Lecture Notes of MATH 230 Calculus III

Hongda Qiu[*][+]

[+] *Department of Mathematics, Penn State University*

April 24, 2022

1 Notations

- $\mathbb{R} = (-\infty, \infty)$: real number set;
- $V = \{(x, y)|x, y \in \mathbb{R}\}$: 2-dimensional vector space;
- $V = \{(x, y, z)|x, y, z \in \mathbb{R}\}$: 3-dimensional vector space;
- $span\{\vec{u}, \vec{v}\}$: the 2-D plane spanned by vector $\vec{u}, \vec{v}$, i.e.,$span\{\vec{u}, \vec{v}\} = \{\lambda\vec{u} + \mu\vec{v}|\lambda, \mu \in \mathbb{R}\}$

2 Vectors

At the beginning, we discuss everything on $\mathbb{R}^2$.

2.1 Definition of Vectors

There are two fundamental methods to describe vectors.

Definition 2.1 (Vectors (Coordinates)) *A vector $\vec{v} = (v_x, v_y)$ is an ordered couple given by $v_x, v_y \in \mathbb{R}$.*

Definition 2.2 (Vectors (Geometry)) *A vector $\vec{v} = \overrightarrow{AB}$ is the position of point B relative to A.*

If we are given the coordinates of $A = (x_A, y_A), B = (x_B, y_B)$, we can get that $\overrightarrow{AB} = (x_B - x_A, y_B - y_A)$.

2.2 Magnitude

Given a vector $\vec{v} = (v_x, v_y)$, we define the magnitude (or the length, the norm) of $\vec{v}$ as:

$$|\vec{v}| = \sqrt{v_x^2 + v_y^2} \tag{1}$$

For vector defined by a pair of point A, B, $\overrightarrow{AB}$, the norm of $\overrightarrow{AB}$ is just the distance between A, B:

$$|\overrightarrow{AB}| = |AB| \tag{2}$$

You may verify the consistency of norm under two different defnitions of vectors via $\overrightarrow{AB} = (x_B - x_A, y_B - y_A)$:

$$|\overrightarrow{AB}| = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2} \tag{3}$$

[*] hqiu@psu.edu

2.3 Addition and Scalar Multiplication

There are two computations for vectors.
For $\vec{v} = (v_x, v_y), \vec{u} = (u_x, u_y)$, we define the vector addition between v, u as

$$\vec{v} + \vec{u} = (v_x + u_x, v_y + u_y) \tag{4}$$

For $\vec{v} = (v_x, v_y)$ and $\lambda \in \mathbb{R}$, we define the scalar multiplication as:

$$\lambda\vec{v} = (\lambda v_x, \lambda v_y) \tag{5}$$

2.4 Parallel

There is also an important concept related to scalar multiplication:

Definition 2.3 (Parallel) *Two vectors v, w are parallel to each other, i.e., $\vec{v}//\vec{w}$ if and only if there exists some $\lambda \in \mathbb{R}$, such that*

$$\vec{v} = \lambda \cdot \vec{w}$$

In case that v_x, v_y, w_x, w_y are nonzero, one may check if $\vec{v}, \vec{w}$ are parallel via checking if $v_x/v_y = w_x/w_y$.
Let us look at some examples:

- $\vec{v} = (1, 2), \vec{w} = (2, 4)$ are parallel since $\vec{v} = \frac{1}{2} \cdot \vec{w}$;
- $\vec{v} = (1, 2), \vec{w} = (2, 5)$ are parallel since $\vec{v} = \frac{1}{2} \cdot \vec{w}$.
- $\vec{v} = (0, 2), \vec{w} = (0, 3)$ are parallel since $\vec{v} = \frac{2}{3} \cdot \vec{w}$.

2.5 Unit Vector

Before going further, we introduce a concept regarding the direction of vector.

Definition 2.4 (Unit Vector) *A vector $\vec{e}$ is called a unit vector if $|\vec{e}| = 1$.*

The main question here is: given $\vec{v}$, what is the unit vector $\vec{e}$ parallel to $\vec{v}$?
The answer is to make use of scalar multiplication:

$$\vec{e} = \frac{1}{|\vec{v}|} \cdot \vec{v} \tag{6}$$

Let us try the computation on an example:

Example 1 *Suppose that $\vec{v} = (2, 3)$. The unit vector parallel to $\vec{v}$ is given by*

$$\vec{e} = \frac{1}{|\vec{v}|} \cdot \vec{v} = \frac{1}{\sqrt{2^2 + 3^2}} \cdot (2, 3) = (\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}) \tag{7}$$

The unit vector e can be regarded as a "note" of direction in the following sense:
For any vector $\vec{v}$, you may always find its parallel unit vector $\vec{e} = \frac{1}{|\vec{v}|} \cdot v$ and re-express $\vec{v}$ as:

$$\vec{v} = |\vec{v}| \cdot \vec{e} \tag{8}$$

Recall that $sin^2\theta + cos^2\theta = 1$, for any unit vector $\vec{e} = (a, b)$, we may take $\theta = arccos a$ and obtain that

$$\vec{e} = (cos\theta, sin\theta) \tag{9}$$

2.6 Dot Product

For $\vec{v} = (v_x, v_y), \vec{u} = (u_x, u_y)$, we define the dot product between v, u as

$$\vec{v} \cdot \vec{u} = v_x u_x + v_y u_y \tag{10}$$

We may verify that $v \perp u \iff v \cdot u = 0$.
However, we need to understand the geometrical meaning of $\vec{v} \cdot \vec{u}$ when it is nonzero, that is,

$$cos\theta_{u,v} = \frac{v \cdot u}{|v| \cdot |u|} \tag{11}$$

where $\theta_{u,v}$ denotes the angle between v, u.
We may prove this formula:

Theorem 1 *Given two vectors $\vec{v}, \vec{u}$ on $\mathbb{R}^2$,*

$$cos\theta_{\vec{v},\vec{u}} = \frac{\vec{v} \cdot \vec{u}}{|\vec{v}| \cdot |\vec{u}|} \tag{12}$$

Proof: First of all, let us "put" $\vec{v}, \vec{u}$ back to the origin, that is, we set $\vec{v} = \overrightarrow{OA}, \vec{u} = \overrightarrow{OB}$ for some points $A = (v_x, v_y), B = (u_x, u_y)$.
Then we shall see that $\theta_{\vec{v},\vec{u}} = \angle AOB$. To finish the proof, we simply need to compute $cos\angle AOB$.
To do so, we draw a line AH from A that is vertical to OB. Suppose AH intersects OB at H. We shall see that $cos\angle AOB = |HO|/|AO|$.
Here $|AO| = |\vec{v}| = \sqrt{v_x^2 + v_y^2}$.
To obtain $|HO|$, note that in the right triangle ΔAOH, $|AO|^2 = |AH|^2 + |HO|^2$. Here $|AH|$ is just the distance from A to the line OB.
We know that the equation of line OB is $u_y x - u_x y = 0$, so

$$|AH| = \frac{|u_y v_x - u_x v_y|}{\sqrt{u_x^2 + u_y^2}} \tag{13}$$

It follows that

$$|HO| = \sqrt{|AO|^2 - |AH|^2} = \frac{|v_x u_x + v_y u_y|}{\sqrt{u_x^2 + u_y^2}} \tag{14}$$

Then we may compute that

$$|cos\angle AOB| = \frac{|HO|}{|AO|} = \frac{|v_x u_x + v_y u_y|}{\sqrt{v_x^2 + v_y^2}\sqrt{u_x^2 + u_y^2}} = \frac{|v_x u_x + v_y u_y|}{|\vec{v}| \cdot |\vec{u}|} \tag{15}$$

We shall see that $cos\angle AOB$ and $v_x u_x + v_y u_y$ have the same sign, thus

$$cos\angle AOB = \frac{v_x u_x + v_y u_y}{|\vec{v}| \cdot |\vec{u}|} \tag{16}$$

□

2.7 3D Vectors

Before discussing 3D geometry, we need to understand 3D vectors.
Given two vectors $v = (v_x, v_y, v_z), u = (u_x, u_y, u_z)$, we define the basic computations of vectors based on the ideas generaliezed from 2-D space:

$$[Addition]\vec{v} + \vec{u} = (v_x + u_x, v_y + u_y,) \tag{17}$$

$$[ScalarMultlication]\lambda\vec{v} = (\lambda v_x, \lambda v_y) \tag{18}$$

$$[DotProduct]\vec{v}\cdot\vec{u} = v_x u_x + v_y u_y + v_z u_z \tag{19}$$

We also have the magnitude, similar to the case in 2-D space:

$$[Magnitude]|\vec{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2} \tag{20}$$

2.8 Lines

To describe some line l, the idea is basically to make use of an initial point $A = (x_0, y_0, z_0)$ and a vector $\vec{v}$. With a parameter t, any point (x, y, z) on a line l can be expressed as"

$$(x(t), y(t), z(t)) = A + t\cdot\vec{v} = (x_0 + tv_x, y_0 + tv_y, z_0 + tv_z) \tag{21}$$

The choice of $A, \vec{v}$ is not unique. You may choose any $\vec{v}'//\vec{v}$.
We can also write l in the following way:

$$\begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \\ z = z_0 + tv_z \end{cases} \tag{22}$$

One shall be aware that

$$\frac{x - x_0}{v_x} = \frac{y - y_0}{v_y} = \frac{z - z_0}{v_z} = t \tag{23}$$

That means we can express l without the usage of parameter t:

$$\frac{x - x_0}{v_x} = \frac{y - y_0}{v_y} = \frac{z - z_0}{v_z} \tag{24}$$

We may also use the combination of two linear equations to express a line. Each equation expresses a plane, so their combination means the intersection of two planes.

$$\begin{cases} a_1x + b_1y + c_1z + d_1 = 0 \\ a_2x + b_2y + c_2z + d_2 = 0 \end{cases} \tag{25}$$

2.9 Planes

The term of a plane S can be given by a linear equation:

$$ax + by + cz + d = 0 \tag{26}$$

When $d \neq 0$, you may divide both sides by d and get:

$$a'x + b'y + c'z + 1 = 0 \tag{27}$$

where $a' = a/d, b' = b/d, c' = c/d$. This is useful when you are required to obtain plane given three points, say $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3)$. By cancelling d, you can just solve the following system for three variables a, b, c:

$$\begin{cases} ax_1 + by_1 + cz_1 + 1 = 0 \\ ax_2 + by_2 + cz_2 + 1 = 0 \\ ax_3 + by_3 + cz_3 + 1 = 0 \end{cases} \tag{28}$$

A more intuitive way is to use a vector that is vertical to plane S:$\vec{v}\perp S$. We have that for each vector of $\overrightarrow{AB}$, $A, B \in S$,

$$\vec{v}\perp\overrightarrow{AB}$$

That means,

$$\vec{v} \cdot \overrightarrow{AB} = 0 \tag{29}$$

Given $A = (x_0, y_0, z_0)$, we may express any point $B = (x, y, z)$ as the following:

$$\vec{v} \cdot \overrightarrow{AB} = v_x(x - x_0) + v_y(y - y_0) + v_z(z - z_0) = 0 \tag{30}$$

Another idea is to make use of linear combinations of two linear independent vectors in S. Given $A = (x_1, y_1, z_1), B = (x_2, y_2, z_2), C = (x_3, y_3, z_3)$ such that A, B, C are not on the same line, we can express any point D on S using the following relation with two parameters λ, μ:

$$\overrightarrow{AD} = \lambda\overrightarrow{AC} + \mu\overrightarrow{AB} \tag{31}$$

In coordinates, this is

$$(x(\lambda, \mu), y(\lambda, \mu), z(\lambda, \mu)) = (x_1, y_1, z_1) + \lambda(x_3 - x_1, y_3 - y_1, z_3 - z_1) + \mu(x_2 - x_1, y_2 - y_1, z_2 - z_1) \tag{32}$$

2.10 Linearity

Some properties of vectors are determined by the linearity of vector space V:

- $\vec{v}, \vec{u} \in V \Rightarrow \vec{v} + \vec{u} \in V$
- $\vec{v} \in V \Rightarrow \lambda\vec{v} \in V$

So in general,

$$\vec{v}, \vec{u} \in V \Rightarrow \forall a, b \in \mathbb{R}, a\vec{v} + b\vec{u} \in V \tag{33}$$

Now we shall have

- $\vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{w} = \vec{v} \cdot (\vec{u} + \vec{w})$
- $ab\vec{v} = a(b\vec{v})$

Given $||\vec{v}||^2 = \vec{v} \cdot \vec{v}$, the important relation you shall have in mind is

$$||\vec{v} \pm \vec{u}||^2 = (\vec{v} \pm \vec{u}) \cdot (\vec{v} \pm \vec{u}) = ||\vec{v}||^2 \pm 2\vec{v} \cdot \vec{u} + ||\vec{u}||^2 \tag{34}$$

2.11 Cross Product

Definition 2.5 (Cross Product) *Given $\vec{v} = (v_x, v_y, v_z), \vec{u} = (u_x, u_y, u_z)$, the cross product $\times : V \times V \to V$ is defined as:*

$$\vec{v} \times \vec{u} = (|\begin{pmatrix} v_y & v_z \\ u_y & u_z \end{pmatrix}|, |\begin{pmatrix} v_z & v_x \\ u_z & u_x \end{pmatrix}|, |\begin{pmatrix} v_x & v_y \\ u_x & u_y \end{pmatrix}|) \tag{35}$$

Remark 2.1 *To remember the correct order, think about the order $x - y - z$ of the three axis.*
Some books may use

$$\vec{v} \times \vec{u} = (|\begin{pmatrix} v_y & v_z \\ u_y & u_z \end{pmatrix}|, -|\begin{pmatrix} v_x & v_z \\ u_x & u_z \end{pmatrix}|, |\begin{pmatrix} v_x & v_y \\ u_x & u_y \end{pmatrix}|) \tag{36}$$

instead. You shall be carefull about the entries.

Remark 2.2 (Right-Hand Principle) *The direction of $\vec{v} \times \vec{u}$ is determined by the Right-Hand Principle: think about v as x-axis (your index finger), u as y-axis (your middle finger), then $\vec{v} \times \vec{u}$ would come to be the z-axis (your thumb).*
In case that you are unable to remember the correct direction, test the simple example:

Example 2

$$(1, 0, 0) \times (0, 1, 0) = (0, 0, 1) \tag{37}$$

For the following content, we need some very basic linear algebra.

Baby Linear Algebra 1 *A two dimensional matrix has the following property:*

- *Determinant of* 2×2 *Matrix:* $|\begin{pmatrix} a & b \\ c & d \end{pmatrix}| = ad - bc$;
- $|\begin{pmatrix} a & b \\ c & d \end{pmatrix}| = -|\begin{pmatrix} b & a \\ d & c \end{pmatrix}| = -|\begin{pmatrix} c & d \\ a & b \end{pmatrix}|$. *For* 3×3 *matrix, we have similar conclusion;*

A three dimensional matrix has the following property:

- *Row Decomposition of* 3×3 *Matrix:* $|\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}| = a|\begin{pmatrix} e & f \\ h & i \end{pmatrix}| + b|\begin{pmatrix} f & d \\ i & g \end{pmatrix}| + c|\begin{pmatrix} d & e \\ g & h \end{pmatrix}|$;
- *Invariance of Determinant:* $|\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}| = |\begin{pmatrix} a & b & c \\ d-a & e-b & f-c \\ g & h & i \end{pmatrix}|$. *This is the example for row 2 minus row 1, i.e.* $r_2 - r_1$. *You can do this for any pair of rows (or columns).*

What you need to realize is that the determinant of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $ad-bc$, *is just the area of the parallelogram spanned by vector* $(a,b),(c,d)$.

Thus, the first entry $|\begin{pmatrix} v_y & v_z \\ u_y & u_z \end{pmatrix}|$, for instance, is just the area of the parallelogram spanned by the vectors $(v_y, v_z), (u_y, u_z)$, the respective projections of $\vec{v}, \vec{u}$ on the yOz plane.
Be careful, the cross product $\times : V\times V \to V$ is different from the dot product $\cdot : V\times V \to \mathbb{R}$. The cross product gives you a vector, while the dot product gives you a real number.
From the term of $\vec{v}\times\vec{u}$, we can see that

- (Asymmetry) $\vec{v}\times\vec{u} = -\vec{u}\times\vec{v}$;
- (Magnitude) $||\vec{v}\times\vec{u}|| = ||\vec{v}||\cdot||\vec{u}||sin\theta_{\vec{v},\vec{u}}$

The first property furthere implies that

$$\vec{v}\times\vec{v} = 0 \tag{38}$$

The second property is important. It implies that the magnitude of $\vec{v}\times\vec{u}$ equals to the area of the parallelogram spanned by $\vec{v},\vec{u}$.
If we compute the dot product of $\vec{v}\times\vec{u}$ with some other $\vec{w}$, we would get, by the Row Decomposition of Matrix:

$$\vec{v}\times\vec{u}\cdot\vec{w} = w_x|\begin{pmatrix} v_y & v_z \\ u_y & u_z \end{pmatrix}| + w_y|\begin{pmatrix} v_z & v_x \\ u_z & u_x \end{pmatrix}| + w_z|\begin{pmatrix} v_x & v_y \\ u_x & u_y \end{pmatrix}| = |\begin{pmatrix} w_x & w_y & w_z \\ v_x & v_y & v_z \\ u_x & u_y & u_z \end{pmatrix}| \tag{39}$$

A quick observation is that, if you take $\vec{w} = \vec{v}$ (or $\vec{w} = \vec{u}$), you would get that

$$\vec{v}\times\vec{u}\cdot\vec{v} = |\begin{pmatrix} v_x & v_y & v_z \\ v_x & v_y & v_z \\ u_x & u_y & u_z \end{pmatrix}| = |\begin{pmatrix} v_x & v_y & v_z \\ 0 & 0 & 0 \\ u_x & u_y & u_z \end{pmatrix}| = 0$$

$\Rightarrow \vec{v}\perp\vec{v}\times\vec{u}$. Similarly, $\vec{u}\perp\vec{v}\times\vec{u}$.

Note that the priority of $\times$ is BEYOND $\cdot$.

$$[Priority]() > \times > \cdot > \pm \tag{40}$$

The term of $\vec{u}\times\vec{v}\cdot\vec{w}$ is called the mixed product of vectors. Based on properties of matrix, you can check that

$$\vec{u}\times\vec{v}\cdot\vec{w} = \vec{u}\cdot\vec{v}\times\vec{w} \tag{41}$$

The geometrical meaning of the value of $\vec{u}\times\vec{v}\cdot\vec{w}$ is just the volume of the parallelepiped spanned by $\vec{u},\vec{v},\vec{w}$. This is from the definition of the determinant of matrices.

2.12 Curves and Surfaces

Whether an object becomes a curve or a surface depends on the number of parameters (or the free dimensions): $\vec{f}(t) = (x(t), y(t), z(t))$ is a curve and $\vec{f}(t,s) = (x(t,s), y(t,s), z(t,s))$, a surface. Both are objects in 3-D space. Here are some examples:

- (Cube) $x^2 + y^2 = 1$;
- (Sphere) $x^2 + y^2 + z^2 = 1$;
- (Curve) $(x(t), y(t), z(t)) = (t, t^2, t^3)$.

2.12.1 Rotations

The idea of rotation around some axis (x, y, z axis) brings us new geometrical objects.
Suppose that you have some curve C in the form of an implicit function $F(u, v) = 0$ on a plane.

- Choose an axis I for the whole object F to rotate around;
- Replace the variable II or III by $\sqrt{II^2 + III^2}$;

Here I, II, III can be one of x, y, z.

Example 3

- *(Hyperboloid, one sheet) The rotation of $x^2 - y^2 = 1$ around x-axis is $x^2 - (y^2 + z^2) = 1$;*
- *(Hyperboloid, two sheets) The rotation of $x^2 - y^2 = 1$ around y-axis is $x^2 + z^2 - y^2 = 1$;*
- *(Cone) The rotation of $z = y$ around z-axis is $z^2 = x^2 + y^2$;*
- *(Paraboloid) The rotation of $y = x^2$ around y-axis is $y = x^2 + z^2$;*
- *(Sphere) The rotation of $x^2 + y^2 = 1$ around x or y-axis is $x^2 + y^2 + z^2 = 1$;*

2.12.2 Quadratic Surfaces

A main class of surfaces in 3-D space are quadratic surfaces. They are surfaces determined by second order implicit functions of x, y, z.

- (Ellipsoid)
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \tag{42}$$
- (Hyperboloid, one sheet)
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \tag{43}$$
- (Hyperboloid, two sheets)
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1 \tag{44}$$
- (Paraboloid, elliptic)
$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2} \tag{45}$$
- (Paraboloid, hyperbolic)
$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2} \tag{46}$$

- (Cone)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \tag{47}$$

How can one obtain a quadratic surface? Let us look at an example.
Let us say, we have a hyperbolic curve

$$y^2 - z^2 = 1 \tag{48}$$

Then we can rotate it around the z-axis and get

$$x^2 + y^2 - z^2 = 1 \tag{49}$$

The last step is to exaggerate this surface in the direction of x, y, z-axis by a, b, c times respectively. We finally obtain that

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \tag{50}$$

In general, a quadratic surface is determined by a second order equation of x, y, z as the following

$$Ax^2 + By^2 + Cz^2 + Dx + Ey + Fz + G = 0 \tag{51}$$

To determine the type (ellipsoid/hyperboloid/...) of this surface, one needs to compare the signs $+/-$ of A, B, C and G. The following example gives a hyperboloid:

$$x^2 + 2y^2 - 3z^2 + 2x + 4y + 6z - 2 = 0 \tag{52}$$

2.12.3 Different Coordinates

Besides rectangular coordinates, there are several different types of coordinate systems.

- (Cylindrical) The cylindrical system

$$\begin{cases} x = rcos\theta \\ y = rsin\theta \\ z = z \end{cases} \tag{53}$$

  and its inverse transformation

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ tan\theta = \dfrac{y}{x} \\ z = z \end{cases} \tag{54}$$

- (Spherical) The spherical system

$$\begin{cases} x = rsin\phi cos\theta \\ y = rsin\phi sin\theta \\ z = rcos\phi \end{cases} \tag{55}$$

  and its inverse transformation

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ tan\theta = \dfrac{y}{x} \\ cos\phi = \dfrac{z}{r} \end{cases} \tag{56}$$

3 Vector Functions

In this section, we mainly consider functions in $\mathbb{R}^3$.

3.1 Vector Functions

A vector function $\vec{f} : \mathbb{R} \to \mathbb{R}^3, t \mapsto (x(t), y(t), z(t))$ is a function that maps a real number t to some vector $(x(t), y(t), z(t))$ in 3-D space.
The graph of a vector function is a curve with a parameter.

Example 4 *Here are some examples of vector functions.*

- *(line)* $\vec{f}(t) = (at + x_0, bt + y_0, ct + z_0)$*;*
- *(curve)* $\vec{f}(t) = (1, t, t^2)$*;*
- *(spiral)* $\vec{f}(t) = (cost, sint, t)$*. Spirals such as* $\vec{f}$ *are an important class of curves. The curve* $\vec{f}$ *rotates and lifts up as* t *increases, in a period* 2π*;*

All the basic computations for vectors also work for vector functions. In particular, we have:

$$[DotProduct]\vec{v}(t) \cdot \vec{u}(t) = v_x(t)u_x(t) + v_y(t)u_y(t) + v_z(t)u_z(t) \tag{57}$$

and

$$[CrossProduct]\vec{v}(t) \times \vec{u}(t) = (|\begin{pmatrix} v_y(t) & v_z(t) \\ u_y(t) & u_z(t) \end{pmatrix}|, |\begin{pmatrix} v_z(t) & v_x(t) \\ u_z(t) & u_x(t) \end{pmatrix}|, |\begin{pmatrix} v_x(t) & v_y(t) \\ u_x(t) & u_y(t) \end{pmatrix}|) \tag{58}$$

3.2 Derivatives of Vector Functions

The derivative of a vector function $\vec{f}(t) = \langle x(t), y(t), z(t)\rangle$ is defined as the following:

Definition 3.1 (Derivative) *Given a vector function* $\vec{f} : \mathbb{R} \to \mathbb{R}^3, t \mapsto \langle x(t), y(t), z(t)\rangle$*, the derivative of* $\vec{f}$ *is*

$$\vec{f}'(t) = \langle x'(t), y'(t), z'(t)\rangle \tag{59}$$

For convenience, in some cases, we also use the notation $\dot{x}(t)$ to express the derivative of $x(t)$.

Example 5 *The derivative of a line* $\vec{f}(t) = \langle at + x_0, bt + y_0, ct + z_0\rangle$ *is just* $\vec{f}'(t) = \langle a, b, c\rangle$

Regarding the derivative and vector computations, we have the following theorems:

Theorem 2 *Given a pair of vector functions* $\vec{f}(t) = (x_1(t), y_1(t), z_1(t)), \vec{g}(t) = (x_2(t), y_2(t), z_2(t))$*, we have:*

$$(\vec{f} \cdot \vec{g})' = \vec{f}' \cdot \vec{g} + \vec{f} \cdot \vec{g}' \tag{60}$$

and

$$(\vec{f} \times \vec{g})' = \vec{f}' \times \vec{g} + \vec{f} \times \vec{g}' \tag{61}$$

Proof: The proof is just some straightforward derivation.
The first formula comes from

$$\begin{aligned}
&(\vec{f} \cdot \vec{g})' = (x_1x_2 + y_1y_2 + z_1z_2)' \\
&= x_1'x_2 + y_1'y_2 + z_1'z_2 + x_1x_2' + y_1y_2' + z_1z_2' \\
&= \vec{f}' \cdot \vec{g} + \vec{f} \cdot \vec{g}'
\end{aligned}$$

The second formula comes from

$$\begin{aligned}
&(\vec{f} \times \vec{g})' = \langle (y_1z_2 - y_2z_1)', (z_1x_2 - z_2x_1)', (x_1y_2 - x_2y_1)'\rangle \\
&= \langle (y_1'z_2 - y_2'z_1) + (y_1z_2' - y_2z_1'), (z_1'x_2 - z_2'x_1) + (z_1x_2' - z_2x_1'), (x_1'y_2 - x_2'y_1) + (x_1y_2' - x_2y_1')\rangle \\
&= \langle (y_1'z_2 - y_2'z_1), (z_1'x_2 - z_2'x_1), (x_1'y_2 - x_2'y_1)\rangle + \langle (y_1z_2' - y_2z_1'), (z_1x_2' - z_2x_1'), (x_1y_2' - x_2y_1')\rangle \\
&= \vec{f}' \times \vec{g} + \vec{f} \times \vec{g}'
\end{aligned}$$

□

Consequencially, we may further obtain the following results:

Corollary 1 *Given a vector function* $\vec{v}(t) = \langle x(t), y(t), z(t) \rangle$,

$$(\vec{v} \cdot \vec{v})' = 2\vec{v} \cdot \vec{v}' \tag{62}$$

One shall be aware that $||\vec{v}||^2 = \vec{v} \cdot \vec{v}$. If $(\vec{v} \cdot \vec{v})' = 0$, it follows that the vector function $\vec{v}(t)$ has constant length. Then we have the further proposition:

Proposition 1 *A vector function* $\vec{v}(t) = \langle x(t), y(t), z(t) \rangle$ *have constant length with respect to* t *if and only if* $\vec{v} \perp \vec{v}'$.

Proof: This proposition follows from the fact that

$$\vec{v} \cdot \vec{v} = 0 \iff \vec{v} \perp \vec{v}'$$

□

Corollary 2 *Given a vector function* $\vec{t} = \langle x(t), y(t), z(t) \rangle$,

$$(\vec{v} \times \vec{v}')' = \vec{v} \times \vec{v}'' \tag{63}$$

Proof: Note that $\vec{v}' \times \vec{v}' = 0$, we can derive this formula via:

$$\begin{aligned} &(\vec{v} \times \vec{v}')' = \vec{v}' \times \vec{v}' + \vec{v} \times \vec{v}'' \\ &= \vec{v} \times \vec{v}'' \end{aligned}$$

□

3.3 Tangent Vectors and Tangent Lines

Once we build the concept of the derivative of vector functions, we need to ask the following question: what is the geometrical meaning of such derivatives?
By regarding a vector function $\vec{v}(t) = \langle x(t), y(t), z(t) \rangle$ as a curve in 3-D space, the answer is that the derivative of $\vec{v}$, $\vec{v}'(t) = \langle x'(t), y'(t), z'(t) \rangle$, is the tangent vector at t.
Then one can obtain the tangent line $\vec{l}$ of $\vec{v}$ at t_0 as

$$\vec{l}(t) = \vec{v}(t_0) + t \cdot \vec{v}'(t_0) \tag{64}$$

3.4 Arc Length and Arc Length Parametrization

3.4.1 Arc Length

When we observe the motion of an object $\vec{r}(t) = \langle x(t), y(t), z(t)$, i.e., the moving of a car on a highway, we may wish to understand how fast it runs along the way. The basic idea is to consider the speed, or, the magnitude of the velocity, $\dot{\vec{r}}(t)$, $||\dot{\vec{r}}(t)|| = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2}$.
With speed $||\dot{\vec{r}}(t)||$, one can further compute how long the object has passed along its path during some time interval $t \in [0, T]$, i.e.,

$$\int_0^T ||\dot{\vec{r}}(t)||dt = \int_0^T \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2}dt \tag{65}$$

This gives rise to the idea of arc length of a vector function $\vec{r}$:

Definition 3.2 *Given a vector function* $\vec{r} : \mathbb{R} \to \mathbb{R}^3, t \mapsto \langle x(t), y(t), z(t) \rangle$, *the integral of* $\vec{r}$ *on* $t \in [a, b]$ *is*

$$\int_a^b ||\dot{\vec{r}}(t)||dt = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2}dt \tag{66}$$

3.4.2 Curves and Parametrized Curves

Curves are one-dimensional objects in $\mathbb{R}^3$. To describe a curve $C \subset \mathbb{R}^3$, the basic method is to build a vector function $\vec{r} : \mathbb{R} \to C \subset \mathbb{R}^3, t \mapsto \langle x(t), y(t), z(t) \rangle$. This process is called **parameterizing**, and the obtained function $\vec{r}$ is called the **parametrization of** C.
One shall be aware that for any curve C, parametrizations are not unique. For instance, both $\vec{r}_1(t) = \langle 2t, 3t, 4t \rangle$ and $\vec{r}_2(t) = \langle 4t, 6t, 8t \rangle$ are representing the same straight line l which has directional vector $\vec{v} = \langle 2, 3, 4 \rangle$.

3.4.3 Arc Length Parametrization

There is a good way of parametrization: we wish to find such a parametrization of a curve $\vec{r}$ so that its derivative is always a unit vector, i.e.,$||\vec{r}|| = 1$.
In physics, this is equivalent to say that we wish to find a way a running along a path so that there is no acceleration.

3.5 Integrals of Vector Functions

It is now a proper point to discuss the reversal of differentiating of vector functions–vector-valued integrals.

Definition 3.3 (Integral) *Given a vector function $\vec{f} : \mathbb{R} \to \mathbb{R}^3, t \mapsto \langle x(t), y(t), z(t)\rangle$, the integral of $\vec{f}$ on $t \in [a, b]$ is*

$$\int_a^b \vec{f}(t)dt = \langle \int_a^b x(t)dt, \int_a^b y(t)dt, \int_a^b z(t)dt\rangle \tag{67}$$

4 Multivariate Functions and Multivariate Calculus

The progress in physics and engineering has given rise to new mathematical models that could hardly be described with single variable functions. Typical examples include the height of a mountain, the distribution of temperature in a room and the pressure on a plane or a surface, etc. These novel objects have inspired the development of the theory of multivariate functions.

4.1 Multivariate Functions

A multivariate function (or, multivariable function, function with several variables) is defined as a map with more than one variable. For instance, a bi-variable function ususally has the following form:

$$f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto f(x, y) \tag{68}$$

Similarly, a tri-variable function has the following form:

$$f : \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto f(x, y, z) \tag{69}$$

Here are some examples:

- (Plane) $f(x, y) = ax + by$;
- (Parabolic surface) $f(x, y) = x^2 + y^2$;
- (Pressure on a 2-D plane) $p(x, y)$;
- (Temperature in a 3-D room) $T(x, y, z)$;

4.1.1 Graph

One main method to understand multivariate functions, especially bi-variate functions, is to visualize them via graphing.
For a bi-variate function $f(x, y)$, the graph could be regarded as a surface $\vec{S}$ with two parameters in $\mathbb{R}^3$:

$$\vec{S}(x, y) := \langle x, y, f(x, y)\rangle \tag{70}$$

4.1.2 Contour

As in three-dimensional geometry, we can further investigate the property of a bi-variate function $f(x, y)$ by considering the projection of its graph $z = f(x, y)$ onto the xOy plane.
Fixing any $z = z_0$, we could obtain a curve $f(x, y) = z_0$ on xOy plane. Such a curve is called a **contour** of f at $z = z_0$. By painting the contours of f (for instance) at $z = 0, 1, 2, 3, 4, ...$, we obtain a graph with a series non-intersecting curves called the **contour plot** of f.

4.2 Limit of Multivariate Functions

The limit of a multivariate function can be defined as the following:

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) \tag{71}$$

The computation of the limit of a multivariate function is similar to that of a single variable function.

Example 6

$$\lim_{(x,y)\to(1,2)} x^2+y^2 = 1^2+2^2 = 5 \tag{72}$$

Nevertheless, we need to understand the significant difference between $x \to x_0$ and $(x,y) \to (x_0,y_0)$:
In the one dimensional case, x may approach x_0 from the left or the right–in other words, from two different directions.
In the two dimensional case, however, (x,y) could approach (x_0,y_0) from infinitely many different directions. For instance, you may think about $(x,y) = (x+x_0, kx+y_0)$ and take $k \to 0$. Whatever k you choose, you can always conclude that $f(x,y)$ is converging to some value as $k \to 0$. However, such value might not be unique. This can happen when a $\frac{0}{0}$ appears in the limit. See, for instance, the following example:

Example 7 *Consider the limit of the function*

$$f(x,y) = \frac{x^2}{x^2+y^2} \tag{73}$$

at $(0,0)$.
Set $y = kx$. *We shall immediately see that*

$$\lim_{(x,y)\to(0,0)} \frac{x^2}{x^2+k^2x^2} = \frac{1}{1+k^2} \tag{74}$$

The result implies that the limit of $f(x,y)$ *at* $(0,0)$ *is dependent on* k*, thus not unique!*

In general, one may follow the following process to judge the existence of the limit of a function:

(a) Try to compute the limit directly. If $\frac{0}{0}$ appears, go to Step b;

(b) (k-test) Set $x = t + x_0, y = kt + y_0$ and compute the limit $\lim_{t\to 0} f(t+x_0, kt+y_0)$. If the result contains k, conclude that the limit does not exist; if k is cancelled, the result is inconclusive.

4.3 Partial Derivatives

The partia derivatives of a function $f(x,y)$ at (x,y) are defined as the following:

$$f_x = \frac{\partial f}{\partial x} = \lim_{\Delta x\to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} \tag{75}$$

$$f_y = \frac{\partial f}{\partial y} = \lim_{\Delta y\to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y} \tag{76}$$

Here $\Delta x, \Delta y$ denote small change of x, y respectively.
The computation of partial derivatives is similar to that of the derivatives of single-variable functions.
The main principle is as the following: to compute the partial derivative with respect to one variable, you may do it by regarding all other variables as constants. For instance, to compute the partial derivative of f with respect to x at (x_0,y_0), i.e.,

$$f_x(x_0,y_0)$$

you shall regard $y = y_0$ as some constant and compute the derivative of

$$h(x) = f(x,y_0)$$

Example 8 *(Computation of partial derivatives) One may compute the partial derivative of f with respect to x by regarding y as a constant. Let us look at some quick examples:*

- $f(x,y) = x^2 + y^2$. $f_x = 2x, f_y = 2y$;
- $f(x,y) = xy$. $f_x = y, f_y = x$. *Furthermore, at the point* $(0,0)$*, we have* $f_x(0,0) = 0, f_y(0,0) = 0$;
- $f(x,y) = e^{x^2+y^2}$. $f_x = 2xe^{x^2+y^2}, f_y = 2ye^{x^2+y^2}$. *Furthermore, at the point* $(1,2)$*, we have* $f_x(1,2) = 2e^5, f_y(1,2) = 4e^5$.

4.4 Geometric Meaning of Partial Derivatives

In this section, we discuss some three-dimensional geometric properties related to partial derivatives of a double-variable function $f(x,y)$.
The concept of derivatives are closely related to changing ratio. For a multivariable function f, its partial derivative reflects the changing ratio of f with respect to one of its variables. For instance, f_x reflects the changing ratio of f in the positive direction of x-axis.
Such idea about changing ratio further inspires the concept of tangential lines on a multivariate function f. This is a generalization of the tangential lines on single variable functions. For a double-variable function $f(x,y)$, each of f_x, f_y determines a tangential line in one direction of x, y-axis, respectively.

Definition 4.1 (Tangential Line, x-axis) *A* ***tangential line*** $\vec{r}$ *of f in the direction of x-axis at some point (x_0, y_0) is a straight line that intersects the graph of f, $z = f(x,y)$ at (x_0, y_0) and has the same slope of f in the direction of x-axis, i.e., with the directional vector $\vec{v} = \langle 1, 0, f_x(x_0, y_0)\rangle$. The equation of $\vec{r}$ is as the following:*

$$\vec{r}(t) = \langle t + x_0, y_0, f_x(x_0, y_0)t + f(x_0, y_0)\rangle \tag{77}$$

Definition 4.2 (Tangential Line, y-axis) *A* ***tangential line*** $\vec{r}$ *of f in the direction of y-axis at some point (x_0, y_0) is a straight line that intersects the graph of f, $z = f(x,y)$ at (x_0, y_0) and has the same slope of f in the direction of y-axis, i.e., with the directional vector $\vec{v} = \langle 0, 1, f_y(x_0, y_0)\rangle$. The equation of $\vec{r}$ is as the following:*

$$\vec{r}(t) = \langle x_0, t + y_0, f_y(x_0, y_0)t + f(x_0, y_0)\rangle \tag{78}$$

Of course, these are not the only two tangential lines of f at (x_0, y_0) (more precisely, the point $(x_0, y_0, f(x_0, y_0))$ on the graph of f). In fact, crossing one fixed point $(x_0, y_0, f(x_0, y_0))$ of the graph $z = f(x,y)$, there are infinitely many different tangential lines all of which stays in the same plane that intersects $z = f(x,y)$ at $(x_0, y_0, f(x_0, y_0))$. Such a plane S is called the **tangential plane** of f.
To determine the equation of S, we need to obtain its perpendicular vector $\vec{u}$. One can do this by using the cross product of the directional vectors of the tangential lines in x, y-axis:

$$\vec{u} = \langle 1, 0, f_x(x_0, y_0)\rangle \times \langle 0, 1, f_y(x_0, y_0)\rangle = \langle -f_x(x_0, y_0), -f_y(x_0, y_0), 1\rangle \tag{79}$$

Then we can define the tangential plane of f as the following:

Definition 4.3 (Tangential Plane) *A* ***tangential plane*** $\vec{S} = \langle x, y, z\rangle$ *of f at (x_0, y_0) is a plane that given by the following equation:*

$$\vec{u} \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0 \tag{80}$$

or, equivalently,

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0 \tag{81}$$

where we denote $\vec{u} = \langle -f_x(x_0, y_0), -f_y(x_0, y_0), 1\rangle, z_0 = f(x_0, y_0)$.

4.5 Gradients and Directional Derivatives

We just mentioned that partial derivatives are related to the changing ratio of a function in some direction. However, one should not be satisfied with the obtaining of chaning ratios in just several directions. This is due to the difference between a surface in $\mathbb{R}^3$ and a curve on plane–a surface contains more information than a curve. At one fixed point, the latter contains only two different directions (back and forward) but the formal contains infinitely many ones.

Such concerning gives rise to the idea of gradient and directional derivatives. A gradient of $f(x,y)$ is a vector defined by

$$\nabla f = \langle f_x, f_y \rangle \tag{82}$$

One may think of the vector $\vec{e}_1 = \langle f_x, 0\rangle, \vec{e}_2 = \langle 0, f_y\rangle$ as a weighted framework on the xOy plane. With this idea, one can define the "general changing ratio" of f denoted by $D_{\vec{u}}f$ in any direction given by some unit vector $\vec{u} = \langle u_x, u_y\rangle, ||\vec{u}|| = 1$ as the following:

$$D_{\vec{u}}f = u_x \cdot \vec{e}_1 + u_y \cdot \vec{e}_2 = u_x f_x + u_y f_y \tag{83}$$

which coincides to be $\nabla f \cdot \vec{u}$.
The discussion above inspires the idea of directional derivative of a function f in some direction $\vec{u}$. We define it rigorously as the following:

Definition 4.4 (Directional Derivative) *The **directional derivative** of a function f in the direction determined by a unit vector $\vec{u} = \langle u_x, u_y\rangle$ at (x_0, y_0) is given by the following:*

$$D_{\vec{u}}f = \nabla f \cdot \vec{u} \tag{84}$$

From this definition, one can also see that $D_{\vec{u}}f$ reaches its minimal when we take $\vec{u} = \frac{\nabla f}{||\nabla f||}$. This is because the length of projection of a vector $\vec{v}$–the dot product between a vector $\vec{v}$ and a unit vector $\vec{e}$ reaches its maximal when $\vec{e}//\vec{v}$.
One may ask for a geometric explanation for directional derivative. It goes as the following:
Let us say that you are approaching the point $(x_0, y_0, f(x_0, y_0))$ on the graph $z = f(x,y)$ along a path whose projection on xOy plane comes to be the line $l : \vec{r}(t) = t \cdot \vec{u} + (x_0, y_0)$. Now let $t \to 0$, you will obtain the changing ratio about how fast your "height" $z = f(x,y)$ changes as you are approaching (x_0, y_0) along the curve $f(\vec{r}(t))$. Such idea offers us an alternative definition of directional derivative:

Definition 4.5 (Directional Derivative, limit form) *The **directional derivative** of a function f in the direction determined by a unit vector $\vec{u} = \langle u_x, u_y\rangle$ at (x_0, y_0) is given by the following:*

$$D_{\vec{u}}f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + h \cdot u_x, y_0 + h \cdot u_y) - f(x_0, y_0)}{h} \tag{85}$$

One can show that the two definitions are equivalent:

Theorem 3

$$\nabla f(x_0, y_0) \cdot \vec{u} = \lim_{h\to 0} \frac{f(x_0 + h \cdot u_x, y_0 + h \cdot u_y) - f(x_0, y_0)}{h} \tag{86}$$

Proof:

$$\begin{aligned}
&\lim_{h\to 0} \frac{f(x_0 + h \cdot u_x, y_0 + h \cdot u_y) - f(x, y)}{h} \\
&= \lim_{h\to 0} \frac{f(x_0 + h \cdot u_x, y_0 + h \cdot u_y) - f(x_0, y_0 + h \cdot u_y) + f(x_0, y_0 + h \cdot u_y) - f(x_0, y_0)}{h} \\
&= \lim_{h\to 0} \frac{f(x_0 + h \cdot u_x, y_0 + h \cdot u_y) - f(x_0, y_0 + h \cdot u_y)}{h} + \lim_{h\to 0} \frac{f(x_0, y_0 + h \cdot u_y) - f(x_0, y_0)}{h} \\
&= \frac{d}{dh} f(x_0 + hu_x, y_0) + \frac{d}{dh} f(x_0, y_0 + hu_y) \\
&= hf_x(x_0, y_0) + hf_y(x_0, y_0)
\end{aligned}$$

□

4.6 Optimizations

One may wish to find the extreme (minimal/maximal) value of a function f on its domain. Global extreme point on $\mathbb{R}^3$ is usually hard to find, but for local ones, mathematicians have developed considerable tools to solve such problems. Basically, these problems are called optimization problems and have been classified into two main classes non-constraint optimization in which some whole domain $\mathcal{D}$ of f is considered and constraint optimization in which the points are limited to some curves (constraints) inside $\mathcal{D}$.
In each of the following sections, we mainly discuss double variable case without further specification.

4.6.1 Non-constraint Optimization

How can we spot extreme values of a function? We need to find its critical points like we once did for single-variable functions.

Definition 4.6 (Critical Points) *A* ***critical point*** *of $f(x, y)$ is a point (x_0, y_0) such that the partial derivatives of f vanishes, i.e.:*

$$f_x(x_0, y_0) = 0, f_y(x_0, y_0) = 0 \tag{87}$$

Then one shall be aware that besides minimal/maximal, there is another class of points on surfaces that are neither minimal nor maximal. Such class of points are called **saddle points**. A point on f is saddle if f is minimal in one direction but maximal in another. For instance, think about the shape of the saddle surface $z = xy$ at $(0, 0)$.
The main process to solve a non-constraint optimization problem is to first find critical points of f and then decide whether it is minimal/maximal, saddle or inconclusive. To determine the property of a critical point of f, we need to introduce the Hessian matrix of f:

Definition 4.7 (Hessian Matrix) *The Hessian Matrix of a function $f(x, y)$ is the following matrix:*

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \tag{88}$$

The determinant of Hessian matrix is usually called Hessian

$$H(f) = |\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}| \tag{89}$$

Once we find any critical point of f, we have the following theorem:

Theorem 4 *Let $A = (x_0, y_0)$ be a critical point of f, i.e., $f_x(x_0, y_0) = 0, f_y(x_0, y_0)$, then its property can be determined using the following process:*

- *$H(f) > 0, f_{xx} > 0$: A is locally minimal;*
- *$H(f) > 0, f_{xx} < 0$: A is locally maximal;*
- *$H(f) < 0$: A is saddle;*
- *$H(f) = 0$: inconclusive;*

Here are some quick examples:

Example 9
- *$z = x^2 + y^2$ are minimal at $(0, 0)$ since $H(f) = 4 > 0, f_{xx}(0, 0) > 0$;*
- *$z = x^2 - y^2$ are saddle at $(0, 0)$ since $H(f) = -4$;*
- *$z = -x^2 - xy - y^2$ are maximal at $(0, 0)$ since $H(f) = 3 > 0, f_{xx} < 0$;*

4.6.2 Constraint Optimization

In this section, we discuss the following problem:
Given a function $f(x, y)$, how can we find its critical point and critical value when (x, y) are constraint to some curve $g(x, y) = 0$?
The idea is to make use of the Lagrangian Multiplier. For a constraint optimization problem, we wish to construct the following function.
Here, we call f, the function to be optimized as the **target function** and $g = 0$, the **constraint**.

Definition 4.8 (Lagrangian Multiplier) *Given a tartget function f and a constraint $g = 0$, the Lagrangian Multiplier is a function F of x, y, λ as the following:*

$$F(x, y, \lambda) = f(x, y) - \lambda g(x, y) \tag{90}$$

With the construction of F, one can solve a constraint optimization problem by computing the critical point of F.

Theorem 5 *For a constraint optimization problem with target function f and constraint $g = 0$, we can use find the solution(s) (x_0, y_0) by solving the following equations for x, y:*

$$F_x(x, y, \lambda) = 0, F_y(x, y, \lambda) = 0 \tag{91}$$

Remark 4.1 *In all these types of problems, the solution(s) may not be unique. A function can have multiple critical points, and in case of multiple solutions, you need to discuss each point separately.*

5 Multivariate Integrals

Welcome to the world of integrals.

5.1 Indefinite Integrals

We introduce the basic idea of integrals on multivariate functions here.
The integrals of a function $f(x, y)$ could be, for instance, in the following terms:

$$\int f(x, y)dx, \int f(x, y)dy \tag{92}$$

Here the notations dx, dy determine with respect to which variable we are integrating.
The idea to compute indefinite integrals with respect to a certain variable is similar to that to compute partial derivatives–you do the integral with one variable while regarding other variable(s) as constant(s).
Let us look at some examples. In these examples, we denote $C \in \mathbb{R}$ as an arbitrary constant.

Example 10
- $\int x + ydx = x^2/2 + xy + C, \int x + ydy = xy + y^2/2 + C$
- $\int xydx = \frac{x^2y}{2} + C$, $\int xydy = \frac{xy^2}{2} + C$;
- $\int e^x sinydx = e^x siny + C, \int e^x sinydy = -e^x cosy + C$.

We may also have multi-integrals:

$$\int \int f(x, y)dxdy, \int \int f(x, y)dydx \tag{93}$$

Here, the sequential notations $dxdy, dydx$ determine the order of integral. We always start the order from the left.
Let us look at an example:

Example 11 *Consider the double integrals of the function $f(x, y) = x^2 + y^2$.*

- $\int x^2 + y^2dx = x^3/3 + xy^2 + C_1, \int \int x^2 + y^2dxdy = x^3y/3 + xy^3/3 + C_1y + C_2$;
- $\int x^2 + y^2dy = x^2y + y^3/3 + C_1, \int \int x^2 + y^2dydx = x^3y/3 + xy^3/3 + C_1x + C_2$;

One should immediately notice from the example above, that "the order matters"–multiple indefinite integrals in different orders give different results. This is different from the case of higher-order partial derivatives.

5.2 Definite Integrals

The definite integrals in multi-variable cases are significantly more complicated than that in single-variable case. We start the discussion from the two-dimensional case. Rigorously speaking, we wish to discuss the integral of a function f on some domain $D \subset \mathbb{R}^2$:

$$\iint_D f dxdy \tag{94}$$

Here, D denotes a set on two-dimensional space.

Remark 5.1 *One can immediately realize that if we take $f \equiv 1$, then this integral equals to the area of D. Furthermore, if f is non-constant, then the geometrical meaning of $\iint_D f dxdy$ is precisely the volume occupied by the three-dimensional region $A = \{(x, y, f(x, y)) | (x, y) \in D\} \subset \mathbb{R}^3$.*

5.2.1 Rectangular Domain

The simplest case stands on that D is a rectangle, i.e., $D = [a,b] \times [c,d]$, or one can write that $D = \{(x,y)|x \in [a,b], y \in [c,d]\}$. The integral may then be rewritten as the following:

$$\int_a^b \int_c^d f(x,y)dydx \tag{95}$$

or

$$\int_c^d \int_a^b f(x,y)dxdy \tag{96}$$

One can verify that in this case, the order of integral does not matter:

$$\int_a^b \int_c^d f(x,y)dydx = \int_c^d \int_a^b f(x,y)dxdy \tag{97}$$

5.2.2 Triangular and Trapezoid Domain

We discuss integrals on a class of domains that is in the shape of right triangular, with legs parallel to the axes. For instance, consider a triangular region:

$$D = \{(x,y)|x \in [a,b], y \in [kx+c, kb+c]\} \tag{98}$$

The corresponding integral is as the following:

$$\int_a^b \int_{kx+c}^{ka+c} f(x,y)dydx \tag{99}$$

We wish to change the order of integral. To do so, we need to have some idea of how the region looks like. Here D has three edges as line segments: $l_1 = (x,y)|x \in [a,b], y = kx+c, l_2 = \{(x,y)|x = a, y \in [ka+c, kb+c]\}, l_3 = \{(x,y)|x \in [a,b], y = kb+c\}$. Then we need to keep the shape of this region while changing the order.
Now think about a scanner scanning through each point in D. The meaning of the current order $dydx$ is that we first fix x, and vary y from $kx+c$ to $kb+c$. Here $x \in [a,b]$.
We switch the order to $dxdy$. This means that we first fix y, and vary x from somewhere to somewhere. In this case, we can first have in mind that $y \in [ka+c, kb+c]$. Then if we fix a y, we know that x starts at a. The ending point of x (for an fixed y) stands on the line $y = kx+c$ which comes to be $x = \frac{1}{k}(y-c)$.
We can then get the following result:

$$\int_{ka+c}^{kb+c} \int_a^{\frac{1}{k}(y-c)} dxdy \tag{100}$$

The case for a region like $D = \{(x,y)|x \in [a,b], y \in [ka+c, kx+c]\}$ or other types would be similar.
Some times the shape of a region may not be a triangle, but a trapezoid or more general shapes.
For the trapezoid, the idea is to first split it into a rectangle whose edges are parallel to the axes and one or two right triangles with legs parallel to the axes, then consider each subregion respectively.

5.2.3 More Complicated Cases

We discuss some more complicated case. First of all, we wish to understand how to change order of integral in the case of "curved right triangle", i.e., a right triangle with a bending hypotenuse determined by some function monotonely increasing function $y = f(x)$:

$$D = \{(x,y)|x \in [a,b], y \in [f(x), f(b)]\} \tag{101}$$

Then the integral could be, for instance,

$$\int_a^b \int_{f(x)}^{f(b)} g(x,y)dydx \tag{102}$$

The idea of changing order is similar to that of a normal right triangle. We then switch from $dydx$ to $dxdy$ and obtain:

$$\int_{f(a)}^{f(b)} \int_{a}^{f^{-1}(y)} dxdy \tag{103}$$

For a monotonely decreasing function, one can also borrow the ideas above and obtain similar results.

5.3 Change of Variables

Changing variables is an elementary method when computing multi-variate integrals. To make use of this method, we start by introducing the basic tool–the Jacobian matrix.

5.3.1 Jacobian Matrices and Jacobian Theorem

Given the following system of reparametrization,

$$x_1(v_1, ..., v_n), ..., x_n(v_1, ..., v_n) \tag{104}$$

we define the **Jacobian Matrix** of system 104 as the following n-by-n matrix:

$$\frac{\partial(x_1, ..., x_n)}{\partial(v_1, ..., v_n)} = \begin{pmatrix} \frac{\partial x_1}{\partial v_1} & \cdots & \frac{\partial x_1}{\partial v_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial v_1} & \cdots & \frac{\partial x_n}{\partial v_n} \end{pmatrix} \tag{105}$$

where the ij-th entry is $\frac{\partial x_i}{\partial v_j}$.

Example 12 (Computation of Jacobian Matrices) *We compute the Jacobian matrices between some frequently used coordinate systems.*

- *(Circular) Let $x = rcos\theta, y = rsin\theta$. Then we have:*

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{pmatrix} cos\theta & -rsin\theta \\ sin\theta & rcos\theta \end{pmatrix} \tag{106}$$

  so $|\frac{\partial(x,y)}{\partial(r,\theta)}| = r$;

- *(Cylindrical) Let $x = rcos\theta, y = rsin\theta z = h$. Then we have:*

$$\frac{\partial(x, y, z)}{\partial(r, \theta, h)} = \begin{pmatrix} cos\theta & -rsin\theta & 0 \\ sin\theta & rcos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{107}$$

  so $\frac{\partial(x,y,z)}{\partial(r,\theta,h)} = r$;

- *(Spherical) Let $x = rcos\theta sin\phi, y = rsin\theta sin\phi, z = rcos\phi$. Then we have:*

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = \begin{pmatrix} cos\theta sin\phi & -rsin\theta sin\phi & rcos\theta cos\phi \\ sin\theta sin\phi & rcos\theta sin\phi & rsin\theta cos\phi \\ cos\phi & 0 & -rsin\phi \end{pmatrix} \tag{108}$$

  so $\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = r^2 sin\phi$.

With the definiton of Jacobian matrix, we have the following theorem for changing variables:

Theorem 6 (Jacobian) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a n-variable function of $x_1, ..., x_n$ and one have the following reparametrization $x_1(v_1, ..., v_n), ..., x_n(v_1, ..., v_n)$. Then the following formula of changing variables holds:*

$$\int ... \int f(x_1, ..., x_n) dx_1 ... dx_n = \int ... \int f(x_1(v_1, ..., v_n), ..., x_n(v_1, ..., v_n)) |\frac{\partial(x_1, ..., x_n)}{\partial(v_1, ..., v_n)}| dv_1 ... dv_n \tag{109}$$

where $|\cdot|$ denotes the determinant of matrices.

6 Vector Calculus

6.1 Vector Fields

The development of modern physics and engineering have brought plenty of novel physical objects to human's attention, i.e., wind, river flows, vortex of liquid, light/magnetic/electrical/gravitational field, and classical theory of calculus appeared to be exhausted when confronted with their physical models. Such challenges inspired the development of many new mathematical tools. What stands at the elementary status among these new tools is vector fields.

The mathematical definition of a vector field is a type of "function" from a vector to a vector. In $\mathbb{R}^3$, we may think about the following term:

$$\vec{F}(x,y,z) = \langle A(x,y,z), B(x,y,z), C(x,y,z)\rangle \tag{110}$$

where each of A, B, C is a scalar function on $\mathbb{R}^3$. In our discussion, we further assume that A, B, C are smooth.

Apparently, $\vec{F}(x,y,z)$ defines a vector at some point (x,y,z). It follows that the term $\vec{F}(\cdot,\cdot,\cdot)$ attach each point (x,y,z) in $\mathbb{R}^3$ a vector $\vec{F}(x,y,z)$. A typical physical object that can be described by vector field is the velocity of wind in some domain of $\mathbb{R}^3$.

Similarly, one can understand the vectors in $\mathbb{R}^2$, that is, a $\mathbb{R}^2$-to-$\mathbb{R}^2$ mapping of the followin term:

$$\vec{F}(x,y) = \langle A(x,y), B(x,y)\rangle \tag{111}$$

Let us look at some examples of vector fields.

Example 13 *Here are some examples of vector fields:*

- *(Sheave Fields) $\vec{S}(x,y) = \langle ay, 0\rangle$ is a vector field that is always parallel ot the x-axis. The magnitude of $\vec{S}(x,y)$ increases linearly as the "height" y increases. This is a frequently-used model in Flow Mechanics;*
- *(Vortex on Plane) We can define a vortex on the plane as $\vec{V}(x,y) = \langle -y, x\rangle$. If we take $x = rcos\theta, y = rsin\theta$, we could rewrite it as $\langle -rsin\theta, rcos\theta\rangle$;*
- *(Gradient Field of Functions) For any function $g(x,y,z)$ on $\mathbb{R}^3$, its gradient $\nabla g = \langle g_x, g_y, g_z\rangle$ is a vector field.*

We introduce some basic attributes of vector fields in the following two sections. Then we turn to the introduction of vector-valued integrals.

6.2 Divergence

In 19th century, physicists wished to learn whether a swarm of gas or liquid would "diverge" or "converge" under the effect of some force fields (electrical, magnetical, etc) in some small time period and how fast such change happens based on merely the mathematical model of these fields. In otherwords, they are concered about the change of the density of gas or liquid at some certain small regions. These concerns gave birth to the concept of divergence.

Here we discuss the case in $\mathbb{R}^3$. The case for $\mathbb{R}^2$ would be similar.

Definition 6.1 (Divergence, Mathematical Definition) *Let $\vec{F}(x,y,z) = \langle A(x,y,z), B(x,y,z), C(x,y,z)\rangle$ be a vector field. The divergence of $\vec{F}$ is defined as*

$$Div\vec{F} = \nabla \cdot \vec{F} = \frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z} \tag{112}$$

6.2.1 Physical Interpretation of Divergence

In Physics, divergence describes the speed of how fast a vector field diverge or converge in each unit area of some certain domain. Visually, such a judgement could be made by observing whether the arrows of the graph of a vector field are becoming more or less crowded. Let $\vec{F}(x,y,z) = \langle A(x,y,z), B(x,y,z), C(x,y,z)\rangle$ be a vector field. We have the following qualitative properties:

- $Div\vec{F} > 0$: $\vec{F}$ is diverging. Visually, arrows are becoming more crowded;
- $Div\vec{F} = 0$: $\vec{F}$ is parallel or solenoidal, that is, neither diverging nor converging. Visually, arrows look parallel locally–if you fix your eye in some small region, the arrows are approximately parallel to each other;

- $Div\vec{F} < 0$: $\vec{F}$ is converging. Visually, arrows are becoming less crowded;

Here are some quick examples:

Example 14
- *$\vec{F}(x,y,z) = \langle x,y,z\rangle$. $Div\vec{F} = 1+1+1 = 3 > 0$, so this is a diverging vector field;*
- *$\vec{F}(x,y,z) = \langle y,0,0\rangle$. $Div\vec{F} = 0$, so this is a parallel vector field;*
- *$\vec{F}(x,y) = \langle -y,x\rangle$. $Div\vec{F} = -1+1 = 0$, so vortex is a parallel vector field. Geometrically, this is because the arrows of a vortex are rotating in a series of concentric circles;*
- *$\vec{F}(x,y,z) = \langle -x,-y,-z\rangle$. $Div\vec{F} = -3 > 0$, so this is a converging vector field;*

We introduced the concept at the begining of this section. Alternatively, one can redefine divergence by making the idea of "the speed of how fast a vector field diverge or converge in each unit area" rigorous as the following:

Definition 6.2 (Divergence, Physical Definition) *Let $\vec{F}(x,y,z) = \langle A(x,y,z), B(x,y,z), C(x,y,z)\rangle$ be a vector field in $\mathbb{R}^3$. Suppose that U is a small region (for instance, a small cubic or ball) containing (x,y,z). Let ∂U be the whole surface of U and $\vec{n}(x,y,z)$, a outward unit vector field always perpendicular to ∂U. Then the divergence of $\vec{F}$ at (x,y,z) is defined as the following surface integral among ∂U:*

$$Div\vec{F}(x,y,z) = \lim_{vol(U)\to 0} \frac{1}{vol(U)} \iint_{\partial U} \vec{F}\cdot\vec{n}dS \tag{113}$$

where $vol(U)$ is the volume of region U.

We will discuss this alternative definition in more details in the next chapter.

6.3 Curl

The concept of curl came into being in the way similar to divergence–flow mechanicists first wished to describe the information of "rotating" in a liquid or gas flow and then founded a physical value to measure the density of circulation of a flow–a new mathematical concept "curl" (in some old terms, "rots").

Definition 6.3 (Curl, Mathematical Definition) *Let $\vec{F}(x,y,z) = \langle A(x,y,z), B(x,y,z), C(x,y,z)\rangle$ be a vector field. The curl of $\vec{F}$ is defined as*

$$Curl\vec{F} = \nabla\times\vec{F} = \begin{pmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A & B & C \end{pmatrix} = \langle \frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}, \frac{\partial A}{\partial z} - \frac{\partial C}{\partial x}, \frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\rangle \tag{114}$$

The physical meaning of curl can be directly described by the following formula.

Theorem 7 *Let $\vec{F}$ be a vector field. Let $\vec{n}$ be a unit vector. Then the average angular velocity $\vec{\omega}$ around $\vec{n}$ (as the rotating axis) in an infinitesimal circle satisfies*

$$|curl\vec{F}\cdot\vec{n}| = 2||\vec{\omega}|| \tag{115}$$

Given $\vec{n} \parallel \vec{\omega}$, one can further detemine that

$$\vec{\omega} = \frac{1}{2}(curl\vec{F}\cdot\vec{n})\cdot\vec{n} \tag{116}$$

Here the sign of $\vec{\omega}$ (direction of rotating, $+/-$–counterclockwise/clockwise when facing the direction of $\vec{n}$) is determined by Right-Hand principle:

- *(Counterclockwise, +) $curl\vec{F}\cdot\vec{n} = 2||\vec{\omega}||$ when $curl\vec{F}\cdot\vec{n} > 0$;*
- *(Clockwise, +) $curl\vec{F}\cdot\vec{n} = -2||\vec{\omega}||$ when $curl\vec{F}\cdot\vec{n} < 0$;*

Remark 6.1 *From the theorem above, one can naturally penetrate that the meaning of $curl\vec{F}$ is the fastest rotating rate of $\vec{F}$ at some certain point (x,y,z). The direction that $\vec{F}$ reaches the fastest rotating rate is parallel to $[curl\vec{F}](x,y,z)$ itself.*

Remark 6.2 *When $curl\vec{F} = 0$, there is no vortex as $\vec{\omega} = 0$. In this case, $\vec{F}$ is said to be "irrotational".*

6.4 Line Integrals

Now we are prepared to discuss line and surface integrals. All these concepts originated in physics. Hence, each of these integrals are correspondent to physical value.

6.4.1 First Class (Scalar) Line Integral

The concept of First Class Line Integral comes from the computation of an uneven rope with a line density. Here we can think about a rope γ with density function f. We parametrize γ as $\vec{r}(t)$ in $\mathbb{R}^2$ or $\mathbb{R}^3$ with $t \in [a, b]$.
The First Class Line Integral, also called the Scalar Line Integral (this names comes from the fact that f is a scalar function), is defined as the following:

$$\int_\gamma f ds = \int_a^b f(\vec{r}(t))||\frac{d\vec{r}(t)}{dt}||dt \tag{117}$$

6.4.2 Second Class (Vector) Line Integrals

The concept of Second Class Line Integrals come from the computation of Work and Flux of some force or orce field along some certain curve.
Here we can think about a path in $\mathbb{R}^2$ or $\mathbb{R}^3$, γ standing in some vector field $\vec{F}$. We also parametrize γ as $\vec{r}(t)$ with $t \in [a, b]$.
The Second Class Line Integral, also called the Vector Line Integral (this names comes from the fact that $\vec{F}$ is a vector field), is defined as the following:

$$\int_\gamma \vec{F} d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \frac{d\vec{r}(t)}{dt} dt \tag{118}$$

There is also a special type of Second Class Line Integral specifically used to compute the flux of some vector field along a curve. To introduce this type of line integral, we first discuss the concept of the perpendicular vector of a curve.
The perpendicular vector $\vec{N}$ of a curve $\vec{r}(t) = \langle x(t), y(t) \rangle$ can be given by the following formula:

$$\vec{N}(t) = \langle \frac{dy(t)}{dt}, -\frac{dx(t)}{dt} \rangle \tag{119}$$

The geometrical meaning of this definition is just **rotating the velocity** $\dot{r}(t) = \langle \frac{dx(t)}{dt}, \frac{dy(t)}{dt} \rangle$ **clockwisely by** 90 **degree**. This direction also gives the positive direction of Flux in physics.
Now we can defined this special type of Second Class Line Integral:

$$\int_\gamma \vec{F} d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{N}(t) dt \tag{120}$$

6.4.3 Conservative Fields and Conservative Law

With the knowlege of Second Class Line Integral, we discuss a special kind of vector fields–conservative fields. Conservative fields are originally from classical mechanics. The key word "conservative" is correspondent to the concept of "energy conservation". Namely, a conservative field $\vec{F}$ can be regarded as a force field that is determined by some potential function p as below:

$$\vec{F} = \nabla p \tag{121}$$

In such a sense, if we consider the work done by force $\vec{F}$ in some closed loop C parametrized as $\vec{r}(t), t \in [a, b]$:

$$\oint_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \dot{\vec{r}}(t) dt$$

one can immediately see that since C is a loop, i.e., $\vec{r}(a) = \vec{r}(b)$, we shall have

$$\int_a^b \vec{F}(\vec{r}(t)) \cdot \dot{\vec{r}}(t) dt = \int_a^b \nabla p(\vec{r}(t)) \cdot \dot{\vec{r}}(t) dt = p(\vec{r}(b)) - p(\vec{r}(a)) = 0$$

Here the notation $\oint$ refers to loop integral–integral on a closed loop. By default, we claim the **the positive orientation of a loop integral is in the counterclockwise direction.**
Rigorously speaking, we have the following theorem–the famous Conservative Law:

Theorem 8 *Let $\vec{F}$ be a conservative vector field on some simply-connected domain $\mathcal{D}$, then for any closed loop $C \subset \mathcal{D}$,*

$$\oint_C \vec{F} \cdot d\vec{r} = 0 \tag{122}$$

Here, the concept "simply-connected" just means that $\vec{F}$ is well-defined on the whole $\mathcal{D}$, i.e., it has no singular point on $\mathcal{D}$. There is a longer story about this concept in the theories of Topology, but we do not have to dive into depth about it at this stage.
In physics, conservative law reflects the fact that the force determined by potential makes no work in any closed loop. In other words, the potential of an item only depends on its position.
A useful variation of the conservative law comes as the following:

Theorem 9 *Let $\vec{F}$ be a conservative vector field on some simply-connected domain $\mathcal{D}$, then for any pair of curves $\gamma_1, \gamma_2 \subset \mathcal{D}$ parametrized as $\vec{r}_1(t), \vec{r}_2(t), t \in [a, b]$ with same endings, i.e., $\vec{r}_1(a) = \vec{r}_2(a), \vec{r}_1(b) = \vec{r}_2(b)$,*

$$\int_{\gamma_1} \vec{F} \cdot d\vec{r} = \int_{\gamma_2} \vec{F} \cdot d\vec{r} \tag{123}$$

This theorem tells us that the change of potential only depends on the initial and ultimate position. One can use this theorem to simplify the computing of second class line integrals by replacing a complicated path by some simplier path (lines, arcs,etc) with same endings.

6.5 Surface Integrals

The ideas of Surface Integrals come from the geometrical problems on **two-dimensional surfaces in three-dimensional space**, i.e., some surface $\mathcal{S}$ in $\mathbb{R}^3$. To start our discussion, we first parametrize $\mathcal{S}$ by function $\vec{G}(u, v) = \langle x(u, v), y(u, v), z(u, v)\rangle$ with $(u, v) \in \mathcal{D} \subset \mathbb{R}^2$.
If $\mathcal{S}$ is given by some double-variable function $z = h(x, y)$, a natural way to parametrize $\mathcal{S}$ is to take $G(u, v) = \langle u, v, h(u, v)\rangle$. However, we shall note that this is not the only method to do parametrization.

6.5.1 First Class (Scalar) Surface Integral

The First Class Surface Integral computes the mass of some shell or mumbrane $\mathcal{S}$ with some density function f.

$$\int\int_{\mathcal{S}} f dA = \int\int_{\mathcal{D}} f(\vec{G}(u, v))||\vec{N}(u, v)||dudv \tag{124}$$

As a quick example, if we take $f \equiv 1$, this integral simply computes the area of $\mathcal{S}$.

6.5.2 Second Class (Vector) Surface Integral

The Second Class Surface Integral computes the Flux among a surface $\mathcal{S}$ in some vector field $\vec{F}$ in $\mathbb{R}^3$.

$$\int\int_{\mathcal{S}} \vec{F} d\vec{S} = \int\int_{\mathcal{D}} \vec{F}(\vec{G}(u, v)) \cdot \vec{N}(u, v)dudv \tag{125}$$

7 Advanced Calculus

7.1 Work and Curl

Let $\mathcal{D}$ be some domain in $\mathbb{R}^2$ or some two-dimensional surface in $\mathbb{R}^3$ and $\partial\mathcal{D}$ its boundary. As a quick remark, a complete spherical surface (or any other closed surface in $\mathbb{R}$) has no boundary.

7.1.1 Physical Intepretation

In flow mechanics, there is an elementary relation regarding a vector field $\vec{F}$ in some domain $\mathcal{D}$:

$$Work\ on\ \partial\mathcal{D} = Total\ Curl\ in\ \mathcal{D}$$

One can restate the relation above in mathematical language as the following:

$$\oint_{\partial\mathcal{D}} \vec{F}\cdot d\vec{r} = \int\int_{\mathcal{D}} curl\vec{F}\cdot d\vec{S} \tag{126}$$

Here $d\vec{r} = \langle dx, dy\rangle$ and $d\vec{S} = dxdy$ in the case of $\mathbb{R}^2$; or $d\vec{r} = \langle dx, dy, dz\rangle$ and $d\vec{S} = \langle dydz, dzdx, dxdy\rangle$ in the case of $\mathbb{R}^3$. For the case of $\mathbb{R}^2$, we can think about $d\vec{S} = dxdy$ since $z \equiv 0$.

7.1.2 Two-dimensional Case: Green's Formula

Theorem 10 (Green) *Let $\vec{F} = \langle A, B\rangle$ be a vector field in $\mathbb{R}^2$ and $\mathcal{D}$ a domain, then*

$$\oint_{\partial\mathcal{D}} \vec{F}\cdot d\vec{r} = \int\int_{\mathcal{D}} curl\vec{F}\cdot d\vec{S} \tag{127}$$

or, explicitly,

$$\oint_{\partial\mathcal{D}} Adx + Bdy = \int\int_{\mathcal{D}} (B'_x - A'_y)dxdy \tag{128}$$

Here the positive orientation of a loop integral is in the counterclockwise direction.

7.1.3 Three-dimensional Case: Stokes Formula

Theorem 11 (Stokes) *Let $\vec{F} = \langle A, B, C\rangle$ be a vector field in $\mathbb{R}^3$ and $\mathcal{D}$ a domain, then*

$$\oint_{\partial\mathcal{D}} \vec{F}\cdot d\vec{r} = \int\int_{\mathcal{D}} curl\vec{F}\cdot d\vec{S} \tag{129}$$

or, explicitly,

$$\oint_{\partial\mathcal{D}} Adx + Bdy + Cdz = \int\int_{\mathcal{D}} (C'_y - B'_z)dydz + (A'_z - C'_x)dzdx + (B'_x - A'_y)dxdy \tag{130}$$

Here the positive orientation of a loop integral is in the counterclockwise direction relative to the positive direction of $curl\vec{F}$. We recall the right-fist principle: take your thumb as the positive direction of curl, then the direction of your fingers is the positive direction of the loop integral.

7.2 Flux and Divergence

In the case of $\mathbb{R}^2$, let $\mathcal{D}$ be some region in $\mathbb{R}^2$ and $\partial\mathcal{D}$ its boundary as a closed loop.
In the case of $\mathbb{R}^3$, let $\mathcal{D}$ be some three-dimensional region in $\mathbb{R}^3$ and $\mathcal{S} = \partial\mathcal{D}$ its surface as a two-dimensional closed surface.

7.2.1 Physical Intepretation

In flow mechanics, there is another elementary relation regarding a vector field $\vec{F}$ in some region $\mathcal{D}$:

$$Flux\ on\ \partial\mathcal{D} = Total\ Divergence\ in\ \mathcal{D}$$

One can restate the relation above in mathematical language as the following:

$$\oint_{\partial\mathcal{D}} \vec{F}\cdot\vec{n}ds = \int_{\mathcal{D}} div\vec{F}\cdot d\mathcal{V} \tag{131}$$

Here the notation $d\mathcal{V}$ refers to the "general volume differential": in $\mathbb{R}^2$, it is the area differential $dA = dxdy$; in $\mathbb{R}^3$, it is the volume differential $dV = dxdydz$.

7.2.2 Two-dimensional Case: Green's Formula in Flux Form

Theorem 12 (Green, flux form) *Let $\vec{F} = \langle A, B \rangle$ be a vector field in $\mathbb{R}^2$ and $\mathcal{D}$ a domain, then*

$$\oint_{\partial\mathcal{D}} \vec{F} \cdot \vec{n} ds = \int\int_{\mathcal{D}} div \vec{F} \cdot dA \tag{132}$$

or, explicitly, note the parametrization of $\partial\mathcal{D}$ as $\vec{r}(t) = \langle x(t), y(t) \rangle, t \in [a, b]$ and $\vec{N}(t) = \langle y'(t), -x'(t) \rangle$,

$$\oint_a^b \vec{F}(\vec{r}(t)) \cdot \vec{N}(t) dt = \int\int_{\mathcal{D}} (A'_x + B'_y) dx dy \tag{133}$$

Here we recall that the positive direction of $\vec{N}$ is determined by the right-flag principle. Geometrically, we obtain the direction of $\vec{N}$ by rotating $\dot{\vec{r}}$ by 90 degrees clockwisely.

7.2.3 Three-dimensional Case: Divergence Theorem

Theorem 13 (Green, flux form) *Let $\vec{F} = \langle A, B, C \rangle$ be a vector field in $\mathbb{R}^2$ and $\mathcal{D}$ a domain, then*

$$\int\int_{\partial\mathcal{D}} \vec{F} \cdot \vec{n} dA = \int\int_{\mathcal{D}} div \vec{F} \cdot dV \tag{134}$$

or, explicitly, note the parametrization of $\partial\mathcal{D}$ as $\vec{G}(u, v) = \langle x(u, v), y(u, v), z(u, v) \rangle, t \in [a, b]$ and $\vec{N}(u, v) = \vec{G}'_u(u, v) \times \vec{G}'_v(u, v)$,

$$\int\int_{\partial\mathcal{D}} \vec{F}(\vec{G}(u, v)) \cdot \vec{N}(u, v) du dv = \int\int\int_{\mathcal{D}} (A'_x + B'_y + C'_z) dx dy dz \tag{135}$$

Here the positive direction of the left-hand side is determined by the direction of perpendicular vector $\vec{N}$ on the surface $\partial\mathcal{D}$. For a closed surface, $\vec{N}$ can be either outward or inward.