. On the other hand, we see that (this is not obvious)

$$C = \left\{ \sum_{i=1}^{\infty} c_i 3^{-i} : c_i \in \{0,2\} \right\},$$

and this shows that C is uncountable, and in fact its cardinality is equal to $\mathbb{R}$ (this is not obvious as well). That is, the cantor set C is "small" in terms of Lebesgue measure, but "large" in terms of cardinality.

The Cantor function $c : [0,1] \to [0,1]$, is defined by (see Figure 1.6.1)

$$c(x) := \begin{cases} \sum_{k=1}^{\infty} a_k 2^{-k} & , x = \sum_{k=1}^{\infty} 2a_k 3^{-k} \in C \text{ for } a_k \in \{0,1\}, \\ \sup_{y \leq x, y \in C} c(y) & , x \in [0,1] \setminus C. \end{cases}$$

Note that $c(0) = 0$, $c(1) = 1$, $0 \leq c(x) \leq 1$ for all $x \in [0,1]$ and c is a monotonic nondecreasing continuous function on $[0,1]$. Note that c is locally constant in each component of $[0,1] \setminus C$, an immediate consequence shows that the derivative c' of c is equal to zero for all points, except a measure zero set, in $[0,1]$. This exhibits a counterexample for the fundamental theorem of calculus for Lebesgue measure if we do not impose any additional assumption. In fact, the fundamental theorem of calculus only holds if and only if the function is absolute continuous function [**WZ15**, Theorem 7.29].

1.7. Pompeiu conjecture for convex domains

Let D be a nonempty bounded convex domain in $\mathbb{R}^n$ ($n \geq 2$) with Lipschitz boundary ∂D. Let Σ denote the set of rigid motions of $\mathbb{R}^n$ onto itself. In other words, each transformation $\sigma \in \Sigma$ is exactly the rotation with respect to the origin and then followed by a translation.

DEFINITION 1.7.1. We say that D has the *Pompeiu property* if and only if the only continuous (complex-valued) function f defined on $\mathbb{R}^n$ for which

$$\int_{\sigma(D)} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \text{ for all } \sigma \in \Sigma, \tag{1.7.1}$$

is the trivial function $f \equiv 0$, where $\sigma(D) := \{\sigma(\boldsymbol{x}) \in \mathbb{R}^n : \boldsymbol{x} \in D\}$.

Note that $\sigma(D)$ is convex since D is convex. It was known that balls do not have the Pompeiu property, that is, there exists a nontrivial continuous function f such that (1.7.1) holds. The well-known Pompeiu problem [**BK82, Pom29, Wil76, Wil81**] asks whether only balls do not have the Pompeiu property. Despite this problem sounds apparent easy, but actually this problem is still open until today (even for the case $n = 2$ and $n = 3$ the problem is still not fully answered yet).

Indeed the assumption in Pompeiu problem is equivalent to the existence of a function v solving the free boundary problem

$$(\Delta + k^2)v = \chi_D \text{ in } \mathbb{R}^n, \quad v = 0 \text{ outside } D \tag{1.7.2}$$

for some $k > 0$, where $\Delta = \sum_{k=1}^n \partial_k^2$ is the Laplacian and

$$\chi_D(\boldsymbol{x}) = \begin{cases} 1 & , \boldsymbol{x} \in D, \\ 0 & \text{otherwise,} \end{cases}$$

see [**Wil76, Wil81**]. The partial differential equation (1.7.2) is understood in distribution sense (see e.g. [**Kow22, Kow24c**] for more details). It was also showed in [**KLSS24**] that the assumption in Pompeiu problem is equivalent to that D is a null k-quadrature domain for some $k > 0$, that is,

$$\int_D w(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \quad \text{for all } w \in L^1(D) \text{ with } (\Delta + k^2)w = 0 \text{ in } D,$$

where the Helmholtz equation $(\Delta + k^2)w = 0$ in D is understood in distribution sense.

It is worth to mention that the assumptions in Pompeiu problem guarantees that ∂D is analytic [**Wil81**]. An explicit example is given in [**KS24**, Example A.2], and the appendix in that paper also explains the difficulties of this problem.

This conjecture demonstrates that the study of geometry in Euclidean space also related to partial differential equations. Hope that one of you can solve this conjecture in the future (either prove it or construct a counterexample), good luck!

CHAPTER 2

Regular lines in Euclidean space

2.1. Some explicit examples of curves in plane

Before we begin our discussions, let us first exhibit some examples of curve in the plane. One can refer the Wikipedia page (https://en.wikipedia.org/wiki/List_of_curves) for a list of curves as well.

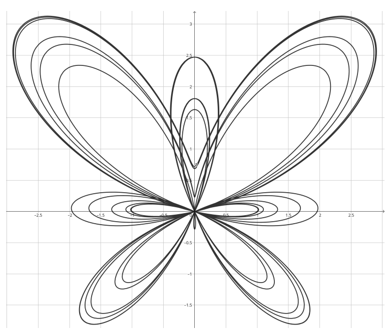


FIGURE 2.1.1. The butterfly curve (https://www.geogebra.org)

EXAMPLE 2.1.1. The butterfly curve [Fay89] can be described by the following parametric equations:

$$x(t) = \sin t \left(e^{\cos t} - 2\cos 4t - \sin^5 \left(\frac{t}{12} \right) \right),$$
$$y(t) = \cos t \left(e^{\cos t} - 2\cos 4t - \sin^5 \left(\frac{t}{12} \right) \right),$$

for $0 \le t \le 12\pi$, or by the polar equation

$$r = e^{\sin \theta} - 2\cos 4\theta + \sin^5 \left(\frac{2\theta - \pi}{24} \right),$$

see Figure 2.1.1 for a plot, see also [GK08] for further discussions.

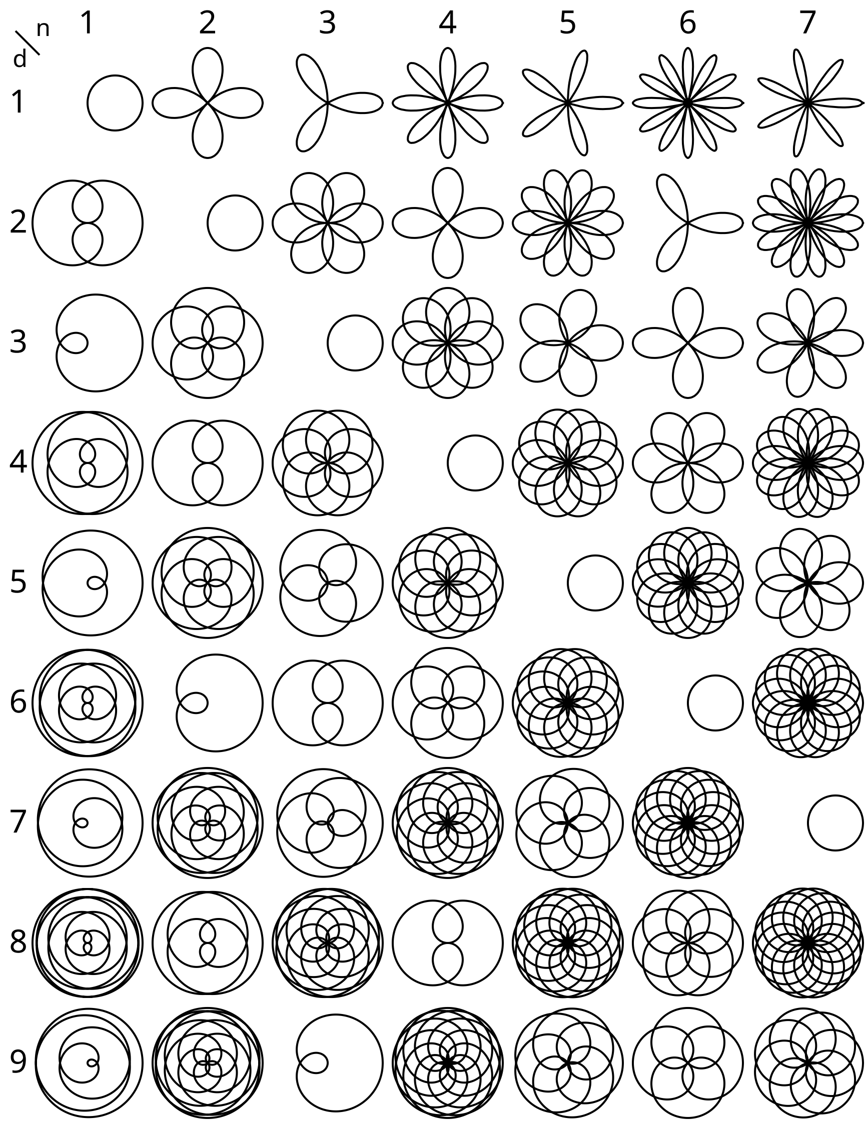


FIGURE 2.1.2. Roses corresponding to some $k = \frac{n}{d}$ with $n, d \in \mathbb{N}$: Jason Davies, CC BY-SA 3.0, via Wikimedia Commons

EXAMPLE 2.1.2. A roses [**CR61**] are sinusoid curves described by the parametric equations:

$$x(\theta) = a\cos(k\theta)\cos\theta,$$
$$y(\theta) = a\cos(k\theta)\sin\theta,$$

with $k \in \mathbb{Q}_{>0}$, or by the polar equation

$$r = a\cos(k\theta),$$

see Figure for plots of some values of $k \in \mathbb{Q}_{>0}$.

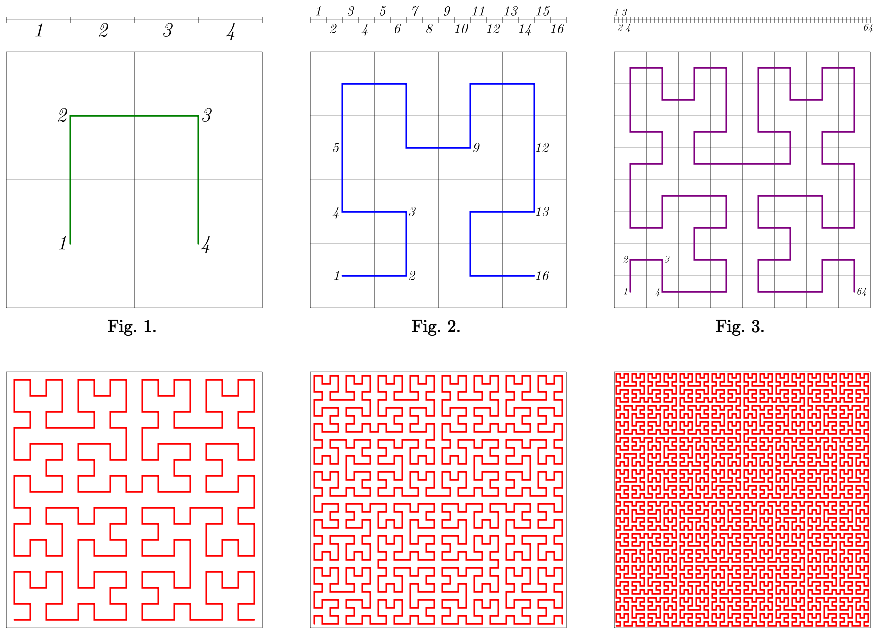


FIGURE 2.1.3. The first 6 iterations of the construction of Hilbert's space filling curve: User:Braindrain0000, CC BY-SA 3.0, via Wikimedia Commons

EXAMPLE 2.1.3 (Hilbert's space filling curve). We now sketch the construction of Hilbert's space filling curve as in Wikipedia page (`https://en.wikipedia.org/wiki/Space-filling_curve`), see Figure 2.1.3. Let C denote the Cantor set in $[0,1]$ and let $c:[0,1]\to[0,1]$ be the Cantor function described in Section 1.6. We now consider the injective continuous function $\phi : C\times C \to [0,1]\times[0,1]$ defined by $\phi(x,y) := (c(x),c(y))$ for all $x,y\in C$. By using the fact that there exists a homeomorphism $\psi : C \to C\times C$ (i.e. $\psi : C\to C$ is a continuous bijection such that its inverse is also continuous), we now consider the injective continuous function

$$\phi\circ\psi : C \to [0,1]\times[0,1].$$

We now use the Tietze extension theorem (Theorem 1.4.20) to see that there exists a continuous function

$$\gamma : [0,1]\to[0,1]\times[0,1]$$

such that $\gamma(x) = \phi\circ\psi(x)$ for all $x\in C$, which is is injective, and this is our desired space filling curve (which is not self-intersecting). One also can take a look on Osgood curve as well, see e.g. the Wikipedia page (`https://en.wikipedia.org/wiki/Osgood_curve`).

The above examples (Example 2.1.1, Example 2.1.2 and Example 2.1.3) show that the curve in a plane can be quite complicated. The study the geometry of Euclidean space is not just plotting lines, but we are interested in mathematical analysis as well, so that we can connect these objects with practical applications. Since this is just an introductory course, we will not going to explain their applications, instead, we will focus on basic mathematical aspects.

2.2. Straight lines and real projective space

We now discuss a fundamental geometric object in $\mathbb{R}^{n+1}$, which is called the straight line in $\mathbb{R}^{n+1}$ (here $n = 1,2,3,\cdots$). We are particular interested in the straight line $L \subset \mathbb{R}^{n+1}$ of the form

$$L \equiv [\boldsymbol{v}] := \left\{ \boldsymbol{x} \in \mathbb{R}^{n+1} : \boldsymbol{x} = t\boldsymbol{v}, t \in \mathbb{R} \right\} \quad \text{for some } \boldsymbol{v} \in \mathbb{R}^{n+1} \setminus \{\boldsymbol{0}\}. \tag{2.2.1}$$

Since $[\boldsymbol{v}] = [\alpha\boldsymbol{v}]$ for all $\alpha \in \mathbb{R} \setminus \{0\}$, we see that the line (2.2.1) can be represented as

$$L \equiv [\boldsymbol{v}] \quad \text{for some } \boldsymbol{v} \in \mathbb{S}^n := \left\{ \boldsymbol{v} \in \mathbb{R}^{n+1} : |\boldsymbol{v}| = 1 \right\}.$$

DEFINITION 2.2.1. [**Lee13**, Example 1.5] The n-dimensional *real projective space* $\mathbb{RP}^n$ is defined as the collection of all straight lines pass through the origin.

For each straight line passing through the origin $L \in \mathbb{RP}^n$, there exists $\boldsymbol{v} \in \mathbb{R}^{n+1} \setminus \{\boldsymbol{0}\}$ such that $L = [\boldsymbol{v}]$. Consequently, there exists $i = 1,\cdots,n+1$ such that $v_i \neq 0$. Now let

$$\tilde{U}_i := \left\{ \boldsymbol{v} \in \mathbb{R}^{n+1} \setminus \{\boldsymbol{0}\} : v_i \neq 0 \right\}, \quad U_i := \left\{ [\boldsymbol{v}] \in \mathbb{RP}^n : \boldsymbol{v} \in \tilde{U}_i \right\}.$$

Now the mapping

$$\varphi_i : U_i \to \mathbb{R}^n, \quad \varphi_i[v_1,\cdots,v_{i-1},v_i,v_{i+1},\cdots,v_n] = \left(\frac{v_1}{v_i},\cdots,\frac{v_{i-1}}{v_i},\frac{v_{i+1}}{v_i},\cdots,\frac{v_{n+1}}{v_i} \right)$$

is well defined because its value is unchanged by multiplying $\boldsymbol{v}$ by a nonzero constant. In fact, $\varphi_i : U_i \to \tilde{U}_i$ is bijective and its inverse function is given by

$$\varphi_i^{-1} : \tilde{U}_i \to U_i, \quad \varphi_i^{-1}(x_1,\cdots,x_n) = [x_1,\cdots,x_{i-1},1,x_{i+1},\cdots,x_n]$$

as you can check. This shows that $\mathbb{RP}^n$ is actually covered by $\{U_1,\cdots,U_{n+1}\}$, and indeed each U_i is simply characterized by some open set $\tilde{U}_i \subset \mathbb{R}^{n+1}$ through the mappings φ_i. In facts, it is also possible to define open subsets of $\mathbb{RP}^n$, that is, one can define a topology of $\mathbb{RP}^n$ from Euclidean topology (called the quotient topology). Accordingly, we say that $(\varphi_i^{-1}, \tilde{U}_i)$ is a chart of $\mathbb{RP}^n$. In fact $\mathbb{RP}^n$ is compact with respect to this topology (despite $\mathbb{R}^{n+1} \setminus \{\boldsymbol{0}\}$ is not compact with respect to Euclidean topology), in other words,

$$\mathbb{RP}^n \text{ is a compact smooth manifold without boundary.}$$

Since this is an introductory course, we will not going too far beyond this point.

Not only the topological/manifold structure, the real projective space $\mathbb{RP}^n$ also has algebraic structure. Let $\mathrm{SO}(3)$ be the collection of all rotation operation about the origin in $\mathbb{R}^n$. In fact, each rotation operation can be uniquely represented by a matrix $R \in \mathbb{R}^{3\times 3}$ with

$$\det(R) = 1, \quad R^{-1} = R^{\intercal},$$

where $\det : \mathbb{R}^{3\times 3} \to \mathbb{R}$ is the determinant and and R^{-1} (resp. $R^{\intercal}$) is the inverse (resp. transpose) matrix of R: Given a vector $\boldsymbol{v} \in \mathbb{R}^3$, then the matrix multiplication $R\boldsymbol{v}$ is the rotated vector. It is interested to mention that there is a one-to-one correspondence between $\mathbb{RP}^3$ and $\mathrm{SO}(3)$: More precisely, this one-to-one correspondence is a diffeomorphism (i.e. the mapping itself and its inverse are C^∞ mapping) by suitable giving a topology/manifold structure on $\mathrm{SO}(3)$. In addition, $\mathrm{SO}(3)$ admits a group (algebraic) structure and it is in fact a Lie group, therefore we call $\mathrm{SO}(3)$ the group of rotation. Therefore, it is possible to pass the group structure of $\mathrm{SO}(3)$ to $\mathbb{RP}^3$. Again, since this is an introductory course, we will not going too far beyond this point.

2.3. Parameterized line

As we saw the space filling curve in Example 2.1.3 above, even the lines in $\mathbb{R}^2$ can be extremely complicated. This is due to the lines are "folded". One way to avoid the line to be "folded" is to consider "smooth/differentiable" curve. We introduce the following definition as in [**dC76**]:

DEFINITION 2.3.1. A *parameterized differentiable curve* is a differentiable map $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ of an open interval $I = (a,b) \subset \mathbb{R}$.

Despite the above definition sounds easy, now we exhibit two examples to show that there is actually no direct relation between the analytic properties of the curve (the differentiability of the mapping $\boldsymbol{\alpha}$) and the geometric properties of the curve $\boldsymbol{\alpha}(I) := \{\boldsymbol{\alpha}(t) \in \mathbb{R}^n : t \in I\}$:

EXAMPLE 2.3.2. The *astroid*, which is smoothly parameterized as

$$x(t) = \cos^3(t), \quad y(t) = \sin^3(t), \quad (0 \leq t \leq 2\pi)$$

has four "folded" points (four outward cusps), see Figure 2.3.1.

EXAMPLE 2.3.3. The heart, which is smoothly parametrized as

$$x(t) = -\frac{1}{2}\sin^3(t), \quad y(t) = \frac{13}{32}\cos(t) - \frac{5}{32}\cos(2t) - \frac{1}{16}\cos(3t) - \frac{1}{32}\cos(4t), \quad (0 \leq t \leq 2\pi)$$

has two "folded" points (one inward cusp and one outward cusp), see Figure 2.3.1. This example was found in Wolfram Mathworld, and also has been used as an example in [**KW21**].

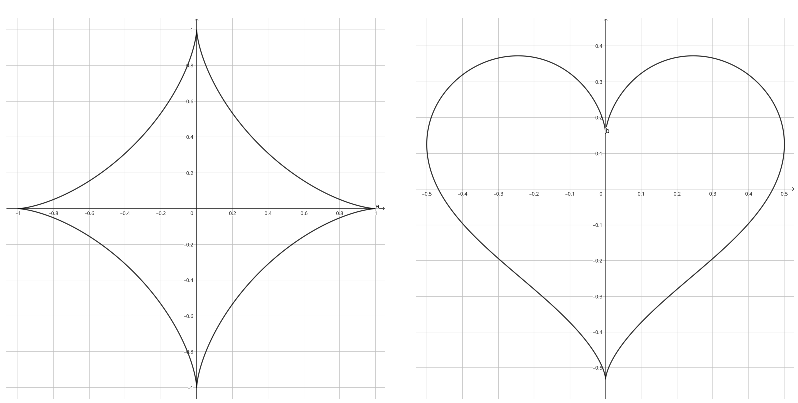


FIGURE 2.3.1. Example 2.3.2 & Example 2.3.3 (`https://www.geogebra.org`)

REMARK 2.3.4. On the other hand, it is possible that a smooth curve has a non-smooth/nondifferentiable parameterization: For example, the straight line $x = y$ also can be parameterized as

$$x(t) = y(t) = t^{1/3}, \quad \text{for all } t \in \mathbb{R}.$$

The above discussions explain the difficulty to guarantee that the curve remains "unfolded", therefore we still follow the notions in Definition 2.3.1 despite this goes against our intuition, at least Definition 2.3.1 rules out the situation described in Remark 2.3.4.

EXERCISE 2.3.5. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ be a parameterized differentiable curve which does not pass through the origin. If there exists $t_0 \in I$ such that $\boldsymbol{\alpha}(t_0)$ is the closest point to the origin, show that $\boldsymbol{\alpha}(t_0) \cdot \boldsymbol{\alpha}'(t_0) = 0$.

EXERCISE 2.3.6. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ be a parameterized differentiable curve with $\boldsymbol{\alpha}'(t) \neq 0$ for all $t \in I$. Show that $|\boldsymbol{\alpha}(t)|$ is a nonzero constant if and only if $\boldsymbol{\alpha}(t) \cdot \boldsymbol{\alpha}'(t) = 0$ for all $t \in I$.

DEFINITION 2.3.7. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ be a parameterized differentiable curve. For each $t \in I$ where $\boldsymbol{\alpha}'(t) \neq 0$, we define the *tangent line* to $\boldsymbol{\alpha}$ at t by

$$[\boldsymbol{\alpha}'(t)] + \boldsymbol{\alpha}(t) \equiv \left\{ \boldsymbol{x} + \boldsymbol{\alpha}(t) \in \mathbb{R}^n : \boldsymbol{x} \in [\boldsymbol{\alpha}'(t)] \right\},$$

where $[\boldsymbol{\alpha}'(t)] \in \mathbb{RP}^{n-1}$ is the straight line passing through the origin is given by (2.2.1).

It is still possible to have tangent line even when $\boldsymbol{\alpha}'(t) = 0$. For example, we consider the parabolic curve

$$x(t) = t^2, \quad y(t) = t.$$

according to Definition 2.3.7, we see that $(x'(0), y'(0)) = (2t, 1)|_{t=0} = (0, 1)$, which means that the tangent line at $(x, y) = (0, 0)$ is given by $[(0, 1)]$. One sees that the parameterization

$$x(t) = t^6, \quad y(t) = t^3$$

produces the exactly same curve (and thus the same tangent line), but with $(x'(0), y'(0)) = (6t^5, 3t^2)|_{t=0} = (0, 0)$. This shows that the the tangent space may exists even the parameterization degenerated, and the way of choosing parameterization is actually important. The following definition suggests us to choose a "good" parameterization in order to guarantee the geometry properties is correctly reflected by analytic representation.

DEFINITION 2.3.8. A parameterized differentiable curve $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ is called a *regular C^k-curve* if all its derivatives of order $\leq k$ are continuous and $\boldsymbol{\alpha}'(t) \neq \mathbf{0}$ for all $t \in I$.

From now on, we will only consider the curves satisfying Definition 2.3.8, and we say that a curve $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ is *regular* if it is a parameterized differentiable regular curve in the sense of Definition 2.3.8.

DEFINITION 2.3.9. The *arc length* of a regular C^1-curve $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ from $t_0 \in I$ is defined by

$$s(t) := \int_{t_0}^{t} |\boldsymbol{\alpha}'(t)| \, \mathrm{d}t.$$

REMARK 2.3.10. Here the arc length is oriented. One sees that $s(t) \le 0$ when $t \le t_0$.

By using fundamental theorem of calculus (see e.g. [**Kow24a**]), one sees that $s : I \to \mathbb{R}$ is differentiable with $s'(t) = |\boldsymbol{\alpha}'(t)|$ for all $t \in I$. If $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ is a regular curve, one has $s'(t) > 0$ for all $t \in I$, which shows that $s : I \to \mathbb{R}$ is strictly increasing. By writing $\tilde{I} := s(I) \subset \mathbb{R}$, we see that $s : I \to \tilde{I}$ is bijective. We slightly abuse the notation by denote its inverse by $t = t(s)$, which maps from $s(I)$ onto I. Now using the chain rule, we see that

$$\frac{\mathrm{d}}{\mathrm{d}s}(\boldsymbol{\alpha}(t(s))) = \boldsymbol{\alpha}'(t)|_{t=t(s)} \frac{\mathrm{d}t}{\mathrm{d}s} = \left.\frac{\boldsymbol{\alpha}'(t)}{s'(t)}\right|_{t=t(s)} = \left.\frac{\boldsymbol{\alpha}'(t)}{|\boldsymbol{\alpha}'(t)|}\right|_{t=t(s)},$$

which implies that

$$\left|\frac{\mathrm{d}}{\mathrm{d}s}(\boldsymbol{\alpha}(t(s)))\right| = 1 \quad \text{for all } s \in \tilde{I}.$$

From now on, we will slightly abuse the notation by writing

$$\boldsymbol{\alpha}(s) := \boldsymbol{\alpha}(t(s)) \quad \text{for all } s \in \tilde{I},$$

and the above discussions showed that $|\boldsymbol{\alpha}'(s)| \equiv \left|\frac{\mathrm{d}}{\mathrm{d}s}(\boldsymbol{\alpha}(t(s)))\right| = 1$ for all $s \in \tilde{I}$.

DEFINITION 2.3.11. Let $\boldsymbol{\alpha} : \tilde{I} \to \mathbb{R}^n$ be a regular C^k-curve. If $|\boldsymbol{\alpha}'(s)| = 1$ for all $s \in \tilde{I}$, then we say that $\boldsymbol{\alpha} : \tilde{I} \to \mathbb{R}^n$ is *parameterized by arc length*.

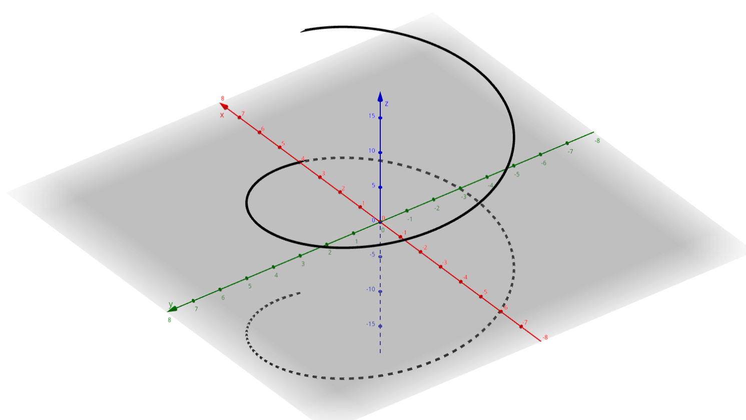


FIGURE 2.3.2. Helix given in Example 2.3.12 (`https://www.geogebra.org`)

EXAMPLE 2.3.12. We consider the helix (see Figure 2.3.2) described by the parameterization $\boldsymbol{\alpha}(t) = (x(t), y(t), z(t))$ with

$$x(t) = 4\cos t, \quad y(t) = 4\sin t, \quad z(t) = 3t.$$

We compute that

$$|\boldsymbol{\alpha}'(t)| = \sqrt{|x'(t)|^2 + |y'(t)|^2 + |z'(t)|^2} = 5 \quad \text{for all } t \in \mathbb{R}$$

which shows that $\boldsymbol{\alpha} : \mathbb{R} \to \mathbb{R}^n$ is a regular curve. By solving the ODE

$$s(t) = |\boldsymbol{\alpha}'(t)|, \quad s(0) = 0, \tag{2.3.1}$$

we see that the unique solution (by the fundamental theorem of ODE, see e.g. [**Kow24b**]) is given by $s(t) = 5t$. Its inverse function is given by

$$t(s) = \frac{s}{5},$$

and thus we see that the helix can be parameterized by arc length as follows:

$$x(s) = 4\cos\left(\frac{s}{5}\right), \quad y(s) = 4\sin\left(\frac{s}{5}\right), \quad z(s) = \frac{3s}{5} \quad \text{for all } s \in \mathbb{R}.$$

Of course, the parameterization is not unique: if we replacing the initial condition $s(0) = 0$ in (2.3.1) by another one still gives an arc length parameterization. So we can choose this wisely to simplify the formula.

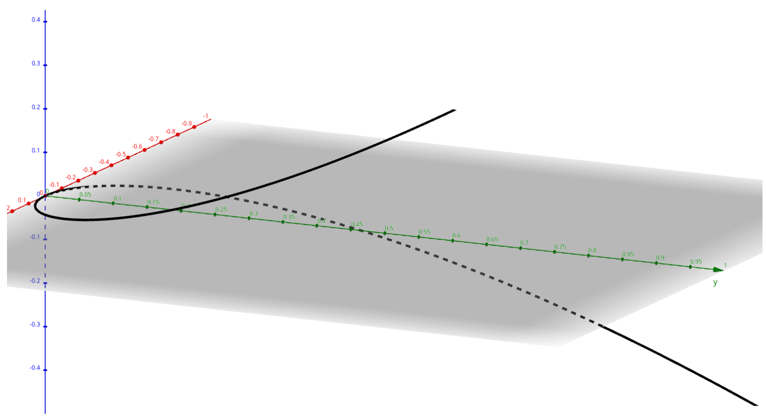


FIGURE 2.3.3. Twisted given in Example 2.3.13 (`https://www.geogebra.org`)

EXAMPLE 2.3.13. We consider the twisted cubic (see Figure 2.3.3) described by the parameterization $\boldsymbol{\alpha}(t) = (x(t), y(t), z(t))$ with

$$x(t) = t, \quad y(t) = t^2, \quad z(t) = \frac{2}{3}t^3.$$

We compute that

$$|\boldsymbol{\alpha}'(t)| = \sqrt{|x'(t)|^2 + |y'(t)|^2 + |z'(t)|^2} = \sqrt{1 + 4t^2 + 4t^4} = 1 + 2t^2 \quad \text{for all } t \in \mathbb{R}$$

which shows that $\boldsymbol{\alpha} : \mathbb{R} \to \mathbb{R}^n$ is a regular curve. By solving the ODE

$$s(t) = |\boldsymbol{\alpha}'(t)|, \quad s(0) = 0,$$

we see that the unique solution (by the fundamental theorem of ODE, see e.g. [**Kow24b**]) is given by $s(t) = t + \frac{2}{3}t^3$. We see that $s : \mathbb{R} \to \mathbb{R}$ is strictly increasing, and thus there exists a unique inverse function $t : \mathbb{R} \to \mathbb{R}$ of $s : \mathbb{R} \to \mathbb{R}$. In fact, the inverse function reads

$$t(s) = \frac{1}{\sqrt[3]{2}\sqrt[3]{\sqrt{9s^2+2}-3s}} - \frac{\sqrt[3]{\sqrt{9s^2+2}-3s}}{2^{2/3}} \quad \text{for all } s \in \mathbb{R},$$

and thus the arc length parameterization of the regular curve reads

$$\begin{aligned} x(s) &= \frac{1}{\sqrt[3]{2}\sqrt[3]{\sqrt{9s^2+2}-3s}} - \frac{\sqrt[3]{\sqrt{9s^2+2}-3s}}{2^{2/3}}, \\ y(s) &= \left(\frac{1}{\sqrt[3]{2}\sqrt[3]{\sqrt{9s^2+2}-3s}} - \frac{\sqrt[3]{\sqrt{9s^2+2}-3s}}{2^{2/3}} \right)^2, \\ z(s) &= \frac{2}{3}\left(\frac{1}{\sqrt[3]{2}\sqrt[3]{\sqrt{9s^2+2}-3s}} - \frac{\sqrt[3]{\sqrt{9s^2+2}-3s}}{2^{2/3}} \right)^3. \end{aligned}$$

This example shows that, despite the arc length is sometimes convenient to prove some results, but it explicit representation can be extremely complicated, and in general we do not expect to write down an explicit formula of arc length parameterization.

2.4. Curvature of a line

DEFINITION 2.4.1. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ be a regular C^2-curve which is parameterized using arc length $s \in I$ (i.e. $|\boldsymbol{\alpha}'(s)| = 1$ for all $s \in I$). The *curvature* $\kappa(s)$ of $\boldsymbol{\alpha}$ at s is defined by

$$\kappa(s) := |\boldsymbol{\alpha}''(s)|.$$

EXAMPLE 2.4.2. We consider the helix in Example 2.3.12, which parameterized using arc length:

$$x(s) = 4\cos\left(\frac{s}{5}\right), \quad y(s) = 4\sin\left(\frac{s}{5}\right), \quad z(s) = \frac{3s}{5} \quad \text{for all } s \in \mathbb{R}.$$

One computes that

$$x'(s) = -\frac{4}{5}\sin\left(\frac{s}{5}\right), \quad y'(s) = \frac{4}{5}\cos\left(\frac{s}{5}\right), \quad z'(s) = \frac{3}{5} \quad \text{for all } s \in \mathbb{R}$$

as well as

$$x''(s) = -\frac{4}{25}\cos\left(\frac{s}{5}\right), \quad y''(s) = -\frac{4}{25}\sin\left(\frac{s}{5}\right), \quad z''(s) = 0 \quad \text{for all } s \in \mathbb{R}.$$

Then its curvature is given by

$$\kappa(s) = \sqrt{|x''(s)|^2 + |y''(s)|^2 + |z''(s)|^2} = \frac{4}{25} \quad \text{for all } s \in \mathbb{R}.$$

Since $1 = |\boldsymbol{\alpha}'(s)|^2 = \boldsymbol{\alpha}'(s) \cdot \boldsymbol{\alpha}'(s)$, then

$$0 = (\boldsymbol{\alpha}'(s) \cdot \boldsymbol{\alpha}'(s))' = 2\boldsymbol{\alpha}'(s) \cdot \boldsymbol{\alpha}''(s),$$

which shows that for each $s \in I$ the vector $\boldsymbol{\alpha}''(s)$ is orthogonal to the *unit tangent vector* given by

$$\boldsymbol{t}(s) := \boldsymbol{\alpha}'(s).$$

Thus it is natural to introduce the *unit normal vector* by

$$\boldsymbol{n}(s) := \frac{\boldsymbol{\alpha}''(s)}{|\boldsymbol{\alpha}''(s)|} = \frac{\boldsymbol{t}'(s)}{\kappa(s)} \quad \text{when } \kappa(s) \neq 0. \tag{2.4.1}$$

2.5. Fundamental theorem of the local theory of curves in $\mathbb{R}^3$

We now restrict ourselves for the case when $n = 3$.

DEFINITION 2.5.1. We now consider the cross product given in Remark 1.2.6 to define the *binormal vector* by

$$\boldsymbol{b}(s) := \boldsymbol{t}(s) \times \boldsymbol{n}(s),$$

so that we see that, for each $s \in I$, the *Frenet trihedron*

$$\{\boldsymbol{t}(s), \boldsymbol{n}(s), \boldsymbol{b}(s)\} \text{ forms an orthogonal basis of } \mathbb{R}^3, \tag{2.5.1}$$

that is:

$$|\boldsymbol{t}(s)| = |\boldsymbol{n}(s)| = |\boldsymbol{b}(s)| = 1, \quad \boldsymbol{t}(s) \cdot \boldsymbol{n}(s) = \boldsymbol{n}(s) \cdot \boldsymbol{b}(s) = \boldsymbol{b}(s) \cdot \boldsymbol{t}(s) = 0$$

and most important:

$$\begin{aligned}&\text{Each } \boldsymbol{v} \in \mathbb{R}^3 \text{ can be uniquely expressed as}\\ &\boldsymbol{v} = (\boldsymbol{v} \cdot \boldsymbol{t}(s))\boldsymbol{t}(s) + (\boldsymbol{v} \cdot \boldsymbol{n}(s))\boldsymbol{n}(s) + (\boldsymbol{v} \cdot \boldsymbol{b}(s))\boldsymbol{b}(s).\end{aligned} \tag{2.5.2}$$

Note that $\mathscr{P}_{\boldsymbol{t}(s)}\boldsymbol{v} := (\boldsymbol{v} \cdot \boldsymbol{t}(s))\boldsymbol{t}(s)$ and $\mathscr{P}_{\boldsymbol{n}(s)}\boldsymbol{v} := (\boldsymbol{v} \cdot \boldsymbol{n}(s))\boldsymbol{n}(s)$ as well as $\mathscr{P}_{\boldsymbol{b}(s)}\boldsymbol{v} := (\boldsymbol{v} \cdot \boldsymbol{b}(s))\boldsymbol{b}(s)$ are the projections of $\boldsymbol{v}$ in the straight line $[\boldsymbol{t}(s)] \in \mathbb{RP}^2$ and $[\boldsymbol{n}(s)] \in \mathbb{RP}^2$ as well as $[\boldsymbol{b}(s)] \in \mathbb{RP}^2$, respectively.

EXERCISE 2.5.2 (Scalar triple product). Show that

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b}) \quad \text{for all } \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3.$$

EXERCISE 2.5.3 (Prouct rule). Let $\boldsymbol{a}, \boldsymbol{b} : I \to \mathbb{R}^3$ be C^1 curves in $\mathbb{R}^3$. Show that

$$(\boldsymbol{a} \times \boldsymbol{b})'(s) = \boldsymbol{a}'(s) \times \boldsymbol{b}(s) + \boldsymbol{a}(s) \times \boldsymbol{b}'(s) \quad \text{for all } s \in I.$$

Now from the product rule (Exercise 2.5.3) we see that

$$\boldsymbol{b}'(s) = \boldsymbol{t}'(s) \times \boldsymbol{n}(s) + \boldsymbol{t}(s) \times \boldsymbol{n}'(s) = \boldsymbol{t}(s) \times \boldsymbol{n}'(s) \quad \text{for all } s \in I,$$

and thus from the scalar triple product (Exercise 2.5.2) we reach

$$\boldsymbol{t}(s) \cdot \boldsymbol{b}'(s) = \boldsymbol{t}(s) \cdot (\boldsymbol{t}(s) \times \boldsymbol{n}'(s)) = \boldsymbol{n}'(s) \cdot (\boldsymbol{t}(s) \times \boldsymbol{t}(s)) = 0.$$

On the other hand, we see that $1 = |\boldsymbol{b}(s)|^2 = \boldsymbol{b}(s) \cdot \boldsymbol{b}(s)$ implies that

$$0 = (\boldsymbol{b}(s) \cdot \boldsymbol{b}(s))' = 2\boldsymbol{b}(s) \cdot \boldsymbol{b}'(s).$$

Since $\boldsymbol{b}'(s)$ is orthogonal to both $\boldsymbol{t}(s)$ and $\boldsymbol{b}(s)$, thus from (2.5.1) we see that $\boldsymbol{b}'(s)$ is parallel to $\boldsymbol{n}(s)$. Accordingly, the following definition is not natural.

DEFINITION 2.5.4. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^3$ be a regular C^2-curve which is parameterized using arc length $s \in I$ (i.e. $|\boldsymbol{\alpha}'(s)| = 1$ for all $s \in I$) such that $\boldsymbol{\alpha}''(s) \neq \mathbf{0}$ for all $s \in I$. The *torsion* $\tau(s)$ of $\boldsymbol{\alpha}$ at s is defined by

$$\boldsymbol{b}'(s) = \tau(s)\boldsymbol{n}(s). \tag{2.5.3}$$

EXERCISE 2.5.5. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^3$ be a regular C^2-curve which is parameterized using arc length $s \in I$ such that $\kappa(s) \neq 0$ for all $s \in I$. Show that $\boldsymbol{\alpha}$ is a plane curve (i.e. $\boldsymbol{\alpha}$ contained in a plane) if and only if $\tau(s) = 0$ for all $s \in I$.

It is worth to mention that the result in Exercise 2.5.5 does not hold true even when there exists only one point s such that $\kappa(s) = 0$, see Exercise 2.5.13 below.

EXERCISE 2.5.6 (Jacobi identity). Show that

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) + \boldsymbol{b} \times (\boldsymbol{c} \times \boldsymbol{a}) + \boldsymbol{c} \times (\boldsymbol{a} \times \boldsymbol{b}) = \mathbf{0} \quad \text{for all } \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3.$$

We now use Jacobi identity (Exercise 2.5.6) to see that

$$\begin{aligned} \boldsymbol{b}(s) \times \boldsymbol{t}(s) &= (\boldsymbol{t}(s) \times \boldsymbol{n}(s)) \times \boldsymbol{t}(s) \\ &= -\boldsymbol{n}(s) \times \overbrace{(\boldsymbol{t}(s) \times \boldsymbol{t}(s))}^{=\mathbf{0}} - \boldsymbol{t}(s) \times \overbrace{(\boldsymbol{t}(s) \times \boldsymbol{n}(s))}^{=\boldsymbol{b}(s)} \\ &= \boldsymbol{b}(s) \times \boldsymbol{t}(s). \end{aligned}$$

Again, from the product rule (Exercise 2.5.3) we see that

$$\boldsymbol{n}'(s) = \boldsymbol{b}'(s) \times \boldsymbol{t}(s) + \boldsymbol{b}(s) \times \boldsymbol{t}'(s) = -\tau(s)\boldsymbol{b}(s) - \kappa(s)\boldsymbol{t}(s).$$

Together with (2.4.1) and (2.5.3), we now reach the following ODE (called the *Frenet formulas*):

$$\boldsymbol{t}'(s) = \kappa(s)\boldsymbol{n}(s), \tag{2.5.4a}$$

$$\boldsymbol{n}'(s) = -\tau(s)\boldsymbol{b}(s) - \kappa(s)\boldsymbol{t}(s), \tag{2.5.4b}$$

$$\boldsymbol{b}'(s) = \tau(s)\boldsymbol{n}(s). \tag{2.5.4c}$$

EXAMPLE 2.5.7. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^3$ be a regular C^3-curve which is parameterized using arc length $s \in I$ such that $\kappa(s) > 0$ and $\tau(s) \neq 0$ for all $s \in I$. From (2.5.4a) we see that

$$(\boldsymbol{\alpha}(s) \cdot \boldsymbol{t}(s))' = \overbrace{\boldsymbol{\alpha}'(s) \cdot \boldsymbol{t}(s)}^{=1} + \boldsymbol{\alpha}(s) \cdot \boldsymbol{t}'(s) = 1 + \kappa(s)\boldsymbol{\alpha}(s) \cdot \boldsymbol{n}(s) \quad \text{for all } s \in I. \tag{2.5.5}$$

We now further differentiate the above identity, and follow by (2.5.4b) to obtain

$$\begin{aligned}\left(\frac{1}{\kappa(s)}(\boldsymbol{\alpha}(s)\cdot\boldsymbol{t}(s))'\right)' &= \left(\frac{1}{\kappa(s)}\right)' + (\boldsymbol{\alpha}(s)\cdot\boldsymbol{n}(s))' \\ &= \left(\frac{1}{\kappa(s)}\right)' + \overbrace{\boldsymbol{\alpha}'(s)\cdot\boldsymbol{n}(s)}^{=\boldsymbol{t}(s)\cdot\boldsymbol{n}(s)=0} + \boldsymbol{\alpha}(s)\cdot\boldsymbol{n}'(s) \\ &= \left(\frac{1}{\kappa(s)}\right)' - \boldsymbol{\alpha}(s)\cdot(\tau(s)\boldsymbol{b}(s)+\kappa(s)\boldsymbol{t}(s)) \\ &= \left(\frac{1}{\kappa(s)}\right)' - \tau(s)\boldsymbol{\alpha}(s)\cdot\boldsymbol{b}(s) - \kappa(s)\boldsymbol{\alpha}(s)\cdot\boldsymbol{t}(s) \quad \text{for all } s\in I,\end{aligned}$$

that is,

(2.5.6)

$$\boldsymbol{\alpha}(s)\cdot\boldsymbol{b}(s) = -\frac{1}{\tau(s)}\left(\frac{1}{\kappa(s)}(\boldsymbol{\alpha}(s)\cdot\boldsymbol{t}(s))'\right)' - \frac{\kappa(s)}{\tau(s)}\boldsymbol{\alpha}(s)\cdot\boldsymbol{t}(s) + \left(\frac{1}{\kappa(s)}\right)'\frac{1}{\tau(s)} \quad \text{for all } s\in I.$$

Now we write (2.5.5) as

$$\boldsymbol{\alpha}(s)\cdot\boldsymbol{n}(s) = \frac{1}{\kappa(s)}(\boldsymbol{\alpha}(s)\cdot\boldsymbol{t}(s))' - \frac{1}{\kappa(s)} \quad \text{for all } s\in I. \tag{2.5.7}$$

EXAMPLE 2.5.8. Let $\boldsymbol{\alpha}: I\to\mathbb{R}^3$ be a regular C^3-curve which is parameterized using arc length $s\in I$ such that $\kappa(s)>0$ and $\tau(s)\neq 0$ for all $s\in I$. If the curve $\boldsymbol{\alpha}$ lies on a sphere of radius $R>0$ about the origin, we see that $|\boldsymbol{\alpha}(s)|=R$ for all $s\in I$, then

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{2}|\boldsymbol{\alpha}(s)|^2\right) = \boldsymbol{\alpha}(s)\cdot\boldsymbol{\alpha}'(s) = \boldsymbol{\alpha}(s)\cdot\boldsymbol{t}(s) \quad \text{for all } s\in I.$$

The from (2.5.6) and (2.5.7) we see that

$$\boldsymbol{\alpha}(s)\cdot\boldsymbol{n}(s) = -\frac{1}{\kappa(s)}, \quad \boldsymbol{\alpha}(s)\cdot\boldsymbol{b}(s) = \left(\frac{1}{\kappa(s)}\right)'\frac{1}{\tau(s)} \quad \text{for all } s\in I.$$

Now using (2.5.2) we see that

$$\begin{aligned}\boldsymbol{\alpha}(s) &= (\boldsymbol{\alpha}(s)\cdot\boldsymbol{t}(s))\boldsymbol{t}(s) + (\boldsymbol{\alpha}(s)\cdot\boldsymbol{n}(s))\boldsymbol{n}(s) + (\boldsymbol{\alpha}(s)\cdot\boldsymbol{b}(s))\boldsymbol{b}(s) \\ &= -\frac{1}{\kappa(s)}\boldsymbol{n}(s) + \left(\frac{1}{\kappa(s)}\right)'\frac{1}{\tau(s)}\boldsymbol{b}(s) \quad \text{for all } s\in I.\end{aligned} \tag{2.5.8}$$

From (2.5.1) we reach the following algebra relation for curvature κ and the torsion τ:

$$\left(\frac{1}{\kappa(s)}\right)^2 + \left(\left(\frac{1}{\kappa(s)}\right)'\frac{1}{\tau(s)}\right)^2 = |\boldsymbol{\alpha}(s)|^2 = R^2 \quad \text{for all } s\in I.$$

In the particular case when $\kappa(s)=\kappa_0$ for some constant $\kappa_0>0$, the above identity immediately gives $\kappa_0 = R^{-1}$.

EXAMPLE 2.5.9. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^3$ be a regular C^3-curve which is parameterized using arc length $s \in I$ such that $\kappa(s) > 0$ and $\tau(s) \neq 0$ for all $s \in I$. Suppose that such curve satisfies

$$\left(\frac{1}{\kappa(s)}\right)^2 + \left(\left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\right)^2 = R^2 \quad \text{for all } s \in I. \tag{2.5.9}$$

We differentiate (2.5.9) to obtain

$$\begin{aligned} 0 &= \left(\left(\frac{1}{\kappa(s)}\right)^2 + \left(\left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\right)^2\right)' \\ &= \frac{2}{\tau(s)}\left(\frac{1}{\kappa(s)}\right)'\left(\frac{\tau(s)}{\kappa(s)} + \left(\left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\right)'\right) \quad \text{for all } s \in I. \end{aligned}$$

Since $\tau(s) \neq 0$ for all $s \in I$, we now reach the following two cases:

(a) $\kappa(s) = \kappa_0$ for some constant $\kappa_0 > 0$; or
(b) $\frac{\tau(s)}{\kappa(s)} + \left(\left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\right)' = 0$ for all $s \in I$.

We now consider case (b). In view of (2.5.8), we define

$$\boldsymbol{\beta}(s) := \boldsymbol{\alpha}(s) + \frac{1}{\kappa(s)}\boldsymbol{n}(s) - \left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\boldsymbol{b}(s) \quad \text{for all } s \in I.$$

We use (2.5.4b) and (2.5.4c) to compute its derivative

$$\begin{aligned} \boldsymbol{\beta}'(s) &= \boldsymbol{t}(s) + \left(\frac{1}{\kappa(s)}\right)'\boldsymbol{n}(s) + \frac{1}{\kappa(s)}\boldsymbol{n}'(s) - \left(\left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\right)'\boldsymbol{b}(s) - \left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\boldsymbol{b}'(s) \\ &= -\left(\frac{\tau(s)}{\kappa(s)} + \left(\left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\right)'\right)\boldsymbol{b}(s) = 0 \quad \text{for all } s \in I, \end{aligned}$$

which shows that $\boldsymbol{\beta}(s) = \boldsymbol{\beta}$ for some constant vector $\boldsymbol{\beta} \in \mathbb{R}^3$, and thus

$$\boldsymbol{\alpha}(s) - \boldsymbol{\beta} = -\frac{1}{\kappa(s)}\boldsymbol{n}(s) + \left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\boldsymbol{b}(s) \quad \text{for all } s \in I.$$

Now from (2.5.9) we conclude that

$$|\boldsymbol{\alpha}(s) - \boldsymbol{\beta}|^2 = \left(\frac{1}{\kappa(s)}\right)^2 + \left(\left(\frac{1}{\kappa(s)}\right)' \frac{1}{\tau(s)}\right)^2 = R^2,$$

which shows that $\boldsymbol{\alpha}$ lies on a sphere of radius $R > 0$ about $\boldsymbol{\beta} \in \mathbb{R}^3$. Here we remark that (a) does not guarantee that $\boldsymbol{\alpha}$ lies on a sphere (because the information on torsion τ is missing, and this intuition also verified by Theorem 2.5.11 below), see for example the helix in Example 2.4.2 above (with $\kappa_0 = \frac{4}{25}$).

EXERCISE 2.5.10. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^3$ be a regular C^3-curve which is parameterized using arc length $s \in I$ such that $\kappa(s) > 0$ for all $s \in I$. Show that the torsion τ is given by

$$\tau(s) = -\frac{(\boldsymbol{\alpha}'(s) \times \boldsymbol{\alpha}''(s)) \cdot \boldsymbol{\alpha}'''(s)}{|\kappa(s)|^2} \quad \text{for all } s \in I.$$

It is worth to mention that the fundamental theorem of ODE (see e.g. [**Kow24b**]) applies to the ODE (2.5.4a)–(2.5.4c) and we reach the following theorem (here we skip the details).

THEOREM 2.5.11 (Fundamental theorem of the local theory of curves in $\mathbb{R}^3$ [**dC76**, Section 1.5]). *For each $\kappa \in C^1(I)$ with $\kappa(s) > 0$ for all $s \in I$ and a $\tau \in C^1(I)$, there exists a regular C^2-curve such that $\kappa(s)$ is its curvature and τ is its torsion, which is unique up to a rigid motion (see Section* (1.7.1)*).*

In other words, under some very good condition, a regular curve in $\mathbb{R}^3$ can be obtained from a straight line by bending (curvature) and twisting (torsion). However, as demonstrated in the twisted cubic in Example 2.3.13 that the arc length parameterization is not easy to compute in general, the following exercise is helpful to compute the curvature and torsion for more general class of parameterization.

EXERCISE 2.5.12. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^3$ be a regular C^3-curve (not necessarily parameterized by arc length) such that $\kappa(s) > 0$ for all $t \in I$. Let $s(t)$ be its arc length from some point $t_0 \in I$, and let $t = t(s)$ be its inverse function and set $\boldsymbol{\alpha}'(t) := \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\alpha}(t)$, $\boldsymbol{\alpha}''(t) := (\frac{\mathrm{d}}{\mathrm{d}t})^2\boldsymbol{\alpha}(t)$ and $\boldsymbol{\alpha}'''(t) := (\frac{\mathrm{d}}{\mathrm{d}t})^3\boldsymbol{\alpha}(t)$.

(a) Show that $\frac{\mathrm{d}}{\mathrm{d}s}t(s) = |\boldsymbol{\alpha}'(t(s))|^{-1}$ and $(\frac{\mathrm{d}}{\mathrm{d}s})^2 t(s) = -|\boldsymbol{\alpha}'(t(s))|^{-4}\boldsymbol{\alpha}'(t(s)) \cdot \boldsymbol{\alpha}''(t(s))$.
(b) Show that the curvature κ is given by

$$\kappa(t) = \frac{|\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)|}{|\boldsymbol{\alpha}'(t)|^3} \quad \text{for all } t \in I.$$

(c) Show that the torsion τ is given by

$$\tau(t) = -\frac{(\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)) \cdot \boldsymbol{\alpha}'''(t)}{|\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)|^2} \quad \text{for all } t \in I.$$

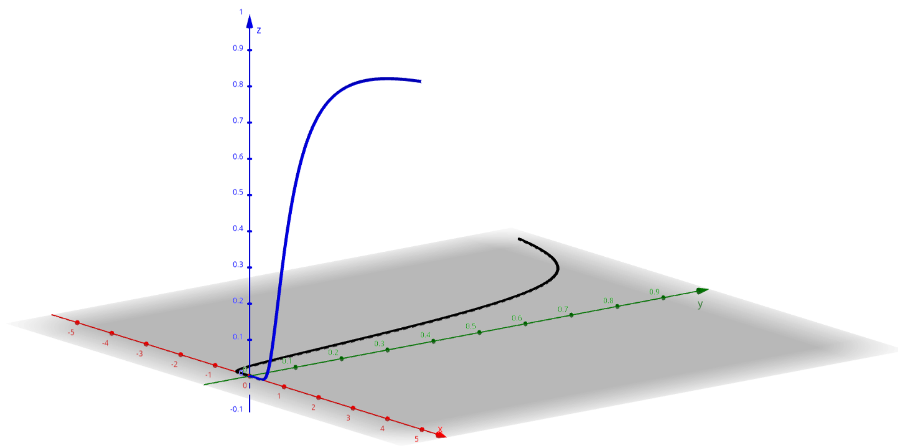


FIGURE 2.5.1. A curve with zero torsion but not contained in a plane (`https://www.geogebra.org`)

EXERCISE 2.5.13. Consider the curve (see Figure 2.5.1)

$$\alpha(t) = \begin{cases} (t, 0, e^{-1/t^2}) & ,t > 0, \\ (t, e^{-1/t^2}, 0) & ,t < 0, \\ (0,0,0) & ,t = 0. \end{cases}$$

(a) Show that $\alpha : \mathbb{R} \to \mathbb{R}$ is C^2.
(b) Show that $\alpha'(t) \neq \mathbf{0}$ for all $t \in \mathbb{R}$.
(c) Show that the curvature $\kappa(t) \neq 0$ for all $t \neq 0$ but $\kappa(0) = 0$.
(d) Show that the torsion τ can be defined so that $\tau(t) = 0$ for all $t \in \mathbb{R}$ even though α is not a plane curve (compare with Exercise 2.5.5).

EXERCISE 2.5.14. One often gives a curve in $\mathbb{R}^2$ in polar coordinates by $r = r(\theta)$ with $a \leq \theta \leq b$ with $r \in C^2((a,b)) \cap C([a,b])$, see e.g. some examples exhibited in Section 2.1 above.

(a) Show that the arc length is

$$\int_a^b \sqrt{|r(\theta)|^2 + |r'(\theta)|^2}\, \mathrm{d}\theta,$$

where $r'(\theta)$ is the derivative of $r(\theta)$ with respect to θ.

(b) Show that the curvature is

$$\kappa(\theta) = \frac{2|r(\theta)|^2 - r(\theta)r''(\theta) + |r(\theta)|^2}{(|r(\theta)|^2 + |r'(\theta)|^2)^{3/2}} \quad \text{for all } a < \theta < b.$$

CHAPTER 3

Regular surfaces in Euclidean space

3.1. Planes and Grassmannian

We now discuss another fundamental geometric object in $\mathbb{R}^3$, which is called the (2-dimensional) planes. Similar to the straight lines, we also have the following definition.

DEFINITION 3.1.1. The Grassmannian $\mathrm{Gr}(2,3)$ is the collection of all (2-dimensional) planes containing the origin.

However, each (2-dimensional) planes in $\mathbb{R}^3$ can be expressed as

$$\left\{\boldsymbol{x} \in \mathbb{R}^3 : \boldsymbol{x} \cdot \boldsymbol{v} = 0\right\} \quad \text{for some } \boldsymbol{v} \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}.$$

This shows that the Grassmannian $\mathrm{Gr}(2,3)$ is actually "identical" to the real projective space $\mathbb{RP}^2$, which is the collection of all straight lines in $\mathbb{R}^3$ containing the origin.

The situation will become complicated if we consider higher dimensional Euclidean space. Each k-dimensional hyperplane in $\mathbb{R}^n$ (with $k < n$) containing the origin can be represented as

$$[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k] = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k \text{ for some } c_i \in \mathbb{R}\} \tag{3.1.1}$$

for some linearly independent $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k \in \mathbb{R}^n$, that is,

$$c_1\boldsymbol{v} + \cdots + c_k\boldsymbol{v}_k = 0 \text{ if and only if } c_1 = \cdots = c_k = 0.$$

DEFINITION 3.1.2. [**Lee13**, Example 1.36] Let $k < n$. The Grassmannian $\mathrm{Gr}(n,k)$ is the collection of all k-dimensional hyperplane in $\mathbb{R}^n$ containing the origin.

We also see that each n dimensional hyperplane in $\mathbb{R}^{n+1}$ can be expressed as

$$\left\{\boldsymbol{x} \in \mathbb{R}^{n+1} : \boldsymbol{x} \cdot \boldsymbol{v} = 0\right\} \quad \text{for some } \boldsymbol{v} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\},$$

which shows that the Grassmannian $\mathrm{Gr}(n+1,n)$ of all k-dimensional hyperplane in $\mathbb{R}^n$ containing the origin is "identical" to $\mathbb{RP}^n = \mathrm{GR}(n+1,1)$. The simplest Grassmannian that is not real projective space is $\mathrm{Gr}(4,2)$. Similar to real projective space, the Grassmannian also have topological/manifold structure as well as algebraic structure. These aspects are currently out of the scope of this course, and we will skip these details.

3.2. Regular surfaces in $\mathbb{R}^3$

In order to guarantee the analytic properties reflect the geometric properties properly, we need some assumptions. We first restrict ourselves for the case when $n = 3$.

DEFINITION 3.2.1 ([**dC76**, Section 2.2, Definition 1]). A subset $S \subset \mathbb{R}^3$ is a *regular surface* if, for each $\boldsymbol{p} \in S$, there exists an open neighborhood $V \subset \mathbb{R}^3$ of such point $\boldsymbol{p}$, and open set $U \subset \mathbb{R}^2$ and a C^∞ bijective mapping $\boldsymbol{x} = (x,y,z) : U \to V \cap S$ such that:

(a) its inverse $\boldsymbol{x}^{-1} : V \cap S \to U$ is continuous in the sense that[1] $\boldsymbol{x}^{-1} : V \cap S \to U$ is the restriction of a continuous map on some neighborhood of $V \cap S$; and

(b) For each $(u,v) \in U$, the matrix

$$\begin{pmatrix} \partial_u \boldsymbol{x}(u,v) & \partial_v \boldsymbol{x}(u,v) \end{pmatrix} \equiv \begin{pmatrix} \partial_u x(u,v) & \partial_v x(u,v) \\ \partial_u y(u,v) & \partial_v y(u,v) \\ \partial_u z(u,v) & \partial_v z(u,v) \end{pmatrix}$$

has linearly independent columns, equivalently, the straight line $[\partial_u \boldsymbol{x}(u,x)]$ is not identical to the straight line $[\partial_v \boldsymbol{x}(u,x)]$ for all $(u,v) \in U$, where

$$\partial_u x(u,v), \partial_u y(u,v), \partial_u z(u,v), \partial_v x(u,v), \partial_v y(u,v), \partial_v z(u,v)$$

are corresponding partial derivatives. This is equivalent to either one of the following Jacobian determinant

$$\begin{aligned} \frac{\partial(x,y)}{\partial(u,v)} &:= \det \begin{pmatrix} \partial_u x(u,v) & \partial_v x(u,v) \\ \partial_u y(u,v) & \partial_v y(u,v) \end{pmatrix} \\ \frac{\partial(x,z)}{\partial(u,v)} &:= \det \begin{pmatrix} \partial_u x(u,v) & \partial_v x(u,v) \\ \partial_u z(u,v) & \partial_v z(u,v) \end{pmatrix} \\ \frac{\partial(y,z)}{\partial(u,v)} &:= \det \begin{pmatrix} \partial_u y(u,v) & \partial_v y(u,v) \\ \partial_u z(u,v) & \partial_v z(u,v) \end{pmatrix} \end{aligned}$$

is nonzero.

The mapping $\boldsymbol{x} : U \to V \cap S$ is called a *local coordinates* near $\boldsymbol{p} \in S$.

REMARK 3.2.2. Let $\boldsymbol{x} : U \to V \cap S$ is called a *local coordinates* near $\boldsymbol{p} \in S$. For each $\boldsymbol{q} = (u,v) \in U$, the *differential* $\mathrm{d}\boldsymbol{x}_{\boldsymbol{q}} : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by (in terms of matrix multiplication)

$$(\mathrm{d}\boldsymbol{x}_{\boldsymbol{q}})[\boldsymbol{\xi}] := \begin{pmatrix} \partial_u x(u,v) & \partial_v x(u,v) \\ \partial_u y(u,v) & \partial_v y(u,v) \\ \partial_u z(u,v) & \partial_v z(u,v) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \quad \text{for all } \boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in \mathbb{R}^2,$$

equivalently,

$$(\mathrm{d}\boldsymbol{x}_{\boldsymbol{q}})(\boldsymbol{e}_1) = \partial_u \boldsymbol{x}(u,v), \quad (\mathrm{d}\boldsymbol{x}_{\boldsymbol{q}})(\boldsymbol{e}_2) = \partial_v \boldsymbol{x}(u,v),$$

[1] See the relative topology (Deinition 1.3.33) and equivalent characterization of continuous functions (Exercise 1.4.13).

where $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We can rephrase Definition 3.2.1(b) as:

(b') For each $\boldsymbol{q} \in U$, the differential $\mathrm{d}\boldsymbol{x}_{\boldsymbol{q}} : \mathbb{R}^2 \to \mathbb{R}^3$ is injective.

The following exercise exhibits another characterization of Definition 3.2.1(b) by using a special structure of the 3-dimensional Euclidean space $\mathbb{R}^3$:

EXERCISE 3.2.3. Show that $\mathrm{d}\boldsymbol{x}_{\boldsymbol{q}} : \mathbb{R}^2 \to \mathbb{R}^3$ is injective if and only if $\partial_u \boldsymbol{x}(u,v) \times \partial_v \boldsymbol{x}(u,v) \neq \mathbf{0}$.

REMARK 3.2.4. This assumption ensures us to define the *tangent plane* at $\boldsymbol{q} \in S$ by

$$T_{\boldsymbol{q}} := \boldsymbol{q} + [\partial_u \boldsymbol{x}(u,v), \partial_v \boldsymbol{x}(u,v)].$$

REMARK 3.2.5. A regular surface (Definition 3.2.1) can be understood as a *2-dimensional manifold which is properly embedded in* $\mathbb{R}^3$. You can imagine while you packing your T-shirt (regular surface) into a luggage (Euclidean space $\mathbb{R}^3$), after you fold your T-shirt, it looks differently, but of course it still a T-shirt. The embedding can be imagined as packing a folded T-shirt into a luggage, without damaging it.

The following lemma exhibits a way to construct regular surfaces.

LEMMA 3.2.6. *Let U be an open set in $\mathbb{R}^2$. If $f : U \to \mathbb{R}$ is a C^∞ function in, then the* graph *of f, which is the set*

$$\mathrm{graph}\,(f) := \{(u,v,f(u,v)) : (u,v) \in U\}$$

is a regular surface.

PROOF. We consider

$$\boldsymbol{x}(u,v) := (u,v,f(u,v)) \quad \text{for all } (u,v) \in U.$$

Note that $\boldsymbol{x} : U \to \mathrm{graph}\,(f)$ is bijective C^∞ function, and note that its inverse

$$\boldsymbol{x}^{-1} = \pi|_{\mathrm{graph}\,(f)} : |_{\mathrm{graph}\,(f)} \to U$$

where $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ is the projection defined by $\pi(u,v,w) := (u,v)$, which is obviously continuous, and thus we verified Definition 3.2.1(a). On the other hand, if we denote $\boldsymbol{x}(u,v) = (x(u,v), y(u,v), z(u,v))$, we also see that

$$\frac{\partial(x,y)}{\partial(u,v)} = 1,$$

which verifies Definition 3.2.1(b). We finally conclude that $\mathrm{graph}\,(f)$ is a regular surface with a single local chart $(\boldsymbol{x}, U)$. $\square$

EXAMPLE 3.2.7. We consider the unit sphere

$$\mathbb{S}^2 = \left\{ (x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + y^2 = 1 \right\}.$$

By using Lemma 3.2.6, one sees that the unit sphere $\mathbb{S}^2$ is a regular surface which can be covered by the local charts $\{\boldsymbol{x}_i, U_i\}_{i=1}^6$ given by

$$\begin{cases} \boldsymbol{x}_1(u,v) = \left(u, v, \sqrt{1-(u^2+v^2)}\right) & ,(u,v) \in U_1 = B_1, \\ \boldsymbol{x}_2(u,v) = \left(u, v, -\sqrt{1-(u^2+v^2)}\right) & ,(u,v) \in U_2 = B_1, \\ \boldsymbol{x}_3(u,v) = \left(u, \sqrt{1-(u^2+v^2)}, v\right) & ,(u,v) \in U_3 = B_1, \\ \boldsymbol{x}_4(u,v) = \left(u, -\sqrt{1-(u^2+v^2)}, v\right) & ,(u,v) \in U_4 = B_1, \\ \boldsymbol{x}_5(u,v) = \left(\sqrt{1-(u^2+v^2)}, u, v\right) & ,(u,v) \in U_5 = B_1, \\ \boldsymbol{x}_6(u,v) = \left(-\sqrt{1-(u^2+v^2)}, u, v\right) & ,(u,v) \in U_6 = B_1. \end{cases}$$

This choice of local chart is not unique, see e.g. the local spherical coordinate given in Exercise 3.2.8. In fact, the unit sphere $\mathbb{S}^2$ can be covered by just only two charts via stereographic projection (Exercise 3.2.9)[2]. However, it is not possible to cover $\mathbb{S}^2$ by just a single chart (Exercise 3.2.10).

EXERCISE 3.2.8. Show that the unit sphere $\mathbb{S}^2$ is a regular surface which can be covered by the local charts $\{\boldsymbol{x}_i, U_i\}_{i=1}^4$ given by the *local spherical coordinate*:

$$\begin{cases} \boldsymbol{x}_1(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) & ,(\theta,\varphi) \in U_1 = (0,\pi)\times(0,2\pi), \\ \boldsymbol{x}_2(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) & ,(\theta,\varphi) \in U_2 = (0,\pi)\times(-\pi,\pi), \\ \boldsymbol{x}_3(\theta,\varphi) = (\sin\theta\cos\varphi, \cos\theta, \sin\theta\sin\varphi) & ,(\theta,\varphi) \in U_3 = (0,\pi)\times(0,2\pi), \\ \boldsymbol{x}_4(\theta,\varphi) = (\sin\theta\cos\varphi, \cos\theta, \sin\theta\sin\varphi) & ,(\theta,\varphi) \in U_4 = (0,\pi)\times(-\pi,\pi). \end{cases}$$

EXERCISE 3.2.9 (Stereographic projection). The stereographic projection $\pi : \mathbb{S}^2 \setminus \{\boldsymbol{e}_3\} \to \mathbb{R}^2$, where $\boldsymbol{e}_3 = (0,0,1)$ is the north pole, carries a point $\boldsymbol{p} \in \mathbb{S}^2 \setminus \{\boldsymbol{e}_3\}$ onto the inversection of the xy plane with the straight line which connects $\boldsymbol{e}_3$ to $\boldsymbol{p}$, that is,

$$\pi(\boldsymbol{p}) := (\boldsymbol{e}_3 + [\boldsymbol{p} - \boldsymbol{e}_3])|_{z=0}.$$

Compute the formula of $\pi^{-1} : \mathbb{R}^2 \to \mathbb{S}^2 \setminus \{\boldsymbol{e}_3\}$ and use this to show that $\pi^{-1}(\mathbb{R}^2) = \mathbb{S}^2 \setminus \{\boldsymbol{e}_3\}$ is a regular surface. From this, one immediately sees that the unit sphere $\mathbb{S}^2$ is a regular surface which can be covered by the local charts $\{\boldsymbol{x}_i, U_i\}_{i=1}^4$ given by

$$\begin{cases} \boldsymbol{x}_1 = \pi^{-1} & ,U_1 = \mathbb{R}^2, \\ \boldsymbol{x}_1 = -\pi^{-1} & ,U_2 = \mathbb{R}^2. \end{cases}$$

EXERCISE 3.2.10. Show that it is not possible to cover $\mathbb{S}^2$ by just a single chart. (**Hint.** Using a contradiction argument)

[2]There are quite a lot of explanations in Wikipedia page (https://en.wikipedia.org/wiki/Stereographic_projection) about this topic.

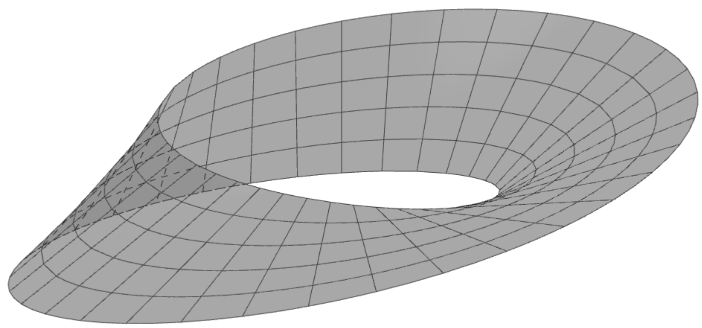


FIGURE 3.2.1. Möbius strip (3.2.1) (https://www.geogebra.org)

EXERCISE 3.2.11 (The Möbius strip [**ConNA**]). Show that the Möbius strip (see Figure 3.2.1), which defined by the local charts $\{\boldsymbol{x}_i, U_i\}_{i=1}^2$ given by

$$\boldsymbol{x}_1(t,\theta) = \boldsymbol{x}_2(t,\theta) = (2\cos 2\theta + t\cos\theta\cos 2\theta, 2\sin 2\theta + t\cos\theta\sin 2\theta, t\sin\theta) \tag{3.2.1}$$

with $U_1 = (-1,1)\times(0,2\pi)$ and $U_2 = (-1,1)\times(-\pi,\pi)$, is a regular surface. It is interested to see that Möbius strip has only one side.

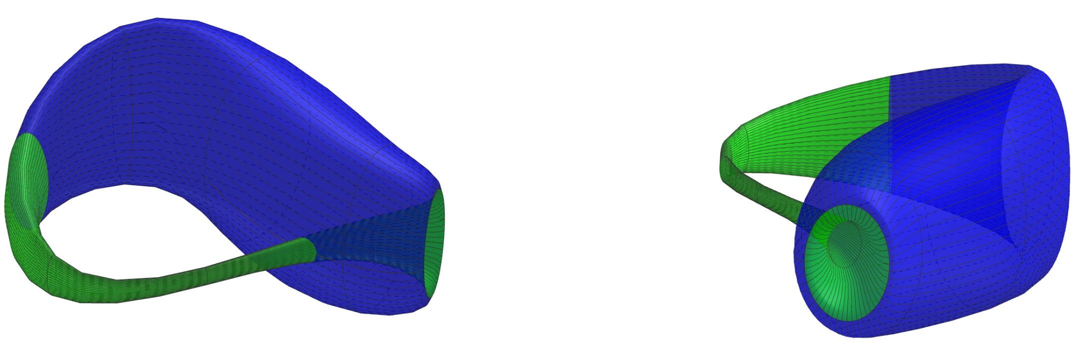


FIGURE 3.2.2. Klein bottle in Example 3.2.12 (https://www.geogebra.org)

EXAMPLE 3.2.12. We now consider the *Klein bottle* (see Figure 3.2.2), which defined by the local charts $\{\boldsymbol{x}_i, U_i\}_{i=1}^4$ given by

$$\boldsymbol{x}_1 = \boldsymbol{x}_2 = \boldsymbol{x}_3 = \boldsymbol{x}_4 = (x,y,z)$$

where (https://en.wikipedia.org/wiki/Klein_bottle#Bottle_shape)

$$x(u,v) = -\frac{2}{15}\cos u\left(3\cos v - 30\sin u + 90\cos^4 u\sin u - 60\cos^6 u\sin u + 5\cos u\cos v\sin u\right)$$

$$\begin{aligned} y(u,v) = -\frac{1}{15}\sin u\Big(&3\cos v - 3\cos^2 u\cos v - 48\cos^4 u\cos v + 48\cos^6 u\cos v - 60\sin u \\ &+5\cos u\cos v\sin u - 5\cos^3 u\cos v\sin u - 80\cos^5 u\cos v\sin u + 80\cos^7 u\cos v\sin u\Big) \end{aligned}$$

$$z(u,v) = \frac{2}{15}\sin v(3+5\cos u\sin u)$$

for all $(u,v) \in U_i$ $(i = 1,2,3,4)$ with $U_1 = (0,\pi)\times(0,2\pi)$, $U_2 = (-\pi/2,\pi/2)\times(0,2\pi)$, $U_3 = (0,\pi)\times(-\pi,\pi)$, $U_4 = (-\pi/2,\pi/2)\times(-\pi,\pi)$. Here we want to point out that Klein bottle is not a regular surface in the sense of Definition 3.2.1 (it is not properly embedded in $\mathbb{R}^3$) since it is self-intersected. It is also interested to see that Klein bottle has only one side.

We see that is is quite inefficient to verify whether the surface is regular by explicitly construct local charts (and in many cases, it cannot be covered by a single chart). Another way to construct surfaces via the level surface of a given C^∞ function $f : U \subset \mathbb{R}^3 \to \mathbb{R}$. Here, the level surface is defined by

$$f^{-1}(a) := \{\boldsymbol{x} \in U : f(\boldsymbol{x}) = a\} \quad \text{for some pre-chosen } a \in \mathbb{R}. \tag{3.2.2}$$

A heuristic dimension count suggests that (3.2.2) produces a surface in many cases. It is obvious that this cannot be done without any additional assumptions, for example, if we consider the trivial function $f(\boldsymbol{x}) \equiv 0$ for all $\boldsymbol{x} \in \mathbb{R}^3$, we see that $f^{-1}(0) = \mathbb{R}^3$ is not a surface.

Before we introduce a sufficient condition so that $f^{-1}(a)$ is a regular surface, we first state the following definition.

DEFINITION 3.2.13. Let U be an open set in $\mathbb{R}^n$ and let $\boldsymbol{F} = (F_1,\cdots,F_m) : U \to \mathbb{R}^m$ be a C^∞ function. Given a point $\boldsymbol{p} \in U$, the *differential* $\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}} : \mathbb{R}^n \to \mathbb{R}^m$ is the linear mapping defined by (see also Remark 3.2.2 above for a special case)

$$(\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}})[\boldsymbol{\xi}] := (\nabla \otimes \boldsymbol{F}(\boldsymbol{p}))\boldsymbol{\xi} \equiv \begin{pmatrix} \partial_1 F_1(\boldsymbol{p}) & \cdots & \partial_n F_1(\boldsymbol{p}) \\ \vdots & \ddots & \vdots \\ \partial_1 F_m(\boldsymbol{p}) & \cdots & \partial_n F_m(\boldsymbol{p}) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

for all $\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{R}^n$. We see that $\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}}$ is represented by the matrix $\nabla \otimes \boldsymbol{F}(\boldsymbol{p})$, and we can slightly abuse the notation by writing

$$\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}} = \begin{pmatrix} \partial_1 F_1(\boldsymbol{p}) & \cdots & \partial_n F_1(\boldsymbol{p}) \\ \vdots & \ddots & \vdots \\ \partial_1 F_m(\boldsymbol{p}) & \cdots & \partial_n F_m(\boldsymbol{p}) \end{pmatrix}. \tag{3.2.3}$$

(a) A point $\boldsymbol{p} \in U$ is called a *regular point* if the differential $\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}} : \mathbb{R}^n \to \mathbb{R}^m$ is a surjective mapping, otherwise, such point is called a *critical point*.
(b) For each regular point (resp. critical point) $\boldsymbol{p} \in U$, we call $\boldsymbol{F}(\boldsymbol{p}) \in \mathbb{R}^m$ the regular value (resp. critical value).

REMARK 3.2.14. In the special case when $m = n$, by using the rank theorem (see e.g. [**Tre17a**, Theorem 7.2]), one sees that the following are equivalent:

- $\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}} : \mathbb{R}^n \to \mathbb{R}^n$ is surjective,
- $\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}} : \mathbb{R}^n \to \mathbb{R}^n$ is injective,
- $\det(\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}}) \neq 0$, where we abuse the notation as in (3.2.3).

In the case, the inverse function theorem (see. e.g. [**Apo74**, Theorem 13.6]) guarantees that $\boldsymbol{F}$ is invertible *near* $\boldsymbol{p}$ with inverse $\boldsymbol{F}^{-1}$ near $\boldsymbol{F}(\boldsymbol{p})$, that is, there exist open neighborhoods $V \ni \boldsymbol{p}$ and $W \ni \boldsymbol{F}(\boldsymbol{p})$ such that $\boldsymbol{F} : V \to W$ is bijective with inverse $\boldsymbol{F}^{-1} \in (C^1(W))^3$ satisfying $\det(\mathrm{d}\boldsymbol{F}^{-1}_{\boldsymbol{F}(\boldsymbol{p})}) \neq 0$.

Now we can give a sufficient condition to guarantee that $f^{-1}(a)$ is a regular surface.

LEMMA 3.2.15. *Let U be an open set in $\mathbb{R}^3$ and let $f : U \to \mathbb{R}$ be a C^∞ function. If $a \in f(U)$ is a regular value of f,then $f^{-1}(a)$ is a regular surface in $\mathbb{R}^3$.*

PROOF. We fix any $\boldsymbol{p} \in f^{-1}(a)$. Since a is a regular value of f, then there exists $i \in \{1,2,3\}$ such that $\partial_i f(\boldsymbol{p}) \neq 0$. Here we remind the readers that such index i depends on $\boldsymbol{p}$ and may not unique. We define the mapping $\boldsymbol{F} : U \subset \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\boldsymbol{F}(x,y,z) := \begin{cases} (f(x,y,z),x,y) & \text{if we choose } i = 1, \\ (x,f(x,y,z),y) & \text{if we choose } i = 2, \\ (x,y,f(x,y,z)) & \text{if we choose } i = 3. \end{cases}$$

Here we demonstrate the case when $i = 3$, as the proof for the cases $i = 1$ and $i = 2$ are similar.

Since

$$\det(\mathrm{d}\boldsymbol{F}_{\boldsymbol{p}}) = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial_1 f(\boldsymbol{p}) & \partial_2 f(\boldsymbol{p}) & \partial_3 f(\boldsymbol{p}) \end{pmatrix} = \partial_3 f(\boldsymbol{p}) \neq 0$$

then we can apply inverse function theorem (as in Remark 3.2.14) to find open neighborhoods $V \ni \boldsymbol{p}$ and $W \ni \boldsymbol{F}(\boldsymbol{p})$ such that $\boldsymbol{F} : V \to W$ is bijective with inverse $\boldsymbol{F}^{-1} \in (C^1(W))^3$ satisfying $\det(\mathrm{d}\boldsymbol{F}^{-1}_{\boldsymbol{F}(\boldsymbol{p})}) \neq 0$. Since $\boldsymbol{F}^{-1}(u,v,t) = (u,v,g(u,v,t))$ for some $g \in C^1(W)$ and

$$\boldsymbol{F}(f^{-1}(a) \cap V) = W \cap \{(u,v,a) : u,v \in \mathbb{R}\},$$

then we conclude our lemma by using Lemma 3.2.6 with local chart $\boldsymbol{x}(u,v) = F^{-1}(u,v,a)$. □

EXAMPLE 3.2.16. We now choose $U = \mathbb{R}^3 \setminus \{(0,0,0)\}$ and $f : U \to \mathbb{R}$ is given by $f(x,y,z) = x^2 + y^2 + z^2$ for all $(x,y,z) \in U$. We see that

$$\begin{aligned} f^{-1}(1) &= \{(x,y,z) \in U : f(x,y,z) = 1\} \\ &= \left\{(x,y,z) \in \mathbb{R}^3 \setminus \{(0,0,0)\} : x^2 + y^2 + z^2 = 1\right\} = \mathbb{S}^2, \end{aligned}$$

and

$$\begin{aligned} \mathrm{d}f_{(x,y,z)} &= (\partial_x f(x,y,z), \partial_y f(x,y,z), \partial_z f(x,y,z)) \\ &= (2x, 2y, 2z) \neq (0,0,0) \quad \text{for all } (x,y,z) \in f^{-1}(1) \subset U, \end{aligned}$$

thus by using Lemma 3.2.15, one again verified that $\mathbb{S}^2$ is a regular surface (see Example 3.2.7 above). If we choose $U = \mathbb{R}^3$, one also sees that the function $f : \mathbb{R}^3 \to \mathbb{R}$ above is well-defined and

$$f^{-1}(1) = \left\{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = 1\right\} = \mathbb{S}^2,$$

but however such choice of U does not guarantee $\mathrm{d}f_{(x,y,z)}$ is nonzero. This also demonstrates that the assumptions Lemma 3.2.15 is not necessarily, but choosing U wisely will be helpful.

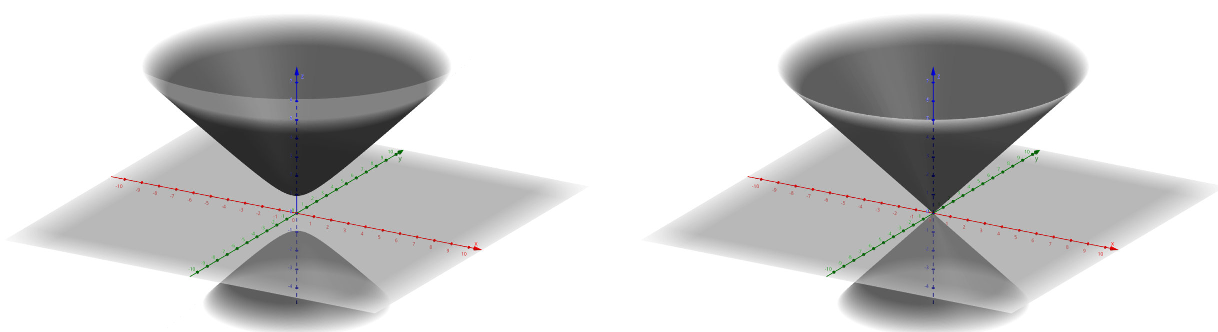


FIGURE 3.2.3. The hyperboloid $-x^2 - y^2 + z^2 = 1$ (left) and the two-sheet cones $-x^2 - y^2 + z^2 = 0$ (right) (https://www.geogebra.org)

EXERCISE 3.2.17. Show that the *hyperboloid* of two sheets $\left\{(x,y,z) \in \mathbb{R}^3 : -x^2 - y^2 + z^2 = 1\right\}$ is a regular surface, but however the *two-sheet cones* $\left\{(x,y,z) \in \mathbb{R}^3 : -x^2 - y^2 + z^2 = 0\right\}$ is *not* a regular surface. (**Hint.** One easy way to show this is by a contradiction argument together with Exercise 1.4.14)

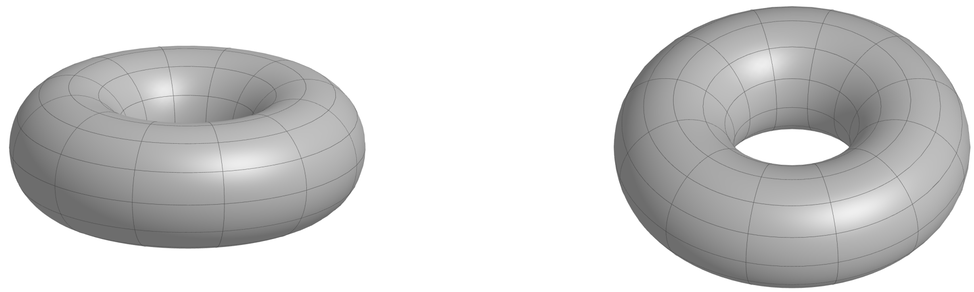


FIGURE 3.2.4. The torus $(\sqrt{x^2+y^2}-1)^2+z^2=\frac{1}{4}$ via parmeterization (https://www.geogebra.org)

EXERCISE 3.2.18. Show that the *torus* $\left\{(x,y,z)\in\mathbb{R}^3:(\sqrt{x^2+y^2}-1)^2+z^2=\frac{1}{4}\right\}$ is a regular surface:

(a) via the parameterization $\boldsymbol{x}(u,v)=((\frac{1}{2}\cos u+1)\cos v,(\frac{1}{2}\cos u+1)\sin v,\frac{1}{2}\sin u)$ for all $(u,v)\in U_i$ $(i=1,2,3,4)$ with $U_1=(0,2\pi)\times(0,2\pi)$, $U_2=(-\pi,\pi)\times(0,2\pi)$, $U_3=(0,2\pi)\times(-\pi,\pi)$, $U_4=(-\pi,\pi)\times(-\pi,\pi)$.

(b) via Lemma 3.2.15.

REMARK 3.2.19. The Geogebra plot (https://www.geogebra.org) works very well for parameterization like Exercise 3.2.18(a), but not so good for the level surface representation like Exercise 3.2.18(b).

We now state some necessary condition for regular surfaces, without giving proofs here.

LEMMA 3.2.20 ([**dC76**, Section 2.2, Proposition 3]). *Let $S\subset\mathbb{R}^3$ be a regular surface and $\mathbf{p}\in S$. Then there exists a neighborhood V of $\mathbf{p}$ in S such that V is the graph of a differentiable function which has one of the following three forms:*

$$z=f(x,y),\quad y=g(x,z),\quad x=h(y,z).$$

LEMMA 3.2.21 ([**dC76**, Section 2.2, Proposition 4]). *Let $S\subset\mathbb{R}^3$ be a regular surface and $\mathbf{p}\in S$. Let $\boldsymbol{x}:U\subset\mathbb{R}^2\to\mathbb{R}^3$ be a C^∞ map with $\mathbf{p}\in\boldsymbol{x}(U)\subset S$ such that $\mathrm{d}\boldsymbol{x}_{\mathbf{p}}:\mathbb{R}^2\to\mathbb{R}^3$ is injective. If $\boldsymbol{x}:U\to\mathbb{R}^3$ is injective, then $\boldsymbol{x}^{-1}:\boldsymbol{x}(U)\to U$ is continuous (i.e. is a restriction of some continuous function defined in an open neighborhood of $\boldsymbol{x}(U)$ in $\mathbb{R}^3$).*

As showed in above, it is easy to find regular surface which cannot be covered by a single chart. The following "change of parameters" lemma ensures that the charts are "properly glued" in an analytic sense.

LEMMA 3.2.22 (Change of parameters, [**dC76**, Section 2.3, Proposition 1]). *Let $S\subset\mathbb{R}^3$ be a regular surface and $\mathbf{p}\in S$. Let $\boldsymbol{x}:U\subset\mathbb{R}^2\to S$ and $\boldsymbol{y}:V\subset\mathbb{R}^2\to S$ be two parameterizations of S*

such that $p \in \boldsymbol{x}(U) \cap \boldsymbol{y}(V) = W$. *Then the* change of coordinate $h = \boldsymbol{x}^{-1} \circ \boldsymbol{y} : \boldsymbol{y}^{-1}(W) \to \boldsymbol{x}^{-1}(W)$ *is a* difffeomorphism *in the sense of both* $h : \boldsymbol{y}^{-1}(W) \to \boldsymbol{x}^{-1}(W)$ *and* $h : \boldsymbol{x}^{-1}(W) \to \boldsymbol{y}^{-1}(W)$ *are* C^∞.

The above lemma also says that a regular surface (Definition 3.2.1) can be understood as a *2-dimensional* C^∞*-manifold which is properly embedded in* $\mathbb{R}^3$. Suggested by Lemma 3.2.22, we also have the following definition.

DEFINITION 3.2.23 ([**dC76**, Section 2.3, Definition 1]). Let $S \subset \mathbb{R}^3$ be a regular surface and let V be a (relative) open subset of S. A function $f : V \to \mathbb{R}$ is said to be *differentiable* (resp. C^k) at $\boldsymbol{p} \in V$ if there exists a parameterization $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S$ with $\boldsymbol{p} \in \boldsymbol{x}(U) \subset V$ such that the composition $f \circ \boldsymbol{x} : U \to \mathbb{R}$ is differentiable (resp. C^k) at $\boldsymbol{x}^{-1}(\boldsymbol{p})$. If f is differentiable at all points in V, then we say that $f : V \to \mathbb{R}$ is said to be *differentiable* (resp. C^k).

REMARK 3.2.24. By using Lemma 3.2.22, the notion in Definition 3.2.23 is actually independent of the choice of parameterization $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S$.

EXERCISE 3.2.25. Let $S \subset \mathbb{R}^3$ be a regular surface and any point $\boldsymbol{v} \in \mathbb{R}^3$. Show that the height function $h : S \to \mathbb{R}$ with respect to the plane $\left\{\boldsymbol{x} \in \mathbb{R}^3 : \boldsymbol{x} \cdot \boldsymbol{v} = 0\right\}$ defined by $h(\boldsymbol{p}) := \boldsymbol{p} \cdot \boldsymbol{v}$ for all $\boldsymbol{p} \in S$ is a C^∞ function.

EXERCISE 3.2.26. Let $S \subset \mathbb{R}^3$ be a regular surface and any point $\boldsymbol{p}_0 \in \mathbb{R}^3$. Show that the squared distance $h : S \to \mathbb{R}$ defined by $h(\boldsymbol{p}) := |\boldsymbol{p} - \boldsymbol{p}_0|^2$ for all $\boldsymbol{p} \in S$ is a C^∞ function.

The definition of differentiability can be easily extended to mappings between surface.

DEFINITION 3.2.27. Let S_1 and S_2 be regular surfaces in $\mathbb{R}^3$ and let V_1 be a (relative) open subset of S_1. A function $\varphi : V_1 \to S_2$ is said to be differentiable (resp. C^k) at $\boldsymbol{p} \in V_1$ if there exist parameterizations

$$\boldsymbol{x}_1 : U_1 \subset \mathbb{R}^2 \to S_1, \quad \boldsymbol{x}_2 : U_2 \subset \mathbb{R}^2 \to S_2 \tag{3.2.4}$$

with $\boldsymbol{p} \in \boldsymbol{x}_1(U)$ and $\varphi(\boldsymbol{x}_1(U_1)) \subset \boldsymbol{x}_2(U_2)$, the composition

$$\boldsymbol{x}_2^{-1} \circ \varphi \circ \boldsymbol{x}_1 : U_1 \to U_2$$

is differentiable (resp. C^k) at $\boldsymbol{q} = \boldsymbol{x}_1^{-1}(\boldsymbol{p})$. If φ is said to be differentiable (resp. C^k) at all point in V_1, then we say that $\varphi : V_1 \to S_2$ is differentiable (resp. C^k) on V_1.

REMARK 3.2.28. By using Lemma 3.2.22, the notion in Definition 3.2.27 is actually independent of the choice of parameterizations (3.2.4).

DEFINITION 3.2.29. Let S_1 and S_2 be regular surfaces in $\mathbb{R}^3$. A C^∞ bijective function $\varphi : S_1 \to S_2$ is called a *diffeomorphism* if its inverse $\varphi^{-1} : S_2 \to S_1$ is also C^∞.

Definition 3.2.27 suggests a way to classify regular surfaces.

EXERCISE 3.2.30. We consider the following relation for regular surfaces S_1 and S_2 in $\mathbb{R}^3$:

$$S_1 \cong S_2 \iff \text{there exists a diffeomorphism } \varphi : S_1 \to S_2.$$

Show that $\cong$ is an equivalence relation, that is:

(a) $S \cong S$ for all regular surface S;
(b) $S_1 \cong S_2 \iff S_2 \cong S_1$ for all regular surfaces S_1 and S_2; and
(c) $S_1 \cong S_2$ and $S_2 \cong S_3$ implies $S_1 \cong S_3$ for all regular surfaces S_1, S_2 and S_3.

In view of Exercise 3.2.30, the following definition is now natural.

DEFINITION 3.2.31. We say that two regular surfaces S_1 and S_2 in $\mathbb{R}^3$ are *diffeomorphic* if $S_1 \cong S_2$.

Roughly speaking, if S_1 and S_2 are diffeomorphic, then one can obtain S_2 by deform S_1, without "damaging" it, see e.g. Figure 3.2.5.

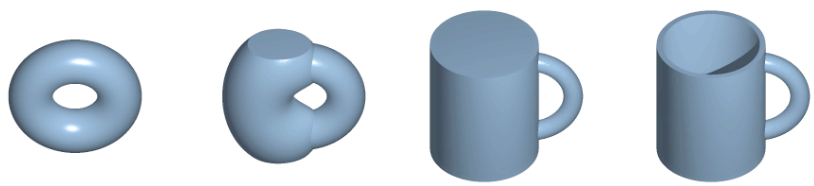


FIGURE 3.2.5. An example of diffeomorphism: obtain a mug by deform a torus without creating "cracks" and "punching a hole": By Lucas Vieira - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=1236079

EXERCISE 3.2.32. Let $A : \mathbb{S}^2 \to \mathbb{S}^2$ be the *antipodal* map $A(\boldsymbol{x}) = -\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{S}^2$. Prove that $A : \mathbb{S}^2 \to \mathbb{S}^2$ is a diffeomorphism.

REMARK 3.2.33. For each $\boldsymbol{x} \in \mathbb{S}^2$, we consider the equivalence class $[\boldsymbol{x}] = \{\boldsymbol{x}, A\boldsymbol{x}\}$ which consists only two points. In fact, the real projective space $\mathbb{RP}^2$ (Section 2.2) can be understood as the set $\left\{[\boldsymbol{x}] : \boldsymbol{x} \in \mathbb{S}^2\right\}$. In proper topological terminology, the real projective space $\mathbb{RP}^2$ is exact the quotient manifold $\mathbb{S}^2/A$, where $A : \mathbb{S}^2 \to \mathbb{S}^2$ is the antipodal map.

EXERCISE 3.2.34. Show that the paraboloid $z = x^2 + y^2$ is diffeomorphic to a plane.

EXERCISE 3.2.35. Construct a diffeomorphism between the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ and the sphere $\mathbb{S}^2$.

3.3. Regular parameterized surfaces in $\mathbb{R}^3$ and tangent plane

We now slightly relax the assumptions on surfaces.

DEFINITION 3.3.1. A *regular parameterized surface* $\boldsymbol{x} : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ is a differentiable map $\boldsymbol{x} : U \to \mathbb{R}^3$ such that $\mathrm{d}\boldsymbol{x}_{\boldsymbol{p}} : \mathbb{R}^2 \to \mathbb{R}^3$ is injective for all $\boldsymbol{p} \in U$. The set $\boldsymbol{x}(U) \subset \mathbb{R}^3$ is called the *trace* of $\boldsymbol{x}$.

By using inverse function theorem (see. e.g. [**Apo74**, Theorem 13.6]), we immediately sees the following lemma.

LEMMA 3.3.2. *Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ is a regular parameterized surface and let $\boldsymbol{q} \in U$. Then there exists a neighborhood V of $\boldsymbol{q}$ in $\mathbb{R}^2$ such that the trace $\boldsymbol{x}(V) \subset \mathbb{R}^3$ is a regular surface.*

COROLLARY 3.3.3. *A regular surface (Definition 3.2.1) must also a regular parameterized surface.*

However, unlike regular surface, we allow the trace of a regular parameterized surface to be self intersected. For example, the Klein bottle (Example 3.2.12) is a regular parameterized surface but not a regular surface. In the following, we exhibit another example of regular parameterized surface which is not a regular surface.

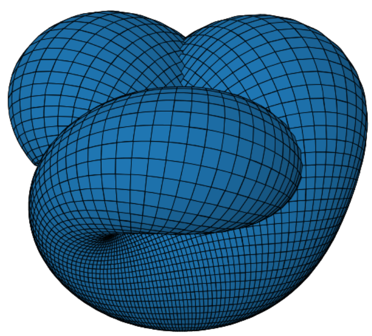


FIGURE 3.3.1. Kusner–Bryant parametrization of the Boy's surface : ggleizer, CC0, via Wikimedia Commons

EXAMPLE 3.3.4. We *formally* write the *imaginary number* $\mathrm{i} := \sqrt{-1}$, one can see e.g. my other lecture note [**Kow23**]. The Boy's surface (discovered by Werner Boy in 1901) can be parameterized by the followings (discovered by Rob Kuzner and Robert Bryant [**Kus87**]), see also the Wikipedia page (`https://en.wikipedia.org/wiki/Boy's_surface#Parametrization`): Write

$$g_1(w) = -\frac{3}{2}\Im\left(\frac{w(1-w^4)}{w^6+\sqrt{5}w^3-1}\right),$$
$$g_2(w) = -\frac{3}{2}\Re\left(\frac{w(1+w^4)}{w^6+\sqrt{5}w^3-1}\right),$$
$$g_3(w) = \Im\left(\frac{1+w^6}{w^6+\sqrt{5}w^3-1}\right) - \frac{1}{2},$$

where $\Re$ and $\Im$ are real part and imaginary part of complex numbers, and set

$$\boldsymbol{x}(u,v) = \frac{(g_1(w), g_2(w), g_3(w))}{|g_1(w)|^2 + |g_2(w)|^2 + |g_3(w)|^2} \quad \text{with } w = u + \mathrm{i}v,$$

for all (u,v) with $u^2+v^2 \leq 1$, see Figure 3.3.1 for a plot. It is important to mention that such Boy's surface is the image of the real projective space $\mathbb{RP}^2$ (Section 2.2) by a C^∞ map (under some suitable sense which is more general than Definition 3.2.27). That is, this parameterization of the Boy's surface is an immersion of the real projective space $\mathbb{RP}^2$ into the Euclidean space $\mathbb{R}^3$.

EXERCISE 3.3.5. Let $\boldsymbol{\alpha} : I \to \mathbb{R}^3$ be a regular parameterized curve (Definition 2.3.8) such that its curvature κ satisfies $\kappa(t) \neq 0$ for all $t \in I$. We consider the open set $U = I \times (\mathbb{R} \setminus \{0\})$ and define

$$\boldsymbol{x}(t,v) := \boldsymbol{\alpha}(t) + v\boldsymbol{\alpha}'(t) \quad \text{for all } (t,v) \in U.$$

Show that $\boldsymbol{x}(U)$ is a regular parameterized curve. (**Hint.** Use Exercise (2.5.9) and Exercise 3.2.3)

A regular parameterized surface can be understood as a 2-dimensional C^∞-manifold which is immersed in $\mathbb{R}^3$, in contrast to a regular surface can be understood as a 2-dimensional C^∞-manifold which is properly embedded in $\mathbb{R}^3$. Lets us recall the T-shirt mentioned in Remark 3.2.5 above: In contrast to embedding, immersion means that you damaged "a bit" the T-shirt (e.g. make some "holes" or "torn"), but it remains untied ("no knot") so that "locally" still cloth.

Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S \subset \mathbb{R}^3$ be a regular parameterized surface (Definition 3.3.1) and let $\boldsymbol{q} \in U$. Similar to Definition 2.3.7, the tangent plane to S at $\boldsymbol{x}(q)$ is given by

$$[\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q})] + \boldsymbol{\alpha}(t) \equiv \{\boldsymbol{x} + \boldsymbol{\alpha}(t) \in \mathbb{R}^n : \boldsymbol{x} \in [\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q})]\}, \tag{3.3.1}$$

where $[\boldsymbol{v}_1, \boldsymbol{v}_2]$ is the plane defined in (3.1.1). It is helpful to see that such tangent plane can be uniquely determined by the differential $\mathrm{d}\boldsymbol{x}_{\boldsymbol{q}}$. In mathematics analysis, it is rather convenient to work with linear space (plane passing origin) rather than affine space (3.3.1), therefore we consider the following definition of shifted tangent plane, called the tangent space:

DEFINITION 3.3.6. Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S \subset \mathbb{R}^3$ be a regular parameterized surface (Definition 3.3.1) and let $\boldsymbol{q} \in U$. The *tangent space* $T_{\boldsymbol{p}}(S)$ to S at $\boldsymbol{x}(q)$ is given by

$$T_{\boldsymbol{p}}(S) := [\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q})] \quad \text{with} \quad \boldsymbol{p} = \boldsymbol{x}(\boldsymbol{q}) \in S.$$

We refer an element in $T_{\boldsymbol{p}}(S)$ a *tangent vector*. A (tangent) *vector field* on S is an assignment of a tangent vector to each point in S.

We now state a result proved by Henri Poincaré in 1885, which says that "it is not possible to comb a hairy ball flat without creasing a cowlick" or "you cannot comb the hair on a coconut":

THEOREM 3.3.7 (Hairy ball theorem). *There is no non-vanishing continuous vector field on* $\mathbb{S}^2$.

It is important to see that $T_{\boldsymbol{p}}(S) = \mathrm{d}\boldsymbol{x}_{\boldsymbol{p}}(\mathbb{R}^2)$, where $\mathrm{d}\boldsymbol{x}_{\boldsymbol{p}}$ is the differential given in Definition 3.2.2, which shows that the tangent plane is actually independent of the choice of parameterization. In view of Exercise 3.2.3, by fixing a parameterization $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S$ at $\boldsymbol{p} \in S$, we can make a definite choice of a unit normal vector at each point $\boldsymbol{q} \in \boldsymbol{x}(U)$ by the rule

$$\boldsymbol{N}(\boldsymbol{q}) = \frac{\partial_u \boldsymbol{x}(\boldsymbol{q}) \times \partial_v \boldsymbol{x}(\boldsymbol{q})}{|\partial_u \boldsymbol{x}(\boldsymbol{q}) \times \partial_v \boldsymbol{x}(\boldsymbol{q})|}. \tag{3.3.2}$$

DEFINITION 3.3.8 ([**dC76**, Section 2.6, Proposition 1]). Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S \subset \mathbb{R}^3$ be a regular parameterized surface (Definition 3.3.1). If there exists a C^∞ field of unit normal vectors $\boldsymbol{N} : S \to \mathbb{S}^2$, then we say that such surface is *orientable*.

REMARK 3.3.9 ([**dC76**, Section 3.2, Definition 1]). Roughly speaking, this means that the surface has two sides. The mapping $N : S \to \mathbb{S}^2$ is also called a Gauss map. In this case, for each $\boldsymbol{p} \in S$, one can always find a suitable local parameterization $(\boldsymbol{x}, U)$ such that $\boldsymbol{N}$ takes the form (3.3.2).

EXAMPLE 3.3.10. The Möbius strip (Example 3.2.11) and Klein bottle (Example 3.2.12) are examples of regular parameterized surface which are not oriented. You can search a meme called “a responsible adult says no to non-orientable shapes” using Google (I am not sure about the copyright, therefore it is safer to not directly include the picture here).

EXAMPLE 3.3.11. We consider a portion of a sphere

$$\boldsymbol{x}(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \quad \text{for all } (\theta,\varphi) \in U.$$

Since $|\boldsymbol{x}(\theta,\varphi)|^2 = 1$ for all $(\theta,\varphi) \in U$, then

$$\begin{aligned} 0 &= \frac{1}{2}\partial_\theta(|\boldsymbol{x}(\theta,\varphi)|^2) = \frac{1}{2}\partial_\theta(\boldsymbol{x}(\theta,\varphi)\cdot\boldsymbol{x}(\theta,\varphi)) = \partial_\theta\boldsymbol{x}(\theta,\varphi)\cdot\boldsymbol{x}(\theta,\varphi), \\ 0 &= \frac{1}{2}\partial_\varphi(|\boldsymbol{x}(\theta,\varphi)|^2) = \frac{1}{2}\partial_\varphi(\boldsymbol{x}(\theta,\varphi)\cdot\boldsymbol{x}(\theta,\varphi)) = \partial_\varphi\boldsymbol{x}(\theta,\varphi)\cdot\boldsymbol{x}(\theta,\varphi), \end{aligned}$$

for all $(\theta,\varphi) \in U$. One also computes that $\boldsymbol{N}(\theta,\varphi) = \boldsymbol{x}(\theta,\varphi)$ for all $(\theta,\varphi) \in U$.

3.4. An application to differential equations

We now consider the quasilinear equation of the form

$$a(x,y,u)\partial_x u + b(x,y,u)\partial_y u = c(x,y,u). \tag{3.4.1}$$

Here we follow the approach in [**Joh78**, Sections 1.4–1.6]. We write (3.4.1) as

$$(a,b,c)\cdot(\partial_x u, \partial_y u, -1) = 0.$$

We represent the function u by a surface $z = u(x,y)$ in $\mathbb{R}^3$, and we write

$$(a,b,c)\cdot\left(\frac{\mathrm{d}z}{\mathrm{d}x}, \frac{\mathrm{d}z}{\mathrm{d}y}, -1\right) = 0.$$

Note that $\left(\frac{\mathrm{d}z}{\mathrm{d}x}, \frac{\mathrm{d}z}{\mathrm{d}y}, -1\right)$ is the normal vector of the surface, thus (a,b,c) is a tangent vector. We now consider a "regular" curve $(x(t),y(t),z(t))$ in that surface, and now we see that $(x'(t),y'(t),z'(t))$ is a tangent vector at the point $(x(t),y(t),z(t))$. This suggests us to consider the characteristic ODE:

$$\begin{cases} x'(t) = a(x(t),y(t),z(t)), \\ y'(t) = b(x(t),y(t),z(t)), \\ z'(t) = c(x(t),y(t),z(t)), \end{cases} \tag{3.4.2}$$

which is a special case of the general standard form of ODE

$$\boldsymbol{u}'(t) = \boldsymbol{f}(t,\boldsymbol{u}), \quad \boldsymbol{u}(t_0) = \boldsymbol{u}_0. \tag{3.4.3}$$

Here, the system is even autonomous, i.e. the coefficients are independent of variable t does not appear explicitly. If we assume that $a,b,c \in C^1$, then one can apply the following fundamental existence theorem to ensure the existence of characteristic curve $(x(t),y(t),z(t))$ which is C^1:

THEOREM 3.4.1 ([**HS99**, Theorem I-2-5]). *Let $a > 0$ and $b > 0$. If $\boldsymbol{f} = \boldsymbol{f}(t,\boldsymbol{y})$ is a (real-valued) continuous function on a closed cylinder*

$$\mathscr{R} = \{(t,\boldsymbol{y}) \in \mathbb{R} \times \mathbb{R}^n : |t-t_0| \le a, |\boldsymbol{y} - \boldsymbol{u}_0| \le b\}$$

such that

$$M := \max_{(t,\boldsymbol{y})\in\mathscr{R}} |\boldsymbol{f}(t,\boldsymbol{y})| > 0,$$

then there exists a function $\boldsymbol{u} \in \left(C^1((t_0-\alpha,t_0+\alpha))\right)^n$ with $\alpha = \min\{a,\frac{b}{M}\}$ satisfying (3.4.3) in $(t_0-\alpha,t_0+\alpha)$.

However, the uniqueness does not hold true in general without further assumption on $\boldsymbol{f}$. We demonstrate this in the following few examples.

EXAMPLE 3.4.2. We define the function

$$u(t) := \begin{cases} 0 & ,t \le 3, \\ \left(\frac{2}{5}(t^2-9)\right)^{5/4} & ,t > 3. \end{cases}$$

By using left and right limits, it is not difficult to check that $u \in C(\mathbb{R})$. By using elementary calculus, one computes that

$$\begin{aligned} u'(t) &= \left(\frac{2}{5}\right)^{1/4} t(t^2-9)^{1/4} \quad \text{for } t > 3, \\ u'(t) &= 0 \quad \text{for } t < 3. \end{aligned}$$

Since

$$\begin{aligned}\lim_{h\to 0_+}\frac{u(3+h)-u(3)}{h} &= \lim_{h\to 0_+}\frac{1}{h}\left(\frac{2}{5}((3+h)^2-9)\right)^{5/4}\\ &= \lim_{h\to 0_+}\frac{1}{h}\left(\frac{2}{5}h(h+6)\right)^{5/4} = \lim_{h\to 0_+} h^{1/5}\left(\frac{2}{5}(h+6)\right)^{5/4} = 0\end{aligned}$$

and

$$\lim_{h\to 0_-}\frac{u(3+h)-u(3)}{h} = \lim_{h\to 0_-}\frac{0}{h} = 0,$$

then

$$u'(3) := \lim_{h\to 0}\frac{u(3+h)-u(3)}{h} = 0.$$

Now we also see that

$$\begin{aligned}\lim_{t\to 3_+} u'(t) &= \lim_{t\to 3_+}\left(\frac{2}{5}\right)^{1/4} t(t^2-9)^{1/4} = 0,\\ \lim_{t\to 3_-} u'(t) &= \lim_{t\to 3_+} 0 = 0,\end{aligned}$$

which concludes that $u' \in C(\mathbb{R})$, and thus $u \in C^1(\mathbb{R})$. One can easily check that

$$\begin{cases} u'(t) = f(t,u(t)) \text{ for all } t\in\mathbb{R}, \quad u(t_0) = 0\\ \text{with } f(t,y) = ty^{1/5} \text{ and } t_0 = 3.\end{cases} \tag{3.4.4}$$

Note that f is continuous in $\mathbb{R}\times\mathbb{R}$, and hence the assumptions in Theorem 3.4.1 satisfy. Since $u\equiv 0$ is also another solution of (3.4.4), one sees that the solution of initial value problem (3.4.4) is *not unique*.

EXERCISE 3.4.3 ([**HS99**, Example III-1-1]). Verify that the initial-value problem

$$u'(t) = (u(t))^{1/3} \text{ for all } t\in\mathbb{R}, \quad u(t_0) = 0$$

has at least two nontivial $C^1(\mathbb{R})$-solutions:

$$u(t) = \begin{cases} 0 & ,t\le t_0,\\ \left(\frac{2}{3}(t-t_0)\right)^{3/2} & ,t > t_0,\end{cases}$$

and

$$u(t) = \begin{cases} 0 & ,t\le t_0,\\ -\left(\frac{2}{3}(t-t_0)\right)^{3/2} & ,t > t_0.\end{cases}$$

EXERCISE 3.4.4 ([**HS99**, Example III-1-3]). Verify that the initial-value problem

$$u'(t) = \sqrt{|u(t)|} \text{ for all } t\in\mathbb{R}, \quad u(t_0) = 0$$

has at least one nontivial $C^1(\mathbb{R})$-solution:

$$u(t) = \begin{cases} -\frac{1}{4}(t-t_0)^2 & ,t \leq t_0, \\ \frac{1}{4}(t-t_0)^2 & ,t > t_0. \end{cases}$$

We now state a sufficient condition to guarantee also the uniqueness of the solution.

THEOREM 3.4.5 (Fundamental theorem of ODE [**HS99**, Theorem I-1-4]). *Suppose that all assumptions in Theorem 3.4.1 hold. If we additionally assume that*

$$|\boldsymbol{f}(t,\boldsymbol{y}_1) - \boldsymbol{f}(t,\boldsymbol{y}_2)| \leq L|\boldsymbol{y}_1 - \boldsymbol{y}_2| \tag{3.4.5}$$

whenever $(t,\boldsymbol{y}_1)$ and $(t,\boldsymbol{y}_2)$ are in $\mathscr{R}$, then the solution described in Theorem 3.4.1 is the unique $\left(C^1((t_0-\alpha,t_0+\alpha))\right)^n$ solution.

EXERCISE 3.4.6. Verify that the ODEs in Example 3.4.2, Exercise 3.4.3 and Exercise 3.4.4 do not satisfy the Lipschitz condition (3.4.5).

We now prove that the above choice of the characteristic ODE really describes the surface $z = u(x,y)$.

LEMMA 3.4.7. *Assume that $a,b,c \in C^1$ near $(x_0,y_0,z_0) \in S$, where S is the surface described by $z = u(x,y)$. If γ is a C^1 curve described by $(x(t),y(t),z(t))$ with $(x(t_0),y(t_0),z(t_0)) = (x_0,y_0,z_0)$, then γ lies completely on S.*

PROOF. For convenience, we write $U(t) := z(t) - u(x(t),y(t))$ so that $U(t_0) = 0$ since $(x_0,y_0,z_0) \in S$. Using chain rule and from (3.4.2) one sees that

$$\begin{aligned} U'(t) &= z'(t) - (\partial_x u)x'(t) - (\partial_y u)y'(t) \\ &= c(x,y,z) - \partial_x u(x,y)a(x,y,z) - \partial_y u(x,y)b(x,y,z) \\ &= c(x,y,U-u(x,y)) - \partial_x u(x,y)a(x,y,U-u(x,y)) \\ &\quad - \partial_y u(x,y)b(x,y,U-u(x,y)). \end{aligned} \tag{3.4.6}$$

From (3.4.1), we see that $U \equiv 0$ is a solution of the ODE (3.4.6). By using the fundamental theorem of ODE (Theorem 3.4.5), we see that $U \equiv 0$ is the unique solution of the ODE (3.4.6), which concludes our lemma. $\square$

We now want to solve the Cauchy problem for (3.4.1) with the Cauchy data

$$h(s) = u(f(s),g(s)) \quad \text{for some } f,g,h \in C^1 \text{ near } s_0. \tag{3.4.7}$$

Note that the initial value problem we previous considered is simply the special case when $f(s) \equiv x_0$ and $g(s) = s$. Now the characteristic ODE (3.4.2) (with suitable parameterization) reads

$$
\begin{cases}
\partial_t X(s,t) = a(X(s,t), Y(s,t), Z(s,t)), \\
\partial_t Y(s,t) = b(X(s,t), Y(s,t), Z(s,t)), \\
\partial_t Z(s,t) = c(X(s,t), Y(s,t), Z(s,t)), \\
\quad \text{with initial conditions} \\
X(s,0) = f(s), \quad Y(s,0) = g(s), \quad Z(s,0) = h(s).
\end{cases}
\tag{3.4.8}
$$

If $a,b,c \in C^1$ near $(f(s_0), g(s_0), h(s_0))$, thus the fundamental theorem of ODE (Theorem 3.4.5) guarantees that there exists a unique solution of (3.4.8):

$$(X(s,t), Y(s,t), Z(s,t))$$

which is C^1 for (s,t) near $(s_0, 0)$. If

$$\text{the matrix} \begin{pmatrix} f'(s_0) & g'(s_0) \\ a(x_0,y_0,z_0) & b(x_0,y_0,z_0) \end{pmatrix} = \begin{pmatrix} \partial_s X(s_0,0) & \partial_s Y(s_0,0) \\ \partial_t X(s_0,0) & \partial_t Y(s_0,0) \end{pmatrix} \text{ is invertible,}$$

then we can use implicit function theorem [**Apo74**, Theorem 13.7] to guarantee that there exists a unique solution $(s,t) = (S(x,y), T(x,y))$ of

$$x = X(S(x,y), T(x,y)), \quad y = Y(S(x,y), T(x,y))$$

of class C^1 is a neighborhood of (x_0, y_0) and satisfying

$$S(x_0, y_0) = s_0, \quad T(x_0, y_0) = 0,$$

so that we finally we conclude that the local solution of the Cauchy problem for (3.4.1) with the Cauchy data (3.4.7) is given by

$$u(x,y) = Z(S(x,y), T(x,y)).$$

The above arguments can be readily extend for higher dimensional case:

THEOREM 3.4.8. *We now consider the Cauchy problem*

$$\sum_{i=1}^{n} a_i(x_1, \cdots, x_n, u) u_{x_i} = c(x_1, \cdots, x_n, u)$$

with Cauchy data

$$h(s_1, \cdots, s_{n-1}) = u(f_1(s_1, \cdots, s_{n-1}), \cdots, f_n(s_1, \cdots, s_{n-1}))$$

for some $f, \cdots, f_n, h \in C^1$ *near* $(s_1^0, \cdots, s_{n-1}^0)$. *If* $a_1, \cdots, a_n, c \in C^1$ *near* $(f_1(s_0), \cdots, f_n(s_0), h(s_0))$ *such that*

$$\text{the matrix} \begin{pmatrix} \partial_{s_1} f_1(s_1^0, \cdots, s_{n-1}^0) & \cdots & \partial_{s_1} f_n(s_1^0, \cdots, s_{n-1}^0) \\ \vdots & & \vdots \\ \partial_{s_n} f_1(s_1^0, \cdots, s_{n-1}^0) & \cdots & \partial_{s_n} f_n(s_1^0, \cdots, s_{n-1}^0) \\ a_1(x_1^0, \cdots, x_n^0, z^0) & \cdots & a_n(x_1^0, \cdots, x_n^0, z^0) \end{pmatrix} \text{is invertible,}$$

where $x_i^0 = f_i(s_1^0, \cdots, s_{n-1}^0)$ *for all* $i = 1, \cdots, n$ *and* $z^0 = h(s_1^0, \cdots, s_{n-1}^0)$, *then there exists a unique* C^1 *solution* $u = u(x_1, \cdots, x_n)$ *near* $(x_1^0, \cdots, x_n^0, z^0)$.

REMARK 3.4.9. The corresponding characteristic ODE is

$$\begin{aligned} \partial_t x_i(s_1, \cdots, s_{n-1}, t) &= a_i(x_1, \cdots, x_n, z) \quad \text{for } i = 1, \cdots, n, \\ \partial_t z(s_1, \cdots, s_{n-1}, t) &= c(x_1, \cdots, x_n, z), \end{aligned}$$

with initial condition

$$\begin{aligned} x_i(s_1, \cdots, s_{n-1}, 0) &= f_i(s_1, \cdots, s_{n-1}) \quad \text{for } i = 1, \cdots, n, \\ z(s_1, \cdots, s_{n-1}, 0) &= h(x_1, \cdots, x_n, z). \end{aligned}$$

EXAMPLE 3.4.10. We now want to solve the initial value problem

$$u \partial_x u + \partial_y u = 0, \quad u(x, 0) = h(x).$$

The characteristic ODE is

$$\partial_t x(s,t) = z, \quad \partial_t y(s,t) = 1, \quad \partial_t z(s,t) = 0$$

with initial condition

$$x(s,0) = s, \quad y(s,0) = 0, \quad z(s,0) = h(s).$$

Solving the ODE yields

$$x = s + zt, \quad y = t, \quad z = h(s).$$

Eliminating s, t yields the implicit equation

$$u(x,y) = h(x - u(x,y)y).$$

It is interesting to see that

$$\partial_x u = h'(x - u(x,y)y)(1 - y\partial_x u),$$

then

$$\partial_x u + yh'(x - u(x,y)y)\partial_x u = h'(x - u(x,y)y),$$

and this implies

$$\partial_x u(x,y) = \frac{h'(x - u(x,y)y)}{1 + yh'(x - u(x,y)y)}.$$

We see, for example when $h(z) = -z$, that

$$\partial_x u(x,y) = \frac{-1}{1-y}.$$

and this quantity will blow up at $y = 1$, which means that there cannot exist a strict solution u of class C^1 beyond $y = 1$. This type of behavior is typical for a nonlinear partial differential equation. In general, we need to consider "weak" solution to study the PDE, but we will not going to go too far beyond this point.

EXERCISE 3.4.11. Solve the initial value problem

$$xu\partial_x u - \partial_y u = 0, \quad u(x,0) = x.$$

For the general first order equation, we refer [**Joh78**, Sections 1.7] for details. Here we will not going to discuss here.

3.5. Some preliminaries on analysis

Before further discussions, now lets recall some facts on computations of integrals as in [**Kow24a**]. For each measurable set E in $\mathbb{R}^n$, we consider the set

$$L^1(E) := \left\{ f \text{ is a measurable function on } E : \int_E |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < +\infty \right\}.$$

THEOREM 3.5.1 (Fubini). *Let $E = E_1 \times E_2$ where E_1 is a measurable set in $\mathbb{R}^n$ and E_2 is a measurable set in $\mathbb{R}^m$. Let $f : E \to [-\infty, +\infty]$ be a measurable function. If either f is non-negative or $f \in L^1(E)$, then*

$$\int_E f(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}(\boldsymbol{x},\boldsymbol{y}) = \int_{E_1} \left(\int_{E_2} f(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \right) \mathrm{d}\boldsymbol{x} = \int_{E_2} \left(\int_{E_1} f(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \right) \mathrm{d}\boldsymbol{y}.$$

REMARK 3.5.2. If all assumptions in Theorem 3.5.1 hold, it is quite often to denote

$$\int_E f(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}(\boldsymbol{x},\boldsymbol{y}) = \int_E f(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y} = \int_E f(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}\boldsymbol{x}.$$

We recall the following change of variable formula (which is even valid for arbitrary dimension Euclidean spaces)

THEOREM 3.5.3 ([**Cha06**, (III.3.1)]). *Let D and Ω are both bounded domains in $\mathbb{R}^n$ with piecewise C^1-boundaries, and let $\varphi : D \to \Omega$ is a C^1-diffeomorphism. Then one has*

$$\int_\Omega f(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \int_D f(\varphi(\boldsymbol{x})) \left| \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} \right| \mathrm{d}\boldsymbol{x} \quad \textit{for all } f \in L^1(\Omega), \tag{3.5.1}$$

where $L^1(\Omega) = \{f : \Omega \to \mathbb{R} : \int_\Omega |f| < +\infty\}$ and the integral is understood in Lebesgue sense.

One may ask why there is an absolute value in the Jacobian. We now consider $n = 1$ and consider the C^1-diffeomorphism $\varphi : D = (c,d) \to \Omega = (a,b)$. In this case, either φ is strictly increasing or φ is strictly decreasing.

- If φ is strictly increasing, then $\varphi^{-1}(a) = c$, $\varphi^{-1}(b) = d$ and the Jacobian is $|\varphi'(x)| = \varphi'(x)$, and the 1-dimensional change of variable formula suggests that
$$\int_a^b f(y)\,\mathrm{d}y = \int_c^d f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(x))\varphi'(x)\,\mathrm{d}x.$$
- If φ is strictly decreasing, then $\varphi^{-1}(a) = d$, $\varphi^{-1}(b) = c$ and the Jacobian is $|\varphi'(x)| = -\varphi'(x)$, and the 1-dimensional change of variable formula suggests that
$$\int_a^b f(y)\,\mathrm{d}y = -\int_c^d f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(x))\varphi'(x)\,\mathrm{d}x.$$

Combining the above two cases, we reach
$$\int_a^b f(y)\,\mathrm{d}y = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(x))\varphi'(x)\,\mathrm{d}x,$$
which is consistent with the 1-dimensional change of variable formula. These computations suggests that the sign of $\det(\nabla \otimes \boldsymbol{\varphi})(\boldsymbol{x})$ is already included in the "orientation" of $\boldsymbol{\varphi}$, and here we remind the readers that *remember to put absolute value in the definition of the Jacobian and remember to check the change of variable is C^1-diffeomorphism.*

EXAMPLE 3.5.4 (Polar coordinate). We now consider the C^1-diffeomorphism $(x,y) : (0,R) \times (0,2\pi) \to B_R(\mathbf{0}) \setminus (\mathbb{R}_{\geq 0} \times \{0\})$ given by
$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta \quad \text{for all } 0 < r < R \text{ and } 0 < \theta < 2\pi.$$

We compute the Jacobian
$$\frac{\partial(x,y)}{\partial(r,\theta)} = \left|\det\begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix}\right| = \left|\det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}\right| = |r(\cos\theta)^2 + r(\sin\theta)^2| = r.$$

By using the change of variable formula with $D = (0,R) \times (0,2\pi)$ and $\Omega = B_R(\mathbf{0}) \setminus (\mathbb{R}_{\geq 0} \times \{0\})$, we see that
$$\int_{B_R(\mathbf{0})\setminus(\mathbb{R}_{\geq 0}\times\{0\})} f(x,y)\,\mathrm{d}(x,y) = \int_{(0,R)\times(0,2\pi)} f(r\cos\theta, r\sin\theta) r\,\mathrm{d}(r,\theta)$$
for all $f \in L^1(B_R(\mathbf{0}) \setminus (\mathbb{R}_{\geq 0} \times \{0\}))$. Since $\mathbb{R}_{\geq 0} \times \{0\}$ has measure zero in $\mathbb{R}^2$, thus
$$L^1(B_R(\mathbf{0})) = L^1(B_R(\mathbf{0}) \setminus (\mathbb{R}_{\geq 0} \times \{0\}))$$
and
$$\int_{B_R(\mathbf{0})\setminus(\mathbb{R}_{\geq 0}\times\{0\})} f(x,y)\,\mathrm{d}(x,y) = \int_{B_R(\mathbf{0})} f(x,y)\,\mathrm{d}(x,y).$$

Since $f \in L^1(B_R(\mathbf{0}))$, by using the Fubini's theorem for L^1 functions (Theorem 3.5.1) we see that

$$\begin{aligned}\int_{(0,R)\times(0,2\pi)} f(r\cos\theta, r\sin\theta) r \,\mathrm{d}(r,\theta) &= \int_0^{2\pi} \left(\int_0^R f(r\cos\theta, r\sin\theta) r\,\mathrm{d}r \right) \mathrm{d}\theta \\ &= \int_0^R \left(\int_0^{2\pi} f(r\cos\theta, r\sin\theta)\,\mathrm{d}\theta \right) r\,\mathrm{d}r.\end{aligned}$$

We summarize the above as:

$$\begin{aligned}\int_{B_R(\mathbf{0})} f(x,y)\,\mathrm{d}(x,y) &= \int_0^{2\pi} \left(\int_0^R f(r\cos\theta, r\sin\theta) r\,\mathrm{d}r \right) \mathrm{d}\theta \\ &= \int_0^R \left(\int_0^{2\pi} f(r\cos\theta, r\sin\theta)\,\mathrm{d}\theta \right) r\,\mathrm{d}r \quad \text{for all } f \in L^1(B_R(\mathbf{0})).\end{aligned}$$

EXAMPLE 3.5.5 (Polar coordinate in the whole Euclidean space). Unlike Riemann integral, one can directly operate unbounded domains for Lebesgue integral. By consider the C^1-diffeomorphism $(x,y) : (0,+\infty)\times(0,2\pi) \to \mathbb{R}^2 \setminus (\mathbb{R}_{\geq 0}\times\{0\})$, the above procedure gives

$$\begin{aligned}\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}(x,y) &= \int_0^{2\pi} \left(\int_0^{+\infty} f(r\cos\theta, r\sin\theta) r\,\mathrm{d}r \right) \mathrm{d}\theta \\ &= \int_0^{+\infty} \left(\int_0^{2\pi} f(r\cos\theta, r\sin\theta)\,\mathrm{d}\theta \right) r\,\mathrm{d}r \quad \text{for all } f \in L^1(\mathbb{R}^2).\end{aligned}$$

We further remark that if f is continuous in $\mathbb{R}^2 \setminus \{0\}$, the integral $\int_0^{+\infty}$ is identical to the improper integral.

3.6. The first fundamental form and the concept of surface area

Let $T_{\boldsymbol{p}}(S)$ be the tangent space of a regular parameterized surface S at $\boldsymbol{p} \in S$ in the sense of Definition 3.3.6. We remind the readers that not to be confused with tangent plane (3.3.1), as the tangent space $T_{\boldsymbol{p}}(S)$ always contains the origin, and its forms a 2-dimensional vector space. By viewing $T_{\boldsymbol{p}}(S)$ as a subspace of $\mathbb{R}^3$, it is natural to define inner product on $T_{\boldsymbol{p}}(S)$ from the inner product of $\mathbb{R}^3$:

$$\langle \boldsymbol{w}_1, \boldsymbol{w}_2 \rangle_{T_{\boldsymbol{p}}(S)} := \boldsymbol{w}_1 \cdot \boldsymbol{w}_2 \quad \text{for all } \boldsymbol{w}_1, \boldsymbol{w}_2 \in T_{\boldsymbol{p}}(S). \tag{3.6.1}$$

The following definition is redundant since S is already immersed in $\mathbb{R}^3$, however here we still introduce it for future generalizations.

DEFINITION 3.6.1. The quadratic form $Q_{T_{\boldsymbol{p}}(s)}$ on $T_{\boldsymbol{p}}(S)$ given by

$$Q_{T_{\boldsymbol{p}}(S)}(\boldsymbol{w}) := \langle \boldsymbol{w}, \boldsymbol{w} \rangle_{T_{\boldsymbol{p}}(S)} = |\boldsymbol{w}|^2 \quad \text{for all } \boldsymbol{w} \in T_{\boldsymbol{p}}(S)$$

is called the *first fundamental form* of the regular parameterized surface S at $\boldsymbol{p} \in S$.

Let $\boldsymbol{p} \in S$ and let $\boldsymbol{w} \in T_{\boldsymbol{p}}(S) \setminus \{\mathbf{0}\}$. In order to motivate us to introduce some geometric notions, we assume that that there exists a regular C^∞-curve $\boldsymbol{\alpha}(t) = \boldsymbol{x}(u(t), v(t))$ in S such that $\boldsymbol{\alpha}(0) = \boldsymbol{p} =$

$\boldsymbol{x}(\boldsymbol{q})$ and $\boldsymbol{\alpha}'(0) = \boldsymbol{w}$. In fact one always can find such a curve, see Theorem 3.8.16 below. We see that

$$\begin{aligned} Q_{\boldsymbol{p}}(\boldsymbol{w}) &= \langle \boldsymbol{\alpha}'(0), \boldsymbol{\alpha}'(0) \rangle_{T_{\boldsymbol{p}}(S)} \\ &= \langle \partial_u \boldsymbol{x}(\boldsymbol{p}) u'(0) + \partial_v \boldsymbol{x}(\boldsymbol{p}) v'(0), \partial_u \boldsymbol{x}(\boldsymbol{p}) u'(0) + \partial_v \boldsymbol{x}(\boldsymbol{p}) v'(0) \rangle_{T_{\boldsymbol{p}}(S)} \\ &= \mathsf{E}(\boldsymbol{q}) |u'(0)|^2 + 2\mathsf{F}(\boldsymbol{q}) u'(0) v'(0) + \mathsf{G}(\boldsymbol{q}) |v'(0)|^2 \end{aligned}$$

where

$$\begin{aligned} \mathsf{E}(\boldsymbol{q}) &= \langle \partial_u \boldsymbol{x}, \partial_u \boldsymbol{x} \rangle_{T_{\boldsymbol{p}}(S)} = |\partial_u \boldsymbol{x}(\boldsymbol{q})|^2, \\ \mathsf{F}(\boldsymbol{q}) &= \langle \partial_u \boldsymbol{x}, \partial_v \boldsymbol{x} \rangle_{T_{\boldsymbol{p}}(S)} = \partial_u \boldsymbol{x}(\boldsymbol{q}) \cdot \partial_v \boldsymbol{x}(\boldsymbol{q}), \\ \mathsf{G}(\boldsymbol{q}) &= \langle \partial_v \boldsymbol{x}, \partial_v \boldsymbol{x} \rangle_{T_{\boldsymbol{p}}(S)} = |\partial_v \boldsymbol{x}(\boldsymbol{q})|^2. \end{aligned}$$

By letting $\boldsymbol{p}$ run in the coordinate neighborhood of $\boldsymbol{q}$ corresponding to $\boldsymbol{x}(u,v)$, we obtains functions

$$\mathsf{E}(u,v), \quad \mathsf{F}(u,v), \quad \mathsf{G}(u,v) \tag{3.6.2}$$

which are C^∞ in that neighborhood of $\boldsymbol{q}$. It is important to point that the functions E,F and G given in (3.6.2) only depends on the local coordinate $\boldsymbol{x}$, which are also well-defined without assuming the existence of the above mentioned regular C^∞-curve $\boldsymbol{\alpha}(t) = \boldsymbol{x}(u(t), v(t))$ in S, therefore we reach the following definition.

DEFINITION 3.6.2. The C^∞ functions E,F and G given in (3.6.2) are called the *coefficients of the first fundamental form*.

EXERCISE 3.6.3. Compute the coefficients of the first fundamental form of the surfaces given above.

EXERCISE 3.6.4. Let $\boldsymbol{\alpha}(t) = \boldsymbol{x}(u(t), v(t))$ for all $t \in I$ be a regular C^∞-curve in S. Recall that the arc length of such curve from $t_0 \in I$ is defined by $s(t_1) = \int_{t_0}^{t_1} |\boldsymbol{\alpha}'(t)| \, \mathrm{d}t$ for all $t_1 \in I$. Show that the arc length s can be expressed in terms of the coefficients of the first fundamental form given in (3.6.2) by the formula

$$s(t_1) = \int_{t_0}^{t_1} \sqrt{\mathsf{E}(u(t),v(t))|u'(t)|^2 + 2\mathsf{F}(u(t),v(t))u'(t)v'(t) + \mathsf{G}(u(t),v(t))|v'(t)|^2} \, \mathrm{d}t.$$

REMARK 3.6.5. By using the fundamental theorem of calculus (see e.g. [**Kow24a**]), one sees that

$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 = \mathsf{E}\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 + 2\mathsf{F}\frac{\mathrm{d}u}{\mathrm{d}t}\frac{\mathrm{d}v}{\mathrm{d}t} + \mathsf{G}\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^2.$$

Some mathematicians abuse the notation by denoting

$$(\mathrm{d}s)^2 = \mathsf{E}\,(\mathrm{d}u)^2 + 2\mathsf{F}\,(\mathrm{d}u)\,(\mathrm{d}v) + \mathsf{G}\,(\mathrm{d}v)^2,$$

and refer ds the "element of the arc length".

Let $\boldsymbol{\alpha} : I \to S$ and $\boldsymbol{\beta} : I \to S$ be two regular C^∞-curve in S which intersects at time $t = t_0$. By using Theorem 1.2.3, one sees the not oriented angle $\theta \in [0,\pi]$ between $\boldsymbol{\alpha} : I \to S$ and $\boldsymbol{\beta} : I \to S$ is given by

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{\alpha}'(t_0)}{|\boldsymbol{\alpha}'(t_0)|} \cdot \frac{\boldsymbol{\beta}'(t_0)}{|\boldsymbol{\beta}'(t_0)|}\right).$$

This suggests the following notion.

DEFINITION 3.6.6. The not oriented angle $\varphi \in [0,\pi]$ of the coordinate curves of a parameterization $\boldsymbol{x}(u,v)$ is

$$\varphi := \cos^{-1}\left(\frac{\partial_u \boldsymbol{x}}{|\partial_u \boldsymbol{x}|} \cdot \frac{\partial_v \boldsymbol{x}}{|\partial_v \boldsymbol{x}|}\right) = \cos^{-1}\left(\frac{\mathsf{F}}{\sqrt{\mathsf{EG}}}\right),$$

where E, F and G are the coefficients of the first fundamental form given in (3.6.2). It follows that the coordinate curves of a parameterization are orthogonal if and only if $F \equiv 0$, and such a parameterization is called an *orthogonal parameterization*.

Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ is a regular C^∞ surface (see Lemma 3.3.2). Recall that the function $|\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}|$, defined in U, measures the area of the parallelogram generated by the vectors $\partial_u \boldsymbol{x}$ and $\partial_v \boldsymbol{x}$. This suggests the following definition.

DEFINITION 3.6.7. Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ is a regular C^∞ surface (see Lemma 3.3.2) and let R be any bounded region in the regular surface $S = \boldsymbol{x}(U)$. The *area* of the region R is defined by

$$\mathrm{area}\,(R) := \int_\Omega |\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}| \,\mathrm{d}u\,\mathrm{d}v, \quad \text{where } \Omega = \boldsymbol{x}^{-1}(R).$$

REMARK 3.6.8. Let $\boldsymbol{y} : V \subset \mathbb{R}^2 \to \mathbb{R}^3$ be another parameterization of with $R \subset \boldsymbol{y}(V)$ and set $\Omega' := \boldsymbol{y}^{-1}(R)$. By using the change of variable formula (Theorem 3.5.3), one sees that

$$\int_{\Omega'} |\partial_{u'} \boldsymbol{y} \times \partial_{v'} \boldsymbol{y}| \,\mathrm{d}u'\,\mathrm{d}v' = \int_{\Omega'} |\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}| \left|\frac{\partial(u,v)}{\partial(u',v')}\right| \mathrm{d}u'\,\mathrm{d}v' = \int_\Omega |\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}| \,\mathrm{d}u\,\mathrm{d}v,$$

which shows that the definition of $\mathrm{area}\,(R)$ is independent of the choice of local coordinates. In fact, it also identical to 2-dimensional Hausdorff measure. By slightly abuse the notation, we refer

$$\mathrm{d}S = |\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}| \,\mathrm{d}u\,\mathrm{d}v \tag{3.6.3}$$

the *surface element*.

In practical, it is inconvenient to compute the cross product $|\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}|$ directly. The following exercise gives another alternative in terms of the coefficients of the first fundamental form.

EXERCISE 3.6.9. Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ is a regular C^∞ surface (see Lemma 3.3.2), show that

$$|\partial_u \boldsymbol{x}(u,v) \times \partial_v \boldsymbol{x}(u,v)| = \sqrt{\mathsf{E}(u,v)\mathsf{G}(u,v) - |\mathsf{F}(u,v)|^2} \quad \text{for all } (u,v) \in U.$$

(**Hint.** One can use (1.2.2) to give an easy proof)

One important consequence of Exercise 3.6.9 is $\mathsf{E}\mathsf{G}-\mathsf{F}^2\neq 0$, that is:

COROLLARY 3.6.10. *The matrix* $\begin{pmatrix} \mathsf{E} & \mathsf{F} \\ \mathsf{F} & \mathsf{G} \end{pmatrix}$ *is invertible.*

3.7. Some geometric interpretation for divergence operator and curl operator

Let $\boldsymbol{F}=(F_1,F_2,F_3)$ be a (smooth) vector field in $\mathbb{R}^3$ (for example, the magnetic field or electric field). We are now interested in the local behavior of the vector field *near*[3] a given point $\boldsymbol{x}_0\in\mathbb{R}^3$.

Let us begin this section by introduce the divergence operator from some heuristic computations. Let Ω be any open neighborhood of $\boldsymbol{x}_0$ in $\mathbb{R}^3$ which is bounded, such that $\partial\Omega$ is a regular C^∞-surface (Definition 3.2.1) which is oriented (Definition 3.3.8), so that for each point $\boldsymbol{x}\in\partial\Omega$ one can assign an unit normal vector $\boldsymbol{\nu}(\boldsymbol{x})$ at $\partial\Omega$ which is pointed outward (in the sense of the well-known Jordan-Brouwer separation theorem[4]). The *total flux* passing through the boundary $\partial\Omega$ is given by

$$\int_{\partial\Omega}\boldsymbol{F}\cdot\boldsymbol{\nu}\,\mathrm{d}S,$$

where $\mathrm{d}S$ is the surface element given in (3.6.3). The divergence theorem [**Str08**, Appendix A.3] says that the total flux passing through the boundary $\partial\Omega$ is equal to the volume integral of the "local change of rate" in the enclosed region Ω:

$$\int_{\partial\Omega}\boldsymbol{F}\cdot\boldsymbol{\nu}\,\mathrm{d}S=\int_{\Omega}\nabla\cdot\boldsymbol{F}\,\mathrm{d}\boldsymbol{x}, \tag{3.7.1}$$

where $\nabla\cdot\boldsymbol{F}:=\partial_1F_1+\partial_2F_2+\partial_3F_3$, see e.g. the advance monograph [**EG15**] for a general result (even holds true for some class of domain with non-Lipschitz boundary). The divergence of the vector field $\boldsymbol{F}$ at a point $\boldsymbol{x}_0$ is defined as the limit of the ratio of the surface integral (3.6.3) out of the volume $|\Omega|$ (equivalent to Lebesgue measure) of Ω as $|\Omega|$ shrink to 0:

$$\operatorname{div}\boldsymbol{F}(\boldsymbol{x}_0):=\lim_{\boldsymbol{x}_0\in\Omega,|\Omega|\to 0_+}\frac{1}{|\Omega|}\int_{\partial\Omega}\boldsymbol{F}\cdot\boldsymbol{\nu}\,\mathrm{d}S=\lim_{\boldsymbol{x}_0\in\Omega,|\Omega|\to 0_+}\frac{1}{|\Omega|}\int_{\Omega}\nabla\cdot\boldsymbol{F}\,\mathrm{d}\boldsymbol{x}.$$

If $\boldsymbol{F}=(F_1,F_2,F_3)$ is C^1 near $\boldsymbol{x}_0$, in view of the mean value theorem for continuous functions, the above discussions suggest the following rigorous definition (which can be easily extend for arbitrary dimension):

DEFINITION 3.7.1. Let $\boldsymbol{x}_0\in\mathbb{R}^3$, and suppose that $\boldsymbol{F}=(F_1,F_2,F_3)$ is C^1 near $\boldsymbol{x}_0$, then we define the *divergence* of $\boldsymbol{F}$ at $\boldsymbol{x}_0$ by

$$\operatorname{div}\boldsymbol{F}(\boldsymbol{x}_0):=\nabla\cdot\boldsymbol{F}(\boldsymbol{x}_0)=\partial_1F_1(\boldsymbol{x}_0)+\partial_2F_2(\boldsymbol{x}_0)+\partial_3F_3(\boldsymbol{x}_0).$$

[3]Recall Definition 1.4.7 and Definition 1.4.12 for some explanations about the term "near".

[4]Any compact, connected hypersurface X in $\mathbb{R}^n$ will divide $\mathbb{R}^n$ into two connected regions; the "outside" D_0 and the "inside" D_1. Furthermore, $\overline{D_1}$ is itself a compact set with boundary $\partial D_1=X$. The Jordan curve theorem is exactly the special case $n=2$.

EXERCISE 3.7.2 (Green's theorem as a special case of divergence theorem). Let Ω be a bounded domain in $\mathbb{R}^2$ such that its boundary $\partial\Omega$ is a regular curve (Definition 2.3.8). The Green's theorem stated that

$$\int_{\Omega} (\partial_x q(x,y) - \partial_y p(x,y))\,\mathrm{d}x\,\mathrm{d}y = \int_{\partial D} (p\,\mathrm{d}x + q\,\mathrm{d}y),$$

for all $p,q \in C^1(\overline{\Omega})$, where the right-hand-side is the line integral defined by

$$\int_{\partial D} (p\,\mathrm{d}x + q\,\mathrm{d}y) := \int_{\partial D} (p(\gamma(s)), q(\gamma(s))) \cdot \boldsymbol{t}(s)\,\mathrm{d}s$$

where $\gamma(s)$ is the arc-lengh parametrization of ∂D and $\boldsymbol{t}(s)$ is the unit tangent vector field (chosen to be counterclockwise oriented). Show that Green's theorem is a special case of the divergence theorem.

We now introduce the definition of curl operator from some heuristic computations. The idea of the curl is to measure the "rotation" of the vector field $\boldsymbol{F}$ near each point $\boldsymbol{x}_0$. Again, let Ω be any open neighborhood of $\boldsymbol{x}_0$ in $\mathbb{R}^3$ which is bounded, such that $\partial\Omega$ is a regular C^∞-surface which is oriented, so that for each point $\boldsymbol{x} \in \partial\Omega$ one can assign an unit normal vector $\boldsymbol{\nu}(\boldsymbol{x})$ at $\partial\Omega$ which is pointed outward. It is not easy to understand the concept of "rotation in $\mathbb{R}^3$", let us restrict ourselves in a plane with normal $\hat{\boldsymbol{n}}$, and we consider the plane $A := \Omega \cap \{\boldsymbol{x} \in \mathbb{R}^3 : \boldsymbol{x} \cdot \hat{\boldsymbol{n}} = 0\}$. Note that its boundary C is a regular curve (Definition 2.3.8) with parameterization $\boldsymbol{r} : [a,b] \to C$ which is counter-clockwise with respect to the orientation $\hat{\boldsymbol{n}}$, see Figure 3.7.1 below.

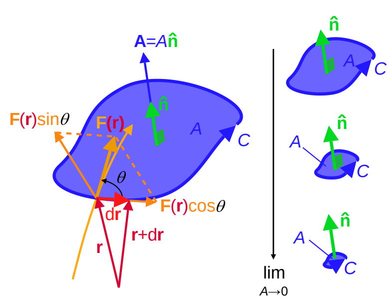


FIGURE 3.7.1. Geometry of the curl operator ("right hand rule"): By Maschen - Own work, CC0, https://commons.wikimedia.org/w/index.php?curid=26769675

Then the "rotation" of the vector field $\boldsymbol{F}$ near each point $\boldsymbol{x}_0$ along the axis $[\hat{\boldsymbol{n}}]$ can be quantified by the line integral

$$\int_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} := \int_a^b \boldsymbol{F}(\boldsymbol{r}(t)) \cdot \boldsymbol{r}'(t)\,\mathrm{d}t$$

If we formally refer $\mathrm{d}\boldsymbol{r}$ the "line element", then not oriented angle between $\boldsymbol{F}$ and the "line element $\mathrm{d}\boldsymbol{r}$ is formally given by

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{F}}{|\boldsymbol{F}|} \cdot \mathrm{d}\boldsymbol{r}\right),$$

so that $\boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = |\boldsymbol{F}|\cos\theta$, which showed in Figure 3.7.1 as well. By using the Kelvin-Stokes theorem (`https://en.wikipedia.org/wiki/Stokes%27_theorem`), one has

$$\int_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = \int_A (\nabla \times \boldsymbol{F}) \cdot \mathrm{d}S,$$

where $\mathrm{d}S$ is the surface element given in (3.6.3) and

$$\nabla \times \boldsymbol{F} = (\partial_2 F_3 - \partial_3 F_2)\boldsymbol{e}_1 + (\partial_3 F_1 - \partial_1 F_3)\boldsymbol{e}_2 + (\partial_1 F_2 - \partial_2 F_1)\boldsymbol{e}_3, \tag{3.7.2}$$

or sometimes we *abuse the notation* by writing

$$\nabla \times \boldsymbol{F} = \det \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ F_1 & F_2 & F_3 \end{pmatrix}.$$

However, $\nabla\times$ is a operator, one should not confuse this with the cross product in Remark 1.2.6 above. Similar to the divergence operator, we would like to define the curl in the same spirit:

$$\operatorname{curl} \boldsymbol{F}(\boldsymbol{x}_0) \cdot \hat{\boldsymbol{n}} := \lim_{\boldsymbol{x}_0 \in A, \operatorname{area}(A) \to 0_+} \frac{1}{\operatorname{area}(A)} \int_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = \lim_{\boldsymbol{x}_0 \in A, \operatorname{area}(A) \to 0_+} \frac{1}{\operatorname{area}(A)} \int_A (\nabla \times \boldsymbol{F}) \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S,$$

where $\operatorname{area}(A)$ dis the surface area (Definition 3.6.7) of the plane A. If $\boldsymbol{F} = (F_1, F_2, F_3)$ is C^1 near $\boldsymbol{x}_0$, in view of the mean value theorem for continuous functions, we expect

$$\operatorname{curl} \boldsymbol{F}(\boldsymbol{x}_0) \cdot \hat{\boldsymbol{n}} := (\nabla \times \boldsymbol{F}) \cdot \hat{\boldsymbol{n}} \quad \text{for all } \hat{\boldsymbol{n}} \in \mathbb{S}^2,$$

and this suggests the following rigorous definition.

DEFINITION 3.7.3. Let $\boldsymbol{x}_0 \in \mathbb{R}^3$, and suppose that $\boldsymbol{F} = (F_1, F_2, F_3)$ is C^1 near $\boldsymbol{x}_0$, then we define the *curl* of $\boldsymbol{F}$ at $\boldsymbol{x}_0$ by

$$\operatorname{curl} \boldsymbol{F}(\boldsymbol{x}_0) := \nabla \times \boldsymbol{F}(\boldsymbol{x}_0),$$

where $\nabla \times \boldsymbol{F}$ is given by (3.7.2).

EXERCISE 3.7.4 (Curl-curl identity). For each vector field $\boldsymbol{F} = (F_1, F_2, F_3) \in (C^1(\mathbb{R}^3))^3$, show that

$$\Delta \boldsymbol{F} - \nabla(\operatorname{div} \boldsymbol{F}) = -\nabla \times (\nabla \times \boldsymbol{F}),$$

where $\Delta \boldsymbol{F} = (\Delta F_1, \Delta F_2, \Delta F_3)$. Here, $\Delta u = \partial_1^2 u + \partial_2^2 u + \partial_3^2 u$ and $\nabla u = (\partial_1 u, \partial_2 u, \partial_3 u)$ for all C^1 scalar functions u.

REMARK 3.7.5. It is possible to extend the definition of curl and the above curl-curl identity to arbitrary dimension $\mathbb{R}^n$, even for symmetric tensors in terms of Saint-Venant operators, see e.g. [**IKS23**] for a proof. See also [**KL19, LUW11**] for an application of the curl-curl identity in Navier-Stokes equation.

3.8. Christoffel symbols, covariant derivative and geodesics

Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S$ be an orientable regular C^∞-surface. Recall that the tangent space (not to be confused with tangent plane) at a point $\boldsymbol{p} = \boldsymbol{x}(\boldsymbol{q}) \in S$ is $[\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q})]$, and the unit normal is given by the Gauss map (3.3.2):

$$\boldsymbol{N}(\boldsymbol{q}) = \frac{\partial_u \boldsymbol{x}(\boldsymbol{q}) \times \partial_v \boldsymbol{x}(\boldsymbol{q})}{|\partial_u \boldsymbol{x}(\boldsymbol{q}) \times \partial_v \boldsymbol{x}(\boldsymbol{q})|} \quad \text{for all } \boldsymbol{q} \in U.$$

Since $|\boldsymbol{N}|^2 = 1$, it is easy to see that

$$\begin{cases} \boldsymbol{N} \cdot \partial_u \boldsymbol{N} = \frac{1}{2}\partial_u(\boldsymbol{N} \cdot \boldsymbol{N}) = \frac{1}{2}\partial_u |\boldsymbol{N}|^2 = 0, \\ \boldsymbol{N} \cdot \partial_v \boldsymbol{N} = \frac{1}{2}\partial_v(\boldsymbol{N} \cdot \boldsymbol{N}) = \frac{1}{2}\partial_v |\boldsymbol{N}|^2 = 0. \end{cases} \tag{3.8.1}$$

This means that, for each $\boldsymbol{q} \in U$, $\{\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q}), \boldsymbol{N}(\boldsymbol{q})\}$ forms a basis of $\mathbb{R}^3$. In other words, the trihedron

$$\{\partial_u \boldsymbol{x}, \partial_v \boldsymbol{x}, N\} \equiv \{\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q}), \boldsymbol{N}(\boldsymbol{q}) : \boldsymbol{q} \in U\}$$

is a collection of basis of $\mathbb{R}^3$, then there exists smooth functions

$$\Gamma_{11}^1, \Gamma_{11}^2, \Gamma_{12}^1, \Gamma_{12}^2, \Gamma_{22}^1, \Gamma_{22}^2, L_1, L_2, \overline{L}_2, L_3, a_{11}, a_{12}, a_{21}, a_{22} : U \to \mathbb{R}. \tag{3.8.2}$$

such that

$$\begin{cases} \partial_u^2 \boldsymbol{x} = \Gamma_{11}^1 \partial_u \boldsymbol{x} + \Gamma_{11}^2 \partial_v \boldsymbol{x} + L_1 \boldsymbol{N}, \\ \partial_v \partial_u \boldsymbol{x} = \Gamma_{12}^1 \partial_u \boldsymbol{x} + \Gamma_{12}^2 \partial_v \boldsymbol{x} + L_2 \boldsymbol{N}, \\ \partial_u \partial_v \boldsymbol{x} = \Gamma_{21}^1 \partial_u \boldsymbol{x} + \Gamma_{21}^2 \partial_v \boldsymbol{x} + \overline{L}_2 \boldsymbol{N}, \\ \partial_v^2 \boldsymbol{x} = \Gamma_{22}^1 \partial_u \boldsymbol{x} + \Gamma_{22}^2 \partial_v \boldsymbol{x} + L_3 \boldsymbol{N}, \\ \partial_u \boldsymbol{N} = a_{11} \partial_u \boldsymbol{x} + a_{21} \partial_v \boldsymbol{x}, \\ \partial_v \boldsymbol{N} = a_{12} \partial_u \boldsymbol{x} + a_{22} \partial_v \boldsymbol{x}. \end{cases} \tag{3.8.3}$$

The functions $\Gamma_{11}^1, \Gamma_{11}^2, \Gamma_{12}^1, \Gamma_{12}^2, \Gamma_{22}^1, \Gamma_{22}^2$ are called the *Christoffel symbols*. Since $\partial_v \partial_u \boldsymbol{x} = \partial_u \partial_v \boldsymbol{x}$, then we immediately see the symmetry relation

$$\Gamma_{12}^1 = \Gamma_{21}^1, \quad \Gamma_{12}^2 = \Gamma_{21}^2, \quad L_2 = \overline{L}_2.$$

We also see that

$$L_1 = \partial_u^2 \boldsymbol{x} \cdot \boldsymbol{N}, \quad L_2 = \partial_v \partial_u \boldsymbol{x} \cdot \boldsymbol{N}, \quad L_3 = \partial_v^2 \boldsymbol{x} \cdot \boldsymbol{N}$$

which are in fact the *coefficients of the second fundamental form* (see Definition 3.9.15 in Section 3.9 below).

The following exercise can be solved by taking the inner product of the relations in (3.8.3) with $\partial_u \boldsymbol{x}$ or $\partial_v \boldsymbol{x}$:

EXERCISE 3.8.1. By denoting E, F and G be the coefficients of the first fundamental form given in (3.6.2), one has

$$\begin{aligned}\Gamma_{11}^{1}\mathsf{E}+\Gamma_{11}^{2}\mathsf{F}&=\frac{1}{2}\partial_u\mathsf{E}, &\Gamma_{11}^{1}\mathsf{F}+\Gamma_{11}^{2}\mathsf{G}&=\partial_u\mathsf{F}-\frac{1}{2}\partial_v\mathsf{E},\\ \Gamma_{12}^{1}\mathsf{E}+\Gamma_{12}^{2}\mathsf{F}&=\frac{1}{2}\partial_v\mathsf{E}, &\Gamma_{12}^{1}\mathsf{F}+\Gamma_{12}^{2}\mathsf{G}&=\frac{1}{2}\partial_u\mathsf{G},\\ \Gamma_{22}^{1}\mathsf{E}+\Gamma_{22}^{2}\mathsf{F}&=\partial_v\mathsf{F}-\frac{1}{2}\partial_u\mathsf{G}, &\Gamma_{22}^{1}\mathsf{F}+\Gamma_{22}^{2}\mathsf{G}&=\frac{1}{2}\partial_v\mathsf{G}.\end{aligned}$$

The equations connecting the Christoffol symbols with the coefficients of the first fundamental form given in (3.6.2), which also equivalent to

(3.8.4a) $$\begin{pmatrix}\mathsf{E}&\mathsf{F}\\ \mathsf{F}&\mathsf{G}\end{pmatrix}\begin{pmatrix}\Gamma_{11}^{1}\\ \Gamma_{11}^{2}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\partial_u\mathsf{E}\\ \partial_u\mathsf{F}-\frac{1}{2}\partial_v\mathsf{E}\end{pmatrix},$$

(3.8.4b) $$\begin{pmatrix}\mathsf{E}&\mathsf{F}\\ \mathsf{F}&\mathsf{G}\end{pmatrix}\begin{pmatrix}\Gamma_{12}^{1}\\ \Gamma_{12}^{2}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\partial_v\mathsf{E}\\ \frac{1}{2}\partial_u\mathsf{G}\end{pmatrix},$$

(3.8.4c) $$\begin{pmatrix}\mathsf{E}&\mathsf{F}\\ \mathsf{F}&\mathsf{G}\end{pmatrix}\begin{pmatrix}\Gamma_{22}^{1}\\ \Gamma_{22}^{2}\end{pmatrix}=\begin{pmatrix}\partial_v\mathsf{F}-\frac{1}{2}\partial_u\mathsf{G}\\ \frac{1}{2}\partial_v\mathsf{G}\end{pmatrix}.$$

which is solvable since $\begin{pmatrix}\mathsf{E}&\mathsf{F}\\ \mathsf{F}&\mathsf{G}\end{pmatrix}$ is invertible by Corollary 3.6.10. This shows that all the Christoffol symbols $\Gamma_{11}^{1},\Gamma_{11}^{2},\Gamma_{12}^{1},\Gamma_{12}^{2},\Gamma_{22}^{1},\Gamma_{22}^{2}$ can be expressed in terms of the coefficients of the first fundamental form given in (3.6.2). We shall not obtain the explicit expressions of the Christoffol symbols $\Gamma_{11}^{1},\Gamma_{11}^{2},\Gamma_{12}^{1},\Gamma_{12}^{2},\Gamma_{22}^{1},\Gamma_{22}^{2}$ since it is more convenient to work with the implicit form in Exercise 3.8.1.

EXAMPLE 3.8.2. We shall compute the Christoffel symbols for a surface of revolution parameterized by

$$\boldsymbol{x}(u,v)=(f(v)\cos u,f(v)\sin u,g(v))\quad\text{with}\quad f(v)\neq 0.$$

Since

$$\mathsf{E}=|f(v)|^2,\quad \mathsf{F}=0,\quad \mathsf{G}=|f'(v)|^2+|g'(v)|^2,$$

we obtain

$$\partial_u\mathsf{E}=0,\quad \partial_v\mathsf{E}=2ff',\quad \partial_u\mathsf{F}=\partial_v\mathsf{F}=0,\quad \partial_u\mathsf{G}=0,\quad \partial_v\mathsf{G}=2(f'f''+g'g'')$$

where prime denotes the derivative with respect to the variable v. From (3.8.4a) we have

$$\begin{pmatrix}\Gamma_{11}^1\\ \Gamma_{11}^2\end{pmatrix} = \begin{pmatrix} |f(v)|^2 & 0 \\ 0 & |f'(v)|^2+|g'(v)|^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ -ff' \end{pmatrix}$$
$$= \begin{pmatrix} |f(v)|^{-2} & 0 \\ 0 & (|f'(v)|^2+|g'(v)|^2)^{-1} \end{pmatrix} \begin{pmatrix} 0 \\ -ff' \end{pmatrix}$$
$$= \begin{pmatrix} 0 \\ -(|f'(v)|^2+|g'(v)|^2)^{-1}ff' \end{pmatrix}.$$

Using similar computations, from (3.8.4b) and (3.8.4c) we can compute that

$$\Gamma_{12}^1 = \frac{ff'}{f^2}, \quad \Gamma_{12}^2 = 0, \quad \Gamma_{22}^1 = 0, \quad \Gamma_{22}^2 = \frac{f'f''+g'g''}{|f'(v)|^2+|g'(v)|^2}.$$

We now introduce the covariant derivative of a vector field, which is the analogue for surfaces of the usual differentiation of vectors in the plane. We recall (Definition 3.3.6) that a tangent vector field in a relative open set $V \subset S$ of a regular surface E is a correspondence $\boldsymbol{w}$ that assigns to each $\boldsymbol{p} \in V$ a vector $\boldsymbol{w}(\boldsymbol{p}) \in T_{\boldsymbol{p}}(S)$.

DEFINITION 3.8.3. The vector field $\boldsymbol{w} : V \to T_{\boldsymbol{p}}(S)$ is said to be *differentiable* at $\boldsymbol{p}$ if there exists a parameterization $\boldsymbol{x}(u,v)$ near $\boldsymbol{p}$, the components $a(u,v)$ and $b(u,v)$ of $\boldsymbol{w} = a(u,v)\partial_u\boldsymbol{x} + b(u,v)\partial_v\boldsymbol{x}$ in the basis $\{\partial_u\boldsymbol{x}, \partial_v\boldsymbol{x}\}$ are differentiable functions at $\boldsymbol{x}^{-1}(\boldsymbol{p})$.

Let $\boldsymbol{w}$ be a differentiable vector field in an open set $U \subset S$ and $\boldsymbol{p} \in U$. Let $\boldsymbol{y} \in T_{\boldsymbol{p}}(S)$. We assume that there exists a curve

$$\boldsymbol{\alpha} : (-\varepsilon, \varepsilon) \to U, \quad \text{with } \boldsymbol{\alpha}(0) = \boldsymbol{p} \text{ and } \boldsymbol{\alpha}'(0) = \boldsymbol{y}.$$

We now consider to the normal projection of $\frac{\mathrm{d}(\boldsymbol{w}\circ\boldsymbol{\alpha})}{\mathrm{d}t}(0)$ onto the plane $T_{\boldsymbol{p}}(S)$. We write

$$\boldsymbol{\alpha}(t) = \boldsymbol{x}(u(t), v(t)),$$

so that

$$(\boldsymbol{w}\circ\boldsymbol{\alpha})(t) = \boldsymbol{w}(u(t), v(t)) = \tilde{a}(t)\partial_u\boldsymbol{x}(u(t),v(t)) + \tilde{b}(t)\partial_v\boldsymbol{x}(u(t),v(t)).$$

with $\tilde{a}(t) := a(u(t),v(t))$ and $\tilde{b}(t) := b(u(t),v(t))$. Hence

$$\frac{\mathrm{d}(\boldsymbol{w}\circ\boldsymbol{\alpha})}{\mathrm{d}t}(t) = \tilde{a}(t)(u'\partial_u^2\boldsymbol{x} + v'\partial_{uv}\boldsymbol{x}) + \tilde{b}(t)(u'\partial_{uv}\boldsymbol{x} + v'\partial_v^2\boldsymbol{x}) + \tilde{a}'\partial_u\boldsymbol{x} + \tilde{b}'\partial_v\boldsymbol{x}$$

where the prime denotes the derivative with respect to t.

EXERCISE 3.8.4. By using (3.8.3),show that the normal projection of $\frac{\mathrm{d}(\boldsymbol{w}\circ\boldsymbol{\alpha})}{\mathrm{d}t}$ onto the plane $T_{\boldsymbol{p}}(S)$ is then given by (equivalently, plugging (3.8.3) into the expression of $\frac{\mathrm{d}(\boldsymbol{w}\circ\boldsymbol{\alpha})}{\mathrm{d}t}$ and ignore the normal component $\boldsymbol{N}$):

$$\begin{aligned}(\tilde{a}' + \Gamma^1_{11}\tilde{a}u' + \Gamma^1_{12}\tilde{a}v' + \Gamma^1_{12}\tilde{b}u' + \Gamma^1_{22}\tilde{b}v')\partial_u\boldsymbol{x} \\ + (\tilde{b}' + \Gamma^2_{11}\tilde{a}u' + \Gamma^2_{12}\tilde{a}v' + \Gamma^2_{12}\tilde{b}u' + \Gamma^2_{22}\tilde{b}v')\partial_v\boldsymbol{x}.\end{aligned}$$

Note that

$$\tilde{a}'(0) = \partial_u a(\boldsymbol{p})u'(0) + \partial_v a(\boldsymbol{p})v'(0) = \nabla a(\boldsymbol{p})\cdot\boldsymbol{y}, \quad \tilde{b}'(0) = \nabla b(\boldsymbol{p})\cdot\boldsymbol{y},$$

which shows that the normal projection of $\frac{\mathrm{d}(\boldsymbol{w}\circ\boldsymbol{\alpha})}{\mathrm{d}t}(0)$ onto the plane $T_{\boldsymbol{p}}(S)$ is thus given by

$$\begin{aligned}(\nabla a(\boldsymbol{q})\cdot\boldsymbol{y} + \Gamma^1_{11}a(\boldsymbol{q})y_1 + \Gamma^1_{12}a(\boldsymbol{q})y_2 + \Gamma^1_{12}b(\boldsymbol{q})y_1 + \Gamma^1_{22}b(\boldsymbol{q})y_2)\partial_u\boldsymbol{x}(\boldsymbol{q}) \\ + (\nabla a(\boldsymbol{q})\cdot\boldsymbol{y} + \Gamma^2_{11}a(\boldsymbol{q})y_1 + \Gamma^2_{12}a(\boldsymbol{q})y_2 + \Gamma^2_{12}b(\boldsymbol{q})y_1 + \Gamma^2_{22}b(\boldsymbol{q})y_2)\partial_v\boldsymbol{x}(\boldsymbol{q}).\end{aligned}$$

with $\boldsymbol{q} = \boldsymbol{x}^{-1}(\boldsymbol{p})$. This leads the following definition, which is well-defined without the existence of the curve $\boldsymbol{\alpha}$ (see Theorem 3.8.16 below):

DEFINITION 3.8.5. Let $\boldsymbol{w} : V \to T_{\boldsymbol{p}}(S)$ be a differentiable vector field in a relative open set $V \subset S$, let $\boldsymbol{p} \in U$ and let $\boldsymbol{y} \in T_{\boldsymbol{p}}(S)$. Let $\boldsymbol{x}(u,v)$ be a parameterization near $\boldsymbol{p}$ with such that $\boldsymbol{w} = a(u,v)\partial_u\boldsymbol{x} + b(u,v)\partial_v\boldsymbol{x}$. The *covariant derivative* of $\boldsymbol{w}$ at $\boldsymbol{p}$ relative to $\boldsymbol{y}$ is defined as

$$\begin{aligned}\mathrm{D}_{\boldsymbol{y}}\boldsymbol{w} := (\nabla a(\boldsymbol{q})\cdot\boldsymbol{y} + \Gamma^1_{11}a(\boldsymbol{q})y_1 + \Gamma^1_{12}a(\boldsymbol{q})y_2 + \Gamma^1_{12}b(\boldsymbol{q})y_1 + \Gamma^1_{22}b(\boldsymbol{q})y_2)\partial_u\boldsymbol{x}(\boldsymbol{q}) \\ + (\nabla a(\boldsymbol{q})\cdot\boldsymbol{y} + \Gamma^2_{11}a(\boldsymbol{q})y_1 + \Gamma^2_{12}a(\boldsymbol{q})y_2 + \Gamma^2_{12}b(\boldsymbol{q})y_1 + \Gamma^2_{22}b(\boldsymbol{q})y_2)\partial_v\boldsymbol{x}(\boldsymbol{q}).\end{aligned}$$

EXAMPLE 3.8.6. We now assume that S is a plane, says $\{(u,v,0) : u,v \in \mathbb{R}\}$, with the natural parameterization

$$\boldsymbol{x}(u,v) = (u,v,0) \quad \text{for all } (u,v) \in \mathbb{R}^2.$$

We now see that $\partial_u\boldsymbol{x} = (1,0,0) = \boldsymbol{e}_1$ and $\partial_v\boldsymbol{x} = (0,1,0) = \boldsymbol{e}_2$ and that

$$\begin{aligned}\mathsf{E} &= |\partial_u\boldsymbol{x}|^2 = 1, \\ \mathsf{F} &= \partial_u\boldsymbol{x}\cdot\partial_v\boldsymbol{x} = 0, \\ \mathsf{G} &= |\partial_v\boldsymbol{x}|^2 = 1.\end{aligned}$$

By direct computations, one sees that all Christoffol symbols

$$\Gamma^1_{11}, \Gamma^2_{11}, \Gamma^1_{12}, \Gamma^2_{12}, \Gamma^1_{22}, \Gamma^2_{22} = 0 \tag{3.8.5}$$

For any vector field

$$\boldsymbol{w} = a(u,v)\partial_u\boldsymbol{x} + b(u,v)\partial_v\boldsymbol{x} = a(u,v)\boldsymbol{e}_1 + b(u,v)\boldsymbol{e}_2,$$

hence

$$a = \boldsymbol{w}\cdot\boldsymbol{e}_1, \quad b = \boldsymbol{w}\cdot\boldsymbol{e}_2$$

we see that

$$\mathrm{D}_{\boldsymbol{y}}\boldsymbol{w} = (\nabla(\boldsymbol{w}\cdot\boldsymbol{e}_1)\cdot\boldsymbol{y})\boldsymbol{e}_1 + (\nabla(\boldsymbol{w}\cdot\boldsymbol{e}_2)\cdot\boldsymbol{y})\boldsymbol{e}_2. \tag{3.8.6}$$

which shows that the covariant derivative agrees with the usual directional derivative of each components in the plane.

The definition of covariant derivative (Definition 3.8.5) may be extended to a vector field which is defined only at the points of a parameterized curve.

DEFINITION 3.8.7 (Curve with endpoints). A parameterized curve (with endpoints) $\boldsymbol{\alpha} : [0,\ell] \to S$ is the restriction to $[0,\ell]$ of a differentiable mapping of $(-\varepsilon, \ell+\varepsilon)$ for some $\varepsilon > 0$ into S. If $\boldsymbol{\alpha}(0) = \boldsymbol{p}$ and $\boldsymbol{\alpha}(\ell) = \boldsymbol{q}$, we say that $\boldsymbol{\alpha}$ joins $\boldsymbol{p}$ to $\boldsymbol{q}$. $\boldsymbol{\alpha}(t)$ is regular if $\boldsymbol{\alpha}'(t) \neq 0$ for all $t \in [0,\ell]$.

In what follows it will be convenient to use the notation $I = [0,\ell]$ whenever the specification of the end point I is not necessary (recall that we denote I be an open interval).

DEFINITION 3.8.8. Let $\boldsymbol{\alpha} : I \to S$ be a parameterized curve in S. A *vector field* $\boldsymbol{w}$ along $\boldsymbol{\alpha}$ is a correspondence that assigns to each $t \in I$ a vector $\boldsymbol{w}(t) \in T_{\boldsymbol{\alpha}(t)}(S)$. The vector field $\boldsymbol{w}$ is said to be *differentiable* at $t_0 \in I$ (t_0 may be the endpoint of I) if for some parameterization $\boldsymbol{x}(u,v)$ in $\boldsymbol{\alpha}(t_0)$ the components $a(t), b(t)$ of $\boldsymbol{w}(t) = a\partial_u\boldsymbol{x} + b\partial_u\boldsymbol{x}$ are differentiable functions of t at t_0. We say that a vector field $\boldsymbol{w}$ is differentiable in I if it is differentiable for all $t \in I$. The covariant derivative of $\boldsymbol{w}$ at t is defined by the formula in Exercise 3.8.4, that is,

$$\begin{aligned}\frac{\mathrm{D}\boldsymbol{w}}{\mathrm{d}t}(t) := {} & (a' + \Gamma^1_{11}au' + \Gamma^1_{12}av' + \Gamma^1_{12}bu' + \Gamma^1_{22}bv')\partial_u\boldsymbol{x} \\ & + (b' + \Gamma^2_{11}au' + \Gamma^2_{12}av' + \Gamma^2_{12}bu' + \Gamma^2_{22}bv')\partial_v\boldsymbol{x}.\end{aligned} \tag{3.8.7}$$

REMARK 3.8.9. By slightly abuse the notation by identifying the vector field $\boldsymbol{w}\circ\boldsymbol{\alpha}$ in Definition 3.8.5 and the vector field $\boldsymbol{w}$ given in Definition 3.8.8, we often write $\frac{\mathrm{D}\boldsymbol{w}}{\mathrm{d}t} = \mathrm{D}_{\boldsymbol{\alpha}'(t)}\boldsymbol{w}$, since the definition of covariant derivative is independent of the choice of the curve $\boldsymbol{\alpha}$ in S.

From a point of view external to the surface, in order to obtain the covariant derivative of a field $\boldsymbol{w}$ along $\boldsymbol{\alpha} : I \to S$ at $t \in I$, we take the usual derivative $\boldsymbol{w}'(t)$ with respect to t and project this vector orthogonally onto the tangent space $T_{\boldsymbol{\alpha}(t)}(S)$. It follows that when two surfaces are tangent along a parameterized curve $\boldsymbol{\alpha}$ the covariant derivative of a field $\boldsymbol{w}$ along $\boldsymbol{\alpha}$ is the same for both surfaces.

We now interpret $\boldsymbol{\alpha}$ be the trajectory of a point which is moving on the surface S, then $\boldsymbol{\alpha}'(t)$ is the velocity and $\boldsymbol{\alpha}''(t)$ is the acceleration. The covariant derivative $\frac{\mathrm{D}\alpha'}{\mathrm{d}t}$ of the field $\boldsymbol{\alpha}'(t)$ is the tangential component of the acceleration $\boldsymbol{\alpha}''(t)$. Intuitively, $\frac{\mathrm{D}\alpha'}{\mathrm{d}t}$ is the acceleration of the point $\boldsymbol{\alpha}(t)$ as seen from the surface S.

DEFINITION 3.8.10. A vector field $\boldsymbol{w}$ along a parameterized curve $\boldsymbol{\alpha}: I \to S$ is said to be *parallel* if $\frac{\mathrm{D}\boldsymbol{w}}{\mathrm{d}t}(t) = \mathbf{0}$ for all $t \in I$. In other words, the vector field $\boldsymbol{w}$ along a parameterized curve $\boldsymbol{\alpha}: I \to S$ is parallel means that its derivative $\boldsymbol{w}'(t)$ is perpendicular to the tangent space $T_{\boldsymbol{\alpha}(t)}(S)$.

EXAMPLE 3.8.11. We now consider the plane S as in Example 3.8.6 and let $\boldsymbol{\alpha}: I \to S$ be a curve in S. If a vector field $\boldsymbol{w} = a\partial_u \boldsymbol{x} + b\partial_u \boldsymbol{x}$ along such curve $\boldsymbol{\alpha}: I \to S$ is parallel, then from (3.8.7) and (3.8.5) we see that

$$0 = \frac{\mathrm{D}\boldsymbol{w}}{\mathrm{d}t}(t) = a'\partial_u \boldsymbol{x} + b'\partial_v \boldsymbol{x}.$$

Since $\{\partial_u \boldsymbol{x}, \partial_v \boldsymbol{x}\}$ is linear independent, then we see that $a' = b' = 0$ in I, which shows that a and b are constant, that is $\boldsymbol{w}$ is a constant vector field, which explains the term "parallel" in usual sense.

LEMMA 3.8.12. *Let $\boldsymbol{w}$ and $\boldsymbol{v}$ be parallel vector fields along $\boldsymbol{\alpha}: I \to S$. Then $\boldsymbol{w} \cdot \boldsymbol{v}$ is constant.*

PROOF. Since both $\boldsymbol{w}'(t)$ and $\boldsymbol{v}'(t)$ are perpendicular to the tangent space $T_{\boldsymbol{\alpha}(t)}(S)$ and both $\boldsymbol{v}(t), \boldsymbol{w}(t) \in T_{\boldsymbol{\alpha}(t)}(S)$, then we see that

$$(\boldsymbol{w}(t) \cdot \boldsymbol{v}(t))' = \boldsymbol{w}'(t) \cdot \boldsymbol{v}(t) + \boldsymbol{w}(t) \cdot \boldsymbol{v}'(t) = 0 \quad \text{for all } t \in I,$$

which completes the proof. □

EXAMPLE 3.8.13. Of course, on an arbitrary surface, parallel fields may look strange to our $\mathbb{R}^3$ intuition. For example, we consider a portion of the great circle

$$\boldsymbol{\alpha}(t) = (\cos t, \sin t, 0) \quad t \in I$$

on the sphere $S = \mathbb{S}^2$. We consider its tangent vector field $\boldsymbol{w}(t) := \boldsymbol{\alpha}'(t)$ for $t \in I$, and we see that

$$\boldsymbol{w}'(t) = -(\cos t, \sin t, 0) = -\boldsymbol{\alpha}(t) \quad t \in I,$$

thus $\boldsymbol{w}'(t)$ is perpendicular to the tangent space $T_{\boldsymbol{\alpha}(t)}(S)$, see Example 2.5.8 and Example 3.3.11. This shows that the $\frac{\mathrm{D}\boldsymbol{w}}{\mathrm{d}t}(t) = \mathbf{0}$ for all $t \in I$, i.e. the tangent vector field $\boldsymbol{w}(t) = (-\sin t, \cos t, 0)$ is parallel along the portion of great circle $\boldsymbol{\alpha}$ (but it is not parallel in our $\mathbb{R}^3$ intuition). One can understood this phenomena as "parallel lines" on the earth does not look "parallel" from outer space. Similarly, a "straight line" does not look "straight" from outer space.

The above example now suggests the following definition.

DEFINITION 3.8.14. A nonconstant, parameterized curve $\boldsymbol{\gamma}: I \to S$ is said to be a *geodesic* at $t \in I$ if the field of its tangent vectors $\boldsymbol{\gamma}'(t)$ is parallel along $\boldsymbol{\gamma}$at t, that is, $\frac{\mathrm{D}\boldsymbol{\gamma}'}{\mathrm{d}t}(t) = 0$. Accordingly, we say that $\boldsymbol{\gamma}: I \to S$ is a (parameterized) geodesic if it is geodesic at all $t \in I$.

By Lemma 3.8.12, we see that $|\boldsymbol{\gamma}'(t)| = c = \text{constant} \neq 0$. Therefore we may introduce the arc length $s = ct$ as a parameter. Observe that a geodesic may admit self-intersections, however its tangent vector is never zero, and thus the parameterization is regular. This consideration allows us to extend the definition of geodesic to regular curves.

DEFINITION 3.8.15. A regular connected curve C in S is said to be a geodesic if, for each $\boldsymbol{p} \in C$, the arc length parameterization $\boldsymbol{\alpha}(s)$ of C near $\boldsymbol{p}$ is a geodesic, i.e. $\boldsymbol{\alpha}'(s)$ is a parallel vector field along $\boldsymbol{\alpha}(s)$.

We shall now introduce the equations of a geodesic in a coordinate neighborhood. Let $\boldsymbol{\gamma} : I \to S$ be a parameterized curve in S and let $\boldsymbol{x}(u,v)$ be a parameterization of S in a neighborhood V of $\boldsymbol{\gamma}(t_0)$ for some $t_0 \in I$. Let $J \subset I$ be an open interval containing t_0 such that $\boldsymbol{\gamma}(J) \subset V$. Let $\boldsymbol{x}(u(t), v(t))$ with $t \in J$ be the expression of $\boldsymbol{\gamma} : J \to S$ in the parameterization $\boldsymbol{x}$. Then, the tangent vector field $\boldsymbol{\gamma}'(t)$ is given by

$$\boldsymbol{w} = u'(t)\partial_u \boldsymbol{x} + v'(t)\partial_v \boldsymbol{x}.$$

Therefore, the fact that $\boldsymbol{w}$ is parallel is equivalent to the system of ODE

$$\begin{cases} u'' + \Gamma_{11}^1 (u')^2 + 2\Gamma_{12}^1 u'v' + \Gamma_{22}^1 (v')^2 = 0, \\ v'' + \Gamma_{11}^2 (u')^2 + 2\Gamma_{12}^2 u'v' + \Gamma_{22}^2 (v')^2 = 0, \end{cases} \tag{3.8.8}$$

which can be obtained from Definition 3.8.8 with the choice $a = u'$ and $b = v'$. In other words,

$\boldsymbol{\gamma} : I \to S$ is a geodesic if and only if the *geodesic equation* (3.8.8) is satisfied locally.

Therefore, by using the theory of ODE (see e.g. [**Kow24b**]), we reach the following important consequence:

THEOREM 3.8.16. *Given a point $\boldsymbol{p} \in S$ and a vector $\boldsymbol{w} \in T_{\boldsymbol{p}}(S) \setminus \{\boldsymbol{0}\}$, there exists an $\varepsilon > 0$ and a unique geodesic $\boldsymbol{\gamma} : (-\varepsilon, \varepsilon) \to S$ such that $\boldsymbol{\gamma}(0) = \boldsymbol{p}$ and $\boldsymbol{\gamma}'(0) = \boldsymbol{w}$.*

EXAMPLE 3.8.17. We now study the geodesics of a surface of revolution in Example 3.8.2 with the parameterization

$$\boldsymbol{x}(u,v) = (f(v)\cos u, f(v)\sin u, g(v)) \quad \text{with} \quad f(v) \neq 0.$$

Recall that the Christoffel symbols are given by

$$\begin{aligned} &\Gamma_{11}^1 = 0, \quad \Gamma_{11}^2 = -\frac{ff'}{|f'(v)|^2 + |g'(v)|^2}, \quad \Gamma_{12}^1 = \frac{ff'}{f^2}, \\ &\Gamma_{12}^2 = 0, \quad \Gamma_{22}^1 = 0, \quad \Gamma_{22}^2 = \frac{f'f'' + g'g''}{|f'(v)|^2 + |g'(v)|^2}. \end{aligned}$$

Now the geodesic equation (3.8.8) reads

$$\begin{aligned} u'' + \frac{2ff'}{f^2} u'v' = 0, \\ v'' - \frac{ff'}{|f'(v)|^2 + |g'(v)|^2} (u')^2 + \frac{f'f'' + g'g''}{|f'(v)|^2 + |g'(v)|^2} (v')^2 = 0. \end{aligned}$$

3.9. Principal curvatures, Gaussian curvature and second fundamental form

Let $\boldsymbol{x} : U \subset \mathbb{R}^2 \to S$ be an orientable regular C^∞-surface. Recall that the tangent space (not to be confused with tangent plane) at a point $\boldsymbol{p} = \boldsymbol{x}(\boldsymbol{q}) \in S$ is $[\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q})]$, and the unit normal is given by the Gauss map (3.3.2):

$$\boldsymbol{N}(\boldsymbol{q}) = \frac{\partial_u \boldsymbol{x}(\boldsymbol{q}) \times \partial_v \boldsymbol{x}(\boldsymbol{q})}{|\partial_u \boldsymbol{x}(\boldsymbol{q}) \times \partial_v \boldsymbol{x}(\boldsymbol{q})|} \quad \text{for all } \boldsymbol{q} \in U.$$

We now consider the differential $\mathrm{d}\boldsymbol{N}_{\boldsymbol{p}} : T_{\boldsymbol{p}}(S) \to T_{\boldsymbol{p}}(S)$ of the Gauss map, which is given by (see Remark 3.2.2 above)

$$\mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\partial_u \boldsymbol{x}) = \partial_u \boldsymbol{N}, \quad \mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\partial_v \boldsymbol{x}) = \partial_v \boldsymbol{N}.$$

By taking derivatives on (3.8.1), we see that

$$\partial_v \boldsymbol{N} \cdot \partial_u \boldsymbol{x} + \boldsymbol{N} \cdot \partial_u \partial_v \boldsymbol{x} = 0, \quad \partial_u \boldsymbol{N} \cdot \partial_v \boldsymbol{x} + \boldsymbol{N} \cdot \partial_u \partial_v \boldsymbol{x} = 0,$$

thus by (3.6.1) we see that

$$\langle \mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\partial_v \boldsymbol{x}), \partial_u \boldsymbol{x} \rangle_{T_{\boldsymbol{p}}(S)} = \partial_v \boldsymbol{N} \cdot \partial_u \boldsymbol{x} = \partial_u \boldsymbol{N} \cdot \partial_v \boldsymbol{x} = \langle \partial_v \boldsymbol{x}, \mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\partial_u \boldsymbol{x}) \rangle_{T_{\boldsymbol{p}}(S)}. \tag{3.9.1}$$

Since $T_{\boldsymbol{p}}(S) = [\partial_u \boldsymbol{x}(\boldsymbol{q}), \partial_v \boldsymbol{x}(\boldsymbol{q})]$, then we reach

$$\langle \mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\boldsymbol{v}), \boldsymbol{w} \rangle_{T_{\boldsymbol{p}}(S)} = \langle \boldsymbol{v}, \mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\boldsymbol{w}) \rangle_{T_{\boldsymbol{p}}(S)} \quad \text{for all } \boldsymbol{v}, \boldsymbol{w} \in T_{\boldsymbol{p}}(S),$$

in other words:

LEMMA 3.9.1. *The differential* $\mathrm{d}\boldsymbol{N}_{\boldsymbol{p}} : T_{\boldsymbol{p}}(S) \to T_{\boldsymbol{p}}(S)$ *of the Gauss map* (3.3.2) *is a self-adjoint operator.*

REMARK 3.9.2. In the language of linear algebra, (3.9.1) means that $\mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}$ can be represented by a 2×2 self-adjoint matrix.

We recall the following important fact in linear algebra:

THEOREM 3.9.3 ([**Tre17b**, Theorem 2.4]). *If $A \in \mathbb{C}^{n \times n}$ is normal (i.e. $AA^* = A^*A$), then there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ (i.e. $U^{-1} = U^*$) such that $D := U^*AU$ is diagonal, i.e. $D =$* $\mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$. *If $A \in \mathbb{C}^{n \times n}$ is self-adjoint, then $D = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$ is real-valued.*

REMARK 3.9.4. The numbers $\lambda_1, \cdots, \lambda_n$ in Theorem 3.9.3 is called the eigenvalues of A. We recall that $U \in \mathbb{C}^{n \times n}$ is unitary if and only if its column vectors forms an orthonormal basis of $\mathbb{C}^n$.

By using Lemma 3.9.1 and Theorem 3.9.3 (with $n = 2$), for each $\boldsymbol{p} \in S$, there exist an orthonormal basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ of $T_{\boldsymbol{p}}(S)$ and $\kappa_1, \kappa_2 \in \mathbb{R}$ such that

$$\mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\boldsymbol{e}_1) = -\kappa_1 \boldsymbol{e}_1, \quad \mathrm{d}\boldsymbol{N}_{\boldsymbol{p}}(\boldsymbol{e}_2) = -\kappa_2 \boldsymbol{e}_2.$$

In order to make our notations consistent, here we consider the choice $\kappa_1 \geq \kappa_2$.

DEFINITION 3.9.5. The eigenvalues $\kappa_1, \kappa_2 \in \mathbb{R}$ of $-\mathrm{d}\boldsymbol{N_p}$ are called the *principal curvatures* at $\boldsymbol{p}$, and the corresponding eigenvectors $\boldsymbol{e}_1, \boldsymbol{e}_2$ are called the *principal directions* at $\boldsymbol{p}$. The number

$$\mathscr{K}_{\boldsymbol{p}} := \det(\mathrm{d}\boldsymbol{N_p}) = \kappa_1 \kappa_2$$

is called the *Gaussian curvature* of S at $\boldsymbol{p}$. A point $\boldsymbol{p}$ of a surface S is called

(a) *elliptic* if $\mathscr{K}_{\boldsymbol{p}} > 0$,
(b) *hyperbolic* if $\mathscr{K}_{\boldsymbol{p}} < 0$,
(c) *parabolic* if $\mathscr{K}_{\boldsymbol{p}} = 0$ with $\mathrm{d}\boldsymbol{N_p} \not\equiv 0$,
(d) *planar* if $\mathrm{d}\boldsymbol{N_p} \equiv 0$.

The number

$$H_{\boldsymbol{p}} = \frac{1}{2}\mathrm{Tr}\left(-\mathrm{d}\boldsymbol{N_p}\right) = \frac{\kappa_1 + \kappa_2}{2}$$

is called the *mean curvature* of S at $\boldsymbol{p}$. A point $\boldsymbol{p}$ of a surface S is called

DEFINITION 3.9.6. Let C be a regular curve in S passing through $\boldsymbol{p} \in S$, and let κ be the curvature of C at $\boldsymbol{p}$. Write $\theta := \cos^{-1}(\boldsymbol{n} \cdot \boldsymbol{N})$ be the not oriented angle between the unit normal vector $\boldsymbol{n}$ to C and the unit normal vector $\boldsymbol{N}$ to S at $\boldsymbol{p}$. The number

$$\kappa_{\boldsymbol{n}} := \kappa \cos\theta \equiv \kappa \boldsymbol{n} \cdot \boldsymbol{N}$$

is called the *normal curvature* of $C \subset S$ at $\boldsymbol{p}$. In other words, $\kappa_{\boldsymbol{n}}$ is the length of the projection of the vector $\kappa\boldsymbol{n}$ over the normal $\boldsymbol{N}$ to the surface S at $\boldsymbol{p}$, with a sign given by the orientation $\boldsymbol{N}$ of S at $\boldsymbol{p}$.

Let $\boldsymbol{\alpha}(s)$ be an arc length parameterization of C. By the definition of curvature (Section 2.4), we see that

$$\kappa_{\boldsymbol{n}} = \boldsymbol{\alpha}''(0) \cdot \boldsymbol{N}(\boldsymbol{\alpha}(0)).$$

Since $\boldsymbol{N}(\boldsymbol{\alpha}(s)) \cdot \boldsymbol{\alpha}'(s) = 0$, then $0 = (\boldsymbol{N} \circ \boldsymbol{\alpha})'(s) \cdot \boldsymbol{\alpha}'(s) + \boldsymbol{N}(\boldsymbol{\alpha}(s)) \cdot \boldsymbol{\alpha}''(s)$, which implies

$$\kappa_{\boldsymbol{n}} = -(\boldsymbol{N} \circ \boldsymbol{\alpha})'(0) \cdot \boldsymbol{\alpha}'(0) = -\mathrm{d}\boldsymbol{N_p}(\boldsymbol{\alpha}'(0)) \cdot \boldsymbol{\alpha}'(0) = -\langle \mathrm{d}\boldsymbol{N_p}(\boldsymbol{\alpha}'(0)), \boldsymbol{\alpha}'(0)\rangle_{T_{\boldsymbol{p}}(S)},$$

where we used (3.6.1). This suggests the following definition.

DEFINITION 3.9.7. The quadratic form $\tilde{Q}_{\boldsymbol{p}} : T_{\boldsymbol{p}}(S) \times T_{\boldsymbol{p}}(S) \to \mathbb{R}$ defined by

$$\tilde{Q}_{\boldsymbol{p}}(\boldsymbol{v}) := -\langle \mathrm{d}\boldsymbol{N_p}(\boldsymbol{v}), \boldsymbol{v}\rangle_{T_{\boldsymbol{p}}(S)} \quad \text{for all } \boldsymbol{v} \in T_{\boldsymbol{p}}(S),$$

is called the *second fundamental form* of S at $\boldsymbol{p}$.

Now let $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ be the principal directions (Definition 2.4), and since $|\boldsymbol{\alpha}'(0)| = 1$ we write

$$\boldsymbol{\alpha}'(0) = \boldsymbol{e}_1 \cos\theta + \boldsymbol{e}_2 \sin\theta \quad \text{for some } \theta \in [0, 2\pi).$$

Thus from the orthonormality of $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ we see that

$$\kappa_{\boldsymbol{n}} = \tilde{Q}_{\boldsymbol{p}}(\boldsymbol{\alpha}'(0)) = \tilde{Q}_{\boldsymbol{p}}(\boldsymbol{e}_1 \cos\theta + \boldsymbol{e}_2 \sin\theta) = \kappa_1 (\cos\theta)^2 + \kappa_2 (\sin\theta)^2. \tag{3.9.2}$$

The identity (3.9.2) is known classically as the *Euler formula*. Combining (3.9.2) with Theorem 3.8.16, we conclude that:

LEMMA 3.9.8. *κ_1 is the maximum normal curvature and κ_2 is the minimum normal curvature.*

DEFINITION 3.9.9. If a regular connected curve C on S is such that for all $\boldsymbol{p} \in C$ the tangent line of C is a principal direction at $\boldsymbol{p}$, then C is said to be a *line of curvature* of S.

EXERCISE 3.9.10 (Olinde Rodrigues). Show that a necessary and sufficient condition for a connected regular curve C on S to be a line of curvature of S is that: given any parameterization $\boldsymbol{\alpha}(t)$ of C, there exists a differentiable function $\lambda(t)$ such that

$$(\boldsymbol{N} \circ \boldsymbol{\alpha})'(t) = \lambda(t)\boldsymbol{\alpha}'(t).$$

In this case, $-\lambda(t)$ is the (principal) curvature along $\boldsymbol{\alpha}'(t)$.

DEFINITION 3.9.11. If at $\boldsymbol{p} \in S$ that $\kappa_1 = \kappa_2$, then $\boldsymbol{p}$ is called an *umbilical point* of S.

REMARK 3.9.12. In particular, the *planar points* (i.e. $\kappa_1 = \kappa_2 = 0$) are umbilical points.

We shall mention an interesting fact (without proof) regarding umbilical points.

THEOREM 3.9.13 ([**dC76**, Section 3.2, Proposition 4]). *If all points of a connected surface S are umbilical points, then S is either contained in a sphere or in a plane.*

REMARK 3.9.14. See e.g. [**IP19**] for an application of Theorem 3.9.13.

Let $\boldsymbol{x}(u,v)$ be a parameterization at a point $\boldsymbol{p} \in S$ of a surface S and let $\boldsymbol{\alpha}(t) = \boldsymbol{x}(u(t),v(t))$ be a parameterized curve on S with $\boldsymbol{\alpha}(0) = \boldsymbol{p}$. One can compute that

$$\tilde{Q}_{\boldsymbol{p}}(\boldsymbol{\alpha}') = \mathsf{e}(u')^2 + 2\mathsf{f}u'v' + \mathsf{g}(v')^2$$

with (since $\boldsymbol{N} \cdot \partial_u \boldsymbol{x} = 0$ and $\boldsymbol{N} \cdot \partial_v \boldsymbol{x} = 0$)

$$\begin{aligned}
\mathsf{e} &= -\partial_u \boldsymbol{N} \cdot \partial_u \boldsymbol{x} = \boldsymbol{N} \cdot \partial_u^2 \boldsymbol{x}, \\
\mathsf{f} &= -\partial_v \boldsymbol{N} \cdot \partial_u \boldsymbol{x} = \boldsymbol{N} \cdot \partial_v \partial_u \boldsymbol{x} = \boldsymbol{N} \cdot \partial_u \partial_v \boldsymbol{x} = -\partial_u \boldsymbol{N} \cdot \partial_v \boldsymbol{x}, \\
\mathsf{g} &= -\partial_v \boldsymbol{N} \cdot \partial_v \boldsymbol{x} = \boldsymbol{N} \cdot \partial_v^2 \boldsymbol{x}.
\end{aligned}$$

DEFINITION 3.9.15. The smooth functions e, f and g are called the *coefficients of the second fundamental form $\tilde{Q}_{\boldsymbol{p}}$*.

We now recall (3.8.3) that

$$\begin{aligned}
\partial_u \boldsymbol{N} &= a_{11} \partial_u \boldsymbol{x} + a_{21} \partial_v \boldsymbol{x}, \\
\partial_v \boldsymbol{N} &= a_{12} \partial_u \boldsymbol{x} + a_{22} \partial_v \boldsymbol{x},
\end{aligned}$$

therefore the Gaussian curvature is given by $\mathscr{K} = \det(\mathrm{d}\boldsymbol{N}) = a_{11}a_{22} - a_{12}a_{21}$. Since the principal curvatures κ_1, κ_2 are eigenvalues of $-\mathrm{d}\boldsymbol{N}$, then it satisfies

$$\det \begin{pmatrix} a_{11}+\kappa & a_{12} \\ a_{21} & a_{22}+\kappa \end{pmatrix} = 0,$$

see e.g. the lecture notes on linear algebra [**Tre17b**, Section 2]. We recall the following theorem in linear algebra.

THEOREM 3.9.16 ([**Tre17b**, Theorem 1.2]). *Let $A \in \mathbb{C}^{n\times n}$ and let $\lambda_1, \cdots, \lambda_n \in \mathbb{C}$ be its eigenvalues. Then*

$$\operatorname{Tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \quad \det(A) = \lambda_1\lambda_2\cdots\lambda_n,$$

where $\operatorname{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$ *is called the* trace *of matrix A.*

This theorems shows that the mean curvature is given by $\mathscr{H} = \frac{1}{2}\operatorname{Tr}(\mathrm{d}\boldsymbol{N}) = \frac{1}{2}(a_{11} + a_{22})$.

EXERCISE 3.9.17 (Weingarten equation). Show that

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\begin{pmatrix} \mathsf{e} & \mathsf{f} \\ \mathsf{f} & \mathsf{g} \end{pmatrix} \begin{pmatrix} \mathsf{E} & \mathsf{F} \\ \mathsf{F} & \mathsf{G} \end{pmatrix}^{-1},$$

where E, F and G are coefficients of the first fundamental form. Use this to show that the Gaussian curvature, the mean curvature and the principal curvatures can be expressed as

$$\mathscr{K} = \frac{\mathsf{eg} - \mathsf{f}^2}{\mathsf{EG} - \mathsf{F}^2}, \quad \mathscr{H} = \frac{1}{2}\frac{\mathsf{eG} - 2\mathsf{fF} + \mathsf{gE}}{\mathsf{EG} - \mathsf{F}^2}, \kappa_1 = \mathscr{H} + \sqrt{\mathscr{H}^2 - \mathscr{K}}, \quad \kappa_2 = \mathscr{H} - \sqrt{\mathscr{H}^2 - \mathscr{K}}.$$

COROLLARY 3.9.18. *$\boldsymbol{p} \in S$ is umbilical if and only if $\mathscr{H}^2 = \mathscr{K}$ at $\boldsymbol{p}$.*

Bibliography

[Apo74] T. M. Apostol. *Mathematical analysis*. Addison-Wesley Publishing Co., second edition, 1974. MR0344384, Zbl:0309.26002.

[BN10] J. Bak and D. Newman. *Complex Analysis*. Undergraduate Texts in Mathematics. Springer, New York, third edition, 2010. MR2675489, Zbl:1205.30001, doi:10.1007/978-1-4419-7288-0.

[BEG85] R. G. Barrera, G. A. Estevez, and J. Giraldo. Vector spherical harmonics and their application to magnetostatics. *Eur. J. Phys.*, 6:287–294, 1985. doi:10.1088/0143-0807/6/4/014.

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004. MR2061575, Zbl:1058.90049.

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011. MR2759829, Zbl:1220.46002, doi:10.1007/978-0-387-70914-7.

[Bro10] L. E. J. Brouwer. Zur Analysis Situs (German). *Math. Ann.*, 68(3):422–434, 1910. MR1511570, JFM:41.0543.02, doi:10.1007/BF01475781.

[BK82] L. Brown and J.-P. Kahane. A note on the Pompeiu problem for convex domains. *Math. Ann.*, 259(1):107–110, 1982. MR0656655, doi:10.1007/BF01456832.

[Cha06] I. Chavel. *Riemannian geometry. A modern introduction*, volume 98 of *Cambridge Stud. Adv. Math.* Cambridge University Press, Cambridge, second edition, 2006. MR2229062, Zbl:1099.53001, doi:10.1017/CBO9780511616822.

[ConNA] B. Conrad. *Differential Geometry handouts*. Stanford University, NA. `http://virtualmath1.stanford.edu/~conrad/diffgeomPage/`.

[CR61] H. M. Cundy and A. P. Rollett. *Mathematical models*. Clarendon Press, Oxford, second edition, 1961. MR0124167, Zbl:0259.00002.

[dC76] M. P. do Carmo. *Differential geometry of curves and surfaces*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1976. MR0394451, Zbl:0326.53001.

[EF14] C. Efthimiou and C. Frye. *Spherical harmonics in p dimensions*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014. MR3290046, Zbl:1296.33001, doi:10.1142/9134, `arXiv:1205.3548`.

[EG15] L. C. Evans and R. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015. MR3409135, doi:10.1201/b18333.

[Fay89] T. H. Fay. The Butterfly Curve. *Amer. Math. Monthly.*, 96(5):442–443, 1989. doi:10.2307%2F2325155, JSTOR:2325155.

[GK08] Y. H. Geum and Y. I. Kim. On the analysis and construction of the butterfly curve using Mathematica. *International Journal of Mathematical Education in Science and Technology*, 39(5):670–678, 2008. doi:10.1080/00207390801923240.

[HS99] P.-F. Hsieh and Y. Sibuya. *Basic theory of ordinary differential equations*. Universitext. Springer-Verlag, New York, 1999. MR1697415, doi:10.1007/978-1-4612-1506-6.

[IKS23] J. Ilmavirta, O.-Z. Kow, and S. K. Sahoo. Unique continuation for the momentum ray transform. *arXiv preprint*, 2023. `arXiv:2304.00327`.

[IP19] J. Ilmavirta and G. P. Paternain. Functions of constant geodesic x-ray transform. *Inverse Problems*, 35(6), 2019. Paper No. 065002, 11 pages. MR3975369, Zbl:1428.35625, doi:10.1088/1361-6420/ab0b6f, `arXiv:1812.03515`.

[Joh78] F. John. *Partial differential equations*, volume 1 of *Appl. Math. Sci.* Springer-Verlag, New York-Berlin, third edition, 1978. MR0514404, Zbl:0426.35002.

[Kow22] P.-Z. Kow. *Fourier analysis and distribution theory*. Lecture notes. University of Jyväskylä, Jyväskylä, 2022. https://puzhaokow1993.github.io/homepage.

[Kow23] P.-Z. Kow. *Complex Analysis*. National Chengchi University, Taipei, 2023. https://puzhaokow1993.github.io/homepage.

[Kow24a] P.-Z. Kow. *Calculus*. National Chengchi University, Taipei, 2024. https://puzhaokow1993.github.io/homepage.

[Kow24b] P.-Z. Kow. *An introduction to ordinary differential equations*. National Chengchi University, Taipei, 2024. https://puzhaokow1993.github.io/homepage.

[Kow24c] P.-Z. Kow. *An introduction to partial differential equations and functional analysis*. Lecture notes. National Chengchi University, Taipei, 2024. https://puzhaokow1993.github.io/homepage.

[KLSS24] P.-Z. Kow, S. Larson, M. Salo, and H. Shahgholian. Quadrature domains for the Helmholtz equation with applications to non-scattering phenomena. *Potential Anal.*, 60(1):387–424, 2024. MR4696043, Zbl:1535.35044, doi:10.1017/s11118-022-10054-5. The results in the appendix are well-known, and the proofs can found at arXiv:2204.13934.

[KL19] P.-Z. Kow and C.-L. Lin. On decay rate of solutions for the stationary Navier-Stokes equation in an exterior domain. *J. Differential Equations*, 266(6):3279–3309, 2019. MR3912683, Zbl:1435.35279, doi:10.1016/j.jde.2018.09.002.

[KS24] P.-Z. Kow and H. Shahgholian. Multi-phase k-quadrature domains and applications to acoustic waves and magnetic fields. *Partial Differ. Equ. Appl.*, 5, 2024. Paper No. 13, 28 pages, MR4732406, Zbl:7856435, doi:10.1007/s42985-024-00283-1, arXiv:2401.13279.

[KW21] P.-Z. Kow and J.-N. Wang. Reconstruction of an impenetrable obstacle in anisotropic inhomogeneous background. *IMA Journal of Applied Mathematics*, 86(2):320–348, 2021. MR4246858, Zbl:1471.76063, doi:10.1093/imamat/hxab002, see also doi:10.1093/imamat/hxab046 for a corrigendum.

[Kus87] R. Kusner. Conformal geometry and complete minimal surfaces. *Bull. Amer. Math. Soc. (N.S.)*, 17(2):291–295, 1987. MR0903735, Zbl:0634.53004, doi:10.1090/S0273-0979-1987-15564-9.

[Lee13] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Grad. Texts in Math.* Springer, New York, second edition, 2013. MR2954043, Zbl:1258.53002, doi:10.1007/978-1-4419-9982-5.

[LUW11] C.-L. Lin, G. Uhlmann, and J.-N. Wang. Asymptotic behavior of solutions of the stationary Navier-Stokes equations in an exterior domain. *Indiana Univ. Math. J.*, 60(6):2093–2106, 2011. MR3008262, Zbl:1295.35355, doi:10.1512/iumj.2011.60.4490, arXiv:1008.3953.

[Pom29] P. Pompeiu. Sur certains systèmes d'équations linéaires et sur une propriété intégrale des fonctions de plusieurs variables. *Comptes Rendus de l'Académie des Sciences. Série I. Mathématique*, 188:1138–1139, 1929. JFM:55.0139.02.

[Rud87] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987. MR0924157, Zbl:0925.00005.

[Str08] W. A. Strauss. *Partial differential equations: An introduction*. John Wiley & Sons, Ltd., Chichester, second edition, 2008. MR2398759.

[Tre17a] S. Treil. *Linear Algebra Done Wrong*. Brown University, 2017. https://www.math.brown.edu/streil/papers/LADW/LADW_2017-09-04.pdf.

[Tre17b] S. Treil. *Linear algebra done wrong*. Brown University, 2017. https://www.math.brown.edu.

[WZ15] R. L. Wheeden and A. Zygmund. *Measure and integral. An introduction to real analysis*. Pure Appl. Math. CRC Press, Boca Raton, FL, second edition, 2015. MR3381284, Zbl:1326.26007.

[Wil76] S. A. Williams. A partial solution of the Pompeiu problem. *Math. Ann.*, 223(2):183–190, 1976. MR0414904, Zbl:0329.35045, doi:10.1007/BF01360881.

[Wil81] S. A. Williams. Analyticity of the boundary for Lipschitz domains without the Pompeiu property. *Indiana Univ. Math. J.*, 30(3):357–369, 1981. MR0611225, Zbl:0439.35046.

[Yon17] K. Yoneyama. Theory of continuous set of points. *Tôhok. Math. J.*, 12:43–158, 1917. JFM:46.0303.01, https://www.jstage.jst.go.jp.