MAT236 Vector Calculus

Author: Qun WANG (qun.wang@utoronto.ca)

Institute: University of Toronto

Date: April 12, 2023

Material for MAT236, 2022-2023 Winter semester

Contents

Chapter 1 Topology in $\mathbb{R}^n$

1.1 Distance

Recall that given two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, with

$$\mathbf{x} = (x_1, x_2, ..., x_n), \mathbf{y} = (y_1, y_2, ..., y_n)$$

The **distance** between the two vectors can be defined as

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}$$

1.2 Open and Closed Sets

This distance function makes it possible to define the open ball.

Definition 1.1. Open Ball

The set

$$B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n | \|y - x\| < r\}$$

*is called the **open ball** centered at $\mathbf{x} \in \mathbb{R}^n$ with radius $r > 0$.*

Remark

- Three things determines the open ball: the center $\mathbf{x} \in \mathbb{R}^n$, the radius $r > 0$, as well as $n \in \mathbb{N}$.
- The inequality in the definition is **strict**.

Example1.1 Open Ball in $\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3$ Consider the open ball $B_r(\mathbf{0})$ in different spaces:

- in $\mathbb{R}^1$, it is the open interval $(-1, 1)$;
- in $\mathbb{R}^2$, it is the open disc, i.e., collection of points within the unit circle;
- in $\mathbb{R}^3$, it is the collection of points within the unit sphere – hence literally an open ball.

Note that when $n \geq 4$ one can no longer visualise it, but the definition is still valid.

With the help of open ball, one can define the open set, which plays a central role in the topology.

Definition 1.2. Open Set

*A set $U \subseteq \mathbb{R}^n$ is called an **open set**, if*

$$\forall \mathbf{x} \in U, \exists r > 0 \text{ s.t. } B_r(\mathbf{x}) \subseteq U.$$

Example1.2 Each open ball is an open set of its own right. To see this consider any open ball $B_R(\mathbf{x}) \subseteq \mathbb{R}^n$. Let's pick any point $\mathbf{y} \in B_R(\mathbf{x})$. Then take $r = \frac{1}{2}(R - \|\mathbf{y} - \mathbf{x}\|)$, it follows that

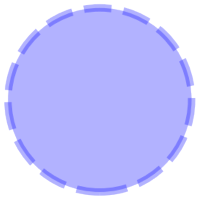


Figure 1.1: Open ball in $\mathbb{R}^2$

$B_r(\mathbf{y}) \subseteq B_R(\mathbf{x})$. As a result $B_R(\mathbf{x})$ is open according to the definition.

Remark

- In this example, one notices that in order to show a set U is open, one only needs to construct an open ball around any given point in U, which is equivalent of specifying a valid r. However, as the example above shows, in general r depends on which point we pick. In the above example, the closer is $\mathbf{y}$ to the boundary, the more $\|\mathbf{y} - \mathbf{x}\|$ is approaching R, hence the less $r = \frac{1}{2}(R - \|\mathbf{y} - \mathbf{x}\|)$ becomes.
- We have shown that an open ball is an open set, but an open set can be more general than an open ball, as we will see below.

Example1.3 The upper half plane

$$H = \{(x, y) \in \mathbb{R}^2 | y > 0\}$$

is an open set in $\mathbb{R}^2$.

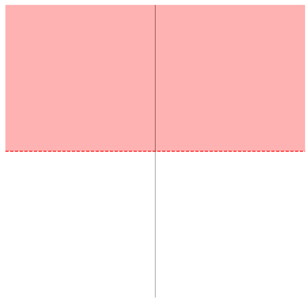


Figure 1.2: Upper Half Plane

Proof Indeed, for any $\mathbf{z} = (x, y) \in H$, by definition of H we know that $y > 0$. Now take $r = \frac{y}{2} > 0$, it follows that $B_r(\mathbf{z}) \in H$, i.e., H is open.

Example1.4 The set

$$C = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 > 1\}$$

is open.

Proof Let $z = (x, y) \in C$, then $\||z\| = \sqrt{x^2 + y^2} > 1$. As a result, take $r = \frac{1}{2}(\|z\| - 1)$, it follows that for any $u \in B_r(z)$, $\|u\| > 1$.

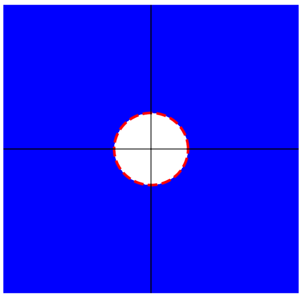


Figure 1.3: Exterior of a (closed) ball

Once we know what is an open set, we can define what is a closed set.

Definition 1.3. Closed Set

*A set $U \in \mathbb{R}^n$ is called a **closed** set if its complement U^c is open.* ♣

Example1.5 The set

$$\overline{B}_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n | \|\mathbf{y} - \mathbf{x}\| \leq 1\}$$

is closed, as we can show that

$$\bar{B}_r(\mathbf{x})^c = \{\mathbf{y} \in \mathbb{R}^n | \|\mathbf{y} - \mathbf{x}\| > 1\}$$

is open (using argument similar to example 1.4). This object is called a **closed ball**.

Example1.6 Similarly, the set

$$H = \{\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n | x_1 \geq 0\}$$

is also closed.

Proof Exercise.

Remark We emphasise the following two facts:

- There exist sets that are both open and closed.
- There exist sets that are neither open nor closed.

Example1.7 Both $\mathbb{R}^n$ and $\emptyset$ are open **and** closed at the same time. (Can you tell the reason?)

Example1.8 Consider the set

$$K_1 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < 1\}, K_2 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 1, x \leq 0\}.$$

Then $K = K_1 \cup K_2$ is neither open nor closed.

Proof Exercise.

Proposition 1.1

The intersection of finitely many open sets is still open. ♠

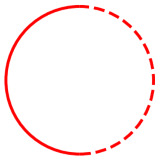


Figure 1.4: Neither open nor closed

Proof Let $U_1, U_2, ..., U_k \subseteq \mathbb{R}^n$ be open sets, Consider

$$U = \bigcap_{i=1}^{k} U_i$$

For any $\mathbf{x} \in U$, by definition of the intersection, $\mathbf{x} \in U_i$ for any $1 \leq i \leq k$. Since each U_i is open, one can finds $r_i > 0$ s.t. $B_{r_i}(\mathbf{x}) \in U_i$ for any $1 \leq i \leq k$. Next, take

$$r = \min_{1 \leq i \leq k} r_i > 0,$$

and one has that $B_r(\mathbf{x}) \in U_i$ for any $1 \leq i \leq k$, i.e., $B_r(\mathbf{x}) \in U$.

Remark The “finitely many" in the proposition cannot be dropped, as is shown in the following counter-example:

Example1.9 Let $B_{\frac{1}{i}}(\mathbf{0})$ be the open ball centered at the origin in $\mathbb{R}^2$ with radius being $\frac{1}{i}$. Each $B_{\frac{1}{i}}$ is open, yet

$$\bigcap_{i \in \mathbb{N}} B_{\frac{1}{i}} = \{\mathbf{0}\}$$

which is a single point, hence closed.

Proposition 1.2

The arbitrary union of open sets is still open.

Proof Let $U_i \subseteq \mathbb{R}^n$ be open sets, $i \in I$ where I is an arbitrary index set (it could be finite, countable, or even uncountable). Now consider

$$U = \bigcup_{i \in \mathbb{I}} U_i.$$

For any point $\mathbf{x} \in U$, $\mathbf{x} \in U_i$ for some $i \in I$. Now since U_i is open, there exists an open ball $B_r(\mathbf{x}) \subseteq U_i \subseteq U$. Hence U is open.

Using the De Morgan law, one can easily prove the following two propositions, the details are left as exercises:

Proposition 1.3

The union of finitely many closed set is still closed.

Proposition 1.4

The arbitrary intersection of closed sets is still closed.

1.3 Bounded and Compact Sets

Definition 1.4. Bounded Set

We say a set $U \subseteq \mathbb{R}^n$ is ***bounded****, if there exists some $M > 0$ s.t.*

$$U \subseteq B_M(\mathbf{0}).$$

Otherwise, we say that the set U is ***unbounded***

Remark The boundedness has nothing to do with being open or closed.

Example1.10 Consider

$$A = \{(x, y) \in \mathbb{R}^2 | \frac{x^2}{9} + \frac{y^2}{4} \leq 1\}$$

then A is bounded. On the other hand,

$$A^c = \{(x, y) \in \mathbb{R}^2 | \frac{x^2}{9} + \frac{y^2}{4} > 1\}$$

is unbounded.

Example1.11 Consider the set

$$H = \{(x, y) \in \mathbb{R}^2 | x > 0\}$$

Then H is unbounded. Moreover,

$$H^c = \{(x, y) \in \mathbb{R}^2 | x \leq 0\}$$

is also unbounded.

Finally we mention a key concept that lies in the central position of analysis:

Definition 1.5. Compact Set

A set $U \subset \mathbb{R}^n$ is called ***compact*** *if it is both closed and bounded.*

Example1.12 The closed ball is a bounded set. Moreover it is closed. As a result it is compact.

Basically speaking, a continuous function defined on a compact set never loses control – The function must be bounded and achieves its (finite) maximum and mininum. We will come back to this point then we have introduced the continuity.

1.4 Connectedness

(Path-)Connected

[1] We will need to discuss the path/curve integral of functions later in this course, so we want the domain of the functions of interest to satisfy some properties, in particular, we need the domain to be (path) connected.

[1]Strictly Speaking the connectedness and the the path-connectedness are not the same concept. Somehow we don't tend to justify the difference between the two in MAT236.

Definition 1.6. (Path-)Connectedness

Let $U \subset \mathbb{R}^n$. We say that U is **(path-) connected**, *if for any two points $A, B \in U$, one can draw a continuous path $\gamma \subseteq U$ connecting A and B.* ♣

Remark Note that to be connected

- One must be able to draw a continuous path for **any** pair A, B.
- The path γ must lie ENTIRELY in U.
- One does not need the path to be unique. As long as there is one connecting the desired points, it is all right.

Example1.13 The open ball is connected, so is the closed ball, and the half-open-half-closed ball.

Example1.14 Consider two open balls $B_1(\mathbf{z})$ and $B_1(-\mathbf{z})$ with $\mathbf{z} = (0, \lambda)$ be a point on the y-axis. Consider $B = B_1(\mathbf{z}) \cup B_1(-\mathbf{z})$. Then

- When $\lambda < 1$, B is connected.
- When $\lambda = 1$, B is not connected.
- When $\lambda > 1$, B is not connected.

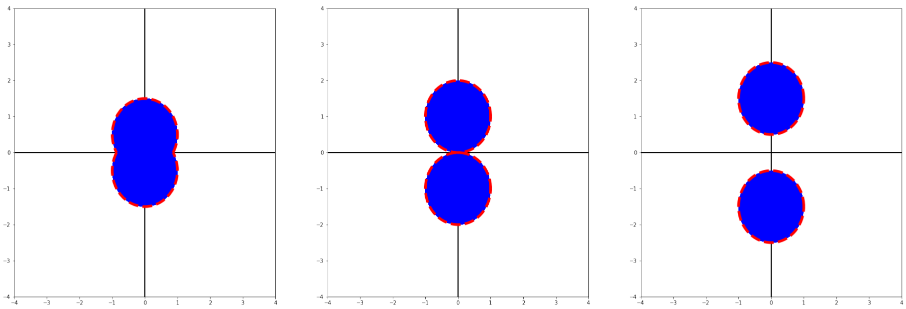


Figure 1.5: Union of two open ball, from left to right: $\lambda = 0.5, 1, 1.5$

Example1.15 Consider two closed balls $\overline{B_1}(\mathbf{z})$ and $\overline{B_1}(-\mathbf{z})$ with $\mathbf{z} = (0, \lambda)$ be a point on the y-axis. Consider $B = \overline{B_1}(\mathbf{z}) \cup \overline{B_1}(\mathbf{z})$. Then

- When $\lambda < 1$, B is connected.
- When $\lambda = 1$, B is connected.
- When $\lambda > 1$, B is not connected.

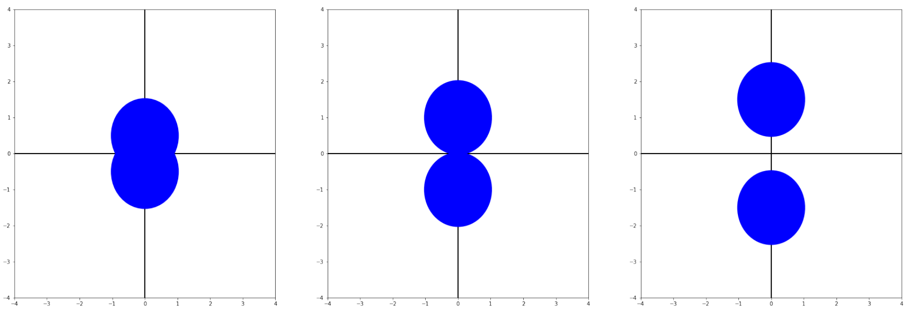


Figure 1.6: Union of two closed ball, from left to right: $\lambda = 0.5, 1, 1.5$

Simply-Connected

Definition 1.7. Simply Connected Domain

*A connected set $D \subseteq \mathbb{R}^n$ is called **simply connected**, if every closed curve can continuously deform into a point.* ♣

Remark It might be too ambitious to understand all the details of simply-connected set in the scope of MAT236. We will not try to do that neither. Somehow students should understand that in $\mathbb{R}^2$, the simply-connectedness is equivalent to non-existence of “holes".

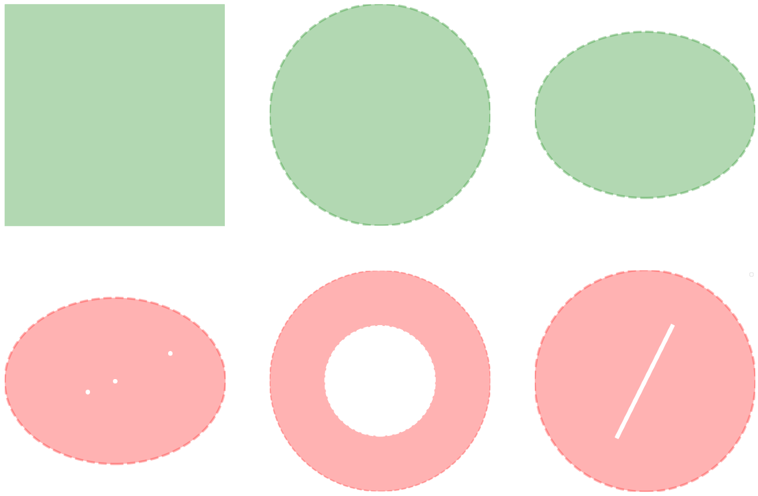


Figure 1.7: Simply Connected (Green) v.s. Connected (Red)

Example1.16

- The whole plane $\mathbb{R}^2$ is simply connected.
- An open disc is simply connected.
- A convex set (for example an ellipse) is simply connected.
- A punctured disc is connected, but **NOT** simply connected
- An open annulus is connected, but **NOT** simply connected
- The disc with a short rift is connected but **NOT** simply connected

♦

Chapter 2 Differentiable Function

In this section we begin to discuss the vector calculus from a differential point of view.

2.1 Limit

Definition 2.1. Limit

We say that a function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ *has a limit equal to* $C \in \mathbb{R}$ *at* $\mathbf{x}_0 \in \mathbb{R}^n$*, if* $\forall \epsilon > 0, \exists \delta > 0$ *s.t.* $\forall \mathbf{x} \in \mathbb{R}^n$ *it holds that*

$$0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta \Longrightarrow |f(\mathbf{x}) - C| < \epsilon.$$

♣

Remark

- In a more intuitive way, it means that as $\mathbf{x}$ is approaching $\mathbf{x}_0$, $f(\mathbf{x})$ will approach C in a **consistent** way, regardless of the direction in which it approaches $\mathbf{x}_0$.
- The function f is **not** necessarily defined on the point $\mathbf{x}_0$ – a function can have limit at a point without being defined there.

Quick Review of 1D Case Before discussing the multi-variable case, let's first review the 1D case quickly to gain some insights.

Example2.1

$$f(x) = \begin{cases} -1 & x < 0 \\ 1 & x > 0 \end{cases}$$

The limit $\lim\limits_{x \to 0} f(x)$ does not exist, since the behaviour of $f(x)$ as x approaches 0 from the left and that from the right are not consistent.

Remark

- Notice also that we don't need $f(0)$ to be defined, in the discussion of the limit.

Roughly speaking, it is a much more difficult for the limit to exists for functions defined on $\mathbb{R}^n$. Indeed

- In $\mathbb{R}^1$, there are only **two** ways to approach a point: from the left or from the right.
- In $\mathbb{R}^n$, there are **infinitely many** ways to approach a point, moreover one don't necessarily need to approach a point in straight line — it could be any curve as well.

As a result,

- For a function to have limit at a given point, we need all the "limit along a given trajectory" to be the same. A rigorous proof of this often requires the $\epsilon - \delta$ formulation;

- For a function to not have limit at a given point, we need only show that
  - Either along a particular trajectory the limit does not exist;
  - Or along two trajectories the limit do not coincide with each other.

Is $\epsilon - \delta$ really a Mission Impossible?

Not at all. Somehow we need to be clear what should be done – you cannot solve your problem without knowing what the problem is (at least in math it is the case).

Example2.2 Consider the function $f : \mathbb{R}^2 \setminus \{\mathbf{0}\}$

$$f(x,y) = \frac{3x^2y}{x^2+y^2}$$

Then

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

1. Apply the definition to find out what is your mission. The definition reads:

   "A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ has a limit equal to $C \in \mathbb{R}$ at $\mathbf{x}_0 \in \mathbb{R}^n$, if $\forall \epsilon > 0, \exists \delta > 0$, s.t. $\forall \mathbf{x} \in A$ it holds that

   $$0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta \Longrightarrow |f(\mathbf{x}) - C| < \epsilon."$$

   Changing it into our setting, we should show that:

   "A function $f(x,y) = \dfrac{3x^2y}{x^2+y^2} : \mathbb{R}^2 \setminus \{\mathbf{0}\} \to \mathbb{R}$ has a limit equal 0 at $\mathbf{0}$, if $\forall \epsilon > 0, \exists \delta > 0$, s.t. $\forall (x,y) \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$,

   $$0 < \sqrt{(x-0)^2 + (y-0)^2} < \delta \Longrightarrow \left|\frac{3x^2y}{x^2+y^2} - 0\right| < \epsilon."$$

   Summary: we simply need to show that

   "$\forall \epsilon > 0, \exists \delta > 0$, s.t. $\forall (x,y) \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$,

   $$0 < \sqrt{x^2+y^2} < \delta \Longrightarrow \left|\frac{3x^2y}{x^2+y^2}\right| < \epsilon."$$

2. The role of ϵ and δ is **NOT** symmetric. In our setting,
   - ϵ is **given**. You can consider it as a constant.
   - δ is **to be chosen** by you. You have your own right to choose δ. There are infinitely ways to choose a proper δ and different person will probably end with different δ.

3. The technical part is how to choose δ. Normally we do this by two steps:

   (a). First, we compute on the **draft paper**, try to find an ideal relation between $|\mathbf{f}(\mathbf{x}) - \mathbf{C}|$(i.e., $|\frac{3x^2y}{x^2+y^2}|$) and $|\mathbf{x} - \mathbf{x_0}|$(i.e., $\sqrt{x^2+y^2}$). **This normally starts by taking $|\mathbf{f}(\mathbf{x}) - \mathbf{C}|$ and perform and operation on it as experiments**. For example, in our example simply computation shows that

$$|\frac{3x^2y}{x^2+y^2}| = \frac{3x^2}{x^2+y^2}|y| \le 3|y| \le 3\sqrt{x^2+y^2}.$$

As a result as long as we can bound $\sqrt{x^2+y^2}$ by $\frac{\epsilon}{3}$, we can bound $|\frac{3x^2y}{x^2+y^2}|$ by ϵ.

(b). Next, claim your choice of δ directly and finish the proof on your **answer sheet**, as is shown below:

Proof We claim that

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

Indeed, $\forall \epsilon > 0$, pick $\delta = \frac{\epsilon}{3}$, then

$$\sqrt{x^2+y^2} < \delta \Rightarrow$$
$$|\frac{3x^2y}{x^2+y^2}| = \frac{3x^2}{x^2+y^2}|y| \le 3|y| \le 3\sqrt{x^2+y^2} < 3\delta = \epsilon,$$

which proved our claim.

Example2.3 Consider the function $f : \mathbb{R}^2 \setminus \{\mathbf{0}\}$

$$f(x,y) = \frac{\sin(x^2+y^2)}{x^2+y^2}$$

We claim that

$$\lim_{(x,y)\to(0,0)} f(x,y) = 1.$$

Proof Indeed, recall that we have seen in single variable calculus that

$$\lim_{t\to 0} \frac{\sin t}{t} = 1$$

which means that

$$\forall \epsilon \in \mathbb{R}, \exists \delta_1, \text{ s.t. } 0 < |t-0| < \delta_1 \Rightarrow |f(t)-1| < \epsilon.$$

Now take $\delta = \sqrt{\delta_1}$ and let $t = x^2+y^2$ it follows that

$$0 < \|\mathbf{x}-\mathbf{0}\| < \sqrt{\delta_1} \Rightarrow 0 < |x^2+y^2| < \delta_1 \Rightarrow 0 < |t| < \delta_1 \Rightarrow |f(t)-1| < \epsilon$$

i.e.,

$$\left|\frac{\sin(x^2+y^2)}{x^2+y^2}\right| < \epsilon.$$

We conclude that the limit exists.

□

Non-Existence of Limit

To show the limit of $f(\mathbf{x})$ does **not** exist at $\mathbf{x}_0$, one can try either of the following two strategies:

1. Show that along a path approaching to the point $\mathbf{x}_0$, the limit does not exist.
2. Show that along two paths approaching to the point $\mathbf{x}_0$, the two limits are not consistent.

We illustrate these situations through examples.

Example2.4 Consider the function $f : \mathbb{R}^2 \setminus \{\mathbf{0}\}$

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

Then $\lim_{(x,y)\to(0,0)} f(x, y)$ does **not** exist. Indeed, note that

$$\lim_{\substack{(x,y)\to(0,0)\\x=0}} f(x, y) = \lim_{x=0,y\to 0} \frac{-y^2}{y^2} = -1;$$

$$\lim_{\substack{(x,y)\to(0,0)\\y=0}} f(x, y) = \lim_{y=0,x\to 0} \frac{x^2}{x^2} = 1.$$

Since $-1 \neq 1$, the limit does not exist.

Example2.5 Consider the function $f : \mathbb{R}^2 \setminus \{\mathbf{0}\}$

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

Then $\lim_{(x,y)\to(0,0)} f(x, y)$ does **not** exist. Indeed, note that

$$\lim_{\substack{(x,y)\to(0,0)\\x=0}} f(x, y) = \lim_{x=0,y\to 0} \frac{0}{y^2} = 0;$$

$$\lim_{\substack{(x,y)\to(0,0)\\x=y}} f(x, y) = \lim_{x\to 0} \frac{x^2}{2x^2} = \frac{1}{2}.$$

Since $0 \neq \frac{1}{2}$, the limit does not exist.

Example2.6 Consider the function $f : \mathbb{R}^2 \setminus \{\mathbf{0}\}$

$$f(x, y) = \frac{xy^2}{x^2 + y^4}$$

Then the limit $\lim_{(x,y)\to(0,0)} f(x, y)$ does **not** exist. Indeed, note that

$$\lim_{\substack{(x,y)\to(0,0)\\x=0}} f(x, y) = \lim_{x=0,y\to 0} \frac{0}{y^4} = 0;$$

$$\lim_{\substack{(x,y)\to(0,0)\\x=y^2}} f(x, y) = \lim_{y\to 0} \frac{y^4}{2y^4} = \frac{1}{2}.$$

Since $0 \neq \frac{1}{2}$, the limit does not exist.

Example2.7 Consider the function $f : \mathbb{R}^2 \setminus \{\mathbf{0}\}$

$$f(x, y) = \frac{x}{x^2 + y^2}$$

Then the limit does not exist. Indeed, take $y = 0$, one sees that

$$\lim_{\substack{(x,y)\to(0,0)\\y=0}} f(x, y) = \lim_{x\to 0} \frac{x}{x^2} = \lim_{x\to 0} \frac{1}{x}.$$

which does not exist.

-

2.2 Continuity

Definition 2.2. Continuity

*A function $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is **continuous** at $\mathbf{x}_0 \in D$, if*

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$$

♣

Remark It is worth mentioning that a function $f(\mathbf{z})$ is continuous at $\mathbf{x}_0$ means 3 things:

1. The function $f(\mathbf{x})$ is defined at $\mathbf{x}_0$. [1]
2. The function $f(\mathbf{x})$ has limit at $\mathbf{x}_0$, i.e.

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$$

   exists.
3. This limit cannot be anyone else, but must be equal to the function value $f(\mathbf{x}_0)$.

As a result, we see that if any of the above three things are violated, then the function $f(\mathbf{x})$ will not be continuous at $\mathbf{x}_0$.

Example2.8 Some continuous functions we often see:

- constant: $f(\mathbf{x}) = 1$, $f(\mathbf{x}) = 2$, $f(\mathbf{x}) = 100$
- polynomial: $f(x, y) = xy$, $f(x, y, z) = x^2y^3z^5$
- exponential: $f(x) = e^x$
- trigonometric: $f(x, y) = \cos y$

and many others.

♦

Proposition 2.1. Arithmetic Operation Preserves Continuity

Suppose that $f(\mathbf{x})$ and $g(\mathbf{x})$ are both continuous at $\mathbf{x}_0$, then

$$f + g, f - g, f \cdot g, \frac{f}{g}(\textit{assume } g(\mathbf{x}_0) \neq 0)$$

are all continuous at $\mathbf{x}_0$ as well.

♠

Proposition 2.2. Composition Preserves Continuity

Suppose that $f(\mathbf{x}) : D_1 \to D_2$ is continuous at $\mathbf{x}_0 \in D_1$, and $g(\mathbf{w}) : D_2 \to \mathbb{R}$ is continuous at $= f(\mathbf{x}_0) \in D_2$, then $g \circ f(\mathbf{x})$ is continuous at $\mathbf{x}_0$.

♠

2.3 Differentiability

Consider the function

$$f(x, y) = x^2 + 2y^2$$

[1]Recall that a function can have limit without being defined on that point. But when we talk about continuity, the function must have definition on that point.

Suppose that know the function value $f(x_0, y_0)$ and would like to compute the function value $f(x_0+\Delta x, y_0+\Delta y)$. One computes directly that

$$\begin{aligned} f(x_0+\Delta x, y_0+\Delta y) &= (x_0+\Delta x)^2 + 2(y_0+\Delta y)^2 \\ &= \underbrace{x_0^2+2y_0^2}_{f(x_0,y_0)} + \underbrace{2x_0\Delta x + 4y_0\Delta y}_{Linear Increment} + \underbrace{(\Delta x)^2 + 2(\Delta y)^2}_{\text{Nonlinear Increment}} \end{aligned}$$

As a result, if we would like to approximate the function value $f(x_0+\Delta x, y_0+\Delta y)$, we can ignore the nonlinear increment and let

$$f(x_0+\Delta x, y_0+\Delta y) \simeq \underbrace{f(x_0.y_0) + 2x_0\Delta x + 4y_0\Delta y}_{Linear Approximation}.$$

Now we would like to generalise this to an arbitrary function, we are interseted in two questions:

- **What is the ideal candidate for the linear approximation.**
- **When is the linear approximation actually valid.**

Partial Derivative

Definition 2.3. Partial Derivative

Let $\mathbf{x} = (z_1, z_2, ...z_n) \in \mathbb{R}^n$ *and* $f(\mathbf{x})$ *be function defined on* $D \subset \mathbb{R}^n \to \mathbb{R}$. *We say that the function admits a partial derivative at* $\mathbf{x}_0 = (x_1, x_2, ..., x_n)$ *w.r.t.* x_i, *denoted by* $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$, *if the quantity*

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = \lim_{h\to 0} \frac{f(x_1, x_2, ..., x_i+h, ..., x_n) - f(x_1, x_2, ..., x_i, ..., x_n)}{h}$$

exist. ♣

In other words, one computes the partial derivatives w.r.t. x_i by considering all the variables as constants.

Example2.9 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x, y) = 3x^2y^3$$

It follows that on a given point $(x_0, y_0) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x_0, y_0) = 6x_0y_0^3,$$
$$\frac{\partial f}{\partial y}(x_0, y_0) = 9x_0^2y_0^2.$$

♦

Example2.10 Consider the function $f : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$,

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

It follows that we can compute the partial deritvatives at any point $(x, y) \neq (0,0)$ easily, yet we need to compute the partial derivatives at $(0,0)$ by using the **definition**. To this end, observe that

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h \to 0} \frac{f(0+h,0) - f(0,0)}{h} = \lim_{h \to 0} \frac{\frac{3 \cdot h^2 \cdot 0^3}{h^2 + 0^2} - 0}{h} = 0,$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h \to 0} \frac{f(0,0+h) - f(0,0)}{h} = \lim_{h \to 0} \frac{\frac{3 \cdot 0^2 \cdot h^3}{0^2 + h^2} - 0}{h} = 0.$$

♦

With the help of partial derivatives we can easily generalise the discussion in the beginning of this section on linear approximations to an arbitrary function in $\mathbb{R}^n$ whose partial derivatives exist. More precisely, let $\mathbf{x}^* = (x_1^*, x_2^*, ..., x_n^*) \in \mathbb{R}^n$ and consider a function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Let $\Delta\mathbf{x} \in \mathbb{R}^n$, we want to ask ourselves if the following approximation holds when $\|\Delta\mathbf{x}\|$ is small:

$$f(\mathbf{x}^* + \Delta\mathbf{x}) \simeq f(\mathbf{x}^*) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$$

i.e., can we use the linear approximation to describe the value of $f(\mathbf{x}^* + \Delta\mathbf{x})$ with a satisfactory accuracy?

It turns out that, to make linear approximation valid, existence of partial derivative is not enough: we need the function to be **differentiable**.

Differentiability

We first give the definition for a function to be differentiable in $\mathbb{R}^n$:

Definition 2.4. Differentiability

Let $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$. *We say that a function* $f(\mathbf{x}) : D \subseteq \mathbb{R}^n \to \mathbb{R}$ *is* ***differentiable*** *at* $\mathbf{x}^* = (x_1^*, x_2^*, ..., x_n^*)$*, if*

$$\lim_{\mathbf{x} \to \mathbf{x}^*} \frac{\left| f(\mathbf{x}) - f(\mathbf{x}^*) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \right|}{\|\mathbf{x} - \mathbf{x}^*\|} = 0$$

♣

Remark The definition of the differentiability is easy to understand if we interpret $\mathbf{x}$ as a variation of $\mathbf{x}$, i.e., Let $\mathbf{x} - \mathbf{x}^* = \Delta\mathbf{x} \in \mathbb{R}^n$, which means when written in coordinates:

$$\mathbf{x} = (x_1, x_2, ..., x_n) = \mathbf{x}^* + \Delta\mathbf{x} = (x_1^* + \Delta x_1, x_2^* + \Delta x_2, ..., x_n^* + \Delta x_n)$$

Hence we see that the definition simply means that the nonlinear increment is negligible w.r.t.

the variation of variable.

$$\lim_{\mathbf{x}\to\mathbf{x}^*} \frac{\overbrace{|\overbrace{f(\mathbf{x})-f(\mathbf{x}^*)}^{\text{Total Increment}} - \overbrace{\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\mathbf{x}^*)(\mathbf{x}-\mathbf{x}^*)}^{\text{Linear Increment}}|}^{\text{Nonlinear Increment=Total Increment-Linear Increment}}}{\underbrace{\|\mathbf{x}-\mathbf{x}^*\|}_{\text{variation of variable}}} = 0$$

2.4 Relations between Continuity and Differentiability

- We have seen that in $\mathbb{R}^1$, that the function is differentiable at a point means exactly that the function has derivative at that point.
- We will see that in $\mathbb{R}^n$, things are totally different: Requesting a function to be differentiable is much stronger than requesting a function to have merely partial derivatives.

Back to 1D

Before we formulate the relations between continuity and differentiability, let's first demonstrate some facts, in 1D case. Similar discussion works for $\mathbb{R}^n$.

Continuous but not Differentiable There exist functions in 1D which are continuous but not differentiable.

Example2.11 Consider the function $f:\mathbb{R}\to\mathbb{R}$,

$$f(x) = |x|.$$

The function is continuous at any point in $\mathbb{R}$, however at $x=0$ the function is not differentiable.

♦

Differentiable but not Continuously Differentiable There exist functions $f(x)$ in 1D which are continuous and is, even better, differentiable. However the derivative $f'(x)$, as a function of its own right, is not continuous.

Example2.12 Consider the function $f:\mathbb{R}\to\mathbb{R}$,

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

This function is clearly differentiable in $\mathbb{R}\setminus\{0\}$. However we can show that the derivative exists at 0 as well. Indeed, using the fact that $\sin\frac{1}{h}$ is a bounded function (i.e. $|\sin\dfrac{1}{h}|\leq 1$) a direct computation shows that

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

Hence $f'(0) = 0$. Somehow when $x \neq 0$, we see that

$$f'(x) = 2x\sin(\frac{1}{x}) - \cos\frac{1}{x}$$

The part $\cos\frac{1}{x}$ oscillates when $x \to 0$, hence

$$\lim_{x\to 0} f'(x)$$

does **NOT** exist, not to mention continuity at 0. In other words, f is differentiable, but its derivative is not continuous.

♦

Typical Counter-Examples in 2D

Now we can discuss the relation between the continuity, partial derivatives and differentiation in 2D. The best way to do this is again to remember some typical (counter-)examples.

First let's consider an example where the partial derivatives $\dfrac{\partial f}{\partial x}$ & $\dfrac{\partial f}{\partial y}$ exist, but f itself is not continuous (not to mention differentiable).

Example2.13 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$:

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

We see that

- The partial derivatives exist at $(0,0)$:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(0+h,0) - f(0,0)}{h} = \lim_{h\to 0}\frac{\frac{h\cdot 0}{h^2+0^2} - 0}{h} = 0$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0}\frac{f(0,0+h) - f(0,0)}{h} = \lim_{h\to 0}\frac{\frac{0\cdot h}{0^2+h^2} - 0}{h} = 0$$

- The function is not continuous at $(0,0)$, since the limit at $(0,0)$ does not exist, as is shown before in example 2.5.

♦

Next, let's consider an example where the partial derivatives $\dfrac{\partial f}{\partial x}$ & $\dfrac{\partial f}{\partial y}$ exist, f itself is also continuous, but still f is not differentiable.

Example2.14 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$:

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

We see that

- The partial derivatives exist at $(0,0)$:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{\frac{h\cdot 0}{\sqrt{h^2+0^2}}-0}{h} = 0$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h} = \lim_{h\to 0}\frac{\frac{0\cdot h}{\sqrt{0^2+h^2}}-0}{h} = 0$$

- The function f is continuous at $(0,0)$. Indeed, for any $\epsilon > 0$, take $\delta = 2\epsilon$, it follows that

$$0 < \sqrt{(x-0)^2+(y-0)^2} < \delta \Rightarrow \left|\frac{xy}{\sqrt{x^2+y^2}}\right| \le \frac{1}{2}\sqrt{x^2+y^2} < \frac{1}{2}\delta = \epsilon$$

  As a result, f has limit equal to 0 at (x,y)=(0,0), which means that

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

  i.e., f is continuous at $(0,0)$.

- The function f is **not** differentiable at $(0,0)$. To see this, using the definition, consider

$$\begin{aligned}
&\lim_{(x,y)\to(0,0)}\frac{f(x,y)-f(0,0)-\frac{\partial f}{\partial x}_{|(0,0)}(x-0)-\frac{\partial f}{\partial y}_{|(0,0)}(y-0)}{\sqrt{(x-0)^2+(y-0)^2}}\\
&= \lim_{(x,y)\to(0,0)}\frac{f(x,y)}{\sqrt{x^2+y^2}}\\
&= \lim_{(x,y)\to(0,0)}\frac{xy}{x^2+y^2},
\end{aligned}$$

  which do not exist, as we have just shown.

♦

We end this subsection by claiming the following theorem without proof.

Theorem 2.1. Differentiability v.s. Partial Derivatives

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Then

- *f is differentiable at $\mathbf{x}^*$ implies that all partial derivatives $\dfrac{\partial f}{\partial x_i}$ exists at $\mathbf{x}^*$.*
- *If all partial derivatives $\dfrac{\partial f}{\partial x_i}$, $1 \le i \le n$ exist and are continuous in an open ball $B_r(\mathbf{x}^*)$, then f is differentiable at $\mathbf{x}^*$.*

♡

We summarise their relations in the following diagram.

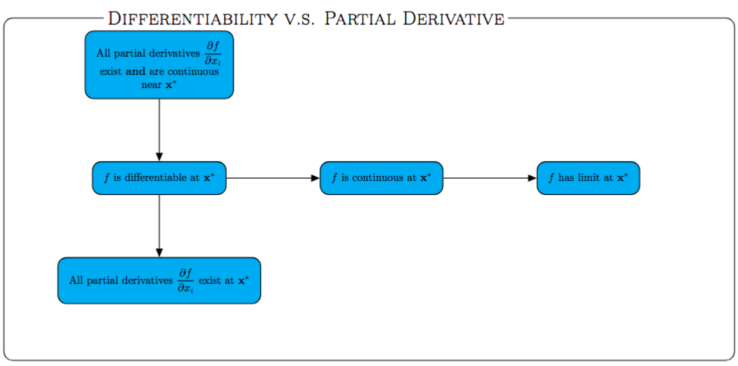


Figure 2.1: Differentiability v.s. Partial Derivatives

2.5 Properties of Continuous Functions

We list some properties of continuous functions without proof.

Proposition 2.3. Pre-image of Open set

Let $f : \mathbb{D} \subset \mathbb{R}$ be a continuous function. Then

- *For any open set $I \subseteq \mathbb{R}$, the pre-image $f^{-1}(I)$ is open as well.*
- *For any closed set $I \subseteq \mathbb{R}$, the pre-image $f^{-1}(I)$ is closed as well.*

Proof See assignment I.

Proposition 2.4. Max/Min

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a continuous function. If D is compact then f achieves its (finite) maximum and minimum on D.

Remark Recall that D is compact if it is both closed and bounded. Both of these two conditions are important, as are illustrated in the following examples.

Example2.15 Consider the function $f : [1, \infty) \to \mathbb{R}$, with

$$f(x) = \frac{1}{x}$$

Then $f(x)$ is continuous and is defined on a closed set (convince yourself that $[1, \infty)$ is a closed set!) Somehow, $f(x)$ can approximate 0 arbitrarily without achieving 0. The reason is that $[1, \infty)$ is **not bounded**.

Example2.16 Consider the function $f : (0, 1] \to \mathbb{R}$, with

$$f(x) = \frac{1}{x}$$

Then $f(x)$ is continuous and is defined on a bounded set. Somehow, $f(x)$ can approximate $+\infty$ arbitrarily without achieving any maximum. The reason is that $(0, 1]$ is **not closed**.

The above properties are already used implicitly in the optimisation with constraint. For example

Example2.17 Consider the optimisation problem

$$\begin{aligned} &\min xyz \\ &\text{s.t. } 2xy + 2yz + yz = 8 \\ &\quad 0 \le x, y, z \le 5 \end{aligned}$$

This is an example we have already studied in MAT232 using Lagrangian multipliers. Now without computation, we would like to show that the problem has a solution. Indeed, note that

- Let $g(x, y, z) : \mathbb{R}^3 \to \mathbb{R}$ be defined as

$$g(x, y, z) = 2xy + 2yz + yz$$

  Then it is a continuous function from $\mathbb{R}^3$ to $\mathbb{R}$, hence $S_8 = g^{-1}(8)$ is closed.

Moreover $A = \{(x, y, z) \in \mathbb{R}^3 | 0 \leq x, y, z \leq 5\}$ is bounded and closed, hence the optimisation for the continuous function $f(x, y, z) = xyz$ should be performed on the region

$$S_8 \cap A,$$

which is thus bounded (because subset of bounded set is bounded) and closed (the intersection of two closed sets is still closed), hence is compact. As a result there must exists a maximum as well as a minimum.

2.6 Properties of Differentiable Functions

The fact that a function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathbf{x}^* \in D$ permits us to study a possibly nonlinear function using two (equivalent) perspectives:

- **Analytical Point of View**: The function value near $\mathbf{x}_0$ permits a valid linear approximation:
$$f(\mathbf{x}) \simeq f(\mathbf{x}^*) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}^*)(x_i - x_i^*) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*).$$
- **Geometric Point of View**: The function graph near $(\mathbf{x}_0, f(\mathbf{x}_0))$ permits a valid tangent (hyper-) plane.
$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}^*)(x_i - x_i^*) - (z - f(\mathbf{x}^*)) = 0$$

Remark In 2D case, if $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$, then the linear approximation becomes:

- The analytical point of view: The valid linear approximation is
$$f(x, y) \simeq f(x^*, y^*) + \frac{\partial f}{\partial x}(x^*, y^*)(x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*)(y - y^*)$$
- The Geometric point of view: The valid tangent plane at $(x^*, y^*, f(x^*, y^*))$ becomes
$$\frac{\partial f}{\partial x}(x^*, y^*)(x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*)(y - y^*) - (z - f(x^*, y^*)) = 0$$

Example2.18 Linear Approximation Using the linear approximation to estimate $(0.99e^{0.02})^8$.

Solution Let $f(x, y) = (xe^y)^8$. then $f(x, y)$ is a differentiable function. Computation shows that

- $f(1, 0) = 1$.
- $\frac{\partial f}{\partial x}(1, 0) = 8(xe^y)^7 e^y|_{(1,0)} = 8$, $\frac{\partial f}{\partial y}(1, 0) = 8(xe^y)^7 xe^y|_{(1,0)} = 8$.

As a result, the linear approximation gives

$$\begin{aligned} f(0.99, 0.02) = f(1 + (-0.01), 0 + (0.02)) &\simeq f(1, 0) + \frac{\partial f}{\partial x}(1, 0)(-0.01) + \frac{\partial f}{\partial y}(1, 0)(0.02) \\ &= 1 + 8 \cdot (-0.01) + 8 \cdot (0.02) = 1.08. \end{aligned}$$

♦

Remark One can see that

- Total Increment $f(0.99, 0.02) - f(1, 0) = \mathbf{0.08285}$
- Linear Increment: $\frac{\partial f}{\partial x}(1, 0)(-0.01) + \frac{\partial f}{\partial y}(1, 0)(0.02) = \mathbf{0.08}$
- Nonlinear Increment: $0.08285 - 0.08 = \mathbf{0.00285}$

Thus the nonlinear increment is much less than the linear increment.

Example2.19 Tangent Plane Find the tangent plane of $f(x, y) = (xe^y)^8$ at $(1, 0)$.

Solution By the discussion in the beginning of this section, one sees that the equation for the plane should be

$$\frac{\partial f}{\partial x}(x^*, y^*)(x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*)(y - y^*) - (z - f(x^*, y^*)) = 0$$

Now $(x^*, y^*) = (1, 0)$ so that the equation is

$$8(x - 1) + 8(y - 0) - (z - 1) = 0$$

i.e.

$$8x + 8y - z = 7$$

♦

2.7 Differential

Definition 2.5. Differential

Given a differentiable function $f(\mathbf{x}) : D \subseteq \mathbb{R}^n \to \mathbb{R}$, its differentiable at a point $\mathbf{x}^* = (x_1^*, x_2^*, ..., x_n^*)$ is

$$df(\mathbf{x}^*) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}^*) dx_i$$

♣

Remark

- dx_i, strictly speaking, is **NOT** the same thing as Δx_i. It should be rather considered as a linear function on Δx_i. However, in this course one can consider it as Δx_i in the computation.
- df should rather be considered as a vector. Indeed, one can consider dx_i as a basis of some vector space, just as the basis $\mathbf{e}_1, \mathbf{e}_2, ...\mathbf{e}_n$ for $\mathbb{R}^n$.
- In $1D$, we have that
$$\frac{df}{dx} = f'(x)dx$$
which makes sense because all vectors in $\mathbb{R}^1$ are collinear, i.e., one can get one vector $\mathbf{v}_1$ from another vector $\mathbf{v}_2$ by multiplying a scalar $\lambda \in \mathbb{R}$. Somehow in $\mathbb{R}^n$, $\dfrac{\mathbf{v}_1}{\mathbf{v}_2}$ might not make sense.
- Note that one can consider $df(\mathbf{x})$ as a function on the increment vector $\Delta\mathbf{x}$, defined as
$$df(\mathbf{x})(\Delta\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \Delta\mathbf{x}.$$

Example2.20 Find the differentials of f and evaluate them at given points.

1. $f(x,y) = y^2 \log(x^2+1)$ at $(x,y) = (0,1)$.
2. $f(x,y,z) = xyz$ at $(x,y,z) = (1,1,2)$.

Solution

1. $df(x,y) = y^2 \dfrac{2x}{x^2+1} dx + 2y\log(x^2+1)dy$. Hence $df(0,1) = 0dx + 0dy = \mathbf{0}$.
2. $df(x,y,z) = yzdx + xzdy + xydz$. Hence $df(1,1,2) = 2dx + 2dy + dz$.

2.8 Higher Order Partial Derivatives

Note that $\frac{\partial f}{\partial x}(x,y)$ is a function of multiple variables of its own right. Hence we can also define higher order derivatives, for example

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)(x,y) = \lim_{h\to 0} \frac{\frac{\partial f}{\partial x}(x+h,y) - \frac{\partial f}{\partial x}(x,y)}{h}$$

Remark Note that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right),$$

Yet

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

Using another equivalent symbols,

$$f_{xy} = (f_x)_y, \quad f_{yx} = (f_y)_x.$$

The following theorem guarantees that we can exchange the order when computing the partial derivatives of the 2nd order without any risk:

Theorem 2.2. Clairaut's Theorem

Suppose that f_{xy} and f_{yx} both exist and are continuous at (x_0, y_0), then

$$f_{xy}(x_0,y_0) = f_{yx}(x_0,y_0)$$

♡

Example2.21 Mixed higher order derivative where order does not matter Let $f(x,y) = e^x \sin(xy)$ which is a differentiable function. One computes that

$$\begin{aligned} f_{xy} &= (e^x \sin(xy) + e^x y \cos xy)_y = xe^x \cos(xy) + e^x(\cos(xy) - xy\sin(xy)) \\ &= e^x\left((x+1)\cos(xy) - xy\sin(xy)\right) \\ f_{yx} &= (xe^x \cos(xy))_x = (e^x + xe^x)\cos(xy) - xye^x \sin(xy) \\ &= e^x\left((x+1)\cos(xy) - xy\sin(xy)\right) \end{aligned}$$

As one can see the order does not matter – thanks to the smoothness of the function f.

♦

Example2.22 Mixed higher order derivative where order does matter Consider the function

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

One verifies that $f_{yx}(0,0) = 1 \neq -1 = f_{xy}(0,0)$. (Do it by yourself using the definitions)

♦

2.9 Taylor's Theorem

We have already seen in MAT232 that sometimes a function's first order derivative is not enough to determine the behaviour of a function near some points. The higher order derivatives might also contain important information that should not be dropped. The Taylor's theorem gives us a systematical way of approximating the function using higher order derivative.

Theorem 2.3. Taylor's Theorem

1. *Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\mathbf{x}_0 \in \mathbb{D}$, then for the increment vector $\mathbf{\Delta x} \in \mathbb{R}^n$,*

$$f(\mathbf{x}_0 + \Delta\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\Delta x_i + R_1(\mathbf{x}_0, \Delta\mathbf{x}),$$

where

$$\lim_{\Delta\mathbf{x}\to\mathbf{0}} \frac{R_1(\mathbf{x}_0, \Delta\mathbf{x})}{\|\Delta\mathbf{x}\|} = 0$$

2. *Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be twice differentiable at $\mathbf{x}_0 \in \mathbb{D}$, then for the increment vector $\mathbf{\Delta x} \in \mathbb{R}^n$,*

$$f(\mathbf{x}_0+\Delta\mathbf{x}) = f(\mathbf{x}_0)+\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\Delta x_i+\frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\Delta x_i \Delta x_j+R_2(\mathbf{x}_0, \Delta\mathbf{x}),$$

where

$$\lim_{\Delta\mathbf{x}\to\mathbf{0}} \frac{R_2(\mathbf{x}_0, \Delta\mathbf{x})}{\|\Delta\mathbf{x}\|^2} = 0$$

♡

Remark The Taylor theorem means that

- If f is differentiable at $\mathbf{x}_0$, then the **best linear approximation** of $f(x)$ near $\mathbf{x}_0$ is

$$f(\mathbf{x}) = f(\mathbf{x}_0 + \Delta\mathbf{x}) \simeq f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\Delta x_i$$

- If f is twice differentiable at $\mathbf{x}_0$, then the **best quadratic approximation** of $f(x)$ near $\mathbf{x}_0$ is

$$f(\mathbf{x}) = f(\mathbf{x}_0 + \Delta\mathbf{x}) \simeq f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\Delta x_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\Delta x_i \Delta x_j$$

Example2.23 Best Linear/Quadratic Approximation Find the linear function $f_1(x, y)$ and the quadratic function on $f_2(x, y)$ which best approximates the function

$$g(x, y) = \sin(xy) - 3x^2 \log y + 1$$

near the point $(\frac{\pi}{2}, 1)$.

Solution

$$g(\frac{\pi}{2}, 1) = 1 - 0 + 1 = 2$$

$$\frac{\partial g}{\partial x}_{|(\frac{\pi}{2},1)} = y\cos(xy) - 6x\log y = 0$$
$$\frac{\partial g}{\partial y}_{|(\frac{\pi}{2},1)} = x\cos(xy) - \frac{3x^2}{y} = -\frac{3}{4}\pi^2$$

Next we compute the second order derivatives

$$\frac{\partial^2 f}{\partial x^2}_{|(\frac{\pi}{2},1)} = -y^2\sin(xy) - 6\log y = -1.$$
$$\frac{\partial^2 f}{\partial y^2}_{|(\frac{\pi}{2},1)} = -x^2\sin(xy) + \frac{3x^2}{y^2} = \frac{\pi^2}{2}.$$
$$\frac{\partial^2 f}{\partial x\partial y}_{|(\frac{\pi}{2},1)} = \cos(xy) - xy\sin(xy) - \frac{6x}{y} = -\frac{7}{2}\pi$$

It follows that

$$f_1(x,y) = \underbrace{2}_{f(x_0,y_0)} + \underbrace{0\cdot(x-\frac{\pi}{2}) - \frac{3}{4}\pi^2(y-1)}_{\text{Linear Increment}}$$

$$f_2(x,y) = \underbrace{2}_{f(x_0,y_0)} + \underbrace{0\cdot(x-\frac{\pi}{2}) - \frac{3}{4}\pi^2(y-1)}_{\text{Linear Increment}}$$
$$+\underbrace{\frac{1}{2}\left(-1\cdot(x-\frac{\pi}{2})^2 + 2\cdot(-\frac{7}{2}\pi)(x-\frac{\pi}{2})(y-1) + \frac{\pi^2}{2}(y-1)^2\right)}_{\text{Quadratic Increment}}$$

♦

Chapter 3 Parametric Curve and Parametric Surfaces

3.1 Motivations

Recall that the graph of a function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is

$$Graph(f) = \{(\mathbf{x}, f(\mathbf{x}))|\mathbf{x} \in D\}$$

The reason to study parametric curves and parametric surfaces are that:

- **Not** every curve/surface is a **graph of a function**.
- If we restrict ourselves to graph of a function, then the power of vector calculus will be too limited.

Example3.1 Consider the circle

$$S = \{(x, y)|x^2 + y^2 = 1\}$$

It is easy to observe that

- It is **not** a graph of any function.
- It can be parameterised as $x = \cos t, y = \sin t, t \in [0, 2\pi]$, hence is a parametric curve.
- With the parametric curve, we can compute its interior area as well as the perimeter arc length using some calculus tool.
- It can be **locally** regarded as a graph near any point (x, y) s.t. $y \neq 0$.

♦

All the following contents in the rest of this chapter aims at systematically generalising the above observations.

3.2 Parametric Curve

Definition 3.1. Parametric Curve

A function $t \to \mathbf{x}(t) = (x_1(t), x_2(t), ..., x_n(t)) \in \mathbb{R}^n$ is called a parametric curve. ♣

Example3.2 Consider the curve $\Gamma : y = x^2, -1 \le x \le 1$. One can parametrise it by

$$\begin{aligned} x(t) &= t \\ y(t) &= t^2, \quad t \in [-1, 1] \end{aligned}$$

♦

Remark It is important to distinguish between the curves and the parametric curves.

- Curve $\mathbf{\Gamma}$: it is a **subset of** $\mathbb{R}^n$.

- Parametric Curve $\gamma(t)$: it is a **Vector-valued Function**.

Example3.3 Consider the curve $\Gamma : y^2 = x, -1 \leq y \leq 1$. One can parameterise it by

$$x(t) = t^2$$
$$y(t) = t, \quad t \in [-1, 1]$$

Compared to the previous example 3.2, here note that $y^2 = x$ is no longer a graph.

♦

Remark This example tells us that although the curve Γ might not be a graph, the parametric curve describing the curve might still exist.

Example3.4 Consider the curve $\Gamma : y = x^2, -1 \leq x \leq 1$. One can parameterise it by

$$x(t) = \sin t$$
$$y(t) = \sin^2 t, \quad t \in [0, 2\pi]$$

If we consider the curve as a subset of $\mathbb{R}^2$, it is exactly the same as example 3.2. We cannot see any difference. Somehow with the parameter t one sees different dynamical information.

- In example 3.2, the point start from (-1,1) and then ends up to $(1, 1)$
- In this example, the points starts from $(0, 0)$, then go to $(1, 1)$, back to $(0, 0)$, and then go to $(-1, 1)$, finally back to $(0, 0)$

♦

Remark This example indicates that the parametric curve does not only provide stationary information about the positions of points on the curve, it reveals useful dynamical information as well.

Example3.5 Now for the circle

$$S = \{(x, y)|x^2 + y^2 = 1\}$$

The curve can be parameterised in infinitely many ways. For example

1. $x(t) = \cos t$, $y(t) = \sin t$. Under this parametrisation, the curve starts from $(1, 0)$, moves in uniform speed in the *counter-clockwise* sense, travels along the whole circle and ends back to $(1, 0)$.
2. $x(t) = \cos(2\pi - t)$, $y(t) = \sin(2\pi - t)$. Under this parametrisation, the curve starts from $(1, 0)$, moves in uniform speed in the *clockwise sense*, and ends back to $(1, 0)$.
3. $x(t) = \cos 2t$, $y(t) = \sin 2t$. Under this parametrisation, the curve starts from $(1, 0)$, moves in uniform speed in the *counter-clockwise* sense but **twice** faster than the first parametrisation, and ends back to $(1, 0)$.

♦

Remark Given a parametric curve, it uniquely determines the curve as a subset of $\mathbb{R}^2$. Somehow for the same curve, one has infinitely many ways to parametrise it.

3.3 Velocity

Definition 3.2. Velocity

Given a parametric curve $\mathbf{x}(t), t \in I$, *the* ***velocity*** *vector at time* t_0 *is defined as*

$$\frac{d\mathbf{x}}{dt}(t_0) = (\frac{dx_1}{dt}(t_0), \frac{dx_2}{dt}(t_0), ..., \frac{dx_n}{dt}(t_0)),$$

if all the coordinates $x_i : I \to \mathbb{R}$ *are differentiable at* $t_0 \in I$. ♣

Remark Velocity is a tangent vector of the **curve** at some point, but not every tangent vector of the curve is velocity of the corresponded **parametric curve**. In particular, once the parametrisation is determined, there is only one vector at t_0 that serves as the velocity.

Example3.6 Velocity vs Tangent Vector Consider the curve

$$\Gamma = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 = 1, z = 0\}$$

Then a parametrisation could be

$$\gamma(t) = (\cos t, \sin t, 0)$$

Under this parametrisation,

$$\frac{d\gamma}{dt}_{|t=\frac{\pi}{2}} = (-1, 0, 0)$$

is the velocity of the parametric curve at time $\frac{\pi}{2}$. It is also a tangent vector of the curve at $(-1, 0, 0)$. However the vector $(a, 0, 0)$ for any $a \in \mathbb{R}$ is also a tangent vector of the curve at the point $(0, 1, 0)$. If we parametrise the curve by

$$\gamma_1(t) = (\cos 2t, \sin 2t, 0),$$

Then the velocity vector at $t = \frac{\pi}{4}$ is

$$\frac{d\gamma_1}{dt}_{|t=\frac{\pi}{4}} = (-2, 0, 0),$$

which is again tangent to the curve Γ at $(0, 1, 0)$, somehow the magnitude is twice as that of the velocity vector $(-1, 0, 0)$ in the previous parametrisation.

♦

Definition 3.3. Speed

Given a parametric curve $\mathbf{x}(t), t \in I$, *the* ***speed*** *at time* t_0 *is defined as*

$$\|\frac{d\mathbf{x}}{dt}(t_0)\| = \sqrt{\left(\frac{dx_1}{dt}(t_0)\right)^2 + \left(\frac{dx_2}{dt}(t_0)\right)^2, ..., \left(\frac{dx_n}{dt}(t_0)\right)^2},$$

if all the coordinates $x_i : I \to \mathbb{R}$ *are differentiable at* $t_0 \in I$. ♣

Example3.7 Velocity vs Speed Still consider the previous example 3.6. For the first parametrisation

$$\frac{d\gamma}{dt}|_{t=\frac{\pi}{2}} = (-1, 0, 0)$$

Hence the speed at time $t = \frac{\pi}{2}$ is $\left\|\frac{d\gamma}{dt}\right\|_{t=\frac{\pi}{2}} = \|\sqrt{(-1)^2 + 0 + 0}\| = 1.$

♦

Remark velocity is a **vector**, while speed is a **scalar**.

3.4 Arc Length

Definition 3.4. Arc Length

For a given curve Γ, let $\gamma(t)$,$t \in [a, b]$ be a parametrisation of the curve. The ***arc length*** *of the curve γ is defined to be*

$$\int_a^b \|\frac{d\gamma}{dt}(t)\|dt$$

♣

Our first question is that: is it a well-defined object? More precisely speaking, we know that given a curve, there exists infinitely many parametrisation. Do they always give the same arc length?

Proposition 3.1. Arc Length is Independent of Parametrisation

Let Γ be a curve, $\gamma_1(t), t \in [a, b]$ is a parametrisation of Γ. Suppose that $s = h(t)$ is a monotonically increasing differentiable function of t, with $h(a) = c, h(b) = d$, then if we consider $\gamma_2(s), s \in [c, d]$ ***defined as***

$$\gamma_2(s) = \gamma_1(h^{-1}(s)),$$

Then the arc length computed in these two parametrisation satisfies that

$$\int_a^b \|\frac{d\gamma_1}{dt}(t)\|dt = \int_c^d \|\frac{d\gamma_2}{ds}(s)\|ds$$

i.e., the arc length is independent of parametrisation.

♠

Remark $\gamma_2(s), s \in [c, d]$ constructed in this way can be seen as a continous time scale change-ment, and is called a **reparametrisation** of $\gamma_1(t), t \in [a, b]$.

Proof By substitution, one has that

$$\int_a^b \|\dot{\gamma}_1(t)\|dt = \int_a^b \sqrt{\left(\frac{dx_1}{dt}(t)\right)^2 + \left(\frac{dx_2}{dt}(t)\right)^2 + ... + \left(\frac{dx_n}{dt}(t)\right)^2} dt$$

Note that

$$\frac{dx_i}{dt} = \frac{dx_i}{ds}\frac{ds}{dt} \qquad \text{(Chain Rule)}$$

Hence

$$
\begin{aligned}
&\int_a^b \sqrt{\left(\frac{dx_i}{dt}\right)^2 + \left(\frac{dx_i}{dt}\right)^2 + \ldots + \left(\frac{dx_i}{dt}\right)^2}\, dt \\
=& \int_c^d \sqrt{\left(\frac{dx_i}{ds}\frac{ds}{dt}\right)^2 + \left(\frac{dx_i}{ds}\frac{ds}{dt}\right)^2 + \ldots + \left(\frac{dx_i}{ds}\frac{ds}{dt}\right)^2}\, \frac{dt}{ds} ds \\
=& \int_c^d \frac{ds}{dt} \sqrt{\left(\frac{dx_i}{ds}\right)^2 + \left(\frac{dx_i}{ds}\right)^2 + \ldots + \left(\frac{dx_i}{ds}\right)^2}\, \frac{dt}{ds} ds \\
=& \int_c^d \sqrt{\left(\frac{dx_i}{ds}\right)^2 + \left(\frac{dx_i}{ds}\right)^2 + \ldots + \left(\frac{dx_i}{ds}\right)^2}\, ds
\end{aligned}
$$

□

Now since we have already justified the definition of arc length is independent of the parametrisation, we can now parametrise the curve using a special parameter, i.e., its arc length.

Definition 3.5. Parametrisation by Arc Length

Let Γ be a curve and $\gamma(t)$ be a parametrisation of the curve. Let s be a function of t s.t.

$$\frac{ds}{dt} = \left\|\frac{d\gamma}{dt}\right\|$$

Then Γ parametrised by s is called the ***parametrisation by arc length****.* ♣

Remark In practice one can parametrise the curve using the arc length in the following steps:

1. Using $\gamma(t)$, $t \in [a, b]$ to compute the partial arc length up to time t:
$$s = h(t) = \int_a^t \left\|\frac{d\gamma}{dt}(u)\right\| du$$
2. From the above, solve the inverse function $t = h^{-1}(s)$.
3. Now define the new parametrisation curve $\gamma^*(s)$, $s \in [0, L]$ by
$$\gamma^*(s) = \gamma(h^{-1}(s)).$$
where
$$L = \int_a^b \left\|\frac{d\gamma}{dt}(u)\right\| du$$
is the total arc length of the curve Γ.

Then $\gamma^*(s)$, $s \in [0, L]$ is the arc-length parametrised curve we are looking.

Example3.8 Consider a helix parametrised by

$$\gamma(t) = (\cos t, \sin t, t), t \in [0, 2\pi]$$

Reparametrise it by arc length.

Solution The arc length from 0 to any $t \in [0, 2\pi]$ is

$$s = h(t) = \int_0^t \|\frac{d\gamma}{dt}(u)\| du = \int_0^t \sqrt{\sin^2 u + \cos^2 u + 1} du = \sqrt{2}t$$

From there one solves that

$$t = h^{-1}(s) = \frac{\sqrt{2}}{2} s$$

Hence the reparametrisation by arc length is given by

$$\gamma^*(s) = \gamma(h^{-1}(s)) = (\cos\frac{\sqrt{2}}{2}s, \sin\frac{\sqrt{2}}{2}s, \frac{\sqrt{2}}{2}s), s \in [0, 2\sqrt{2}\pi]$$

♦

Example3.9 Consider the cycloid parametrised by

$$x(t) = t - \sin t, \quad y(t) = 1 - \cos t, t \in [0, 2\pi]$$

Parametrise the curve using the arc length.

Solution One has that the arc length from 0 to any $t \in [0, 2\pi]$ is

$$s = h(t) = \int_0^t \sqrt{(1-\cos t)^2 + \sin^2 t}dt = \int_0^t \sqrt{2 - 2\cos t}dt = \int_0^t 2\sin\frac{t}{2}dt = -4\cos\frac{t}{2}|_0^t$$
$$= 4(1 - \cos\frac{t}{2})$$

From there one solves that

$$t = h^{-1}(s) = 2\arccos(1 - \frac{s}{4}).$$

Hence the parametrisation by arc length is

$$x(s) = 2\arccos(1 - \frac{s}{4}) - \sin(2\arccos(1 - \frac{s}{4}))$$
$$y(s) = 1 - \cos(2\arccos(1 - \frac{s}{4}))$$

♦

3.5 Curvature

An Intuitive Measure of the Shape

Suppose that we want a measure for how fast the curve is changing its direction. Since the direction is closely related to the tangent vector, thus to the velocity, a natural candidate is to parametrise the curve by $\gamma(t)$ and compute $\|\ddot{\gamma}(t)\| = \|\frac{d}{dt}\dot{\gamma}(t)\|$. Somehow there is an evident problem: how fast the curve is changing its measure should not depend on the parametrisation of the curve.

Example3.10 Consider you go along a sharp turn. There are two situations:

- You go slowly with your dog as a promenade after dinner. You will probably not feel the sharpness of the turn, as the acceleration is very small.
- You go rapidly with your Ducati Panigale like a flash. Be careful you will definitely feel the sharpness of the turn, as your acceleration is very huge (hopefully not huge enough to cause an accident).

However, the turn is just there, how sharp it is does not really depend the way you go across it.

As a result, we should take a very special parametrisation. This parametrisation takes the information of arc-length itself into consideration. i.e., we should consider the parametrisation by arc-length.

- we replace dt by ds.
- The unit velocity vector $T(t) = \frac{\frac{d\gamma}{dt}}{\|\frac{d\gamma}{dt}\|}$, which evaluates the direction of the velocity, becomes $T(s) = \frac{\frac{d\gamma}{ds}}{\|\frac{d\gamma}{ds}\|} = \dfrac{d\gamma}{ds}$ — because $\|\frac{d\gamma}{ds}\| = 1$;

The acceleration vector now becomes $\frac{d^2\gamma}{ds^2}$ — that sounds like a promising candidate now.

The Definition of Curvature

From now on, we use $\kappa(t_0)$ to represent the curvature of the point $\gamma(t_0)$ on a parametric curve $\gamma(t), t \in [a, b]$.

Definition 3.6. Curvature (defined via parametrisation by arc length)

Let $\gamma(s)$ be the parametrisation of the curve by arc length. The curvature of a curve at $\gamma(s_0)$ is defined to be

$$\kappa(s_0) = \left\| \frac{dT}{ds} \right\|_{|s=s_0}$$

♣

Remark The above definition, although explained the meaning of curvature in a fairly straightforward way, is not suitable for computation in practice. Indeed, you might even want to avoid use this definition in your computation of curvature for a particular curve. The reason is that, to use this definition, one must first parametrise the curve by arc-length, which involves in computing a definite integral explicitly, which could be a nightmare! One rule of thumb in calculus is that : differentiation is easier to compute that integration.

As a result, we should look for an approach that permits us to compute the curvature at a given point (or equivalently, at a given time spot in the time interval of the parametric curve)

Indeed, by considering everything as a function of t (the original parameter), one sees that

$$\kappa(t) = \left\| \frac{dT}{ds}(s(t)) \right\| = \left\| \frac{dT/dt}{ds/dt}(t) \right\|,$$

i.e., the following equivalent definition:

Definition 3.7. Curvature (defined via any parametrisation)

Let $\gamma(t)$ be an arbitrary parametrisation of the curve. The curvature of a curve at $\gamma(t_0)$ is defined to be $\kappa(t_0) = \dfrac{\left\|\frac{d\mathbf{T}}{dt}\right\|}{\left\|\frac{d\gamma}{dt}\right\|}_{|t=t_0}$.

♣

Computation of Curvature

The second definition of curvature is still not what could use directly in the computation. Although we now have the expression of the curvature with the help of the arbitrary given parametrisation, and we no longer need to solve the function of reparametrisation $t = h^{-1}s$ through a definite integral. Somehow the information we need is $\frac{dT}{dt}$, which is not directly expressed in terms of $\gamma(t)$ and its derivatives (w.r.t. the parameter t). To get a more practical formula for computation, let's first realised the following general fact:

Lemma 3.1

Suppose that $\mathbf{a}(t) : \mathbb{R} \to \mathbb{R}^n$ is such that

$$\|\mathbf{a}(t)\| = c,$$

where $c \geq 0$ is a constant. Then $\dfrac{d\mathbf{a}}{dt}$ and $\mathbf{a}$ are orthogonal at each $t \in \mathbb{R}$. ♡

Proof Let $\mathbf{a}(t) = (a_1(t), a_2(t), ..., a_n(t)) \in \mathbb{R}^n$. Then $\|\mathbf{a}(t)\|^2 = c^2$, which means that

$$\sum_{i=1}^{n} a_i^2(t) = c^2.$$

Differentiate with respect to t on both hand sides results in

$$2\sum_{i=1}^{n} a_i(t)\frac{d}{dt}a_i(t) = 0,$$

i.e., $\mathbf{a}(t) \cdot \dfrac{d\mathbf{a}(t)}{dt} = 0.$

□

Now by definition

$$\frac{d\gamma}{dt} = \frac{\frac{d\gamma}{dt}}{\|\frac{d\gamma}{dt}\|}\|\frac{d\gamma}{dt}\| = \mathbf{T}\frac{ds}{dt}$$
$$\frac{d^2\gamma}{dt^2} = \frac{d\mathbf{T}}{dt}\frac{ds}{dt} + \mathbf{T}\frac{d^2s}{dt^2}$$

It follows that

$$\frac{d\gamma}{dt} \times \frac{d^2\gamma}{dt^2} = \mathbf{T} \times \frac{d\mathbf{T}}{dt}\left(\frac{ds}{dt}\right)^2 \qquad (\mathbf{T} \times \mathbf{T} = 0)$$

Now taking norms on both hand sides, one has that

$$\|\frac{d\gamma}{dt} \times \frac{d^2\gamma}{dt^2}\| = \|\mathbf{T} \times \frac{d\mathbf{T}}{dt}\left(\frac{ds}{dt}\right)^2\| = \|\mathbf{T}\|\|\frac{d\mathbf{T}}{dt}\left(\frac{ds}{dt}\right)^2\| = \|\frac{d\mathbf{T}}{dt}\|\|\left(\frac{ds}{dt}\right)^2\|$$

Thus

$$\frac{\|\frac{d\mathbf{T}}{dt}\|}{\|\frac{d\gamma}{dt}\|} = \frac{\|\frac{d\gamma}{dt} \times \frac{d^2\gamma}{dt^2}\|}{\|\left(\frac{ds}{dt}\right)^3\|} = \frac{\|\frac{d\gamma}{dt} \times \frac{d^2\gamma}{dt^2}\|}{\|\frac{d\gamma}{dt}\|^3}$$

We conclude that

Proposition 3.2. Computation of Curvature in $3D$

Let $\gamma(t)$ be a parametric curve in $3D$. Then

$$\kappa(t) = \frac{\|\frac{d\gamma}{dt} \times \frac{d^2\gamma}{dt^2}\|}{\|\frac{d\gamma}{dt}\|^3}$$

♠

We see that the above formula is useful in practice because it uses only the information on $\gamma(t)$ and its derivatives, hence is easy to compute. We discuss a few special cases, before going to examples.

When the curve is 2D

If the curve is in 2D, i.e., the parametrisation is given as $\gamma(t) = (x(t), y(t))$, then we can embeds this vector into 3D by writing it as (still denoted by $\gamma(t)$)

$$\gamma(t) = (x(t), y(t), 0)$$

In this case, since

$$\begin{aligned}
\|\frac{d\gamma}{dt} \times \frac{d^2\gamma}{dt^2}\| &= \left\|(\frac{dx}{dt}, \frac{dy}{dt}, 0) \times (\frac{d^2x}{dt^2}, \frac{d^2y}{dt^2}, 0)\right\| = \left\|(0, 0, \frac{d^2y}{dt^2}\frac{dx}{dt} - \frac{d^2x}{dt^2}\frac{dy}{dt})\right\| \\
&= \left|\frac{d^2y}{dt^2}\frac{dx}{dt} - \frac{d^2x}{dt^2}\frac{dy}{dt}\right|, \\
\left\|\frac{d\gamma}{dt}\right\| &= \sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}
\end{aligned}$$

One has the following simplified formula:

Proposition 3.3. Computation of Curvature in $2D$

Let $\gamma(t)$ be a parametric curve in $2D$. Then

$$\kappa(t) = \frac{\left|\frac{d^2y}{dt^2}\frac{dx}{dt} - \frac{d^2x}{dt^2}\frac{dy}{dt}\right|}{\left(\sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}\right)^3}$$

Finally if $y = F(x)$, the parametric curve is a graph and we can use x itself as the parameter, i.e. let

$$x = x, y = F(x)$$

The above formula then becomes an even simpler form:

Proposition 3.4. Computation of Curvature in $2D$ for Graph

Let $\gamma(t)$ be a parametric curve in $2D$ which is the graph of $y = F(x)$. Then

$$\kappa(t) = \frac{\left|\frac{d^2F}{dx^2}\right|}{\left(\sqrt{1 + (\frac{dF}{dx})^2}\right)^3}$$

Example3.11 Straight Line Let l be a straight line, one can parametrise it by

$$x = t, \quad y = 0, \quad t \in \mathbb{R}$$

Now

$$\frac{dx}{dt} = 1, \frac{dy}{dt} = 0, \frac{d^2x}{dt^2} = 0, \frac{d^2y}{dt^2} = 0.$$

It follows that

$$\kappa(t) = 0.$$

This is reasonable as the straight line is not curved at all.

◆

Example3.12 Circle Let $\gamma(t) = (x = r\cos t, y = r\sin t), t \in [0, 2\pi)$, which represents a circle of radius r. Direct computation shows that

$$\frac{dx}{dt} = -r\sin(t), \frac{dy}{dt} = r\cos t, \frac{d^2x}{dt^2} = -r\cos t, \frac{d^2y}{dt^2} = -r\sin t.$$

It follows that

$$\kappa(t) = .\frac{r|\sin^2 t + \cos^2 t|}{(r\sqrt{\cos^2 t + \sin^2 t})^3} = \frac{1}{r}$$

As a result,

- When $r \to \infty$, the curvature is going to 0.
- When $r \to 0$, the curvature is going to $+\infty$.

Intuitively this is also as expected: after all a straight line can be considered as a circle with infinitely huge radius.

♦

Remark The straight lines and the circles both have curvatures that are constant along the curve. These curves are called **curves of constant curvature**. Indeed, it can be shown that these two are also the only curves of constant curvature on the plane.

Example3.13 Ellipse Consider the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

A parametrisation could be

$$x = a\cos t, y = b\sin t, t \in [0, 2\pi).$$

Similar computation as that in the circle shows that

$$\frac{dx}{dt} = -a\sin t, \frac{dy}{dt} = b\cos t, \frac{d^2x}{dt^2} = -a\cos t, \frac{d^2y}{dt^2} = -b\sin t.$$

It follows that

$$\kappa(t) = .\frac{ab|\sin^2 t + \cos^2 t|}{(\sqrt{a^2\cos^2 t + b^2\sin^2 t})^3} = \frac{ab}{(\sqrt{a^2\cos^2 t + b^2\sin^2 t})^3}$$

One sees that

- When $a = b$, the ellipse becomes the circle, and the curvature becomes $\frac{1}{a}$ and is thus a constant.
- When $a \neq b$, we may assume without loss of generality that $a > b$. Then
  - The curvature is minimised at $t \in \{\frac{\pi}{2}, \frac{3\pi}{2}\}$, with $\kappa(\frac{\pi}{2}) = \kappa(\frac{3\pi}{2}) = \frac{b}{a^2}$.
  - The curvature is maximised at $t \in \{0, \pi\}$, with $\kappa(0) = \kappa(\pi) = \frac{a}{b^2}$.

  This also very reasonable as one sees clearly that an ellipse becomes more sharp when it comes to vertices on the long ends and more flat when it comes to vertices on the short ends.

♦

Example3.14 Find the curvature of $\mathbf{r}(t) = (t^2, \ln t, t\ln t)$ at the point $(1, 0, 0)$.

Solution First, note that the curve passes the point $(1,0,0)$ when $t=1$. Direct computation shows that

$$\dot{\mathbf{r}}(t) = (2t, \frac{1}{t}, \ln t + 1)$$
$$\ddot{\mathbf{r}}(t) = (2, -\frac{1}{t^2}, \frac{1}{t})$$

Now one has that

$$\kappa(1) = \frac{\|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}\|}{\|\dot{\mathbf{r}}\|^3}\Big|_{t=1} = \frac{\|(2,1,1)\times(2,-1,1)\|}{\|\sqrt{2^2+(-1)^2+1^2}\|^3} = \frac{\sqrt{20}}{6\sqrt{6}} = \frac{\sqrt{30}}{18}.$$

♦

Remark If you want to use the original definition, i.e. parametrisation by arc-length to compute the curvature, you will need to solve the integral

$$s = \int_0^t \|\dot{\mathbf{r}}(u)\| du = \int_0^t \sqrt{4u^2 + \frac{1}{u^2} + (\ln u + 1)^2} du$$

which is not an easy task.

3.6 Parametric Surface

Now we generalise the idea of parametric curve to parametric surface. For a curve we need one parameter, for example t, as well as two functions $(x(t), y(t))$. For a surface we need two parameters, for example u, v and will have three functions $(x(u,v), y(u,v), z(u,v))$.

Example3.15 Sphere Consider the surface given by the equation

$$x^2 + y^2 + z^2 = a^2. \quad a > 0$$

One can use the following parametrisation (θ, ϕ)

$$x = a\sin\theta\cos\phi, y = a\sin\theta\sin\phi, z = a\cos\theta, \quad \theta \in [0,\pi], \phi \in [0, 2\pi]$$

note that θ is essentially the latitude and ϕ is essentially the longitude.

♦

Example3.16 Ellipsoid Consider the surface given by the equation

$$4x^2 + 4y^2 + z^2 = a^2. \quad a > 0$$

One can use the following parametrisation (θ, ϕ)

$$x = \frac{a}{2}\sin\theta\cos\phi, y = \frac{a}{2}\sin\theta\sin\phi, z = a\cos\theta, \quad \theta \in [0,\pi], \phi \in [0, 2\pi]$$

♦

Example3.17 Hyperboloid Consider the surface given by the equation

$$x^2 + y^2 - z^2 = a^2. \quad a > 0, z > 0$$

One can use the following parametrisation by (θ, ϕ)

$$x = a\sec\theta\cos\phi, y = a\sec\theta\sin\phi, z = a\tan\theta, \quad \theta \in [0, \frac{\pi}{2}), \phi \in [0, 2\pi]$$

♦

Example3.18 Cone Consider the surface given by the equation

$$z = \sqrt{x^2 + y^2}$$

One can use the following parametrisation by (z, θ)

$$x = z\cos\theta\, y = z\sin\theta, z = z, \quad \theta \in [0, \pi], \phi \in [0, 2\pi]$$

♦

Tangent Plane

A point and a (normal) direction determines a plane. To this end, consider a surface $S = \{(x, y, z) \in \mathbb{R}^3 | x = x(u, v), y = y(u, v), z = z(u, v), (u, v) \in D \subset \mathbb{R}^2\}$. If we want to find the tangent plane of S at the point

$$(x_0 = x(u_0, v_0), y_0 = y(u_0, v_0), z_0 = z(u_0, v_0))$$

, then we can follow the steps below:

1. Compute a tangent vector correspond to the horizontal direction variation:
$$v_1(u_0, v_0) = (\frac{\partial x}{\partial u}(u_0, v_0), \frac{\partial y}{\partial u}(u_0, v_0), \frac{\partial z}{\partial u}(u_0, v_0)).$$
2. Compute a tangent vector correspond to the vertical direction variation:
$$v_2(u_0, v_0) = (\frac{\partial x}{\partial v}(u_0, v_0), \frac{\partial y}{\partial v}(u_0, v_0), \frac{\partial z}{\partial v}(u_0, v_0)).$$
3. Compute the normal vector
$$\mathbf{n} = \mathbf{v}_1(u_0, v_0) \times \mathbf{v}_2(u_0, v_0)$$
4. Write down the defining equation of the tangent plane at x_0, y_0, z_0, i.e.
$$\mathbf{n} \cdot (x - x_0, y - y_0, z - z_0) = 0$$

We give the detailed description of the above procedure in the following example

Example3.19 Compute the tangent plane of the sphere

$$x^2 + y^2 + z^2 = 1$$

at the point $(x_0, y_0, z_0) = (\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2})$

1. Find the parametrisation. As one sees in the previous example, a parametrisation could be
$$x = \sin\theta\cos\phi, y = \sin\theta\sin\phi, z = \cos\theta$$
2. Find the parameter value corresponding to the point of interest:
$$\begin{cases} \frac{\sqrt{2}}{4} = \sin\theta\sin\phi \\ \frac{\sqrt{2}}{4} = \sin\theta\cos\phi \\ \frac{\sqrt{3}}{2} = \cos\theta \end{cases} \to \theta = \frac{\pi}{6}, \phi = \frac{\pi}{4}$$
3. Find the first tangent vector
$$v_1 = (\frac{dx}{d\theta}, \frac{dy}{d\theta}, \frac{dz}{d\theta})_{|(\theta,\phi)=(\frac{\pi}{6},\frac{\pi}{4})} = (\cos\theta\sin\phi, \cos\theta\cos\phi, -\sin\theta)_{|(\theta,\phi)=(\frac{\pi}{6},\frac{\pi}{4})} = (\frac{\sqrt{6}}{4}, \frac{\sqrt{6}}{4}, -\frac{1}{2})$$

4. Find the second tangent vector
$$v_2 = (\frac{dx}{d\phi}, \frac{dy}{d\phi}, \frac{dz}{d\phi})_{|(\theta,\phi)=(\frac{\pi}{6},\frac{\pi}{4})} = (\sin\theta\cos\phi, -\sin\theta\sin\phi, 0)_{|(\theta,\phi)=(\frac{\pi}{6},\frac{\pi}{4})} = (\frac{\sqrt{2}}{4}, -\frac{\sqrt{2}}{4}, 0)$$
5. Compute the normal vector $\mathbf{n}$ by taking the cross product of the two tangent vectors $v_1 \times v_2$:
$$\mathbf{n} = v_1 \times v_2 = (\frac{\sqrt{6}}{4}, \frac{\sqrt{6}}{4}, -\frac{1}{2}) \times (\frac{\sqrt{2}}{4}, -\frac{\sqrt{2}}{4}, 0) = (-\frac{\sqrt{2}}{8}, -\frac{\sqrt{2}}{8}, -\frac{\sqrt{3}}{4})$$
6. Write down the equation of the tangent plane:
$$\mathbf{n} \cdot (x - \frac{\sqrt{2}}{4}, y - \frac{\sqrt{2}}{4}, z - \frac{\sqrt{3}}{2}) = 0,$$
which can be simplified into the form
$$\sqrt{2}(x - \frac{\sqrt{2}}{4}) + \sqrt{2}(y - \frac{\sqrt{2}}{4}) + 2\sqrt{3}(z - \frac{\sqrt{3}}{2}) = 0$$

Alternative Solution We rewrite the defining equation of the curve as $F(x,y,z) = 0$, where
$$F(x,y,z) = x^2 + y^2 + z^2 - 1$$
Note that
$$\frac{\partial F}{\partial z}(\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2}) = \sqrt{3} \neq 0$$
, hence the implicit function theorem (see next section) implies that locally the surface can be seen as a graph of $z = g(x,y)$. In this problem, since $\frac{\sqrt{3}}{2} > 0$, one easily solves
$$z = \sqrt{1 - x^2 - y^2}$$
Hence another parametrisation could be
$$x = x, y = y, z = \sqrt{1 - x^2 - y^2}, x, y \in D \subset \mathbb{R}^2, (x_0, y_0) \in D$$
Now

- For the first tangent vector

♦

Example3.20 Compute the tangent plane of $z = \sqrt{x^2 + y^2}$ at the point $(x_0, y_0, z_0) = (1, 1, \sqrt{2})$.

Solution

1. Find the parametrisation. As one sees in the previous example, a parametrisation of the surface could be done by picking $(z, \theta) \in [0, \infty) \times [0, 2\pi]$
$$x = z\cos\theta, y = z\sin\theta, z = z$$
2. Find the parameter value corresponding to the point of interest: the point $(x_0, y_0, z_0) = (1, 1, \sqrt{2})$ corresponds to $(z_0, \theta_0) = (\sqrt{2}, \frac{\pi}{4})$.

3. Find the first tangent vector
$$v_1 = (\frac{\partial x}{\partial \theta}, \frac{\partial y}{\partial \theta}, \frac{\partial z}{\partial \theta})_{|(z,\theta)=(\sqrt{2},\frac{\pi}{4})} = (-z\sin\theta, z\cos\theta, 0)_{|(z,\theta)=(\sqrt{2},\frac{\pi}{4})} = (-1,1,0).$$
4. Find the second tangent vector
$$v_2 = (\frac{\partial x}{\partial z}, \frac{\partial y}{\partial z}, \frac{\partial z}{\partial z})_{|(z,\theta)=(\sqrt{2},\frac{\pi}{4})} = (\cos\theta, \sin\theta, 1)_{|(z,\theta)=(\sqrt{2},\frac{\pi}{4})} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 1).$$
5. Compute the normal vector $\mathbf{n}$ by taking the cross product of the two tangent vectors $v_1 \times v_2$:
$$\mathbf{n} = v_1 \times v_2 = (-1,1,0) \times (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 1) = (1,1,-\sqrt{2})$$
6. Write down the equation of the tangent plane:
$$\mathbf{n} \cdot (x-1, y-1, z-\sqrt{2}) = 0,$$
which can be simplified into the form
$$x + y - \sqrt{2}z = 0$$

♦

3.7 Implicit Function Theorem

Let's consider a function $y = f(x), x \in [a,b]$ and consider the graph
$$Graph(f) = \{(x, f(x)) \subseteq \mathbb{R}^2 | x \in [a,b]\}.$$
Similarly one can consider a function $z = g(x,y), x, y \in D$ and consider the graph
$$Graph(g) = \{(x, y, g(x,y)) \subseteq \mathbb{R}^3 | (x,y) \in D\}.$$
Clearly one has that

Proposition 3.5

Graph of $f(x)$ is a parametric curve/surface. ♠

Proof To see this,

- In single variable case, simply take t as the parameter, so that
$$x = t, y = f(t), t \in [a,b]$$
which is thus a parametric curve.
- In multiple variable case, simply take $(u,v) \in D$ as the parameter, so that
$$x = u, y = v, z = g(u,v), \quad (u,v) \in D$$
which is thus a parametric surface.

□

We have already mentioned examples where the parametric curve/surfaces fail to be a graph, for example

$$S^1 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$$
$$S^2 = \{(x, y, z) \in \mathbb{R}^2 | x^2 + y^2 + z^2 = 1\}$$

Now we introduce the following theorem which serves to determine whether **locally** a curve/surface can be seen as a function graph near the points of interest.

Theorem 3.1. Implicit function theorem

Consider m equations of $m + n$ variables

$$F_1(x_1, x_2, ..., x_n, z_1, ..z_m) = 0$$
$$F_2(x_1, x_2, ..., x_n, z_1, ..z_m) = 0$$
$$F_3(x_1, x_2, ..., x_n, z_1, ..z_m) = 0$$
$$\cdots$$
$$F_m(x_1, x_2, ..., x_n, z_1, ..z_m) = 0$$

Consider the Jacobian

$$\left|\frac{\partial \mathbf{F}_i}{\partial \mathbf{z}_j}\right| = \begin{vmatrix} \frac{\partial F_1}{z_1} & \frac{\partial F_1}{z_2} & \cdots & \frac{\partial F_m}{z_m} \\ \frac{\partial F_2}{z_1} & \frac{\partial F_2}{z_2} & \cdots & \frac{\partial F_m}{z_m} \\ \frac{\partial F_3}{z_1} & \frac{\partial F_3}{z_2} & \cdots & \frac{\partial F_3}{z_m} \cdots \\ \frac{\partial F_m}{z_1} & \frac{\partial F_3}{z_2} & \cdots & \frac{\partial F_m}{z_m} \end{vmatrix}$$

If $\left|\frac{\partial \mathbf{F}_i}{\partial \mathbf{z}_j}\right| \neq 0$ at point $(x_1^*, x_2^*, ..., x_n^*, z_1^*, z_1^* ..., z_m^*)$, then near $x_1^*, x_2^*, ..., x_n^*$ there exists m functions $g_1, g_2, ...g_m : \mathbb{R}^n \to \mathbb{R}$ s.t.

$$z_1 = g_1(x_1, x_2, ..., x_n),$$
$$z_2 = g_2(x_1, x_2, ..., x_n),$$
$$\cdots$$
$$z_m = g_m(x_1, x_2, ..., x_n),$$

satisfy

$$F_1(x_1, x_2, ..., x_n, g_1(x_1, x_2, ..., x_n), ..., g_m(x_1, x_2, ..., x_n)) = 0$$
$$F_2(x_1, x_2, ..., x_n, g_1(x_1, x_2, ..., x_n), ..., g_m(x_1, x_2, ..., x_n)) = 0$$
$$F_3(x_1, x_2, ..., x_n, g_1(x_1, x_2, ..., x_n), ..., g_m(x_1, x_2, ..., x_n)) = 0$$
$$\cdots$$
$$F_m(x_1, x_2, ..., x_n, g_1(x_1, x_2, ..., x_n), ..., g_m(x_1, x_2, ..., x_n)) = 0$$

♡

Remark The full detail and proof of this theorem is out of the scope of our course, yet it is one of the most important theorem in calculus and we encourage students who got true passion in math to explore more on it. Somehow for the sake of MAT236, we just need some of the most simple cases of the above theorem, as are investigated below.

Curves as graphs: $n = 1, m = 1$

In this case we have a curve defined by

$$F(x, z) = 0$$

the Jacobian just becomes the partial derivative $\frac{\partial F}{\partial z}$. Hence

Proposition 3.6. Curves as graphs

Suppose (x_0, z_0) *satisfies that* $F(x_0, z_0) = 0$ *and* $\frac{\partial F}{\partial z}_{|(x_0,z_0)} \neq 0$, *then locally near* x_0 *there is a function* $z = g(x)$ *s.t.* $F(x, g(x)) = 0$. *In other words, the curve is locally a graph near* (x_0, z_0).

Example3.21 Consider the circle $x^2 + y^2 = 1$, which is not a graph. Somehow notice that we can write it as $F(x, y) = 0$ where

$$F(x, y) = x^2 + y^2 - 1$$

Since $\dfrac{\partial F}{\partial y} = 2y$, one sees that around any point (x_0, y_0) with $y_0 \neq 0$, $x^2 + y^2 = 1$ is locally a graph. This can even be explicitly computed. For example one can take

$$y = \begin{cases} \sqrt{1 - x^2}, & y_0 > 0 \\ -\sqrt{1 - x^2}, & y_0 < 0 \end{cases}$$

♦

Example3.22 Consider the parabola $y^2 = x$, which is not a graph. Somehow notice that we can write it as $F(x, y) = 0$ where

$$F(x, y) = y^2 - x$$

Since $\dfrac{\partial F}{\partial y} = 2y$, one sees that around any point (x_0, y_0) with $y_0 \neq 0$, $y^2 = x$ is locally a graph. This can even be explicitly computed. For example one can take

$$y = \begin{cases} \sqrt{x}, & y_0 > 0 \\ -\sqrt{x}, & y_0 < 0 \end{cases}$$

♦

Surfaces as graphs: $n = 2, m = 1$

In this case we have a surface defined by

$$F(x, y, z) = 0$$

the Jacobian just becomes the partial derivative $\frac{\partial F}{\partial z}$. Hence

Proposition 3.7. Surfaces as graphs

Suppose (x_0, y_0, z_0) *satisfies that* $F(x_0, y_0, z_0) = 0$ *and* $\frac{\partial F}{\partial z}_{|(x_0,y_0,z_0)} \neq 0$, *then locally near* (x_0, y_0) *there is a function* $z = g(x, y)$ *s.t.* $F(x, y, g(x, y)) = 0$. *In other words, the curve is locally a graph near* (x_0, y_0, z_0).

♠

Example3.23 Consider the sphere $x^2 + y^2 + z^2 = 1$, which is not a graph. Somehow notice that we can write it as $F(x, y, z) = 0$ where

$$F(x, y, z) = x^2 + y^2 + z^2 - 1$$

Since $\dfrac{\partial F}{\partial z} = 2z$, one sees that around any point (x_0, y_0, z_0) with $z_0 \neq 0$, $x^2 + y^2 + z^2 = 1$ is locally a graph. This can even be explicitly computed. For example one can take

$$z = \begin{cases} \sqrt{1 - x^2 - y^2}, & z_0 > 0 \\ -\sqrt{1 - x^2 - y^2}, & z_0 < 0 \end{cases}$$

♦

Example3.24 Near the point $(2, 4, 1)$ can surface

$$0 = F(x, y, z) = x^3 z^2 - 2z^3 y$$

be represented as a graph $z = g(x, y)$?

Solution Compute the derivative with respect to z, one sees that

$$\frac{\partial F}{\partial z}_{|(2,4,1)} = (2x^3 z - 6z^2 y)_{|(2,4,1)} = -8 \neq 0$$

Hence near the point $(2, 4, 1)$ the surface can be represented as a graph $z = g(x, y)$ (Although in this example it will be more complicated to figure out what the function g should be).

Inverse Function Theorem

A direct corollary is the following theorem regarding the existence of a inverse function from $\mathbb{R}^n$ to $\mathbb{R}^n$

Theorem 3.2

Consider the function

$$z_1 = F_1(x_1, x_2, ..., x_n),$$
$$z_2 = F_2(x_1, x_2, ..., x_n),$$
$$\cdots$$
$$z_n = F_n(x_1, x_2, ..., x_n)$$

Suppose that the Jacobian

$$\left|\frac{\partial z_i}{\partial x_j}\right|_{(x_1^*, x_2^*, \ldots x_n^*)} \neq 0$$

Then near $(x_1^*, x_2^*, ...x_n^*)$ *locally the inverse exists, i.e., there exists functions* $G_1, G_2, ...G_n$ *s.t.*

$$x_1 = G_1(z_1, z_2, ..., z_n)$$
$$x_2 = G_2(z_1, z_2, ..., z_n)$$
$$\cdots$$
$$x_n = G_n(z_1, z_2, ..., z_n).$$

♡

In most situations we are only interested in the $\mathbb{R}^2$ to $\mathbb{R}^2$ functions.

Example3.25 The polar coordinates Consider the change of variables by polar coordinates

$$x = r\cos\theta$$
$$y = r\sin\theta$$

When does the inverse change of variable exist?

Solution Compute the Jacobian, one sees that

$$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = r$$

Hence if $r \neq 0$, i.e., except on the origin of the plane, the inverse change of variable exists. This is easy to see: actually at the origin, $r = 0$ and θ can be anything. Hence the function is not bijective thus not having an inverse.

Chapter 4 Techniques in Double Integrals

4.1 Change of Order of Iterated Integral

Let's first recall the Fubini theorem

Theorem 4.1. Fubini Theorem

Suppose that

- $F : \mathbb{R}^2 \to \mathbb{R}$ *is continuous on*

$$R = \{(x,y)|a \le x \le b, c \le y \le d\}$$

-

Then

$$\int\int_R F(x,y)dA \overset{①}{=\!=} \int_a^b \int_c^d F(x,y)dydx \overset{②}{=\!=} \int_c^d \int_a^b F(x,y)dxdy$$

♡

Note that

- The equality ① means the **double** integral is equal to the **iterated** integral
- The equality ② means the **order** in the computation of the iterated integral is irrelevant.

This suggests the following strategy of computing an iterated integral:

Change of Order for Iterated Integral

Suppose that we want to compute the iterated integral of type I region

$$\int_a^b \int_{g_1(x)}^{g_2(x)} F(x,y)dydx$$

One can follow the steps below:

1. Find the domain

$$D = \{(x,y)|a \le x \le b, g_1(x) \le y \le g_2(x)\}$$

2. Check if the domain D is also a type II domain, i.e.

$$D = \{(x,y)|c \le y \le d, h_1(y) \le x \le h_2(y)\}$$

   and write the double integral again as an iterated integral

$$\int_c^d \int_{h_1(y)}^{h_2(y)} F(x,y)dxdy$$

Remark If one was given an iterated integral of type II region and would like to switch it into an iterated integral of type I region, in that case one can follow the similar steps as indicated above.

Example4.1 Compute $\displaystyle\int_0^9 \int_{\sqrt{y}}^3 \sin(x^3)dxdy$.

This integral is not easy to compute directly in the given order since

- It involves the anti-derivative of $\sin x^3$ which is not easy to get explicitly
- Even if you are lucky to get the anti-derivative, you need to then evaluate it at $\sqrt{y}$, making it tougher to get the anti-derivative as a function of y in the sequel.

As a result we consider the change of order. First, note that the domain can be written as a type II region.

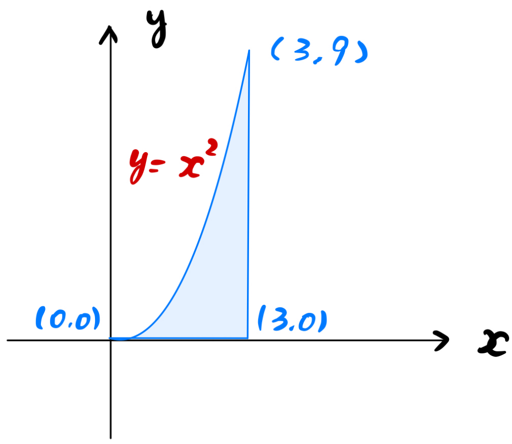


Figure 4.1: Example 4.1

$$D = \{(x,y)|0 \le y \le 9, \sqrt{y} \le x \le 3\}$$

We saw from the figure that it is indeed also a type I region,

$$D = \{(x,y)|0 \le x \le 3, 0 \le y \le x^2\}$$

Hence by Fubini theorem the original integral can be written as

$$\int_0^9 \int_{\sqrt{y}}^3 \sin(x^3)dxdy = \int_0^3 \int_0^{x^2} \sin(x^3)dydx = \int_0^3 \sin(x^3)x^2dx \overset{u=x^3}{=\!=} \frac{1}{3}\int_0^{27} \sin udu$$
$$= \frac{1}{3}(-\cos u|_0^{27}) = \frac{1}{3}(1-\cos 27).$$

♦

Example4.2 Evaluate the iterated integral $\displaystyle\int_0^6 \int_{\frac{x}{3}}^2 x\sqrt{1+y^3}dydx$

Again the difficulty here is that the inner integral is difficult (actually impossible) to evaluate explicitly. But if we look at the domain of this type I region

$$D = \{(x,y)|0 \le x \le 6, \frac{x}{3} \le y \le 2\}$$

Then we noticed that it is indeed a type II region as well, and we can rewrite D as

$$D = \{(x,y)|0 \le y \le 2, 0 \le x \le 3y\}$$

In this case the iterated integral becomes

$$\int_0^2 \int_0^{3y} x\sqrt{1+y^3}dxdy = \int_0^2 \frac{x^2}{2}\sqrt{1+y^3}\Big|_0^{3y} dy = \frac{9}{2}\int_0^2 y^2\sqrt{1+y^3}dy$$
$$\overset{u=y^3}{=\!=} \frac{3}{2}\int_0^8 \sqrt{1+u}du = (1+u)^{\frac{3}{2}}\Big|_0^8 = 26$$

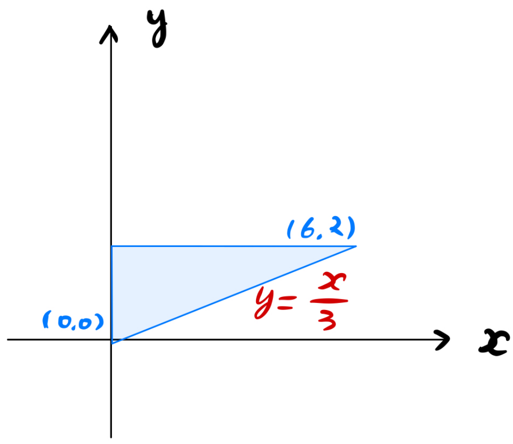


Figure 4.2: Example 4.2

♦

Example4.3 Compute the iterated integral $\int_1^e \int_{lny}^1 \cos{(e^x - x)} dx dy$
Again this integral on the type II region seems to be hopeless to compute in the given order. Yet we can check the region of integral

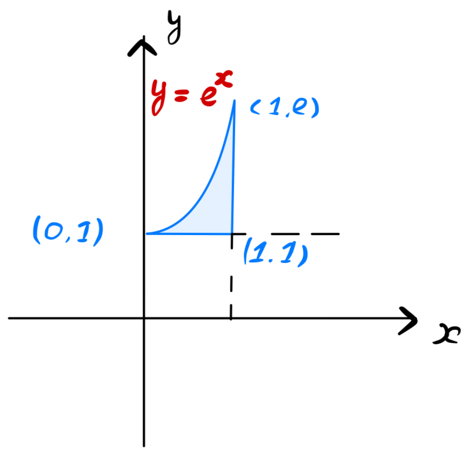


Figure 4.3: Example 4.3

$$D = \{(x, y)|1 \le y \le e, \ln y \le x \le 1\}$$

It is indeed also a type I region:

$$D = \{(x, y)|0 \le x \le 1, 1 \le y \le e^x\}$$

As a result, by Fubini theorem,

$$\int_1^e \int_{lny}^1 \cos{(e^x - x)} dx dy = \int_0^1 \int_1^{e^x} \cos{(e^x - x)} dy dx = \int_0^1 \cos{(e^x - x)} y|_1^{e^x} dx$$
$$= \int_0^1 (e^x - 1)\cos(e^x - x) dx \xlongequal{u=e^x-x} \int_1^{e-1} \cos u du$$
$$= \sin u|_1^{e-1} = \sin(e-1) - \sin 1$$

♦

4.2 Change of Variables

To apply change of variables, one first needs to understand what kind of change of variables we are looking for. Suppose we are studying the double integral

$$\int\int_D f(x,y) dA$$

Then if we want to replace x, y by u, v, are we looking for systems like

$$\begin{cases} x = x(u,v) \\ y = y(u,v) \end{cases}$$

Now the original integral will undergo three changes

$$\int\int_D f(x,y) dA \longrightarrow \underbrace{\int\int_{\tilde{D}}}_{\text{1. New domain in the } (u,v)\text{-plane}} \overbrace{f(x(u,v), y(u,v))}^{\text{2. function of } (u,v)} \underbrace{\left|\det\left(\frac{\partial(x,y)}{\partial(u,v)}\right)\right| d\tilde{A}}_{\text{3. New representation of the infinitesimal area element}}.$$

We first deal with a basic example to exhibit the change of variable technique.

Example4.4 Compute $\int\int_D 1 dA$, where

$$D = \{(x,y)| -1 \le x \le 1, -1 \le y \le 1\}.$$

- Without change of variable, we know that this is just the volume of a cuboid, the base area is $2 \times 2 = 4$, the height at point (x, y) is $f(x, y) = 1$. Hence the volume $2 \times 2 \times 1 = 4$.

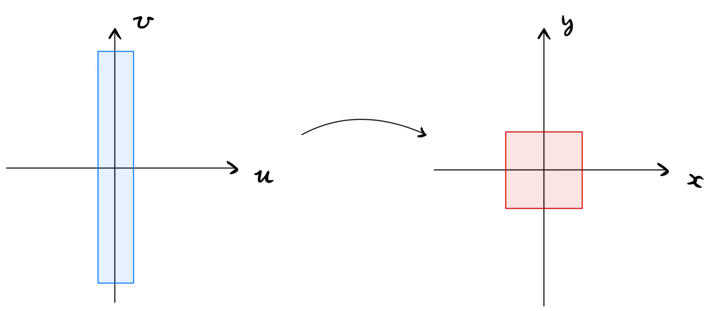


Figure 4.4: Example 4.4

- Now let's consider a change of variable

$$(x,y) \xleftarrow{F} (u,v) : \begin{cases} x &= 3u, \\ y &= \frac{1}{2}v. \end{cases}$$

Under this change of variable F now introduces the following three facts:

1. The new domain becomes the pre-image
$$\tilde{D} = F^{-1}(D) = \{(u,v) | \frac{1}{3} \leq u \leq \frac{1}{3}, -2 \leq v \leq 2\}$$
2. The new function becomes $f(u,v) = 1$ (it is a constant function hence now change).
3. The new area form becomes now
$$\left| det \left(\frac{\partial(x,y)}{\partial(u,v)} \right) \right| d\tilde{A} = \left| det \begin{pmatrix} 3 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \right| d\tilde{A} = \frac{3}{2} d\tilde{A}$$

Hence the new integral becomes
$$\int\int_{\tilde{D}} \frac{3}{2} d\tilde{A}$$
which by definition of the double integral is $\frac{2}{3} \times 4 \times \frac{3}{2} = 4$.

♦

How to choose the correct change of variable?

The short answer is: We don't always know. There is no such universal approach of picking up the correct change of variable. Even the word correct here is not defined, as the change of variable does not change the integral, hence in this sense it is always correct.

Somehow one can see that the difficulty of the integral might in general come from two parts:

- The function $f(x,y)$ might be complicated.
- The region D might be complicated.

As a result, a good change of variable should aim at making at least one of the two points above easier. We illustrate through some examples.

Polar coordinates and its variants We first study the polar coordinates.

Example4.5 Compute $\int\int_D \ln(1+x^2+y^2) dA$, where

$$D = \{(x,y) | 0 \leq x \leq 1, 0 \leq y \leq 1, x^2+y^2 \leq 1\}$$

The region is the unit disc in the first quadrant, and the integrand $f(x,y) = 1 + x^2 + y^2$ has the term $x^2 + y^2$. Thus a natural choice is the polar coordinates. Let

$$(x,y) \xleftarrow{F} (r,\theta) : \begin{cases} x = r\cos\theta, \\ y = r\sin\theta. \end{cases}$$

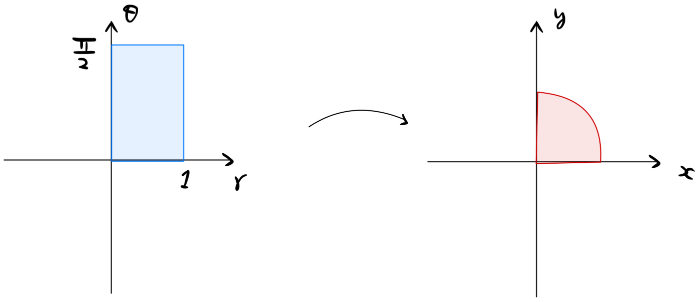


Figure 4.5: Example 4.5

It follows that

$$\tilde{D} = F^{-1}(D) = \{(r,\theta) | 0 \le r \le 1, 0 \le \theta \le \frac{\pi}{2}\}.$$

The Jacobian can be computed directly:

$$det\left(\frac{\partial(x,y)}{\partial(r,\theta)}\right) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = r > 0$$

Hence the integral becomes

$$\int\int_D f(x,y)dA = \int\int_{\tilde{D}} \ln(1+r^2) r\tilde{A} \overset{\text{Fubini}}{=\!=} \int_0^{\frac{\pi}{4}} \int_0^1 \ln(1+r^2) r dr d\theta$$
$$= \frac{\pi}{2}\int_0^1 \ln(1+r^2) r dr \overset{u=1+r^2}{=\!=} \frac{\pi}{2}\cdot\frac{1}{2}\int_0^1 \ln u du \overset{IBP}{=\!=} \frac{\pi}{4}\left(u \ln u|_1^2 - \int_1^2 1 du\right)$$
$$= \frac{\pi}{4}(2\ln 2 - 1).$$

♦

Remark The polar coordinates and its variants naturally work for following shapes:

- Disc
- Donut
- A slice of Pizza
- A slice of Pizza, with a bite...

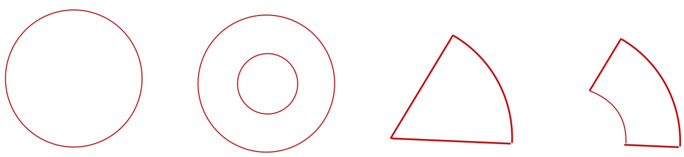


Figure 4.6: Some typical regions for polar coordinates

Example4.6 Polar coordinates for "Pizza with a bite" Compute $\int\int_D (x^2+y^2)^{\frac{3}{2}} dA$, where

$$D = \{(x,y) | 1 \le x^2+y^2 \le 9, 0 \le y \le \sqrt{3}x\}$$

Again such domain and the integrand suggests a natural choice is to take the polar coordinates, then

$$(x,y) \xleftarrow{F} (r,\theta) : \begin{cases} x = r\cos\theta, \\ y = r\sin\theta. \end{cases}$$

It follows that

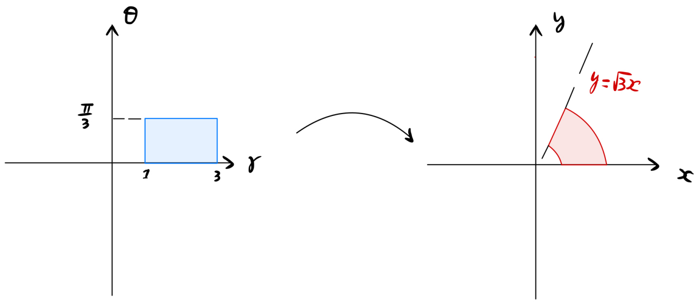


Figure 4.7: Example 4.6

$$\tilde{D} = F^{-1}(D) = \{(r,\theta) | 1 \le r \le 3, 0 \le \theta \le \frac{\pi}{3}\}.$$

Hence the integral becomes

$$\int\int_D (x^2+y^2)^{\frac{3}{2}} dA = \int\int_{\tilde{D}} r^3 r d\tilde{A} \xlongequal{Fubini} \int_0^{\frac{\pi}{3}} \int_1^3 r^4 dr d\theta = \frac{\pi}{3}\frac{r^5}{5}|_1^3 = \frac{242}{15}\pi$$

♦

Example4.7 Polar coordinates for ellipses Evaluate $\int\int_D \sin(9x^2+4y^2)dA$, where

$$D = \{(x,y)|9x^2+4y^2 \le 1\}$$

Note that the domain of integral is the interior of an ellipse, hence we can use an adaptive version of polar coordinates:

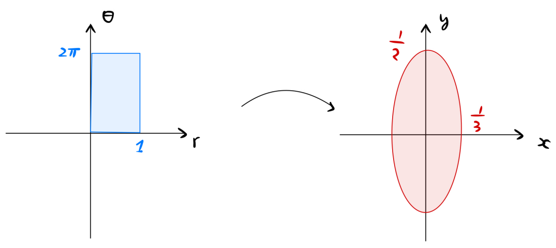


Figure 4.8: Example 4.7

$$(x,y) \xleftarrow{F} (r,\theta) : \begin{cases} x = \frac{1}{3} r\cos\theta, \\ y = \frac{1}{2} r\sin\theta. \end{cases}$$

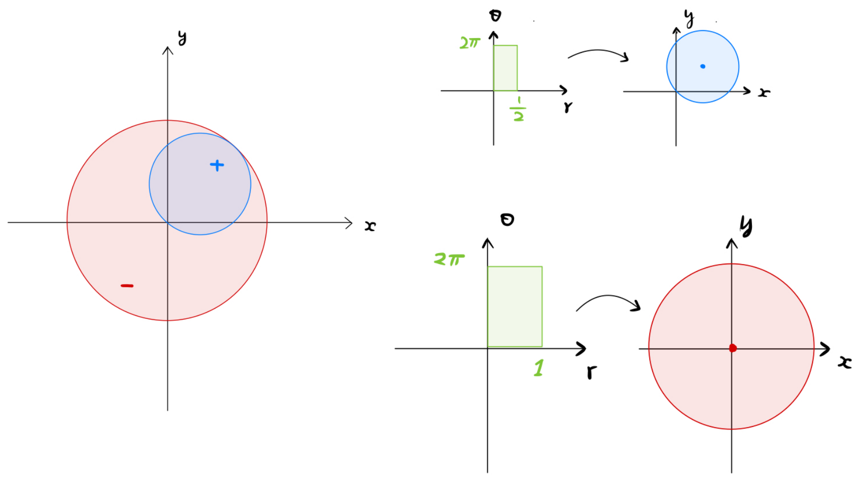


Figure 4.9: Example 4.8

It follows that the Jacobian is

$$det\left(\frac{\partial(x,y)}{\partial(r,\theta)}\right)=\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{vmatrix}=det\begin{pmatrix}\frac{1}{3}\cos\theta & -\frac{1}{3}r\sin\theta\\ \frac{1}{2}\sin\theta & \frac{1}{2}r\cos\theta\end{pmatrix}=\frac{1}{6}r>0$$

Hence the integral becomes

$$\int\int_D \sin(9x^2+4y^2)dA=\int\int_{\tilde{D}}\sin(r^2)\frac{1}{6}\tilde{A}\overset{Fubini}{=\!=\!=}\int_0^{2\pi}\int_0^1\sin(r^2)\frac{1}{6}drd\theta$$
$$\overset{u=r^2}{=\!=\!=}\frac{\pi}{6}\int_0^1\sin udu=\frac{\pi}{6}-\cos u|_0^1=\frac{\pi}{6}(1-\cos 1).$$

♦

Example4.8 Polar coordinates with a translation Compute $\int\int_D\left|\frac{x+y}{\sqrt{2}}-x^2-y^2\right|dA$, where

$$D=\{(x,y)|x^2+y^2\le 1\}.$$

Note that the difficulty is mainly on the absolute value. To this end, define

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

We first figure out in which region $f>0$ and in which region $f<0$. Since f is clearly a continuous function, these two regions should be separated by the curve $f(x,y)=0$. Next, by completing the square,

$$0=f(x,y)=\frac{x+y}{\sqrt{2}}-x^2-y^2=-(x-\frac{1}{2\sqrt{2}})^2-(y-\frac{1}{2\sqrt{2}})^2+\frac{1}{4}.$$

i.e., We can divide the region D into two sub-regions. To this end, let

$$D_1=\{(x,y)|(x-\frac{1}{2\sqrt{2}})^2+(y-\frac{1}{2\sqrt{2}})^2<\frac{1}{4}\}$$

- In D_1, $|f(x,y)|=f(x,y)$.
- In $D\setminus D_1$, $|f(x,y)|=-f(x,y)$.

It turns out that

$$\begin{aligned}\int\int_D |f(x,y)|dA &= \int\int_{D_1} f(x,y)dA - \int\int_{D\setminus D_1} f(x,y)dA \\ &= 2\int\int_{D_1} f(x,y)dA - \left(\int\int_{D_1} f(x,y)dA + \int\int_{D\setminus D_1} f(x,y)dA\right) \\ &= 2\int\int_{D_1} f(x,y)dA - \int\int_D f(x,y)dA\end{aligned}$$

- To compute $\int\int_D f(x,y)dA$, using the standard polar coordinates

$$x = r\cos\theta, y = r\sin\theta, \quad 0 \le r \le 1, 0 \le \theta \le 2\pi,$$

  and one computes directly that

$$\int\int_D f(x,y) = -\frac{\pi}{2}$$

- To compute $\int\int_{D_1} f(x,y)dA$, using the standard polar coordinates **with a translation**, i.e.

$$x = r\cos\theta + \frac{1}{2\sqrt{2}}, y = r\sin\theta + \frac{1}{2\sqrt{2}}, 0 \le r \le \frac{1}{2}, 0 \le \theta \le 2\pi.$$

It follows that the Jacobian is still r and by Fubini theorem,

$$\int\int_{D_1} f(x,y)dA = \int_0^{2\pi}\int_0^{\frac{1}{2}} (\frac{1}{4} - r^2)rdrd\theta = 2\pi\,(\frac{1}{8}r^2 - \frac{1}{4}r^4)\Big|_0^{\frac{1}{2}} = \frac{\pi}{32}.$$

Hence the final value is

$$\int\int_D |f(x,y)|dA = 2\int\int_{D_1} f(x,y)dA - \int\int_D f(x,y)dA = \frac{\pi}{16} + \frac{\pi}{2} = \frac{9\pi}{16}$$

♦

Region and Integrand, which to simplify As we discussed, for the same double integral there might exist multiple efficient change of variables. Different person has different tastes/computational skills/determination/time constraints, so as long as you can solve your problem at a cost you can afford, it is enough. Somehow a natural way as we have suggested in the beginning of the discussion, is that you can aim at two missions:

1. Simplify D.
2. Simplify $f(x,y)$.

We give an example and suggests two methods of change of variables, aiming at each of these missions above.

Example4.9 Compute $\int\int_D e^{\frac{x+y}{x-y}}dA$, where

$$D = \{(x,y)|1 \le x - y \le 2, y \le 0 \le x.\}$$

1. **Simplify the integrand**. To this end introduce the following change of variables

$$(x,y) \xleftarrow{F} (u,v) : \begin{cases} x = \frac{1}{2}(u+v) \\ y = \frac{1}{2}(u-v). \end{cases}$$

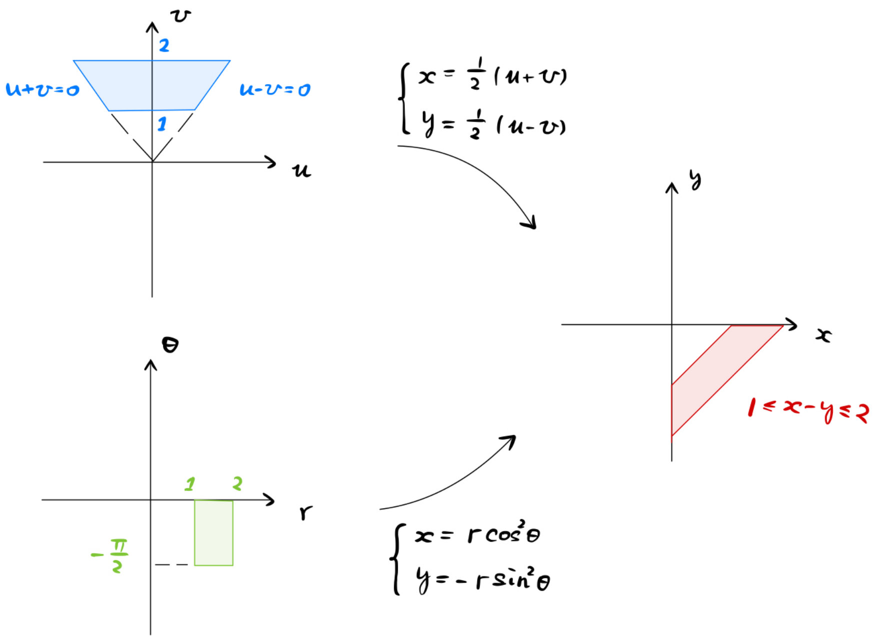


Figure 4.10: Example 4.9

Now $e^{\frac{x+y}{x-y}} \to e^{\frac{u}{v}}$ is simpler than before, on the other hand

$$\tilde{D} = F^{-1}(D) = \{(u,v)|1 \leq v \leq 2, -v \leq u \leq v\}.(\textbf{ Check this by yourself!})$$

The Jacobian becomes

$$det\left(\frac{\partial(x,y)}{\partial(u,v)}\right) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \|det\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}| = |-\frac{1}{2}| = \frac{1}{2}$$

The integral now becomes easier to evaluate:

$$\int\int_D e^{\frac{x+y}{x-y}} dA = \int\int_{\tilde{D}} e^{\frac{x+y}{x-y}} \frac{1}{2} d\tilde{A} \overset{Fubini}{=\!=\!=} \int_1^2 \int_{-v}^{v} \frac{1}{2} e^{\frac{u}{v}} dudv = \frac{1}{2}\int_1^2 ve^{\frac{u}{v}}\Big|_{-v}^{v} dv$$

$$= \frac{1}{2}(e - e^{-1})\left.\frac{1}{2}v^2\right|_1^2 = \frac{3}{4}(e - e^{-1}).$$

2. **Simplify the domain**. Another method is focusing on the change of variables which could simplify the domain. Note that the domain looks pretty much let the one in example 4.6, except that we have stripe of line segments instead of arcs. Such suggests the following change of variable (known as the generalised polar coordinates in some textbooks)

$$\begin{cases} x = r\cos^2\theta \\ y = -r\sin^2\theta, \theta \in [-\frac{\pi}{2}, 0], 1 \leq r \leq 2. \end{cases}$$

The good point of this transform is that

$$\tilde{D} = F^{-1}(D) = \{(r,\theta)|1 \leq r \leq 2, -\frac{\pi}{2} \leq \theta \leq 0\}$$

which is a rectangle! So from this point of view the domain cannot be simpler anymore.

Next, the Jacobian is

$$det\left(\frac{\partial(x,y)}{\partial(r,\theta)}\right)=\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{vmatrix}=-2r\sin\theta\cos\theta>0(\textbf{why? Look at the domain of } \theta)$$

Hence

$$\begin{aligned}\int\int_D e^{\frac{x+y}{x-y}}dA &= \int\int_{\tilde{D}} e^{\frac{\cos^2\theta-\sin^2\theta}{1}}\cdot(-2r\sin\theta\cos\theta)d\tilde{A}\\ &\xlongequal{Fubini} -\int_{-\frac{\pi}{2}}^{0}\int_{1}^{2} e^{\cos 2\theta}\sin 2\theta d\theta dr\\ &\xlongequal{u=\cos 2\theta} e^u|_{-1}^{1}\cdot\frac{3}{4}\\ &=\frac{3}{4}(e^u-e^{-u}).\end{aligned}$$

♦

Some Tips for the change of variable Before we ends this section, We give some tips for the change of variable for domains of some special types.

Change of variables for Curved Parallelogram

Suppose that the domain is of the form

$$D=\{(x,y)|a\le f(x,y)\le b, c\le g(x,y)\le d\}$$

Then one can consider the (inverse) change of variable

$$u=f(x,y), v=g(x,y)$$

The domain D will be transformed to $\tilde{D}=\{(u,v)|a\le u\le b, c\le v\le d\}$, which is a rectangle in the uv-plane

Example4.10 Suppose the domain is

$$D=\{(x,y)|1\le xy\le 3, x\le y\le 3x, 0<x, 0<y\}$$

Then one can let

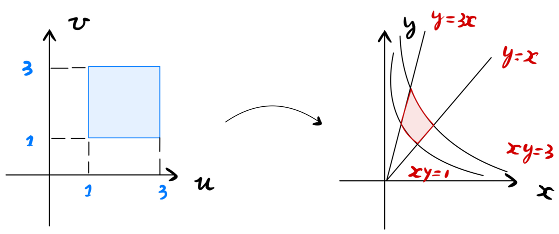


Figure 4.11: Example 4.10

$$u=xy, v=\frac{y}{x},$$

i.e.,

$$x = \sqrt{\frac{u}{v}}, y = \sqrt{uv}$$

From where we see that D is switched to $\tilde{D} = \{(u,v)|1 \le u \le 3, 1 \le v \le 3\}$, which is a rectangle.

Computation of Jacobian

One sees that given a change of variable

$$(x,y) \xleftarrow{F} (u,v)$$

we need to evaluate the Jacobian, i.e., to compute the determinant of the matrix

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

Sometimes it is much more simpler to have F^{-1} rather than F. In this case, one can compute

$$|J^{-1}| = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

Then using the fact that $|J^-1| = |J|^{-1}$

Example4.11 Consider the previous example, we need to compute the Jacobian for

$$(x,y) = F(u,v), \quad x = \sqrt{\frac{u}{v}}, y = \sqrt{uv}$$

Note that $(u,v) = F^{-1}(x,y) = (xy, \frac{y}{x})$. Hence the Jacobian is

$$|J^{-1}| = \begin{vmatrix} y & x \\ -\frac{y}{x^2} & \frac{1}{x} \end{vmatrix} = 2\frac{y}{x} = 2v$$

Hence $|J| = \dfrac{1}{2v}$. In this way, we avoid differentiation involving square root.

4.3 Area of Parametric Surface

Suppose that we are given a parametric surface defined on $D \subset \mathbb{R}^2$,

$$\gamma(u,v) = (x(u,v), y(u,v), z(u,v)),$$
$$D = \{(u,v) \in \mathbb{R}^2 | a \le u \le b, c \le v \le d\}.$$

and we want to compute the area of the surface, then a natural idea is that

1. We divide the rectangle D into small rectangles $R_{ij}, 1 \le i \le m, 1 \le j \le n$.
2. We look at the curved surface $\gamma(R_{ij})$ and compute its area **by approximation**.
3. We take limit as m, n goes to infinity and hope such gives us the exact number.

This idea definitely works, the key step is indeed step 2: How do we approximate the area of $\gamma(R_{ij})$, i.e., the image of small rectangles R_{ij}?

Approximation for area of small piece of curved surface In this paragraph we discuss how to deduce the area formula, by explain in detail how to get the approximation of area for $\gamma(R_{ij})$. To simplify the notation we drop the index i, j. Note that the four end points of the small piece of curved surface is

$$
\begin{aligned}
Q_0 &= \{x(u,v), y(u,v), z(u,v)\} \\
Q_1 &= \{x(u+\Delta u,v), y(u+\Delta u,v), z(u+\Delta u,v)\} \\
Q_2 &= \{x(u,v+\Delta v), y(u,v+\Delta v), z(u,v+\Delta v)\} \\
Q_3 &= \{x(u+\Delta,v+\Delta v), y(u+\Delta,v+\Delta v), z(u+\Delta,v+\Delta v)\}
\end{aligned}
$$

- **First Approximation** We take $\overrightarrow{Q_0Q_1}$ and $\overrightarrow{Q_0Q_2}$, to form a parallelogram. The two vectors can be computed readily:

$$
\begin{aligned}
\overrightarrow{Q_0Q_1} &= (x(u+\Delta u,v)-x(u,v), y(u+\Delta u,v)-y(u,v), z(u+\Delta u,v)-z(u,v)) \\
\overrightarrow{Q_0Q_2} &= (x(u,v+\Delta v)-x(u,v), y(u,v+\Delta v)-y(u,v), z(u,v+\Delta v)-z(u,v))
\end{aligned}
$$

  and we compute its area by using the cross product:

$$\left|\overrightarrow{Q_0Q_1} \times \overrightarrow{Q_0Q_2}\right|$$

  The area of the parallelogram is already a good approximation, but not convenient enough for computation using calculus yet.

- **Second Approximation** Next we approximate the vectors $\overrightarrow{Q_0Q_1}$ and $\overrightarrow{Q_0Q_2}$ by using the linear approximation, then putting the vectors thus obtained into the cross product. More precisely speaking:

$$
\begin{aligned}
&(x(u+\Delta u,v)-x(u,v), y(u+\Delta u,v)-y(u,v), z(u+\Delta u,v)-z(u,v)) \\
\approx &\left(\frac{\partial x}{\partial u}(u,v)\Delta u, \frac{\partial y}{\partial u}(u,v)\Delta u, \frac{\partial z}{\partial u}(u,v)\Delta u\right) = \gamma_u \Delta u \\
&(x(u,u,v+\Delta v)-x(u,v), y(u,v+\Delta v)-y(u,v), z(u,v+\Delta v)-z(u,v)) \\
\approx &\left(\frac{\partial x}{\partial v}(u,v)\Delta u, \frac{\partial y}{\partial v}(u,v)\Delta u, \frac{\partial z}{\partial v}(u,v)\Delta u\right) = \gamma_v \Delta v
\end{aligned}
$$

  Hence the new parallelogram has the area

$$|\gamma_u \Delta u \times \gamma_v \Delta v| = |\gamma_u \times \gamma_v|\,\Delta u \Delta v$$

As a result, we can compute the area by

$$\sum_{i=1}^{m}\sum_{j=1}^{n} |\gamma_u(u_i,v_j) \times \gamma_v(u_i,v_j)|\,\Delta u \Delta v \xrightarrow{n,m\to\infty} \int\int_D |\gamma_u \times \gamma_v| dA$$

Computation and Examples In practice, we can compute the area following the steps below:

Step by Step Guidance on Surface Area Computation

Given a surface $F(x, y, z) = 0$, to compute its surface

1. Find a parametrisation $\gamma(u, v)$ for the surface. (For example using trigonometric functions, using implicit function theorem, and so on)
2. Compute the vectors γ_u and γ_v.
3. Compute the $\|\gamma_u \times \gamma_v\|$ (which will in general be a function of u and v)
4. Compute the double integral

$$\int\int_D \|\gamma_u \times \gamma_v\| dA$$

Remark The double integral

$$\int\int_D \|\gamma_u \times \gamma_v\| dA,$$

after all, is just a double integral. So you can use all techniques we have learnt in the previous part to compute it (change of variables/change of orders/...)

Example4.12 Compute the area of the unit sphere

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

Solution We have already studied the parametrisation, so that we provide two methods for the computation.

- **Method 1** Parametrise the surface by

$$x = \sin\theta\cos\phi, y = \sin\theta\sin\phi, z = \cos\theta,$$
$$0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$$

Now we compute the partial derivatives

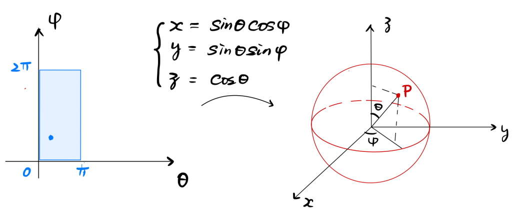


Figure 4.12: Sphere Surface Area: Method 1

$$\gamma_\theta = \left(\frac{\partial x}{\partial\theta}, \frac{\partial y}{\partial\theta}, \frac{\partial z}{\partial\theta}\right) = (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta)$$
$$\gamma_\phi = \left(\frac{\partial x}{\partial\phi}, \frac{\partial y}{\partial\phi}, \frac{\partial z}{\partial\phi}\right) = (-\sin\theta\sin\phi, \sin\theta\cos\phi, 0)$$

Next, we compute the cross product and its norm, i.e.

$$\gamma_\theta \times \gamma_\phi = (\sin^2\theta\cos\phi, \sin^2\theta\sin\phi, \cos\theta\sin\theta)$$

Then

$$\|\gamma_\theta \times \gamma_\phi\| = \sqrt{\sin^4\theta\cos^2\phi + \sin^4\theta\sin^2\phi + \cos^2\theta\sin^2\theta} = \sin\theta$$

Note that we have used the fact that $0 \le \sin\theta$ for $\theta \in [0, \pi]$. Finally, we compute the double integral

$$\int\int_D \sin\theta dA \overset{Fubini}{=\!=} \int_0^{2\pi}\int_0^{\pi} \sin\theta d\theta d\phi = 2\pi\left(-\cos\theta\right)|_0^{\pi} = 4\pi.$$

- **Method 2** By symmetry the surface area of the north hemsphere is the same as that of the south hemsphere. Thus we can compute the area above the equator and then multiply the result by 2. Since for $z > 0$ one has that

$$\frac{\partial}{\partial z}(x^2 + y^2 + z^2) = 2z > 0$$

It follows by the implicit function theorem, we can consider the (half copy of) surface as a graph, and parametrise it by the natrual parametrisation

$$x = x, y = y, z = \sqrt{1 - x^2 - y^2}$$

Next we compute the tangent vectors

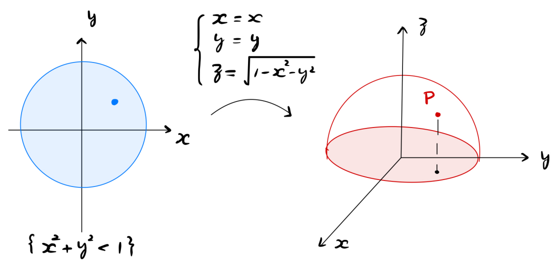


Figure 4.13: Sphere Surface Area: Method 2

$$\gamma_x = \left(1, 0, -\frac{x}{\sqrt{1 - x^2 - y^2}}\right)$$

$$\gamma_y = \left(0, 1, -\frac{y}{\sqrt{1 - x^2 - y^2}}\right)$$

Thus the cross product is

$$\gamma_x \times \gamma_y = \left(-\frac{x}{\sqrt{1 - x^2 - y^2}}, -\frac{y}{\sqrt{1 - x^2 - y^2}}, 1\right)$$

and the norm is then

$$\|\gamma_x \times \gamma_y\| = \sqrt{1 + \frac{x^2 + y^2}{1 - x^2 - y^2}} = \sqrt{\frac{1}{1 - x^2 - y^2}}$$

Finally we evaluate the integral

$$\int\int_{x^2+y^2<1} \frac{1}{\sqrt{1 - x^2 - y^2}} dA \overset{\text{Polar Coordinates}}{=\!=} \int_0^{2\pi}\int_0^1 \frac{r}{\sqrt{1 - r^2}} dr d\theta$$

$$= 2\pi \int_0^1 \frac{r}{\sqrt{1 - r^2}} dr \overset{u=r^2}{=\!=} \pi \int_0^1 \frac{1}{\sqrt{1 - u}} du = -2\pi \sqrt{1 - u}\Big|_0^1 = 2\pi.$$

Recall that this is the area for half of the whole sphere, hence the total area is $2 \times 2\pi = 4\pi$.

♦

Example4.13 Area of the surface of revolution Find the area of the surface S obtained by rotating the curve

$$y = f(x), a \leq x \leq b,$$

about the x-axis, where $f(x) \geq 0$ and f' is continuous. We have the following parametrisation

$$\mathbf{r}(x,\theta) : x = x, \quad y = f(x)\cos\theta, \quad z = f(x)\sin\theta. \quad a \leq x \leq b, \quad ,0 \leq \theta \leq 2\pi$$

It follows that

$$\mathbf{r}_x = (1, f'(x)\cos\theta, f'(x)\sin\theta).$$
$$\mathbf{r}_\theta = (0, -f(x)\sin\theta, f(x)\cos\theta).$$

Hence

$$\mathbf{r}_x \times \mathbf{r}_\theta = (f(x)f'(x), -f(x)\cos\theta, -f(x)\sin\theta).$$

As a result,

$$|\mathbf{r}_x \times \mathbf{r}_\theta| = f(x)\sqrt{1 + (f'(x))^2}$$

We conclude that the area is given by

$$\begin{aligned} A = \int\int_D |\mathbf{r}_x \times \mathbf{r}_\theta| dA &= \int_0^{2\pi} \int_a^b f(x)\sqrt{1+(f'(x))^2} dx d\theta \\ &= 2\pi \int_a^b f(x)\sqrt{1+(f'(x))^2} dx \end{aligned}$$

As an illustration for the application of the formula, we can provide a third method for the computation of the sphere area:

- **Method 3** Note that the sphere is indeed a surface of revolution, by rotating

$$f(x) = \sqrt{1 - x^2}, \quad -1 \leq x \leq 1$$

The parametrisation should then be

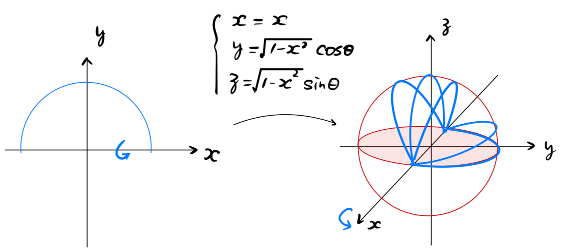


Figure 4.14: Sphere Surface Area: Method 3

$$x = x, y = \sqrt{1 - x^2}\cos\theta, z = \sqrt{1 - x^2}\sin\theta, \quad -1 \leq x \leq 1, 0 \leq \theta \leq 2\pi$$

-

Now by using the above formula, we have

$$A = 2\pi \int_{-1}^{1} \sqrt{1-x^2}\sqrt{1+\left(-\frac{x}{\sqrt{1-x^2}}\right)^2}dx = 2\pi \int_{-1}^{1} 1dx = 2\pi \int_{-1}^{1} \sqrt{1-x^2}\frac{1}{\sqrt{1-x^2}}dx$$

$$= 4\pi$$

Chapter 5 Vector Fields

5.1 Basic Definitions and Examples

From now until the end of this lecture, we always assume that

- A domain $D \subseteq \mathbb{R}^n$ is an **open** and **connected** subset of $\mathbb{R}^n$
- All functions discussed are smooth – they have continuous (partial) derivatives of all orders.

Definition 5.1. Vector Field

*Let $D \subseteq \mathbb{R}^n$ be a domain in $\mathbb{R}^n$. A **vector field** is a multi-variable vector valued function $\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$.* ♣

Remark Before we go further, let's first compare the following items:

- A **parametric curve** can be seen as a vector valued function
$$\gamma(t) : [a, b] \in \mathbb{R} \to \mathbb{R}^n, n \geq 1;$$
- A **parametric surface** can be seen as a vector valued function
$$\gamma(u, v) : D \in \mathbb{R}^2 \to \mathbb{R}^n, n \geq 2;$$
- A **multi-variable (scalar) function** is a real-valued function
$$f : D \subseteq \mathbb{R}^n \to \mathbb{R};$$
- A **vector field** is a vector valued function
$$\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n.$$

Geometric Meaning

A vector field $\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ can be seen as a rule of associating a vector $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$ to each point $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$.

How to draw a vector field

Here is a step-by-step guidance on how to draw a vector field $F(x, y) = (u(x, y), v(x, y))$ in 2D:

1. Make two planes, one is xy plane, another is the uv plane.
2. Take a point $(x_0, y_0) \in D$
   - Using form of the vector field **F** to find $(u_0, v_0) = F(x_0, y_0)$, draw this vector in the uv-plane ,with starting point being $(0, 0)$.
   - Translate this vector (u_0, v_0) to xy-plane, with starting point being (x_0, y_0).
3. Repeat the process for as many as possible points in D.

Remark The same process works for vector fields in $\mathbb{R}^3$ as well. For $\mathbb{R}^4$ and above, unfortunately we cannot visualise it directly.

Example5.1 Expansion The vector field

$$\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2, \qquad \mathbf{F}(x, y) = (x, y)$$

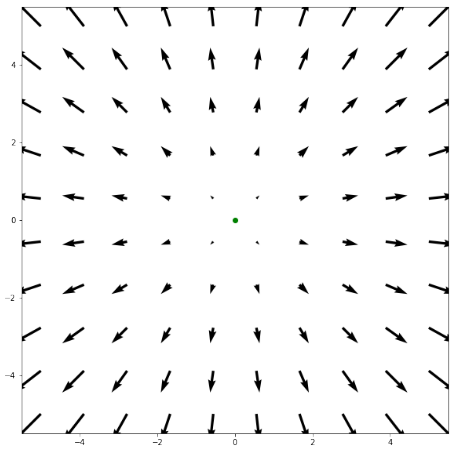


Figure 5.1: Vector field of expansion

If you consider this vector field as the velocity vector field, then particle in this field will run away faster and faster from the origin.

♦

Example5.2 Rotation The vector field

$$\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2, \qquad \mathbf{F}(x, y) = (-y, x)$$

-

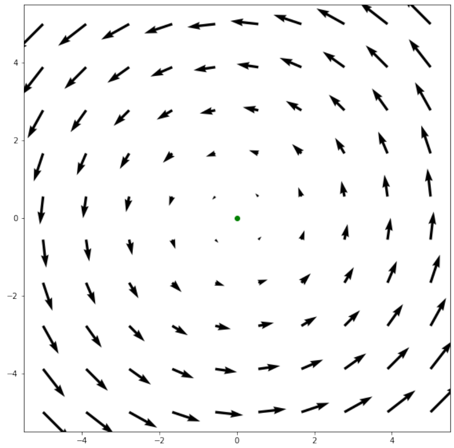


Figure 5.2: Vector field of rotation

If you consider this vector field as the velocity vector field, then particle in this field will rotate around the center at a constant speed.

♦

Example5.3 Gravity Field The vector field

$$\mathbf{F} : \mathbb{R}^2 \setminus \{\mathbf{0}\} \to \mathbb{R}^2, \qquad \mathbf{F}(x, y) = \left(\frac{-x}{(x^2+y^2)^{\frac{3}{2}}}, \frac{-y}{(x^2+y^2)^{\frac{3}{2}}} \right)$$

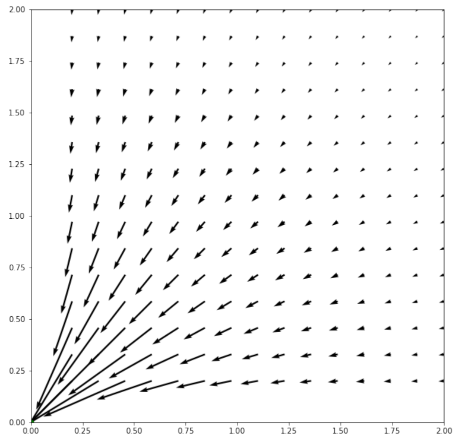


Figure 5.3: Vector field of gravity

This is the gravity force field in 2D, particle in this field will be under the force of attraction from the origin, in particular, the closer it is to origin, the greater the force becomes.

♦

Curve Integral of Vector Field

Let's consider the curve integral along a vector field. Let C be an **oriented** (piecewise) smooth curve contained in $D \subset \mathbb{R}^n$, and $\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a vector field. We use the symbol

$$\int_C \mathbf{F} \cdot d\mathbf{r}$$

to represent the integral of vector field $\mathbf{F}$ along the oriented curve C.

Remark

- Note that the above notation is just a **symbol**, we need to give it a meaning, so that it can be interpreted as a real number. To this end, let $\mathbf{r}(t) = (x_1(t), x_2(t), ..., x_n(t)), t \in [a, b]$ be a parametrisation of C **consistent with the orientation**. Then the above curve integral is computed as

$$\int_C \mathbf{F} \cdot d\mathbf{r} \stackrel{\text{def}}{=\!=} \int_a^b \underbrace{\mathbf{F} \cdot \dot{\mathbf{r}}}_{\text{inner product}} \; dt \tag{$\star$}$$

where

$$\dot{\mathbf{r}}(t) = (\dot{x}_1(t), \dot{x}_2(t), ..., \dot{x}_n(t))$$

is the velocity vector along the parametric curve at time t. Note that $\mathbf{F} \cdot \dot{\mathbf{r}}(t)$ is a scalar function, and the above integral is just a usual **Riemann Integral** that you have learnt in MAT136.

- Recall that the vector field can be written componentwisely:

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), ...F_n(\mathbf{x})).$$

Thus another equivalent way of writing $\int_C \mathbf{F} \cdot d\mathbf{r}$ is to denote it by

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C F_1(\mathbf{x})dx_1 + F_2(\mathbf{x})dx_2 + ... + F_n(\mathbf{x})dx_n$$

Again this is a **symbol**. To determine the value, one should compute the associated Riemann integral ($\star$) as described above.

How to compute $\int_C \mathbf{F} \cdot d\mathbf{r}$

Here is a step-by-step guidance on how to compute $\int_C \mathbf{F} \cdot d\mathbf{r}$. Suppose that the vector field

$$\mathbf{F}(x_1, x_2, ..., x_n) = (F_1(x_1, x_2, ..., x_n), F_2(x_1, x_2, ..., x_n), ..., F_n(x_1, x_2, ..., x_n))$$

is given. To compute $\int_C \mathbf{F} \cdot d\mathbf{r}$,

1. Parametrise C by $\mathbf{r}(t) : [a, b] \to C$ in a **consistent** way.
2. Compute the velocity vector field
$$\dot{\mathbf{r}}(t) = (\dot{x}_1(t), \dot{x}_2(t), ..., \dot{x}_n(t))$$
3. Compute the inner product
$$\mathbf{F}(x_1(t), x_2(t), ..., x_n(t)) \cdot \dot{\mathbf{r}}(t) = \sum_{i=1}^{n} F_i(x_1(t), x_2(t), ..., x_n(t))\dot{x}_i(t)$$
which gives you a **scalar function of t**.
4. Compute the usual Riemann integral
$$\int_a^b \mathbf{F}(x_1(t), x_2(t), ..., x_n(t)) \cdot \dot{\mathbf{r}}(t) dt$$
and ends with a real number.

Remark In most cases we will only study curves that are in either $2D$ or $3D$, and the formula becomes simpler:

- **In 2D**: Let
$$\mathbf{F}(x, y) = (P(x, y), Q(x, y))$$
be the vector field and let
$$\mathbf{r}(t) = (x(t), y(t))$$
be the parametrisation, then
$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_a^b P(x(t), y(t))\dot{x}(t) + Q(x(t), y(t))\dot{y}(t)dt.$$
- **In 3D**: Let
$$\mathbf{F}(x, y, z) = (P(x, y, z), Q(x, y, z), R(x, y, z))$$
be the vector field and let
$$\mathbf{r}(t) = (x(t), y(t), z(t))$$
be the parametrisation, then
$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_a^b P(x(t), y(t), z(t))\dot{x}(t) + Q(x(t), y(t), z(t))\dot{y}(t) + R(x(t), y(t), z(t))\dot{z}(t)dt.$$

Example5.4 Evaluate the integral

$$\int_{C_1} y^2 dx + x dy$$

where C_1 is part of the parabola $x = 4 - y^2$, starting from $(-5, -3)$ and ending at $(0, 2)$.

Solution First, note that the words "starting from" and "ending at" suggest the orientation of the curve. We solve it step by step.

1. The parametrisation can be easily done: after all we are dealing with a graph $x = f(y)$ so the natural parametrisation works, i.e.

$$\mathbf{r}(t) : \begin{cases} x(t) & = 4 - t^2 \\ y(t) & = t \quad -3 \leq t \leq 2 \end{cases}$$

2. The velocity vector is computed as

$$\dot{\mathbf{r}}(t) : \begin{cases} \dot{x}(t) & = -2t, \\ \dot{y}(t) & = 1 \quad -3 \leq t \leq 2 \end{cases}$$

3. The inner product is computed as

$$\begin{aligned} \mathbf{F}(x(t), y(t)) \cdot \dot{\mathbf{r}}(t) &= y(t)^2 \dot{x}(t) + x(t)\dot{y}(t)) \\ &= t^2 \cdot (-2t) + 1 \cdot (4 - t^2) \\ &= -2t^3 - t^2 + 4 \end{aligned}$$

4. The Riemann integral is computed as

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_{-3}^{2} \mathbf{F}(x(t), y(t)) \cdot \dot{\mathbf{r}}(t) dt \\ &= \int_{-3}^{2} -2t^3 - t^2 + 4dt \\ &= -\frac{1}{2}t^4 - \frac{1}{3}t^3 + 4t \Big|_{-3}^{2} = \frac{245}{6}. \end{aligned}$$

Example5.5 Evaluate the integral

$$\int_{C_2} y^2 dx + x dy$$

where C_2 is part of the straight line starting from $(-5, -3)$ and ending at $(0, 2)$.

Solution First, note that the words "starting from" and "ending at" suggest the orientation of the curve. We solve it step by step.

1. The parametrisation can be easily done: it is a segment of a straight line, i.e.

$$\mathbf{r}(t) : \begin{cases} x(t) & = -5 \cdot (1 - t) + 0 \cdot t = 5t - 5, \\ y(t) & = -3 \cdot (1 - t) + 2 \cdot t = 5t - 3. \quad 0 \leq t \leq 1 \end{cases}$$

2. The velocity vector is computed as

$$\dot{\mathbf{r}}(t) : \begin{cases} \dot{x} &= 5, \\ \dot{y} &= 5. \quad 0 \le t \le 1 \end{cases}$$

3. The inner product is computed as

$$\begin{aligned} \mathbf{F}(x(t), y(t)) \cdot \dot{\mathbf{r}}(t) &= y(t)^2 \dot{x}(t) + x(t)\dot{y}(t)) \\ &= (5t-3)^2 \cdot 5 + (5t-5) \cdot 5 \\ &= 125t^2 - 125t + 20. \end{aligned}$$

4. The Riemann integral is computed as

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_{-3}^{2} \mathbf{F}(x(t), y(t)) \cdot \dot{\mathbf{r}}(t) dt \\ &= \int_0^1 125t^2 - 125t + 20 dt \\ &= \frac{125}{3}t^3 - \frac{125}{2}t^2 + 20t \Big|_0^1 = -\frac{5}{6}. \end{aligned}$$

♦

Remark In the two examples above, we have computed the integral of the **same** vector field

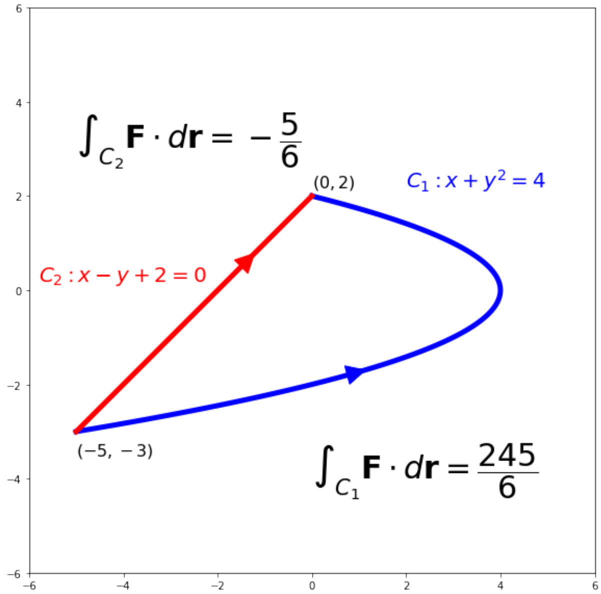


Figure 5.4: Integral of Vector Fields along different curves

along **different** curves sharing the **same** starting & ending points. The results are **different**. This tells us that:

Proposition 5.1. Dependence on the whole curve

Let C be a curve starting from $\mathbf{x}_0$ ending at $\mathbf{x}_1$. The integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ ***in general*** *depends on* ***the whole curve*** *C.*

However, if the vector fields is a conservative (gradient) vector field, then the situation would become much easier. This will be explored in the next section.

5.2 Conservative Vector Field

We first give the definition of a conservative vector field, also known as a gradient vector field.

Definition 5.2. Conservative Vector Fields

*A vector field $\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is called a **conservative vector field**, if there exists some **scalar** function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ s.t.*

$$\nabla f = \mathbf{F}.$$

In this case,

- *$\mathbf{F}$ is called the **gradient vector field** of f, and*
- *f is called the **potential** of $\mathbf{F}$.*

♣

Example5.6 Let $f(x, y, z) = xyz^2$, then

$$\mathbf{F}(x, y, z) = (yz^2, xz^2, 2xyz)$$

is the gradient vector field, obtained by computing the gradient of f.

♦

Example5.7 Consider the vector field

$$\mathbf{F}(x, y) = \left(-\frac{x}{(x^2+y^2)^{\frac{3}{2}}}, -\frac{y}{(x^2+y^2)^{\frac{3}{2}}}\right), \quad (x, y) \in D = \mathbb{R}^2 \setminus \{\mathbf{0}\}.$$

It is conservative since we know that the function

$$\mathbf{f}(x, y) = \frac{1}{(x^2+y^2)^{\frac{1}{2}}}$$

serves as the potential of $\mathbf{F}$: indeed, one verifies by direct computation (do it!) that $\nabla f = \mathbf{F}$.

♦

Remark The potential in the above example is called the **Newton's potential** for gravity and plays an essential role in the development of celestial mechanics.

The gradient vector field has a nice property concerning the curve integral:

Proposition 5.2. Dependence on the initial point and the end point

Let $\mathbf{F} = \nabla f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a gradient vector field. Then for any two fixed points $\mathbf{x}_0, \mathbf{x}_1 \in D$, and any curve C starting from $\mathbf{x}_0$ ending at $\mathbf{x}_1$, one has that

$$\int_C \mathbf{F} \cdot d\mathbf{r} = f(\mathbf{x}_1) - f(\mathbf{x}_0)$$

♠

Proof Let $\mathbf{r}(t) = (\mathbf{x}(t)), t \in [a, b]$ be a parametrisation of C, s.t. $\mathbf{r}(a) = \mathbf{x}_0, \mathbf{r}(b) = \mathbf{x}_1$. If follows that

$$\int_C \mathbf{F} \cdot d\mathbf{r} \stackrel{\text{def}}{=\!=} \int_a^b \mathbf{F} \cdot \dot{\mathbf{r}} dt \stackrel{\mathbf{F}=\nabla f}{=\!=} \int_a^b \nabla f \cdot \dot{\mathbf{r}} dt \stackrel{\text{Chain Rule}}{=\!=} \int_a^b \frac{d}{dt}(f \circ \mathbf{r}(t))\, dt \stackrel{\text{FTC}}{=\!=} f \circ r(b) - f \circ r(a)$$
$$= f(\mathbf{x}_1) - f(\mathbf{x}_0).$$

□

Remark Several remerks should be mentioned here:

- This result is in sharp contrast to general vector fields, as we have already seen that the integral of a vector field along different curves connecting two fixed points might be different (review example 5.4 and example 5.5 in the previous section). However, for gradient vector field, the integral only depends on the two endpoints, indeed the above proposition leads to the following two corollaries directly:

  Corollary 5.1. Independence of Curves

  Let $\mathbf{F} = \nabla f$ be a conservative (gradient) vector in $D \subseteq \mathbb{R}^n$, and let $\mathbf{x}_0, \mathbf{x}_1$ be two points. For any two curves C_1 and C_2 that both start from $\mathbf{x}_0$ and end at $\mathbf{x}_1$, one has that

  $$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}$$

  ♡

  Proof This follows from

  $$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = f(\mathbf{x}_1) - f(\mathbf{x}_0) = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

  □

- In particular, if we are dealing with a closed curve C, i.e., $\mathbf{x}_0 = \mathbf{x}_1$, then the above proposition actually implies that

  Corollary 5.2. Vanished integral along closed curves

  Let $\mathbf{F} = \nabla f$ be a conservative (gradient) vector in $D \subseteq \mathbb{R}^n$, and let $C \subseteq D$ be a closed curve(i.e. the initial point and the terminal point are both $\mathbf{x}_0$). Then one has that

  $$\int_C \mathbf{F} \cdot d\mathbf{r} = 0.$$

  ♡

  Proof This follows from

  $$\int_C \mathbf{F} \cdot d\mathbf{r} = f(\mathbf{x}_0) - f(\mathbf{x}_0) = 0.$$

  □

When is F conservative?

We have seen that the conservative vector fields possess nice properties, in particular the an integral of a gradient vector field depends only on the potential values on endpoints themselves, ranther than the path connecting them. This provides us significant convenience as we no longer need to parametrise the curve C in computing the integral. Nevertheless, we have to answer the following important question:

Important Question

Given a vector field $\mathbf{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$, how do we know if $\mathbf{F}$ is conservative or not?

In this section we focus mainly in $\mathbb{R}^2$. The result developed here can essentially be generalised to $\mathbb{R}^3$ (and indeed $\mathbb{R}^n$), although some extra effort need to be made.

Three conditions and their relations Let $\mathbf{F}(x, y) = (P(x, y), Q(x, y))$ be a vector field in $D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$. Let's discuss the following three conditions:

Gradient Vector Field Condition

$$\exists f : D \to \mathbb{R}, \text{ s.t.} \quad \mathbf{F}(x, y) = \nabla f(x, y). \qquad \text{(Condition I)}$$

Closed Curve Integral Vanishing Condition

$$\forall \text{ closed curve } C \subset D, \quad \int_C \mathbf{F} \cdot d\mathbf{r} = 0. \qquad \text{(Condition II)}$$

Partial Derivative Conditions

$$\forall (x, y) \in D, \quad P_y(x, y) = Q_x(x, y). \qquad \text{(Condition III)}$$

Remark If we are given a vector field $\mathbf{F}$ in $3D$, i.e.

$$\mathbf{F}(x, y, z) = (P(x, y, z), Q(x, y, z), R(x, y, z)),$$

then condition III becomes slightly more complicated:

Partial Derivative Conditions

$$\forall (x, y, z) \in D, \quad \begin{cases} P_y(x, y, z) = Q_x(x, y, z), \\ Q_z(x, y, z) = R_y(x, y, z), \\ R_x(x, y, z) = P_z(x, y, z) \end{cases} \qquad \text{(Condition III)}$$

We will see in the next section such a condition can be concisely written as $\mathrm{curl}(\mathbf{F}) = \mathbf{0}$.

We give the criterion for a vector field being conservative by the theorem below:

Theorem 5.1. Criterion of being gradient vector field

Let D be a domain, i.e., open and connected. Given a vector field

$$\mathbf{F} = (P, Q) : D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$$

. Then

- *Condition I and condition II are equivalent.*
- *Condition I and II each implies condition III.*
- *If moreover D is **simply connected**, then condition III implies condition I and II.*
- *If D is **not simply connected**, then condition III is not enough to either conclude or exclude the conservativity of the vector field.*

♡

Proof Discussed in detail in lecture and homework, we ignore it here.

□

We also summarise the mutual relations among these conditions in figure below.

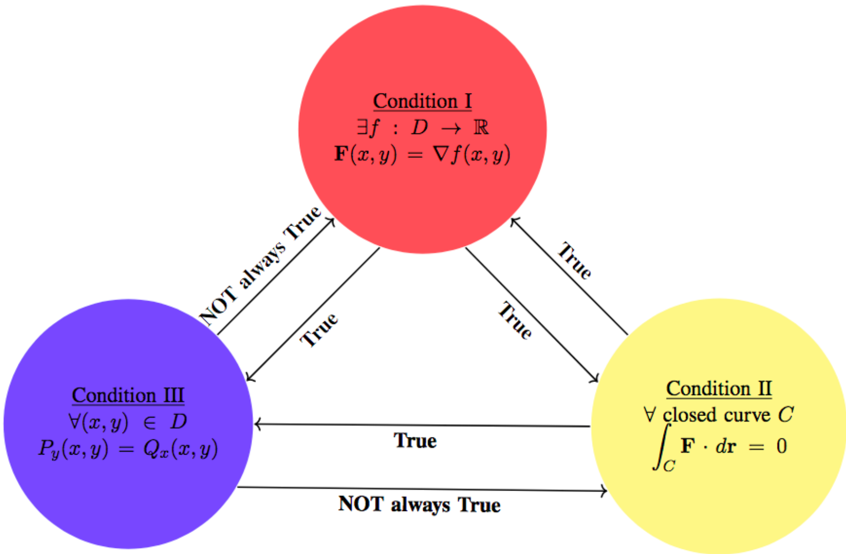


Figure 5.5: Relations among condtions I,II,III

Important Remark

Let D be a domain, i.e., open and connected. Given a vector field

$$\mathbf{F} = (P, Q) : D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$$

which satisfies that

$$\forall (x, y) \in D, \quad P_y(x, y) = Q_x(x, y).$$

If furthermore the domain is **not** simply connected, then

- $\mathbf{F}$ might be conservative, see example 5.7.
- $\mathbf{F}$ might not be conservative, see example 5.8.

Example5.8 Consider the vector field $\mathbf{F} : \mathbb{R}^2 \setminus \{\mathbf{0}\} \to \mathbb{R}^2$, defined by

$$\mathbf{F}(x, y) = (P(x, y), Q(x, y)) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right).$$

- **Partial derivative condition is satisfied**:
$$P_y = \frac{-(x^2 + y^2) - (-y)2y}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{(x^2 + y^2) - x(2x)}{(x^2 + y^2)^2} = Q_x.$$
- **Closed curve integral vanishing condition is violated**: To this end, consider the unit circle oriented counter-clockwisely. One finds the parametrisation
$$\mathbf{r}(t) : \begin{cases} x(t) = \cos t, \\ y(t) = \sin t. \quad 0 \leq t \leq 2\pi. \end{cases} \Rightarrow \dot{\mathbf{r}}(t) : \begin{cases} \dot{x}(t) = -\sin t, \\ \dot{y}(t) = \cos t. \quad 0 \leq t \leq 2\pi. \end{cases}$$
One can then compute the line integral by definition:

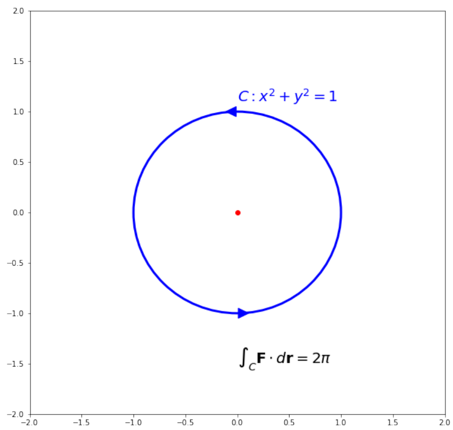


Figure 5.6: A non-conservative vector field

$$\int_C \mathbf{F} d\mathbf{r} = \int_0^{2\pi} \left(\frac{-\sin t}{\sin^2 t + \cos^2 t}, \frac{\cos t}{\sin^2 t + \cos^2 t}\right) \cdot (-\sin t, \cos t) dt = \int_0^{2\pi} 1 dt = 2\pi \neq 0.$$

We conclude that $\mathbf{F}$ is **NOT** a conservative vector field in $D = \mathbb{R}^2 \setminus \{\mathbf{0}\}$.

♦

From Gradient Vector Field to Potential

We end this session by discussing the computational aspects, namely:

-

- Given a potential function, how can we find the gradient vector field? — This is trivial, just compute

$$\nabla f = (f_{x_1}, f_{x_2}, \ldots, f_{x_n})$$

- Given a gradient vector field, how do we find its potential? — This is less trivial, and should be done via several operations of integral.

How to find the potential

Given a vector field $\mathbf{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$:

1. Check the partial derivative condition III. If it does not hold, stop — the field cannot have any potential.
2. If partial derivative condition III holds, then check the domain: if the domain is simply connected, then the potential exists.
3. If the potential exists, then carry out a sequence of integrals of find the potential.

Example5.9 From gradient vector field to potential Decide if the vector field $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$F(x, y, z) = (y^2, 2xy + e^{3z}, 3ye^{3z})$$

is conservative or not. In case yes, find its potential.

Solution We solve this problem step by step.

1. First, check the derivative condition:

$$\frac{\partial}{\partial y} 3ye^{3z} = 3e^{3z} = \frac{\partial}{\partial z}(2xy + e^{3z});$$
$$\frac{\partial}{\partial z} y^2 = 0 = \frac{\partial}{\partial x} 3ye^{3z};$$
$$\frac{\partial}{\partial x}(2xy + e^{3z}) = 2y = \frac{\partial}{\partial y} y^2.$$

   As a result the partial derivatives condition is satisfied.

2. Second, note that the domain $\mathbb{R}^3$ is simply connected. We conclude that the vector field is indeed conservative.
3. Last but not least, let's find out f s.t.

$$f = \nabla \mathbf{F}.$$

By definition of gradient, one has that

$$\begin{aligned}
&\frac{\partial}{\partial x} f = y^2 \Rightarrow f(x,y,z) = xy^2 + g(y,z). \\
&\frac{\partial}{\partial y} f = 2xy + e^{3z} = 2xy + \frac{\partial}{\partial y} g \\
&\quad \Rightarrow g(y,z) = ye^{3z} + h(z) \\
&\quad \Rightarrow f(x,y,z) = xy^2 + ye^{3z} + h(z). \\
&\frac{\partial}{\partial z} f = 2xy + 3ye^{3z} + \frac{\partial}{\partial z} h = 2xy + 3ye^{3z} \\
&\quad \Rightarrow h(z) = C \\
&\quad \Rightarrow f(x,y,z) = xy^2 + ye^{3z} + C,
\end{aligned}$$

where C is an arbitrary constant.

We conclude that $\mathbf{F} = \nabla f$ is a gradient vector field, with a potential being, for instance (by choosing $C = 0$),

$$f(x,y,z) = xy^2 + ye^{3z}.$$

♦

5.3 Curl and Divergence

Consider a given smooth vector field $\mathbf{F}$ in $\mathbb{R}^3$. We associate two operations to this vector field.

Definition 5.3. Curl and Divergence

Let $\mathbf{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$*. Then one defines*

$$\operatorname{curl}\mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right).$$

$$\operatorname{div}\mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

♣

Remark The natural of curl and that div are not the same:

- The curl$\mathbf{F}$ is a **vector field**.
- The div$\mathbf{F}$ is a **scalar function**.

Remark If we consider $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ as a **symbolic** vector, then the curl and div can be seen as results of the following symbolic operation:

$$\operatorname{curl}\mathbf{F} = \nabla \times \mathbf{F}, \quad \operatorname{div}\mathbf{F} = \nabla \cdot \mathbf{F}.$$

Warning : Don't get confused between the two notations ∇f and $\nabla \cdot \mathbf{F}$. The former means taking the gradient, which is applied to a scalar function; the latter means taking the divergence, which is applied to a vector field. Note that there is a "·" in the divergence notation.

Example5.10 Consider the vector field $\mathbf{F}(x,y,z) = (z+\sin x\cos y, xyz, x^2+z^2)$, then

$$\begin{aligned}\mathrm{curl}\mathbf{F} &= (0-xy, 1-2z, yz+z\sin x\sin y) = (-xy, 1-2z, z(y+\sin x\sin y)).\\ \mathrm{div}\mathbf{F} &= \frac{\partial}{\partial x}(z+\sin x\cos y) + \frac{\partial}{\partial y}xyz + \frac{\partial}{\partial z}(x^2+z^2) = \cos x\cos y + xz + 2z.\end{aligned}$$

♦

Example5.11 Consider the vector field $\mathbf{F}(x,y,z) = (x^2, \sin y, e^z)$, then

$$\begin{aligned}\mathrm{curl}\mathbf{F} &= \left(\frac{\partial e^z}{\partial y} - \frac{\partial \sin y}{\partial z}, \frac{\partial x^2}{\partial z} - \frac{\partial e^z}{\partial x}, \frac{\partial \sin y}{\partial x} - \frac{\partial x^2}{\partial y}\right) = (0-0, 0-0, 0-0) = \mathbf{0}.\\ \mathrm{div}\mathbf{F} &= \frac{\partial}{\partial x}x^2 + \frac{\partial}{\partial y}\sin y + \frac{\partial}{\partial z}e^z = 2x + \cos y + e^z.\end{aligned}$$

♦

Remark In general, if a vector field is of the form $\mathbf{F}(x,y,z) = (P(x), Q(y), R(z))$, then we see that $\mathrm{curl}\mathbf{F} = \mathbf{0}$.

Example5.12 Consider the vector field $\mathbf{F}(x,y,z) = (y^2+z^2, \sin x\cos z, e^{x+y})$, then

$$\begin{aligned}\mathrm{curl}\mathbf{F} &= \left(\frac{\partial e^{x+y}}{\partial y} - \frac{\partial \sin x\cos z}{\partial z}, \frac{\partial (y^2+z^2)}{\partial z} - \frac{\partial e^{x+y}}{\partial x}, \frac{\partial \sin x\cos z}{\partial x} - \frac{\partial (y^2+z^2)}{\partial y}\right)\\ &= (e^{x+y} + \sin x\sin z, 2z - e^{x+y}, \cos x\cos z - 2y).\\ \mathrm{div}\mathbf{F} &= \frac{\partial}{\partial x}(y^2+z^2) + \frac{\partial}{\partial y}\sin x\cos z + \frac{\partial}{\partial z}e^{x+y} = 0+0+0 = 0.\end{aligned}$$

Remark In general, if a vector field is of the form $\mathbf{F}(x,y,z) = (P(y,z), Q(x,z), R(x,y))$, then we see that $\mathrm{div}\mathbf{F} = 0$.

Example5.13 Linear Vector Fields Consider the vector field $\mathbf{F}(x,y,z) = A\begin{bmatrix} x\\ y\\ z\end{bmatrix}$, where

$$A = \begin{bmatrix} 1 & 3 & 2\\ 2 & 0 & 2\\ 3 & 1 & 1\end{bmatrix}$$

In this case, one sees that

$$F(x,y,z) = (x+3y+2z, 2x+2z, 3x+y+z).$$

Now we see that

$$\begin{aligned}\mathrm{curl}\mathbf{F} &= (1-2, 2-3, 2-3) = (-1,-1,-1).\\ \mathrm{div}\mathbf{F} &= 1+0+1 = 2.\end{aligned}$$

Remark Vector fields of the above form are called linear vector field. In particular, we see that the curl of a linear vector field is a constant vector field, and the div of linear vector field is a constant function.

5.4 Helmholtz Decomposition

The Helmholtz decomposition plays a key role in the study of 3D vector fields, and the div and curl are precisely the terms we need to describe this result. To this end, we give the following definitions:

Definition 5.4. Irrotational vector field

A vector field $\mathbf{F} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ *is called* ***irrotational****, if*

$$\forall (x, y, z) \in D, \mathrm{curl}\mathbf{F} = \mathbf{0}.$$

♣

Definition 5.5. Irrotational vector field

A vector field $\mathbf{F} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ *is called* ***incompressible****, if*

$$\forall (x, y, z) \in D, \mathrm{div}\mathbf{F} = 0.$$

♣

Remark In literature, “irrotational" is also called “**curl free**", and “incompressible" is also called “**divergence free**".

Now note that we can generate new vector fields using scalar function a g or a vector field $\mathbf{G}$ by the following operations:

- We can generate new vector field by taking ∇g.
- We can generate new vector field by taking $\mathrm{curl}\mathbf{G}$.

The following proposition in general shows that the vector fields generated by the above two operations are, in some sense, “orthogonal" to each other. Of course, we always assume the objects involved in the discussion is smooth enough so that such operations make sense.

Proposition 5.3

- *Let* $\mathbf{F}$ *be any* $3D$ *vector field, then* $\mathrm{curl}F$ *is incompressible, i.e.* $\mathrm{div}\mathrm{curl}\mathbf{F} = 0$.
- *Let* g *be any scalar function, then* ∇g *is irrotational, i.e.* $\mathrm{curl}\nabla g = \mathbf{0}$.

♠

Proof The proof is done by direct computation.

- Suppose that $\mathbf{F} = (P, Q, R)$ is a $3D$ vector field, then

$$\begin{aligned}
\mathrm{div}\mathrm{curl}\mathbf{F} &= \mathrm{div}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \\
&= \frac{\partial^2 R}{\partial x \partial y} - \frac{\partial^2 R}{\partial y \partial x} + \frac{\partial^2 Q}{\partial z \partial x} - \frac{\partial^2 Q}{\partial x \partial z} + \frac{\partial^2 R}{\partial y \partial x} - \frac{\partial^2 R}{\partial x \partial y} \\
&= 0 + 0 + 0 \\
&= 0.
\end{aligned}$$

  where the second to last equality is by the Clairaut’s theorem, which permits us to exchange the order of taking partial derivatives.

- Suppose that $f : D \subseteq \mathbb{R}^3 \to \mathbb{R}$ is a scalar function, then

$$\begin{aligned}\operatorname{curl}(\nabla f) &= \operatorname{curl}\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)\\ &= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y}, \frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z}, \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x}\right)\\ &= (0,0,0)\\ &= \mathbf{0}.\end{aligned}$$

  where the second to last equality is again by the Clairaut's theorem, which permits us to exchange the order of taking partial derivatives.

Now we give the central result of this section without providing the proof.

Theorem 5.2. Helmholtz Decomposition

Let $\mathbf{F} : D \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, then there exists a decomposition of $\mathbf{F}$ into the sum of vector fields $\mathbf{F_1}$ and $\mathbf{F}_2$, i.e.

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2$$

s.t.

- $\mathbf{F}_1$ *is irrotational,*
- $\mathbf{F}_2$ *is incompressible.*

♡

Example5.14 Consider the vector field

$$\mathbf{F}(x,y,z) = (x-y, y+x, 0) = \mathbf{F}_1(x,y,z) + \mathbf{F}_2(x,y,z),$$

where

$$\begin{aligned}\mathbf{F}_1(x,y,z) &= (x,y,0), \quad \operatorname{div}\mathbf{F}_1 = 2, \operatorname{curl}\mathbf{F}_1 = (0,0,0).\\ \mathbf{F}_2(x,y,z) &= (-y,x,0), \quad \operatorname{div}\mathbf{F}_2 = 0, \operatorname{curl}\mathbf{F}_1 = (0,0,2).\end{aligned}$$

Hence the vector field $\mathbf{F}_1$ is curl-free, while the vector field $\mathbf{F}_2$ is divergence free.

♦

Remark In this particular example, the irrotational part and the incompressible part can be seen from the vector field drawing directly, as is shown in example 5.1 and example 5.2, respectively, hence justified the names of these two types of vector fields.

Remark The decomposition in general is **not unique**: Given a vector field $\mathbf{F}$, it might happen that

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 = \mathbf{G}_1 + \mathbf{G}_2$$

where

- $\mathbf{F}_1$, $\mathbf{G}_1$ are both irrotational,and $\mathbf{F}_1 \neq \mathbf{G}_1$;
- $\mathbf{F}_2$, $\mathbf{G}_2$ are both incompressible, and $\mathbf{F}_2 \neq \mathbf{G}_2$.

Example5.15 Non-uniqueness of Helmholtz decomposition Suppose that $\mathbf{F}$ has the Helmholtz decomposition:

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2, \quad \mathrm{curl}\mathbf{F}_1 = \mathbf{0}, \mathrm{div}\mathbf{F}_2 = 0.$$

Consider the linear vector field:

$$\mathbf{H}(x, y, z) = (z, 0, x).$$

It follows that

$$\mathrm{div}\mathbf{H} = 0, \mathrm{curl}\mathbf{H} = \mathbf{0}.$$

It follows that

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 = \underbrace{(\mathbf{F}_1 + \mathbf{H})}_{\mathbf{G}_1} + \underbrace{(\mathbf{F}_2 - \mathbf{H})}_{\mathbf{G}_2} = \mathbf{G}_1 + \mathbf{G}_2$$

Note that

$$\mathrm{curl}\mathbf{G}_1 = \mathrm{curl}\mathbf{F}_1 + \mathrm{curl}\mathbf{H} = \mathbf{0}$$

$$\mathrm{div}\mathbf{G}_2 = \mathrm{div}\mathbf{F}_2 + \mathrm{div}\mathbf{H} = 0.$$

Hence $\mathbf{F} = \mathbf{G}_1 + \mathbf{G}_2$ is another Helmholtz decomposition.

5.5 Surface Integral of Vector fields

Oriented Surface

Let's first define what is a line integral. To this end, we need the idea of oriented surface.

Definition 5.6. Oriented Surface

We say that a surface S is an oriented surface if it is possible to choose a unit normal vector $\mathbf{n}(x, y, z)$ at every point $(x, y, z) \in S$ s.t. $\mathbf{n}(x, y, z)$ varies continuously over S. ♣

Remark A typical example for the **non-orientable surface** is the **Möbius stripe**. One can consider a person holding an umbrella walking along the stripe. At the beginning the umbrella is pointing upward, but when it comes back to the same position the umbrella is pointing downward. You can watch the following link: Animation of walking on a Möbius stripe.

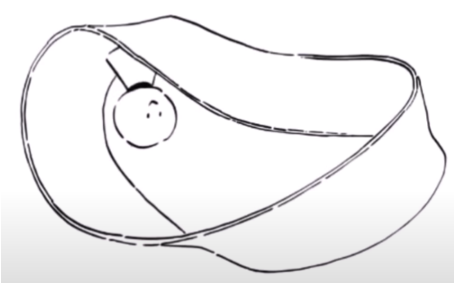


Figure 5.7: Möbius stripe (screen shot taken from the above YouTube link)

Remark Every surface is **locally** oriented. The problem arises only when we consider the surface **globally** Indeed, if we just move within a small piece of the Möbius stripe,we will not see any abnormal thing, and we can always choose an orientation in that small region. It is only when one moves along the whole stripe and move back to the origin that we found the inconsistency.

Induced Orientation Suppose that S is an oriented surface with normal vectors $\mathbf{n}(x, y, z)$ at each given point x, y, z. Then it induces an orientation on the boundary ∂S: consider a person walking along the beach of an ocean and the normal vectors are the lighthouses. If the person always finds the lighthouses on the left while walking, then we say ∂S is equipped with the induced orientation from S.

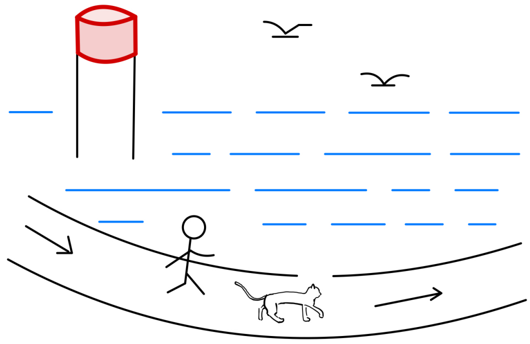


Figure 5.8: Feel the induced orientation by walking around Oakville lighthouse along the coast of Ontario lake (with a cat)

Now let's define the following two operations.

Integral of scalar function on a surface

Let $f : U \subseteq \mathbb{R}^3 \to \mathbb{R}$ be a **scalar function**. The integral of f along the surface $S \subseteq U$, is denoted by

$$\iint_S f(x, y, z)dS$$

Remark As before, so far the expression $\iint_S f(x, y, z)dS$ is nothing more than a **symbol**, and we need to give it a meaning (as a real number). This is done as the following:

1. **Parametrisation**: Let

$$D \subseteq \mathbb{R}^2 \to S,$$
$$(u, v) \in D \to \mathbf{r}(u, v) = (x(u, v), y(u, v), z(u, v)) \in S$$

   be a parametrisation of the surface.

2. **Switch the integral into a usual Riemann (double) integral**:

$$\iint_S f(x, y, z)dS \overset{\text{def}}{=\!=} \iint_D f(x(u, v), y(u, v), z(u, v))|\mathbf{r}_u \times \mathbf{r}_v|dA$$

Note that the right hand side is a usual **double integral** of variable $(u, v) \in D$, which we already know how to compute.

Integral of vector fields on a surface

Let $\mathbf{F} : U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ be a **vector field**. The integral of $\mathbf{F}$ along the surface $S \subseteq U$, is denoted by

$$\iint_S \mathbf{F}(x,y,z) \cdot d\mathbf{S}.$$

Again, so far the expression $\iint_S \mathbf{F}(x,y,z) \cdot d\mathbf{S}$ is nothing more than a **symbol**, and we need to give it a meaning (as a real number). This is done as the following:

1. **Parametrisation**: Let

$$D \subseteq \mathbb{R}^2 \to S,$$
$$(u,v) \in D \to \mathbf{r}(u,v) = (x(u,v), y(u,v), z(u,v)) \in S$$

be a parametrisation of the surface.

2. **Switch the integral into a usual Riemann (double) integral**: Let $\mathbf{n}(x,y,z)$ be the **unit normal vector**, then

$$\begin{aligned}\iint_S \mathbf{F}(x,y,z) \cdot d\mathbf{S} &\stackrel{\text{def}}{=\!=} \iint_S \underbrace{\mathbf{F}(x,y,z) \cdot \mathbf{n}(x,y,z)}_{\text{a scalar function}} dS \\ &= \iint_D (\mathbf{F}(x,y,z) \cdot \mathbf{n}(x,y,z))|\mathbf{r}_u \times \mathbf{r}_v| dA \\ &= \iint_D (\mathbf{F}(x,y,z) \cdot \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}|\mathbf{r}_u \times \mathbf{r}_v| dA \\ &= \iint_D \underbrace{\mathbf{F}(x,y,z) \cdot (\mathbf{r}_u \times \mathbf{r}_v)}_{\text{a scalar function}} dA\end{aligned}$$

So again it becomes a Riemann (double) integral of scalar function on variables $(u,v) \in D$, which we know how to compute (by using Fubini theorem for example).

Remark If the parametric surface is indeed a graph of a function, i.e., $z = g(x,y)$, then we have a natural parametrisation:

$$\mathbf{r}(x,y) = (x, y, g(x,y)),$$

under which the normal vector and unit normal vector reads

$$\mathbf{r}_x \times \mathbf{r}_y = \left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right) \qquad \text{(normal vector)}$$

$$\mathbf{n}(x,y,z) = \frac{1}{\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}} \left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right) \qquad \text{(unit normal vector)}$$

It follows that for $\mathbf{F} = (P, Q, R)$,

$$\begin{aligned}\iint_S \mathbf{F}(x,y,z) \cdot d\mathbf{S} &= \iint_D \mathbf{F}(x,y,z) \cdot (\mathbf{r}_x \times \mathbf{r}_y) dA \\ &= \iint_D \underbrace{\left(-P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R\right)}_{\text{a scalar function}} dA\end{aligned}$$

5.6 Green's Theorem, Stokes' Theorem and Gauss' Theorem

Theorem 5.3. Stokes' Theorem

Consider the following setting:

- S: *oriented piecewise smooth* ***surface***;
- ∂S: *simple, closed, piecewise smooth* ***boundary curve*** *of* S *with induced positive orientation;*
- $\mathbf{F} : U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$: *smooth* ***vector field*** *defined on* U, $S \subseteq U$.

Then

$$\iint_S \mathrm{curl}\mathbf{F} \cdot d\mathbf{S} = \int_{\partial S} \mathbf{F} \cdot d\mathbf{r}$$

♡

Remark

- Note the different nature of the two objects:
  - LHS: *integral of vector field on a surface*;
  - RHS: *integral of vector field along a curve*.
- Stokes' theorem can simplify the computation if
  - either $\mathrm{curl}\mathbf{F} \cdot \mathbf{n}$ is easy to compute in S ($\mathrm{curl}\mathbf{F}$ itself could be complicated),
  - or $\mathbf{F} \cdot \dot{\mathbf{r}}$ is easy to compute on ∂S ($\mathbf{F}$ itself could be complicated).
- Given a simple closed curve C, there are **infinitely many** surfaces S s.t. $\partial S = C$. It follows that the integral $\iint_S \mathrm{curl}\mathbf{F} \cdot d\mathbf{S}$ is independent on the choice of such S: Indeed, if S_1, S_2 are two surfaces, and both the orientations of S_1 and S_2 induce the same orientation on ∂S, then

$$\iint_{S_1} \mathrm{curl}\mathbf{F} \cdot d\mathbf{S} = \int_{C=\partial S_1=\partial S_2} \mathbf{F} \cdot d\mathbf{r} = \iint_{S_2} \mathrm{curl}\mathbf{F} \cdot d\mathbf{S}$$

Theorem 5.4. Gauss' Theorem

Consider the following setting:

- E: *simple* ***solid***;
- ∂E: ***boundary surface*** *of* E *with positive (outward) orientation;*
- $\mathbf{F} : D \subseteq \mathbb{R}^3 \to \mathbb{R}^3$: *smooth* ***vector field*** *defined on* D, $E \subseteq D$.

Then

$$\iint_{\partial E} \mathbf{F} \cdot d\mathbf{S} = \iiint_E \mathrm{div}\mathbf{F} dV$$

♡

Remark

- Note the different nature of the two objects:
  - LHS: *integral of vector field on a surface*;

- RHS: *integral of scalar function in a solid region*.
- Gauss' theorem can simplify the computation if
  - either $\mathbf{F}\cdot\mathbf{n}$ is simple on ∂E,
  - or $\mathrm{div}\mathbf{F}$ is simple in E.
- In 3D, given a closed surface ∂S, there is **only one** solid region E s.t. $\partial E = S$.

Remark In general,

$$\mathbf{F} \underset{\text{less trivial}}{\overset{\text{trivial}}{\rightleftarrows}} \mathrm{curl}\mathbf{F}, \mathrm{div}\mathbf{F}$$

Because $\rightarrow$ involves **differentiation**, yet $\leftarrow$ involves **integration**, and in general the former is much easier than the latter.

Warning

There might be cases where the fancy theorem like Green, Stokes, or Gauss' theorem does not really provide much convenience. In such cases, one should never forget that we can still try to compute the integral by using its definition from the scratch!

In the following examples, we try to solve the problems always by using two ways: with fancy formula / without using fancy formula.

Example5.16 Evaluate $\displaystyle\int_C \mathbf{F}\cdot d\mathbf{r}$, where

$$\mathbf{F}(x,y,z) = (-y^2, x, z^2),$$
$$C = \{(x,y,z)\in\mathbb{R}^3 | y+z=2\} \cap \{(x,y,z)\in\mathbb{R}^3 | x^2+y^2=1\}.$$

Solution We present two methods of solving it:

- **Method I: Direct Computation**

  - **Parametrisation of curves**

    The curve can be parametrised as:

    $$\mathbf{r}(t) : x(t) = \cos t, y(t) = \sin t, z(t) = 2 - \sin t, 0 \le t \le 2\pi.$$

  - **Switch to the Riemann integral**

    Given

    $$\mathbf{F}(x,y,z) = (-y^2, x, z^2),$$
    $$\dot{\mathbf{r}}(t) = (-\sin t, \cos t, -\cos t),$$

    one has that

    $$\mathbf{F}\cdot\dot{\mathbf{r}}(t) = \sin^3 t + \cos^2 t - (2-\sin t)^2\cos t.$$

    Hence the integral we are looking for is

    $$\int_0^{2\pi} \sin^3 t + \cos^2 t - (2-\sin t)^2 \cos t dt.$$

- **Evaluation of the Riemann integral:** By definition of line integral of vector field,

$$\begin{aligned}\int_C \mathbf{F}\cdot d\mathbf{r} &= \int_0^{2\pi} \mathbf{F}(\mathbf{x}(\mathbf{t}),\mathbf{y}(\mathbf{t}),\mathbf{z}(\mathbf{t}))\cdot\dot{\mathbf{r}}(t)dt \\ &= \int_0^{2\pi}(\sin^3 t+\cos^2 t-(2-\sin t)^2\cos t)dt \\ &= \int_0^{2\pi}\sin^3 tdt+\int_0^{2\pi}\cos^2 tdt-\int_0^{2\pi}(2-\sin t)^2\cos tdt \\ &= 0+\pi+0 \\ &= \pi.\end{aligned}$$

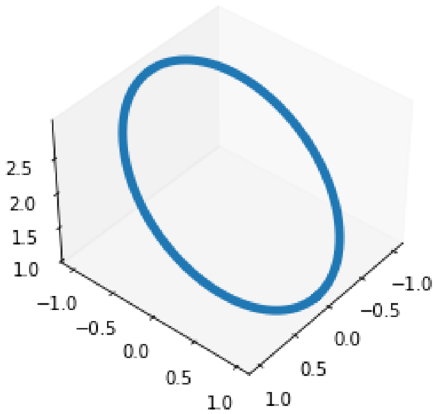


Figure 5.9: The parametric ellipse $(\cos t, \sin t, 2-\sin t), \quad 0\le t\le 2\pi$

- **Method II: Stokes' Theorem**
  - **Parametrisation of the surface**

Choosing the section of

$$\{(x,y,z)\in\mathbb{R}^3|y+z=2\} \quad \text{(Plane)}$$
$$\{(x,y,z)\in\mathbb{R}^3|x^2+y^2=1\} \quad \text{(Cylinder)}$$

which is an ellipse that serves as a natural candidate of the surface in search.(You can for sure find **infinitely many** other surfaces S having C as their boundary. However, the computation of its normal vector on an arbitrary S might be a difficult task!) The surface is achieved indeed as a graph of (x,y), i.e.

$$z(x,y):\{x^2+y^2\le 1\}\to\mathbb{R}, \quad z=g(x,y)=2-y$$

and can be parametrised naturally by

$$\mathbf{S}(x,y):(x,y,2-y)$$

It follows that

$$\left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right) = (0,1,1);$$

$$\text{curl}\mathbf{F} = \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y^2 & x & z^2 \end{vmatrix} = (0,0,1+2y).$$

Note that the vector $\mathbf{n}$ indeed induces the correct orientation on C (why?).

- **Switch to the Riemann (double) integral:**

$$\text{curl}\mathbf{F} \cdot \left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right) = (0,0,1+2y)\cdot(0,1,1) = 1+2y.$$

Since x, y are the parameters of the parametrisation, the domain D is just

$$D = \{(x,y)|x^2+y^2<1\}.$$

The double integral we are looking for is then

$$\iint_D 1+2ydA.$$

- **Evaluation of the Riemann (double) integral:** The double integral is then evaluated directly:

$$\iint_D 1+2ydA = \pi.$$

♦

Example5.17 Compute the integral

$$\iint_S \text{curl}\mathbf{F}\cdot d\mathbf{S}$$

where

$$\mathbf{F} = (xz, yz, xy)$$

and

$$S = \{(x,y,z)\in\mathbb{R}^3|x^2+y^2+z^2=4\}\cap\{(x,y,z)\in\mathbb{R}^3|z\geq\sqrt{3}\}.$$

Solution We present two methods of solving it:

- **Method I: Direct Computation**
  - **Parametrisation of surface** The surface can be parametrised as:

$$\mathbf{S}(x,y) : x = x, y(t) = y, z(t) = \sqrt{4-x^2-y^2}, 0\leq x^2+y^2\leq 1.$$

  - **Switch to the Riemann (double) integral:** Direct computation shows that

$$\text{curl}\mathbf{F} = \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xz & yz & xy \end{vmatrix} = (x-y, x-y, 0)$$

Note that the surface is a graph of the function $z = g(x, y) = \sqrt{4 - x^2 - y^2}$, hence

$$\begin{aligned}\mathrm{curl}\mathbf{F}(x, y, z) \cdot (\mathbf{S}_x \times \mathbf{S}_y) &= (P, Q, R) \cdot \left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right) \\ &= (x - y, x - y, 0) \cdot \left(\frac{x}{\sqrt{4 - x^2 - y^2}}, \frac{y}{\sqrt{4 - x^2 - y^2}}, 1\right) \\ &= \frac{x^2 - y^2}{\sqrt{4 - x^2 - y^2}}.\end{aligned}$$

one has that

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_{x^2+y^2\leq 1} \frac{x^2 - y^2}{\sqrt{4 - x^2 - y^2}} dA$$

- **Evaluation of the Riemann (double) integral:** By symmetry of the disk, the above integral can be written as

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_D \frac{x^2}{\sqrt{4 - x^2 - y^2}} dA - \iint_D \frac{y^2}{\sqrt{4 - x^2 - y^2}} dA = 0.$$

(because x and y are symmetric.)

- <u>**Method II: Stokes' Theorem**</u>
  - **Parametrisation of the boundary curve**: The boundary of the surface S is

$$\partial S = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 = 1, z = \sqrt{3}\}.$$

which is an circle that can be parametrised as

$$\mathbf{r}(t) : x(t) = \cos t, y(t) = \sin t, z(t) = \sqrt{3}.$$

  - **Switch to Riemann integral:** Note that $\dot{\mathbf{r}}(t) = (-\sin t, \cos t, 0) = (-y, x, 0)$

$$\mathbf{F} \cdot \dot{\mathbf{r}} = (xz, yz, xy) \cdot (-y, x, 0) = -xyz + xyz + 0 = 0.$$

The Riemann integral is

$$\int_0^{2\pi} 0 dt.$$

  - **Evaluation of the Riemann integral:** Clearly

$$\int_0^{2\pi} 0 dt = 0.$$

♦

Example5.18 Given a curve $C \subset \mathbb{R}^3$ parametrised by

$$\gamma(t) = (\sin(t), \cos t, \sin(2t)), \quad 0 \leq t \leq 2\pi,$$

calculate

$$\int_C \mathbf{F} \cdot d\mathbf{r}$$

where

$$\mathbf{F} = (y + \sin x + 3zx^2, x + \cos y, x^3).$$

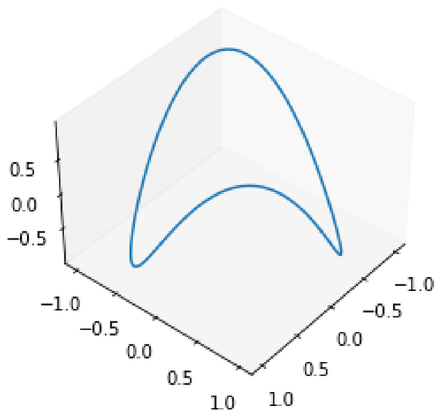


Figure 5.10: The parametric curve $\gamma(t) = (\sin(t), \cos t, \sin(2t)), \quad 0 \le t \le 2\pi$.

Solution We present two methods of solving it:

- **Method I: Direct Computation**
  - **Parametrisation of curve** The curve is already parametrised:

$$\mathbf{r}(t) = (\sin(t), \cos t, \sin(2t)), \quad 0 \le t \le 2\pi$$

  - **Switch to the Riemann (double) integral:** Direct computation shows that

$$\mathbf{F} \cdot \dot{\mathbf{r}} = (\cos t + \sin(\sin t) + 3\sin(2t)\sin^2 t)\cos t - (\sin t + \cos(\cos t))\sin t + 2\sin^3(t) 2\cos 2t$$

Hence the Riemann integral is

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} \mathbf{F} \cdot \dot{\mathbf{r}} dt = A_1 + A_2 + A_3,$$

where

$$A_1 = \int_{-\pi}^{\pi} (\cos t + \sin(\sin t) + 3\sin(2t)\sin^2 t)\cos t dt$$

$$A_2 = \int_{-\pi}^{\pi} -(\sin t + \cos(\cos t))\sin t dt$$

$$A_3 = \int_{-\pi}^{\pi} 2\sin^3(t) 2\cos 2t dt$$

where we have used the fact that both $\sin t$ and $\cos t$ are 2π periodic.

  - **Evaluation of the Riemann (double) integral:** Now notice that

$$\sin(\sin t\cos t, 3\sin(2t)\sin^2 t\cos t, \cos(\cos t)\sin t, 2\sin^3(t)2\cos 2t$$

are all **odd** functions, hence

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_{-\pi}^{\pi} \cos^2 t - \sin^2 t dt = \pi - \pi = 0.$$

- **Method II: Stokes' Theorem**
  - **Parametrisation of the surface** Note that the curve C is a simple closed curve (try to draw it! Either by hand or use software). As a result Stokes theorem works. We know the curve bounds a surface $S = (x, y, g(x, y))$ defined on $D = \{(x, y)|x^2 + y^2 = 1\}$. Somehow in this problem setting there is no natural candidate of the surface that could provide considerable convenience for computation. **Don't worry, be happy!**

- **Switch to Riemann (double) integral:** Direct computation shows that
$$\text{curl}\mathbf{F} = (0-0, 3x^2-3x^2, 1-1) = (0,0,0).$$
Hence no matter which surface S (enclosed by C) we choose (thus the normal vector $\mathbf{n}$), one always has that
$$\text{curl}\mathbf{F}\cdot\mathbf{n} = 0.$$
- **Evaluation of the Riemann (double) integral:** The Riemann double integral is then
$$\int_S \mathbf{F}\cdot d\mathbf{S} = 0.$$

♦

Example5.19 Compute the integral
$$\iint_S \mathbf{F}_1\cdot d\mathbf{S}, \iint_S \mathbf{F}_2\cdot d\mathbf{S}$$
where $S = \{(x,y,z)\in\mathbb{R}^3 | x^2+y^2+z^2=1\}$, for

1. $\mathbf{F}_1 = (x,y,0)$;
2. $\mathbf{F}_2 = (-y,x,0)$.

Solution Note that $S = \partial E$, where E is the solid unit ball. As a result, by Gauss theorem,
$$\iint \mathbf{F}_1\cdot d\mathbf{S} = \iiint_E \text{div}\mathbf{F}_1 dV = \iiint_E 2dV = 2\text{Vol}(E) = \frac{8}{3}\pi.$$
$$\iint \mathbf{F}_2\cdot d\mathbf{S} = \iiint_E \text{div}\mathbf{F}_2 dV = \iiint_E 0dV = 0.$$

♦

Remark Recall we have already seen that these two simple vector fields represents important two classes of vector fields (see example 5.1 and example 5.2), namely irrotational vector fields and incompressible vector fields, respectively. In general, one sees that if a vector field is incompressible, and E is any smooth solid region, then
$$\iint_{\partial E} \mathbf{F}\cdot d\mathbf{S} = \iiint_E \text{div}\mathbf{F}dV = 0.$$

Example5.20 Let $\mathbf{F}(x,y,z) = (0,y,-z)$ be a vector field. Consider the closed surface S formed by the two parts:

- the paraboloid $y = x^2+z^2$, $0\le y\le 1$, and
- the disk $x^2+z^2\le 1$, $y=1$

Calculate (using the **positive outward orientation** of the surface)
$$\iint_S \mathbf{F}\cdot d\mathbf{S}.$$

Solution We present two methods of solving it:

- **Method I: Direct Computation**

- **Parametrisation of surface** The surface S_1, S_2 can be parametrised as:

$$\mathbf{S}_1(x,z) = (x, x^2+z^2, z), \quad x^2+z^2 \le 1.$$
$$\mathbf{S}_1(x,z) = (x, 1, z), \quad x^2+z^2 \le 1.$$

- **Switch to the Riemann (double) integral:**

For S_1, the surface is a graph of the function $y = g_1(x,z) = x^2+z^2$, hence

$$\mathbf{S}_{1x} \times \mathbf{S}_{1z} = \begin{vmatrix} i & j & k \\ 1 & 2x & 0 \\ 0 & 2z & 1 \end{vmatrix} = (2x, -1, 2z).$$

$$\mathbf{F}(x,y,z) \cdot (\mathbf{S}_x \times \mathbf{S}_y) = (0, y, -z) \cdot (2x, -1, 2z) = -y - 2z^2 = -x^2 - 3z^2.$$

one has that

$$\iint_{S_1} \mathbf{F} \cdot d\mathbf{S} = \iint_{x^2+z^2\le 1} -x^2 - 3z^2 dA$$

For S_2, the surface is a graph of the (constant) function $y = g_2(x,z) = 1$, hence

$$\mathbf{S}_{2x} \times \mathbf{S}_{2z} = \begin{vmatrix} i & j & k \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = (0, -1, 0).$$

However, the question told us we should choose **outward** direction as normal direction, hence we should revert this normal vector and take $(0,1,0)$. It follows that

$$\iint_{S_2} \mathbf{F} \cdot d\mathbf{S} = \iint_{x^2+z^2\le 1} (0, y, -z) \cdot (0,1,0) dA = \iint_{x^2+z^2\le 1} 1 dA.$$

where we have used the fact that on S_2, $\mathbf{F} = (0, y, -z)|_{y=1} = (0, 1, -z)$.

- **Evaluation of the Riemann (double) integral:**

For S_1, by polar coordinates,

$$\iint_{S_1} -x^2 - 3z^2 dA = \int_0^{2\pi} \int_0^1 (-r^2\cos^2\theta - 3r^2\sin^2) r dr d\theta = -\pi.$$

For S_2, by polar coordinates,

$$\iint_{x^2+z^2\le 1} 1 dA = \int_0^{2\pi} \int_0^1 r dr d\theta = \pi.$$

Hence

$$\iint_{S_1} \mathbf{F} \cdot d\mathbf{S} + \iint_{S_2} \mathbf{F} \cdot d\mathbf{S} = -\pi + \pi = 0.$$

- **Method II: Gauss' Theorem**
  - **Specify the Solid Region**: The solid region E is enclosed by S_1 and S_2, given by

$$E = \{(x,y,z) \in \mathbb{R}^3 | x^2+z^2 \le 1, x^2+z^2 \le y \le 1\}.$$

As a result, we can consider Gauss' theorem.

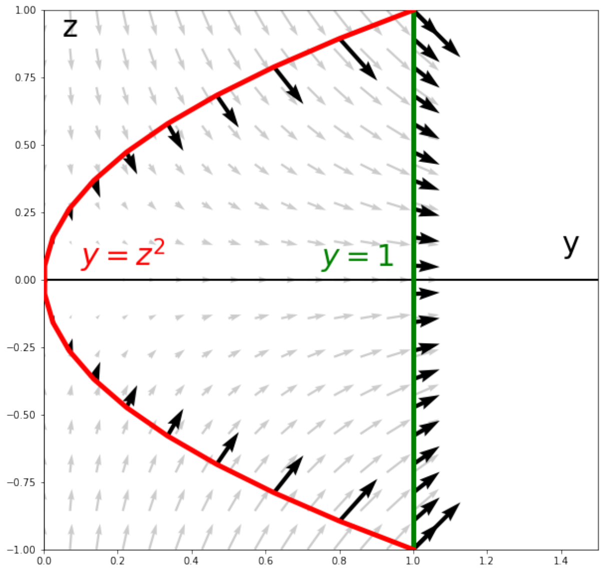


Figure 5.11: The flux is positive at the green boundary, negative at the red boundary, and total flux is 0.

- **Switch to Riemann (triple) integral:** Compute the divergence:
$$\text{div}\mathbf{F} = \frac{\partial}{\partial x}0 + \frac{\partial}{\partial y}y + \frac{\partial}{\partial z}(-z) = 0 + 1 - 1 = 0.$$
The Riemann (triple) integral is
$$\iiint_E 0dV.$$
- **Evaluation of the Riemann integral:** Clearly
$$\iiint_E 0dV = 0.$$

♦

Green's theorem Revisited

Green's theorem can be seen as the special case of the stokes theorem and Gauss' theorem, by embedding the $2d$ vector field $\mathbf{F}(x, y) = (P(x, y), Q(x.y))$ into the $3d$ vector field $\tilde{\mathbf{F}}(x, y, z) = (P(x, y), Q(x, y), 0)$. Recall the following two forms of Green's theorem:

Theorem 5.5. Green's Theorem

Let D be a domain in $\mathbb{R}^2$ with outward unit normal vector $\mathbf{n}(x, y) \in \mathbb{R}^2$, and ∂D be a positively oriented closed curve. Then

$$\int_{\partial D} \mathbf{F} \cdot d\mathbf{r} = \iint_D \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dA \qquad \text{(First Green Theorem)}$$
$$\int_{\partial D} \mathbf{F} \cdot \mathbf{n} ds = \iint_D div\mathbf{F}(x, y) dA \qquad \text{(Second Green Theorem)}$$

♡

Remark In the first Green theorem,

- LHS: Integral of vector field along a curve
- RHS: Double integral.

In the second Green theorem,

- LHS: Integral of scalar function along a curve w.r.t. arc-length
- RHS: Double integral.

Proof We prove these two forms of Green's theorem one by one:

1. Embed the $2d$ vector field $\mathbf{F}(x, y) = (P(x, y), Q(x.y))$ into the $3d$ vector field $\tilde{\mathbf{F}}(x, y, z) = (P(x, y), Q(x, y), 0)$. Since the surface D is already in the xy-plane, the normal vector is then $\mathbf{m} = (0, 0, 1)$ everywhere in D. By Stokes' theorem,
$$\int_{\partial D} \tilde{\mathbf{F}} \cdot d\mathbf{r} = \iint_D \text{curl}\tilde{\mathbf{F}} \cdot d\mathbf{S} = \iint_D \text{curl}\tilde{\mathbf{F}} \cdot \mathbf{m} dA$$
Note that for the RHS
$$\text{curl}\tilde{\mathbf{F}} = (0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}).$$
Thus
$$\text{curl}\tilde{\mathbf{F}} \cdot \mathbf{m} = (0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}) \cdot (0, 0, 1) = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$$
On the other hand, for the LHS,
$$\int_{\partial D} \tilde{\mathbf{F}} \cdot d\mathbf{r} = \int_{\partial D} P(x, y)dx + Q(x, y)dy + 0dz = \int_{\partial D} P(x, y)dx + Q(x, y)dy$$
Hence
$$\int_{\partial D} P(x, y)dx + Q(x, y)dy = \iint_D \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dA.$$
We have thus proved the first Green's theorem.
2. - Embed the $2d$ vector field $\mathbf{F}(x, y) = (P(x, y), Q(x.y))$ into the $3d$ vector field $\tilde{\mathbf{F}}(x, y, z) = (P(x, y), Q(x, y), 0)$;
   - Embed the $2d$ outward normal vector field $\mathbf{n}(x, y) = (n_1(x, y), n_2(x, y))$ into the $3d$ vector $\tilde{\mathbf{n}}(x, y, z) = (n_1(x, y), n_2(x, y), 0)$.
   - Embed the surface D into the curved cylinder $E = D \times [0, 1]$, which is a solid region. Note that
$$\partial E = \underbrace{S_-}_{\text{the floor}} \cup S_0 \cup \overbrace{S_+}^{\text{the roof}}.$$

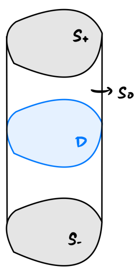


Figure 5.12: Green theorem 2^{nd} form as special case of Gauss' theorem

By Gauss's theorem

$$\iint_{\partial E} \tilde{\mathbf{F}} \cdot d\mathbf{S} = \iiint_E \operatorname{div}\tilde{\mathbf{F}} dV$$

For the LHS, the integral on the roof surface S_+ and the floor surface S_- cancel each other (same vector field with opposite outward normal vector). We only need to compute the integral on S_0. Note that if we parametrise ∂D by $(x(t), y(t), a \le t \le b)$, then S_0 can be parametrised by

$$\mathbf{r}(t,s) = (x = x(t), y = y(t), z(l) = l), a \le t \le b, 0 \le l \le 1.$$

It follows that $\mathbf{r}_t \times \mathbf{r}_l = (\dot{y}, \dot{x}, 0)$, and

$$\begin{aligned} \iint_{S_0} \tilde{\mathbf{F}} \cdot d\mathbf{S} &= \int_0^1 \int_a^b (P, Q, 0)(-\dot{y}(t), \dot{x}(t), 0) dt dl \\ &= \int_a^b (P, Q) \cdot (-\dot{y}(t), \dot{x}(t)) dt \\ &= \int_{\partial D} \mathbf{F} \cdot \mathbf{n} ds. \end{aligned}$$

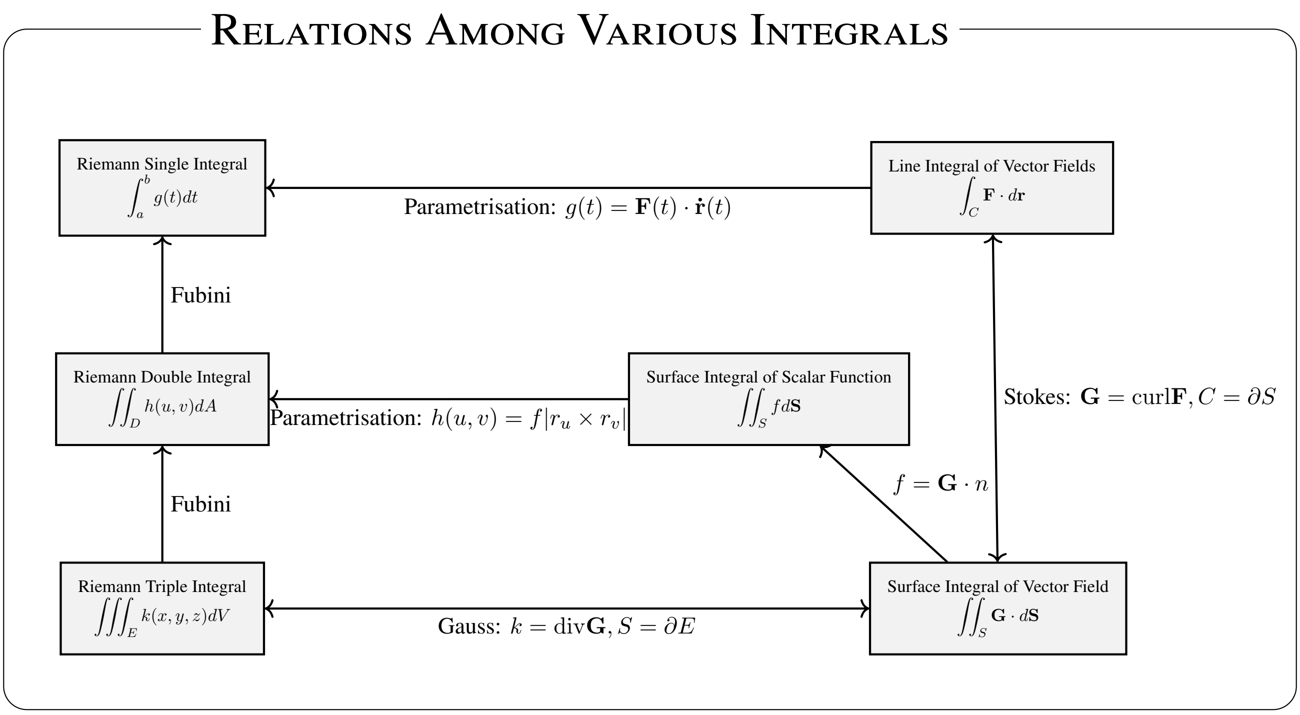
For the RHS, the Fubini theorem implies that

$$\iiint_E \operatorname{div}\tilde{\mathbf{F}} dV = \iint_D (\int_0^1 \operatorname{div}\tilde{\mathbf{F}} dz) dA = \iint_D \operatorname{div}\tilde{\mathbf{F}} dA = \iint_D \operatorname{div}\mathbf{F} dA.$$

where the second and the last equality are due to the fact that $\operatorname{div}\tilde{\mathbf{F}}$ does not depend on z.

$\square$

Relations among Various Integrals



Length, Area, and Volume

Quantity	**Formula**
Arc-length of curve C in $\mathbb{R}^3$	$\int_C 1ds$
Area of surface S in $\mathbb{R}^3$	$\iint_S 1dS$
Volumn of solid region E in $\mathbb{R}^3$	$\iiint_E 1dV$

Some Examples and Counter-Examples

Example	Description
Example 2.13	A function in $2D$ with partial derivatives, but not differentiable.
Example 2.14	A function in $2D$ with partial derivatives, continuous, but not differentiable.
Example 2.22	A function in $2D$ whose mixed partial derivatives are not equal.
Example 3.21,3.22,3.23	A non-graph parametric Curve/surface being a graph locally
Example 5.4,5.5	A line integral of a vector field in general depends on the whole trajectory.
Example 5.7	A vector field defined on a non-simply-connected domain being conservative.
Example 5.8	A vector field defined on a non-simply-connected domain not being conservative.
Example 5.1,5.2,5.14	A vector field and its Helmholtz decomposition
Example 5.15	A vector field having distinct Helmholtz decomposition.