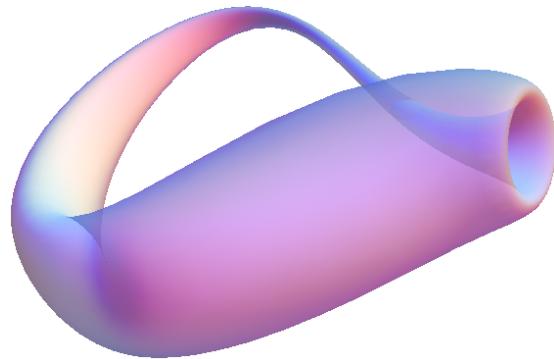


Multivariable and Vector Calculus

Lecture Notes for
MATH 0200
(Spring 2015)

Frederick Tsz-Ho Fong

Department of Mathematics
Brown University





Contents

| | | |
|----------|---|-----------|
| 1 | Three-Dimensional Space | 5 |
| 1.1 | Rectangular Coordinates in \mathbb{R}^3 | 5 |
| 1.2 | Dot Product | 7 |
| 1.3 | Cross Product | 9 |
| 1.4 | Lines and Planes | 11 |
| 1.5 | Parametric Curves | 13 |
| 2 | Partial Differentiations | 19 |
| 2.1 | Functions of Several Variables | 19 |
| 2.2 | Partial Derivatives | 22 |
| 2.3 | Chain Rule | 26 |
| 2.4 | Directional Derivatives | 30 |
| 2.5 | Tangent Planes | 34 |
| 2.6 | Local Extrema | 36 |
| 2.7 | Lagrange's Multiplier | 41 |
| 2.8 | Optimizations | 46 |
| 3 | Multiple Integrations | 49 |
| 3.1 | Double Integrals in Rectangular Coordinates | 49 |
| 3.2 | Fubini's Theorem for General Regions | 53 |
| 3.3 | Double Integrals in Polar Coordinates | 57 |
| 3.4 | Triple Integrals in Rectangular Coordinates | 62 |
| 3.5 | Triple Integrals in Cylindrical Coordinates | 67 |
| 3.6 | Triple Integrals in Spherical Coordinates | 70 |

| | | |
|----------|--|------------|
| 4 | Vector Calculus | 75 |
| 4.1 | Vector Fields on \mathbb{R}^2 and \mathbb{R}^3 | 75 |
| 4.2 | Line Integrals of Vector Fields | 83 |
| 4.3 | Conservative Vector Fields | 88 |
| 4.4 | Green's Theorem | 98 |
| 4.5 | Parametric Surfaces | 105 |
| 4.6 | Stokes' Theorem | 120 |
| 4.7 | Divergence Theorem | 127 |
| 5 | Topics in Physics and Engineering | 133 |
| 5.1 | Coulomb's Law | 133 |
| 5.2 | Introduction to Maxwell's Equations | 137 |
| 5.3 | Heat Diffusion | 141 |
| 5.4 | Dirac Delta Functions | 144 |



1 — Three-Dimensional Space

1.1 Rectangular Coordinates in \mathbb{R}^3

Throughout the course, we will use an ordered triple (x, y, z) to represent a point in the three dimensional space. The real numbers x , y and z in an ordered triple (x, y, z) are respectively the x -, y - and z -coordinates which, by convention, are defined according to the following diagram:

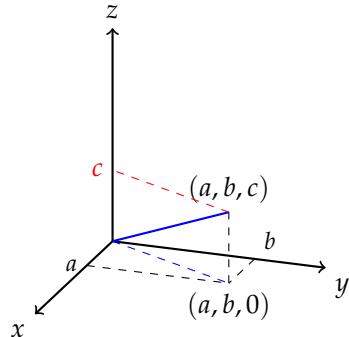


Figure 1.1: Rectangular coordinates system in 3-space

Notation We will use the notation \mathbb{R}^3 to denote the entire three dimensional space.

Any point on the x -axis has the form $(x, 0, 0)$, i.e. $y = 0$ and $z = 0$. Similarly, points on the y -axis are of the form $(0, y, 0)$, and points on the z -axis are of the form $(0, 0, z)$. The three coordinate axes meet at a point with coordinates $(0, 0, 0)$ which is called the **origin**.

A vector in \mathbb{R}^3 is an arrow which is based at one point and is pointing at another point. If a vector \mathbf{v} is based at (x_0, y_0, z_0) and points toward (x_1, y_1, z_1) , then the vector is written as:

$$\mathbf{v} = (x_1 - x_0)\mathbf{i} + (y_1 - y_0)\mathbf{j} + (z_1 - z_0)\mathbf{k}.$$

For example, the vector based at $(3, 2, -1)$ pointing at $(5, 2, 0)$ is expressed as

$$(5 - 3)\mathbf{i} + (2 - 2)\mathbf{j} + (0 - (-1))\mathbf{k} = 2\mathbf{i} + \mathbf{k}.$$

- i** Consequently, any two vectors that are **pointing at the same direction** and have **the same length** are considered to be **equal**, even though they may have different base points. For instance, a vector \mathbf{v} based at $(1, 2, 3)$ pointing at $(4, 3, 2)$, i.e.

$$\mathbf{v} = (4 - 1)\mathbf{i} + (3 - 2)\mathbf{j} + (2 - 3)\mathbf{k} = 3\mathbf{i} + \mathbf{j} - \mathbf{k},$$

is considered to be equal to the vector \mathbf{w} based at $(0, 0, 0)$ pointing at $(3, 1, -1)$. In other words, we can write $\mathbf{w} = \mathbf{v}$.

An alternative notation for a vector is the angle bracket $\langle a, b, c \rangle$ (to save the hassle of writing down \mathbf{i} , \mathbf{j} and \mathbf{k}):

Notation $\langle a, b, c \rangle = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$.

In this course, we make very little conceptual distinction between a *point* (x, y, z) and a *vector based at $(0, 0, 0)$ pointing at the point (x, y, z)* . However, speaking of notations, one should use $\langle x, y, z \rangle$ or $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ to denote a vector and (x, y, z) to denote a point so as to avoid confusion.

Vector additions and scalar multiplications are defined as follows:

Definition 1.1 — Vector Additions and Scalar Multiplications. Let $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ be two vectors in \mathbb{R}^3 , and c be a real scalar, then:

$$\begin{aligned}\mathbf{a} + \mathbf{b} &= \langle a_1 + b_1, a_2 + b_2, a_3 + b_3 \rangle && \text{(vector addition)} \\ c\mathbf{a} &= \langle ca_1, ca_2, ca_3 \rangle && \text{(scalar multiplication)}\end{aligned}$$

The negative of a vector is defined as: $-\mathbf{a} = (-1)\mathbf{a}$. The difference between vectors is defined as $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$.

Geometrically, these vector operations can be represented by the following diagrams:

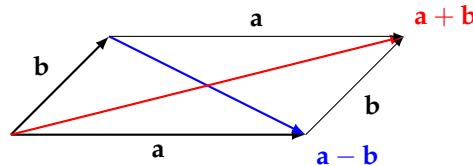


Figure 1.2: Geometric representations of various vector operations

Property Vector additions and scalar multiplications have the following algebraic properties

1. commutative rule: $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
2. associative rule: $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$
3. distributive rules: $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$ and $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$

1.2 Dot Product

There are two types of products for vectors in \mathbb{R}^3 , namely the *dot product* and the *cross product*. The former outputs a scalar whereas the latter outputs a vector. In this section, We first talk about the dot product.

Definition 1.2 — Dot Product. Let $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, then dot product between the vectors \mathbf{a} and \mathbf{b} are defined as:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

It is important to note that the dot product of a vector $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ by itself is given by:

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2$$

which is incidently the **square of the length** of the vector \mathbf{a} (by the Pythagoreas' Theorem in \mathbb{R}^3).

Notation Let $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$. We denote the length of a vector \mathbf{a} by $|\mathbf{a}|$, which is given by:

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

It is important to note that $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$.

The *length* of a vector, i.e. $|\mathbf{a}|$, is also called the *norm*, or the *magnitude*, of the vector.

Property It can easily be verified that the dot product satisfies the following algebraic properties:

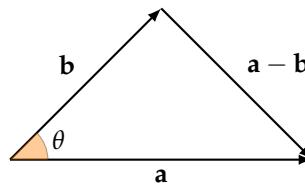
1. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.
2. $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$.
3. $(\lambda \mathbf{a}) \cdot \mathbf{b} = \lambda(\mathbf{a} \cdot \mathbf{b})$
4. $\mathbf{0} \cdot \mathbf{a} = \mathbf{a} \cdot \mathbf{0} = 0$.

The following theorem gives the geometric meaning of the dot product:

Theorem 1.1 Let $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ be two vectors in \mathbb{R}^3 , and θ be the angle between these two vectors. Then we have:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos \theta. \quad (1.1)$$

Proof. The proof uses the Law of Cosines. Consider the triangle in the diagram below:



The side opposite to the angle is represented by the vector $\mathbf{a} - \mathbf{b}$. Using the Law of Cosines:

$$\begin{aligned} |\mathbf{a} - \mathbf{b}|^2 &= |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2 |\mathbf{a}| |\mathbf{b}| \cos \theta \\ (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) &= \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} - 2 |\mathbf{a}| |\mathbf{b}| \cos \theta \\ \mathbf{a} \cdot \mathbf{a} - \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} &= \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} - 2 |\mathbf{a}| |\mathbf{b}| \cos \theta \\ \mathbf{a} \cdot \mathbf{a} - 2\mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{b} &= \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} - 2 |\mathbf{a}| |\mathbf{b}| \cos \theta \\ -2\mathbf{a} \cdot \mathbf{b} &= -2 |\mathbf{a}| |\mathbf{b}| \cos \theta \\ \mathbf{a} \cdot \mathbf{b} &= |\mathbf{a}| |\mathbf{b}| \cos \theta. \end{aligned}$$

■

One immediate consequence of Equation (1.1) is that it allows us to use the dot product to find the angle between two vectors. Precisely, the angle θ between two vectors \mathbf{a} and \mathbf{b} is given by:

$$\theta = \cos^{-1} \left(\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|} \right).$$

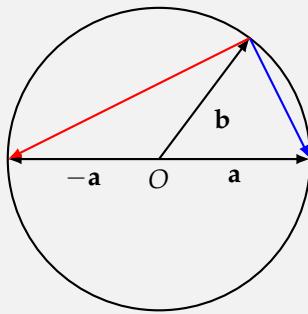
The most important case is that the two vectors \mathbf{a} and \mathbf{b} are **perpendicular**, also known as **orthogonal**. The angle between the vectors is $\frac{\pi}{2}$ and so we have the following important fact:

Corollary 1.2 Two non-zero vectors \mathbf{a} and \mathbf{b} are orthogonal if and only if $\mathbf{a} \cdot \mathbf{b} = 0$.

This corollary is particularly useful to determine whether two vectors are perpendicular.

■ **Example 1.1** Show that any triangle which is inscribed in a circle and has one of its side coincides the diameter of the circle must be a right-angled triangle.

■ **Solution** Let O be the center of the circle. Define vectors \mathbf{a} and \mathbf{b} as in the diagram below.



We would like to show the vectors in red and blue are orthogonal to each other. By basic vector additions and subtractions:

$$\text{Red vector} = -\mathbf{a} - \mathbf{b} \quad \text{Blue vector} = \mathbf{a} - \mathbf{b}.$$

Their dot product equals to

$$\begin{aligned} (-\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) &= -\mathbf{a} \cdot \mathbf{a} + \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} \\ &= -|\mathbf{a}|^2 + |\mathbf{b}|^2 \end{aligned} \quad (\text{recall } \mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2)$$

Since both \mathbf{a} and \mathbf{b} represent the radii of the circle, they have the same magnitude. Therefore $|\mathbf{a}| = |\mathbf{b}|$ and we have $(-\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = 0$. This shows the red and blue vectors are orthogonal.

1.3 Cross Product

The cross product is another important vector operations. In contrast to the dot product, the cross product outputs a **vector** instead of a scalar. A vector is characterized by its length and direction, we define the cross product by declaring these two attributes:

Definition 1.3 — Cross Product. Given two vectors $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ in \mathbb{R}^3 with angle θ between them, the cross product between \mathbf{a} and \mathbf{b} , denoted by $\mathbf{a} \times \mathbf{b}$, is defined as the vector such that:

- the length is given by $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin \theta$, i.e. the area of the parallelogram formed by vectors \mathbf{a} and \mathbf{b} ;
- the cross product $\mathbf{a} \times \mathbf{b}$ is orthogonal to both \mathbf{a} and \mathbf{b} ;
- the direction of $\mathbf{a} \times \mathbf{b}$ is determined by the right-hand grab rule illustrated by the Figure 1.3.

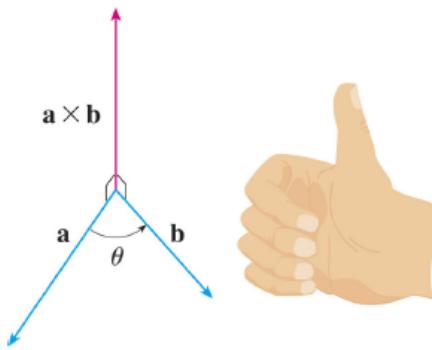
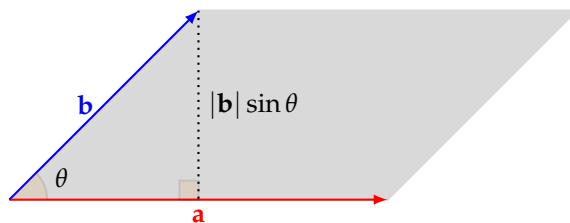


Figure 1.3: right-hand grab rule

From the right-hand grab rule, we can clearly see that $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{a}$ are vectors with the same length but in **opposite** direction, i.e. $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.

The magnitude of $|\mathbf{a} \times \mathbf{b}|$, which is defined to be $|\mathbf{a}| |\mathbf{b}| \sin \theta$, is the **area of the parallelogram** formed by \mathbf{a} and \mathbf{b} :



The following are algebraic properties of the cross products which are all useful. Based on the definition of cross products we presented above, the proofs are purely geometric and are omitted here.

Property The cross product satisfies:

1. $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
2. $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$
3. $\mathbf{a} \times \mathbf{0} = \mathbf{0}$
4. $\mathbf{a} \times \mathbf{a} = \mathbf{0}$

For simple vectors such as \mathbf{i} , \mathbf{j} and \mathbf{k} , their cross products can be easily found from the definition:

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}.$$

For more complicated vectors, the cross product can be computed using the following determinant formula:

Theorem 1.3 — Determinant Formula of Cross Product. Given two vectors $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, their cross product is given by:

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \\ &= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}.\end{aligned}\tag{1.2}$$

Proof. The proof follows from expanding:

$$\mathbf{a} \times \mathbf{b} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})$$

using the algebraic properties of the cross product. It is left as an exercise for readers. ■

The cross product will be extremely useful for us to find a vector which is orthogonal to a plane.

■ **Example 1.2** Given three points in the xyz -space:

$$A(0, 2, -1), \quad B(4, 0, -1), \quad C(7, -3, 0)$$

Find a vector \mathbf{n} which is orthogonal to the plane passing through A , B and C . Moreover, find the area of the triangle $\triangle ABC$.

■ **Solution** A vector \mathbf{n} is orthogonal to the plane if and only if it is orthogonal to any two (non-parallel) vectors on the plane. We will first find two vectors on the plane and then take the cross product. The outcome will give a vector orthogonal to the two vectors, and hence to the plane.

The following two vectors lie on the plane:

$$\begin{aligned}\overrightarrow{AB} &= \langle 4, 0, -1 \rangle - \langle 0, 2, -1 \rangle = \langle 4, -2, 0 \rangle \\ \overrightarrow{AC} &= \langle 7, -3, 0 \rangle - \langle 0, 2, -1 \rangle = \langle 7, -5, 1 \rangle\end{aligned}$$

Taking the cross product: $\overrightarrow{AB} \times \overrightarrow{AC} = \langle -2, -4, -6 \rangle$.

Therefore, the required vector \mathbf{n} can be taken to be any scalar multiple of $\langle -2, -4, -6 \rangle$, such as $\langle 1, 2, 3 \rangle$ or $\langle 2, 4, 6 \rangle$.

The length of the cross product $\overrightarrow{AB} \times \overrightarrow{AC}$ is equal to the area of the parallelogram formed by \overrightarrow{AB} and \overrightarrow{AC} . The area of the triangle $\triangle ABC$ is $\frac{1}{2}$ of the area of this parallelogram. Therefore,

$$\text{Area of } \triangle ABC = \frac{1}{2} |\overrightarrow{AB} \times \overrightarrow{AC}| = \frac{1}{2} \sqrt{(-2)^2 + (-4)^2 + (-6)^2} = \sqrt{14}.$$

1.4 Lines and Planes

1.4.1 Parametric Equations of Lines

In the three dimensional space, lines are no longer represented by an equation like $x + 2y = 1$ in the two dimensional plane. In order to represent a straight-line (and a curve as well), we need to introduce the **time variable** t , and think of a straight-line or a curve as the **path of a particle** travelling as t varies.

Suppose the line L is through the point $P_0(x_0, y_0, z_0)$ and is parallel to the vector $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$. For any variable point $P(x, y, z)$, the vector $\overrightarrow{P_0P}$ is parallel to the vector \mathbf{v} , meaning that $\overrightarrow{P_0P} = t\mathbf{v}$ for some real number t . Therefore,

$$\begin{aligned}\langle x, y, z \rangle - \langle x_0, y_0, z_0 \rangle &= t\mathbf{v} \\ \langle x, y, z \rangle &= \langle x_0, y_0, z_0 \rangle + t \langle v_1, v_2, v_3 \rangle\end{aligned}$$

Therefore, we have:

$$\begin{aligned}x &= x_0 + tv_1 \\ y &= y_0 + tv_2 \\ z &= z_0 + tv_3\end{aligned}$$

which is called the **parametric equation** of the line L . It is called this way because the variable t is called the **parameter** of the line. While intuitively t can be thought as the time, it needs not always be.

Notation In this course, the vector \mathbf{r} is ‘reserved’ to denote the position vector $\langle x, y, z \rangle$.

Using this notation, we can also write the parametric equation of the line L in vector form:

$$\mathbf{r}(t) = \overrightarrow{OP_0} + t\mathbf{v}.$$

The t -variable in $\mathbf{r}(t)$ emphasizes the fact that the position vector \mathbf{r} depends on t . It can be omitted if it is clear that t is the parameter letter.

To summarize, in order to write down the parametric equation of a straight-line, we need two “ingredients”:

1. a given point P_0 on the line, and
2. the direction \mathbf{v} of the line.

1.4.2 Equation of Planes

In three dimensions, equation of a plane can be represented in the form of $Ax + By + Cz = D$, as compared to $Ax + By = C$ for straight-lines in two dimensions.

For a plane through a given point $P_0(x_0, y_0, z_0)$ with a normal vector $\mathbf{n} = \langle A, B, C \rangle$, the equation of the plane is given by:

$$Ax + By + Cz = Ax_0 + By_0 + Cz_0. \quad (1.3)$$

Equation (1.3) can be proved by considering a variable point $P(x, y, z)$. As illustrated in Figure 1.4, the vector $\overrightarrow{P_0P}$ lies on the plane and therefore is orthogonal to the normal vector \mathbf{n} . Therefore, we have:

$$\begin{aligned}\mathbf{n} \cdot \overrightarrow{P_0P} &= 0 \\ \langle A, B, C \rangle \cdot (\langle x, y, z \rangle - \langle x_0, y_0, z_0 \rangle) &= 0 \\ \langle A, B, C \rangle \cdot \langle x, y, z \rangle - \langle A, B, C \rangle \cdot \langle x_0, y_0, z_0 \rangle &= 0 \\ Ax + By + Cz - Ax_0 - By_0 - Cz_0 &= 0 \\ Ax + By + Cz &= Ax_0 + By_0 + Cz_0.\end{aligned}$$

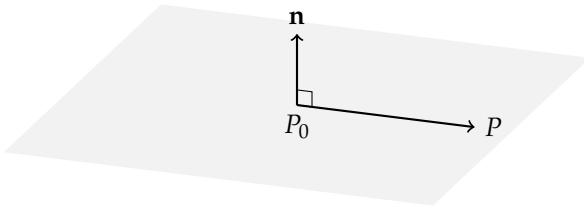


Figure 1.4: equation of a plane

■ **Example 1.3** Determine whether the following points are co-planar, i.e. contained in a single plane:

$$A(0, 2, -1), \quad B(4, 0, -1), \quad C(7, -3, 0), \quad D\left(\frac{1}{3}, \frac{1}{6}, \frac{1}{9}\right).$$

■ **Solution** Since any given three points determine a plane, we will first find an equation of the plane through points A , B and C , then we will check whether D lies on the plane by substituting its coordinates into the equation found.

The two "ingredients" of finding the equation of a plane are

1. a given point on the plane; and
2. a normal vector to the plane.

In order to find the normal vector to the plane through A , B and C , we take the cross product of \vec{AB} and \vec{AC} . Incidentally, these three points are exactly the same as in Example 1.2, in which we have worked out that a normal vector to the plane through A , B and C is given by $\mathbf{n} = \langle 1, 2, 3 \rangle$.

Take $A(0, 2, -1)$ to be the given point, then the equation of the plane through A , B and C is given by:

$$1x + 2y + 3z = (1) + 2(2) + 3(-1).$$

After simplification: $x + 2y + 3z = 1$.

Substitute the coordinates of D into the equation:

$$\text{LHS} = \frac{1}{3} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{9} = 1 = \text{RHS}.$$

Therefore, D lies on the plane through A , B and C . In other words, these four points are co-planar.

(i) Given a plane $Ax + By + Cz = D$, the vanishing of some of the A , B , C and D has the following geometric significance (assuming A , B and C are not all zero in all cases):

- When $D = 0$, the plane passes through the origin $(0, 0, 0)$.
- When $C = 0$, the normal vector is of the form $\langle A, B, 0 \rangle$ and hence is horizontal. In this case, the plane vertical.
- When $A = B = 0$, the normal vector is of the form $\langle 0, 0, C \rangle$ and hence is vertical. In this case, the plane is horizontal.

1.5 Parametric Curves

1.5.1 Parametric Equations of Curves

In two dimensions, there are two ways to represent a curve, namely in the form of $x^2 + y^2 = 1$, or of the form

$$\begin{aligned}x &= \cos t \\y &= \sin t.\end{aligned}$$

The former is called the **Cartesian equation** and the second one is called the **parametric equation**.

However, in three dimensions, a single Cartesian equation such as $x^2 + y^2 + z^2 = 1$ represents a **surface** instead. Therefore, we will only use the parametric equations to present curves in three dimensions.

Definition 1.4 — Parametric Equation of a Curve. The parametric equation of a curve is of the form:

$$\begin{aligned}x &= f(t) \\y &= g(t) \\z &= h(t)\end{aligned}$$

where $f(t)$, $g(t)$ and $h(t)$ are differentiable functions of t . In vector notations, the parametric equation of this curve is written as:

$$\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}.$$

The parametric equation of a straight-line is a special case of parametric equation of a "curve".

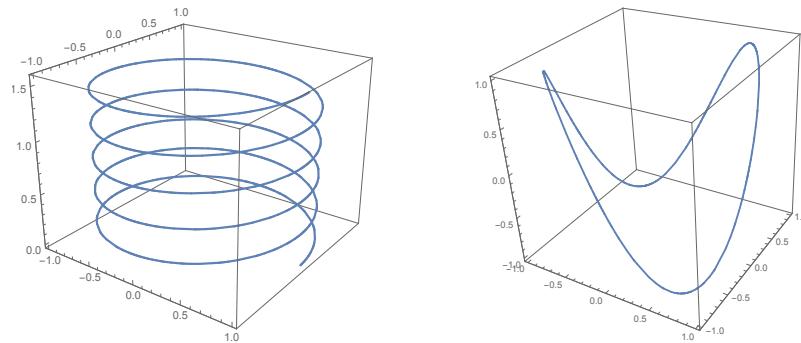
An interesting example of a parametric curve is the **helix**:

$$\mathbf{r}_1(t) = (\cos t)\mathbf{i} + (\sin t)\mathbf{j} + \frac{t}{20}\mathbf{k}.$$

It is a curve that goes around the circle but the altitude is constantly increasing. See Figure 1.5a for the computer sketch.

Here is another example of a parametric curve. See Figure 1.5b for the sketch.

$$\mathbf{r}_2(t) = (\sin t)\mathbf{i} + (\cos t)\mathbf{j} + (\sin 2t)\mathbf{k}.$$



(a) sketch of $\mathbf{r}_1(t)$, a helix

(b) sketch of $\mathbf{r}_2(t)$

Figure 1.5: Sketches of two parametric curves

1.5.2 Derivatives of Parametric Curves

One reason for using vector notations for a parametric curve is that their derivatives carry various physical and geometric meanings.

Given a parametric curve $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, regarded as the position vector of a particle at time t , then:

- the first derivative, $\mathbf{r}'(t) = f'(t)\mathbf{i} + g'(t)\mathbf{j} + h'(t)\mathbf{k}$, represents the **velocity vector** of the particle,
- geometrically, the first derivative $\mathbf{r}'(t)$ (if non-zero) is a **tangent vector** of the curve,
- the length $|\mathbf{r}'(t)| = \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2}$ represents the **speed** of the particle, and
- the second derivative, $\mathbf{r}''(t) = f''(t)\mathbf{i} + g''(t)\mathbf{j} + h''(t)\mathbf{k}$, represents the **acceleration vector** of the particle.

Conservation of Angular Momentum

In physics, given a particle with mass m travelling along the $\mathbf{r}(t)$, the following vector is defined to be the **angular momentum** of the particle:

$$\mathbf{L}(t) = \mathbf{r}(t) \times m\mathbf{r}'(t).$$

When $\mathbf{L}(t)$ is a non-zero constant vector (independent of t), we say that the angular momentum is *conserved*. The conservation of angular momentum implies that the path of the particle is contained in a plane. It can be explained as follows:

By the definition of cross product, the angular momentum $\mathbf{L}(t)$ is always orthogonal to $\mathbf{r}(t)$ (and to $\mathbf{r}'(t)$ too, but we do not need this). Therefore, at any time t , we have:

$$\mathbf{L}(t) \cdot \mathbf{r}(t) = 0.$$

Let $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$. If $\mathbf{L}(t)$ is a constant vector, it can be expressed as $\mathbf{L} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ where A, B and C are fixed numbers. Then:

$$\begin{aligned} (A\mathbf{i} + B\mathbf{j} + C\mathbf{k}) \cdot (x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}) &= 0 \\ Ax(t) + By(t) + Cz(t) &= 0. \end{aligned}$$

Therefore, the point $(x(t), y(t), z(t))$ lies on the plane $Ax + By + Cz = 0$, which is plane with normal vector \mathbf{L} and through the origin. In other words, the path of the particle is confined in this plane.

Product Rules of Differentiating Curves

When differentiating the dot or cross product of two curves $\mathbf{u}(t)$ and $\mathbf{v}(t)$, you may apply the product rule as like in single variable calculus.

Property Given two curves $\mathbf{u}(t)$ and $\mathbf{v}(t)$, and a scalar function $f(t)$, we have:

1. $\frac{d}{dt}(f(t)\mathbf{u}(t)) = f'(t)\mathbf{u}(t) + f(t)\mathbf{u}'(t)$
2. $\frac{d}{dt}(\mathbf{u}(t) \cdot \mathbf{v}(t)) = \mathbf{u}'(t) \cdot \mathbf{v}(t) + \mathbf{u}(t) \cdot \mathbf{v}'(t)$
3. $\frac{d}{dt}(\mathbf{u}(t) \times \mathbf{v}(t)) = \mathbf{u}'(t) \times \mathbf{v}(t) + \mathbf{u}(t) \times \mathbf{v}'(t).$

Here is a good example on the use of one of the above product rules.

■ **Example 1.4** Given $\mathbf{r}(t)$ represents a particle travelling at uniform speed C . Show that its velocity and acceleration vectors are always orthogonal.

■ **Solution** The particle is travelling at uniform speed C . Therefore, $|\mathbf{r}'(t)| \equiv C$. We want to show that $\mathbf{r}'(t) \cdot \mathbf{r}''(t) \equiv 0$, so it is natural to differentiate $|\mathbf{r}'(t)|$ with respect to t so that the RHS vanishes and the LHS perhaps may be related to $\mathbf{r}''(t)$.

However, since $|\mathbf{r}'(t)|$ is the form of a square root so it is cumbersome to differentiate it.

Instead, we differentiate $|\mathbf{r}'(t)|^2 = C^2$ using the fact that $|\mathbf{r}'(t)|^2 = \mathbf{r}'(t) \cdot \mathbf{r}'(t)$:

$$\begin{aligned} |\mathbf{r}'(t)|^2 &= C^2 \\ \mathbf{r}'(t) \cdot \mathbf{r}'(t) &= C^2 \\ \frac{d}{dt} (\mathbf{r}'(t) \cdot \mathbf{r}'(t)) &= 0 \\ \mathbf{r}''(t) \cdot \mathbf{r}'(t) + \mathbf{r}'(t) \cdot \mathbf{r}''(t) &= 0 \\ 2\mathbf{r}'(t) \cdot \mathbf{r}''(t) &= 0. \end{aligned}$$

Therefore, we have $\mathbf{r}'(t) \cdot \mathbf{r}''(t) = 0$, which is desired.

1.5.3 Arc Lengths of Curves

For a particle travelling at uniform speed, the distance (i.e. arc length) travelled is simply:

$$\text{distance} = \text{speed} \times \text{time lapsed}$$

As compared to the area of a rectangle: height \times width. However, if one is asked to calculate the area under a curve $y = f(x)$, $a \leq x \leq b$, one should consider the integral

$$\int_a^b y dx = \int_a^b f(x) dx$$

under the rationale of Riemann sum: $\int_a^b y dx \approx \sum_i y_i \Delta x_i$.

At the same token, if the particle is not travelling at uniform speed, one should calculate the distance by integration:

$$\text{distance} = \int (\text{speed}) \times d(\text{time}),$$

as compared to area $= \int \text{height} \times d(\text{width})$. Precisely, we have:

Theorem 1.4 Given a parametric curve $\mathbf{r}(t)$. The arc length of the curve from the point at $t = a$ to the point at $t = b$ is given by:

$$\text{arc length} = \int_a^b |\mathbf{r}'(t)| dt. \quad (1.4)$$

■ **Example 1.5** Find the arc length of the curve:

$$\mathbf{r}(t) = \frac{1}{2}t^2 \mathbf{i} + \frac{2\sqrt{2}}{3}t^{\frac{3}{2}} \mathbf{j} + t \mathbf{k}$$

from $(0, 0, 0)$ to $(2, \frac{8}{3}, 2)$.

■ **Solution** It is simple to verify that the initial point corresponds to $t = 0$ since $\mathbf{r}(0) = \langle 0, 0, 0 \rangle$, whereas the final point corresponds to $t = 2$ since $\mathbf{r}(2) = \langle 2, \frac{8}{3}, 2 \rangle$.

$$\begin{aligned} \mathbf{r}'(t) &= \left\langle t, \sqrt{2}t^{\frac{1}{2}}, 1 \right\rangle \\ |\mathbf{r}'(t)| &= \sqrt{t^2 + 2t + 1} \\ &= \sqrt{(t+1)^2} \\ &= t+1 \quad (\text{note that } t+1 > 0 \text{ in our case}) \end{aligned}$$

The desired arc length is given by:

$$\begin{aligned}\int_0^2 |\mathbf{r}'(t)| dt &= \int_0^2 (t+1) dt \\ &= \frac{t^2}{2} + t \Big|_0^2 \\ &= 10.\end{aligned}$$

1.5.4 Arc-Length Parametrization

Let's begin our discussion by considering the three curves:

$$\begin{aligned}\mathbf{r}_1(t) &= (\cos t)\mathbf{i} + (\sin t)\mathbf{j} + t\mathbf{k}, \quad 0 \leq t \leq 2\pi \\ \mathbf{r}_2(t) &= (\cos 2t)\mathbf{i} + (\sin 2t)\mathbf{j} + 2t\mathbf{k}, \quad 0 \leq t \leq \pi\end{aligned}$$

If you plot them using Mathematica, you should find out that these two curves are the same, although their speeds are different. The curve $\mathbf{r}_2(t)$ is obtained by replacing every t in $\mathbf{r}_1(t)$ by $2t$. The initial and final times are adjusted so that the end-points of both \mathbf{r}_1 and \mathbf{r}_2 are $(0, 0, 0)$ and $(0, 0, 2\pi)$. We say \mathbf{r}_2 is a **reparametrization** of \mathbf{r}_1 .

If $\mathbf{r}(s)$ is a parametric curve such that $|\mathbf{r}'(s)| = 1$ for any s , we say the curve is **parametrized by arc-length**. For such a parametrization, it is conventional to use s to denote the parameter. Given a parametric curve $\mathbf{r}(t)$, *in theory* one can reparametrize the curve by arc-length, such that with the new parameter s , the curve $\mathbf{r}(s)$ travels at unit speed. To find the arc-length parametrization, you may follow the procedure:

- Given a curve $\mathbf{r}(t) : [a, b] \rightarrow \mathbb{R}^3$, compute the following integral:

$$s = \int_a^t |\mathbf{r}'(\tau)| d\tau.$$

- Since the upper limit of the above integral is t , the function s should be a function of t . Express t in terms of s whenever it is possible, so that t is a function of s , i.e. $t = t(s)$.
- Finally, replace all t 's by this function of s in the curve $\mathbf{r}(t)$.

The new parametrization $\mathbf{r}(s)$ will be arc-length parametrized. Let's see some examples before we learn why it works:

- Example 1.6** Find the arc-length parametrization of the curve:

$$\mathbf{r}(t) = (\cos t)\mathbf{i} + (\sin t)\mathbf{j} + t\mathbf{k}, \quad t \in [0, 2\pi].$$

- Solution** By straight-forward computations, we get

$$|\mathbf{r}'(t)| = \sqrt{2}.$$

Therefore,

$$s(t) = \int_0^t |\mathbf{r}'(\tau)| d\tau = \int_0^t \sqrt{2} d\tau = \sqrt{2}t.$$

Express t in terms of s , we get $t = \frac{s}{\sqrt{2}}$. Replace all t 's in $\mathbf{r}(t)$ by $\frac{s}{\sqrt{2}}$, we get an arc-length parametrization:

$$\mathbf{r}(s) = \left(\cos \frac{s}{\sqrt{2}} \right) \mathbf{i} + \left(\sin \frac{s}{\sqrt{2}} \right) \mathbf{j} + \frac{s}{\sqrt{2}} \mathbf{k}.$$

■ **Example 1.7** Find the arc-length parametrization of the curve:

$$\mathbf{r}(t) = \frac{1}{2}t^2\mathbf{i} + \frac{2\sqrt{2}}{3}t^{\frac{3}{2}}\mathbf{j} + t\mathbf{k}, \quad t \geq 0.$$

■ **Solution** By straight-forward computations (done before), we get:

$$|\mathbf{r}'(t)| = t + 1.$$

Consider:

$$s = \int_0^t |\mathbf{r}'(\tau)| d\tau = \int_0^t (\tau + 1) d\tau = \frac{\tau^2}{2} + \tau \Big|_0^t = \frac{t^2}{2} + t.$$

To solve t in terms of s , we use the quadratic equation. One should get:

$$t = \frac{-2 + \sqrt{4 + 8s}}{2} = -1 + \sqrt{1 + 2s}.$$

Finally, replace all t 's in $\mathbf{r}(t)$ by this function of s , we get an arc-length parametrization:

$$\mathbf{r}(s) = \frac{1}{2}(-1 + \sqrt{1 + 2s})^2 \mathbf{i} + \frac{2\sqrt{2}}{3}(-1 + \sqrt{1 + 2s})^{3/2} \mathbf{j} + (-1 + \sqrt{1 + 2s}) \mathbf{k}.$$

To see why this procedure gives an arc-length parametrization, we need to show $|\mathbf{r}'(s)| = 1$. We use the Fundamental Theorem of Calculus and the chain rule:

$$\begin{aligned} |\mathbf{r}'(s)| &= \left| \frac{d\mathbf{r}}{ds} \right| = \left| \frac{d\mathbf{r}}{dt} \frac{dt}{ds} \right| \\ &= |\mathbf{r}'(t)| \frac{1}{\frac{ds}{dt}}. \end{aligned}$$

Recall that s is defined to be

$$s = \int_a^t |\mathbf{r}'(\tau)| d\tau.$$

The Fundamental Theorem of Calculus tells us that:

$$\frac{ds}{dt} = |\mathbf{r}'(t)|$$

and so,

$$|\mathbf{r}'(s)| = |\mathbf{r}'(t)| \cdot \frac{1}{|\mathbf{r}'(t)|} = 1.$$

Although the above procedure of finding arc-length parametrization works in the two examples we have seen, in general it may be hard to find an arc-length parametrization. Since both steps – integration and solving t in terms of s – can be difficult if the given curve $\mathbf{r}(t)$ is not *nice*.

The arc-length parametrization plays a significant role in finding curvature and torsion of a curve. Interested readers may consult Briggs-Cochran-Gillett's book Section 12.9 for detail. In this course, the arc-length parametrizations will come up later in the line integral chapters.



2 — Partial Differentiations

2.1 Functions of Several Variables

As the name implies, a function of several variables, or a multivariable function, is one that not only depends on one but several quantities. Examples of which include:

$$\text{volume of a cylinder} = \pi r^2 h$$

where r is the radius of the cylinder and h is the height.

Symbolically, we denote the function for the volume of a cylinder by $V(r, h)$, which indicates V depends on both r and h . We can write:

$$V(r, h) = \pi r^2 h.$$

In this chapter, we will extend theory and applications of single-variable differentiation to multivariable differentiation. Multivariable integration will be discussed in the next chapter.

2.1.1 Graphs of Two-Variable Functions

In single-variable calculus, we visualize a function $y = f(x)$ through its graph. The horizontal x -axis stands for the inputs, and the height of the graph above x represents the output $f(x)$. Many concepts in single-variable calculus, such as derivatives, integrals, critical points, etc. are introduced using the graph of a function.

For functions of two variables, i.e. $f(x, y)$, the graph is no longer a *curve* in \mathbb{R}^2 , but a *surface* in \mathbb{R}^3 . The inputs involve two variables x and y , and are represented by points on the xy -plane. The value of the function $f(x, y)$ is now represented by the height z of the surface above the point (x, y) .

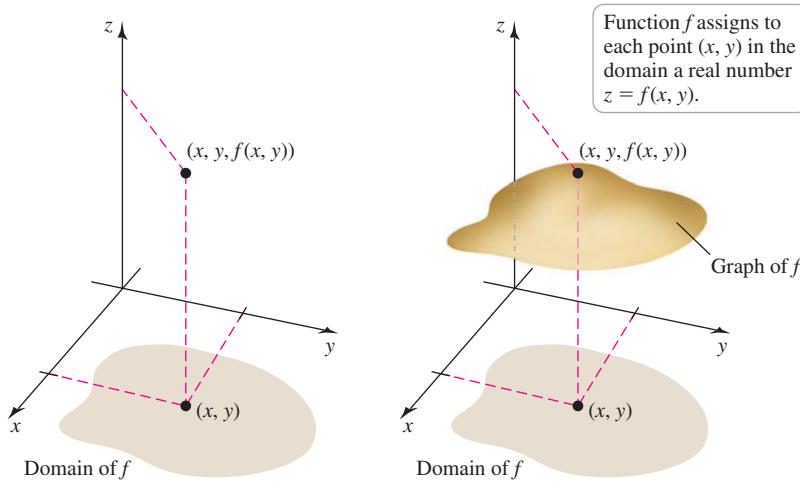


Figure 2.1: value of $f(x_0, y_0)$ is the height of the surface above the point $(x_0, y_0, 0)$

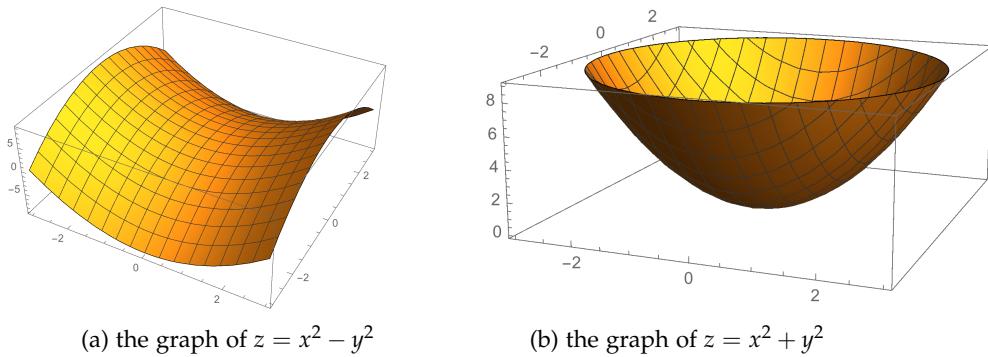


Figure 2.2: graphs of two-variables functions

2.1.2 Level Set Diagrams

Another common way to visualize a two-variable function is through its level sets:

Definition 2.1 — Level Sets. Given a function $f(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$, a level set of the function f is a subset of \mathbb{R}^n of the form:

$$f(x_1, \dots, x_n) = c$$

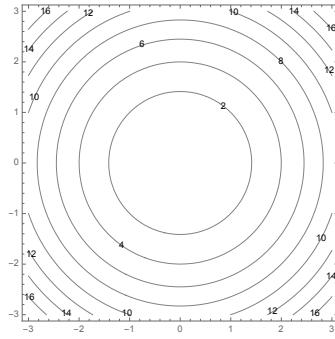
where c is a constant.

Given $f(x, y) = x^2 + y^2$, which is a function from \mathbb{R}^2 to \mathbb{R} . An example of a level set of f is $\underbrace{x^2 + y^2}_{f(x,y)} = 1$, which is a unit circle on \mathbb{R}^2 centered at the origin.

By taking c to be different values, we get several level sets on the plane. They are circles centered at the origin with varying radii depending on the value of c chosen:

| | | |
|---------|-----------------|---------------------|
| $c = 0$ | $x^2 + y^2 = 0$ | the origin only |
| $c = 1$ | $x^2 + y^2 = 1$ | radius = 1 |
| $c = 2$ | $x^2 + y^2 = 2$ | radius = $\sqrt{2}$ |
| $c = 3$ | $x^2 + y^2 = 3$ | radius = $\sqrt{3}$ |

The **level set diagram** of the two-variable function $f(x, y)$ consists of some representative level sets of the function on \mathbb{R}^2 . See Figure 2.3.

Figure 2.3: level set diagram of $f(x, y) = x^2 + y^2$

For three-variable functions $f(x, y, z)$, we do not attempt to visualize its **graph**, but we can visualize its level set diagram. The former requires the fourth dimension while the latter can be visualized in \mathbb{R}^3 . A *generic* level set of a three-variable function $f(x, y, z)$ is a **surface** in \mathbb{R}^3 .

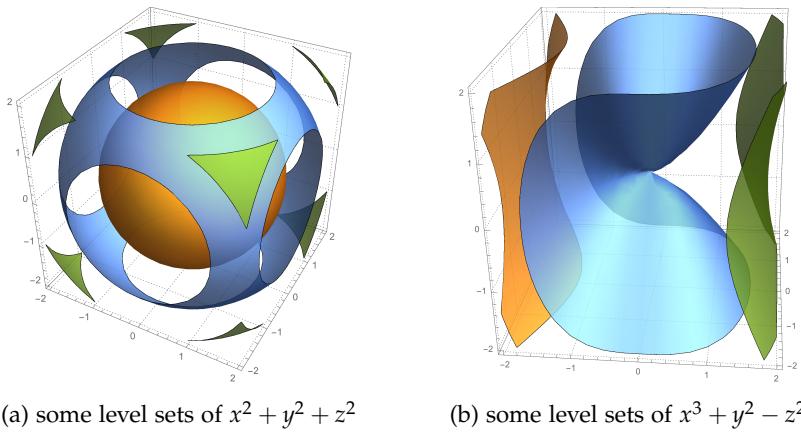


Figure 2.4: level sets of three-variable functions

2.2 Partial Derivatives

2.2.1 First Derivatives

Given a multivariable function such as $f(x, y)$, one can talk about derivatives with respect to both variables x and y . Taking partial derivatives means differentiating $f(x, y)$ with respect to one of the variables while keeping the other variables fixed.

Definition 2.2 — Partial Derivatives. Given a multivariable function $f(x, y)$, we define

$$\frac{\partial f}{\partial x}(x, y) := \text{the derivative of } f(x, y) \text{ with respect to } x \text{ regarding } y \text{ constant}$$

$$= \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h};$$

$$\frac{\partial f}{\partial y}(x, y) := \text{the derivative of } f(x, y) \text{ with respect to } y \text{ regarding } x \text{ constant}$$

$$= \lim_{h \rightarrow 0} \frac{f(x, y + h) - f(x, y)}{h}.$$

Notation Alternatively, we sometimes denote $\frac{\partial f}{\partial x}$ by f_x , and $\frac{\partial f}{\partial y}$ by f_y . Note also that we do not use $f'(x, y)$ for multivariable functions, since it is ambiguous to whether it means f_x or f_y .

Computations of partial derivatives are as easy as single-variable derivatives, as illustrated in the following example:

■ **Example 2.1** Find $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for the function:

$$f(x, y) = x^2 \sin(xy).$$

■ **Solution** To calculate $\frac{\partial f}{\partial x}$, we regard y as a constant:

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial}{\partial x} (x^2 \sin(xy)) && \text{(regarding } y \text{ constant)} \\ &= \frac{\partial x^2}{\partial x} \sin(xy) + x^2 \frac{\partial}{\partial x} \sin(xy) && \text{(product rule)} \\ &= 2x \sin(xy) + x^2 \cdot \cos(xy) \frac{\partial}{\partial x} xy \\ &= 2x \sin(xy) + x^2 y \cos(xy). \end{aligned}$$

Similarly, to calculate $\frac{\partial f}{\partial y}$, regard x as a constant:

$$\begin{aligned} \frac{\partial f}{\partial y} &= \frac{\partial}{\partial y} (x^2 \sin(xy)) \\ &= x^2 \frac{\partial}{\partial y} \sin(xy) && \text{(here } x^2 \text{ is regarded as a constant)} \\ &= x^2 \cos(xy) \cdot \frac{\partial}{\partial y} xy \\ &= x^2 \cos(xy) \cdot x \\ &= x^3 \cos(xy). \end{aligned}$$

- i** Similar to single-variable calculus, to evaluate f_x at a fixed point say $(x, y) = (1, \pi)$, one should perform the differentiation *first*, and *then* substitute $(x, y) = (1, \pi)$ into the derivative. Not the other way round! For example,

$$f_x(1, \pi) = 2x \sin(xy) + x^2 y \cos(xy) \Big|_{(x,y)=(1,\pi)} = 2 \cdot 1 \sin \pi + 1^2 \cdot \pi \cos(\pi) = -\pi.$$

Geometric interpretation of partial derivatives

The geometric meaning of the partial derivative f_x is illustrated in Figure 2.5. By keeping y constant (say we let $y = b$) and let x varies, the path traced on the surface is the curve of intersection between the plane $y = b$ and the graph of the function. This curve is sometimes called an x -curve. The partial derivative $f_x(a, b)$ is the slope of the tangent to this x -curve at (a, b) :

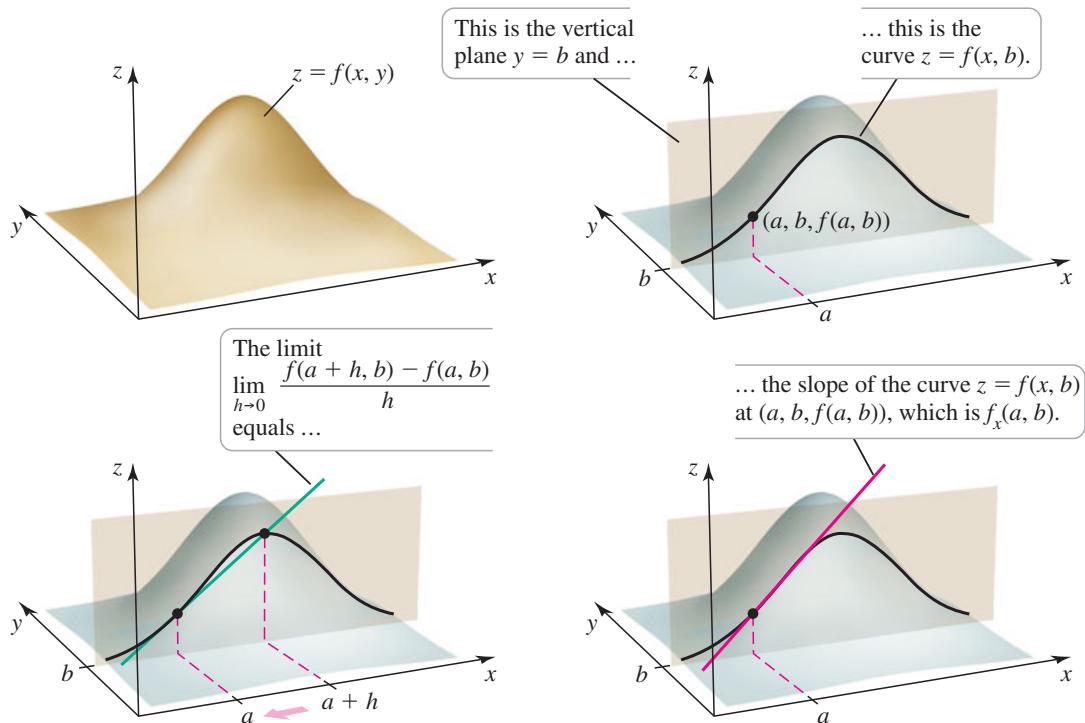


Figure 2.5: geometric interpretation of $\frac{\partial f}{\partial x}$

Partial derivatives of function with more than two variables

Partial derivatives of functions with more than two variables, say $f(x, y, z)$, are defined in an analogous way. For instance, $\frac{\partial f}{\partial x}$ is the derivative with respect to x regarding all other variables, i.e. y and z , constant. Although it is not easy to interpret $\frac{\partial f}{\partial x}$ geometrically since the graph of $f(x, y, z)$ sits inside a 4-dimensional space, the way to compute $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$ are exactly the same as two-variable functions.

■ **Example 2.2** Given $f(x, y, z) = e^{x^2+y^3+xyz}$. Compute $\frac{\partial f}{\partial z}$.

■ **Solution**

$$\begin{aligned}\frac{\partial f}{\partial z} &= \frac{\partial}{\partial z} e^{x^2+y^3+xyz} \\ &= e^{x^2+y^3+xyz} \cdot \frac{\partial(x^2+y^3+xyz)}{\partial z} \quad (\text{chain rule}) \\ &= e^{x^2+y^3+xyz} \cdot (0+0+xy) \\ &= xy e^{x^2+y^3+xyz}.\end{aligned}$$

2.2.2 Second Derivatives

As in single-variable calculus, one can also talk about second derivatives for multivariable functions. Given a two-variable function $f(x, y)$, its first partial derivatives f_x and f_y are also functions of x and y . Therefore, we can further differentiate them with respect to either x or y .

■ **Example 2.3** Let $f(x, y) = 3x^4y - 2xy + 5xy^3$. Compute all first and second partial derivatives.

■ **Solution** It is easy to see that

$$\begin{aligned}\frac{\partial f}{\partial x} &= 12x^3y - 2y + 5y^3 \\ \frac{\partial f}{\partial y} &= 3x^4 - 2x + 15xy^2\end{aligned}$$

Then, the second derivatives are:

$$\begin{aligned}\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) &= \frac{\partial}{\partial x} (12x^3y - 2y + 5y^3) = 36x^2y \\ \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) &= \frac{\partial}{\partial y} (12x^3y - 2y + 5y^3) = 12x^3 - 2 + 15y^2 \\ \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) &= \frac{\partial}{\partial x} (3x^4 - 2x + 15xy^2) = 12x^3 - 2 + 15y^2 \\ \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) &= \frac{\partial}{\partial y} (3x^4 - 2x + 15xy^2) = 30xy.\end{aligned}$$

Notation Since it is a bit clumsy to write $\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right)$ every time, we can use the following short-hand:

$$\begin{array}{ll}\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) & \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \\ \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) & \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right)\end{array}$$

Similarly for the subscript notations, we write $f_{xx} = (f_x)_x$, and $f_{xy} = (f_x)_y$. The latter means to differentiate by x first and then by y . Therefore, it is related to the fraction notation by:

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

The above remark *seems* to suggest that we should be very careful when converting $\frac{\partial^2 f}{\partial y \partial x}$ into the subscript notation f_{xy} . The order of x and y needs to be switched in the conversion.

However, thanks to the following important theorem, we don't need to worry about this too much, since in *many cases*, we have $f_{xy} = f_{yx}$.

Theorem 2.1 — Mixed Partial Theorem. Consider the function $f(x, y)$, if at least one of the second partials f_{xy} and f_{yx} exists and is continuous, then we must have $f_{xy} = f_{yx}$.

Proof. Beyond the scope of this course. ■

Although the theorem requires that f_{xy} or f_{yx} needs to be continuous, most functions we will encounter in this course are continuous and so this theorem applies.

In Example 2.3, you may have already noticed that f_{xy} and f_{yx} are equal. The Mixed Partial Theorem tells us that it is not a coincident!

■ **Example 2.4** Consider the function:

$$f(x, y) = \sqrt{\frac{e^{\sin x}}{x^{2014} + \sqrt{x^{2012} + 1}}} + \cos(xy).$$

Find the second partial derivative $\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right)$.

■ **Solution** Needless to say, it is very tedious and time-consuming to compute $\frac{\partial f}{\partial x}$. However, by the Mixed Partial Theorem, we can try to find $\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right)$, and if it is continuous, then $\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right)$.

It is much easier to compute $\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right)$ since the “monster” term is gone after differentiating the function by y :

$$\begin{aligned} \frac{\partial f}{\partial y} &= 0 - \sin(xy) \cdot \frac{\partial}{\partial y} xy \\ &= -x \sin(xy) \\ \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) &= -\frac{\partial}{\partial x} (x \sin(xy)) \\ &= -\sin(xy) - xy \cos(xy), \end{aligned}$$

which is a continuous function. Therefore, $f_{yx} = f_{xy}$ and so

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = -\sin(xy) - xy \cos(xy).$$

(i) The Mixed Partial Theorem also applies to functions of more than two variables and to higher-order derivatives. Let's say given a function of four variables $f(x, y, z, t)$, we have:

$$f_{xyyyzt} = f_{xyztyy} = f_{yyzxtt} = \dots$$

2.3 Chain Rule

2.3.1 Multivariable Chain Rule

In single-variable calculus, we apply the chain rule when there is a chain of relations between variables. For example, if $f(x)$ is a function of x , and x is in turn a function of t , then f is ultimately a function of t . The derivative $\frac{df}{dt}$ can be calculated by:

$$\frac{df}{dt} = \frac{df}{dx} \frac{dx}{dt}.$$

One may represent this chain of relations by the schematic diagram:

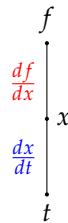


Figure 2.6: the schematic diagram for chain rule formula $\frac{df}{dt} = \frac{df}{dx} \frac{dx}{dt}$.

For multivariable functions, the relation between variables can be more complicated. For example, let the function $u(x, y, z)$ be the temperature at the point (x, y, z) in the space. Suppose a particle moves along the path

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$$

Then, the coordinates x, y and z all depend on t , and so u is ultimately a function of t . See Figure 2.7 for the tree diagram of the variables.

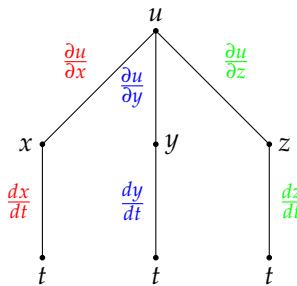


Figure 2.7: tree diagram of variables for $u(x, y, z)$.

The derivative $\frac{du}{dt}$ is the rate of change of the temperature that the particle “feels” as it travels. The multivariable chain rule can be read off from the tree diagram 2.7:

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt} + \frac{\partial u}{\partial z} \frac{dz}{dt}.$$

Precisely, to write down the chain rule for $\frac{du}{dt}$, we find all possible paths from u to t in the tree diagram. Each path consists of some segments. A segment, say, from u to x represents the derivative $\frac{\partial u}{\partial x}$. To write down the chain rule, we “multiply” all segments of each path, and “add” up all the paths.

- (i) In this course, our focus is on how to use the chain rule and how to write it down, but not the proof of the chain rule. The proof is usually taught in advanced real analysis courses.

Let’s see a concrete example of the use of the multivariable chain rule:

■ **Example 2.5** Let $u(x, y, z) = x^2 + y^2 - 2z^2$ and $\langle x, y, z \rangle = \langle \cos t, \sin t, t \rangle$. Compute $\frac{du}{dt}$.

■ **Solution** Although it can be done by substituting $x = \cos t$, $y = \sin t$ and $z = t$ into $u(x, y, z) = x^2 + y^2 - 2z^2$ and then compute $\frac{du}{dt}$ directly, let's try to use the chain rule to do it. From the tree diagram (Figure 2.7), we have:

$$\begin{aligned}\frac{du}{dt} &= \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt} + \frac{\partial u}{\partial z} \frac{dz}{dt} \\ &= \underbrace{2x}_{\frac{\partial u}{\partial x}} \cdot \underbrace{(-\sin t)}_{\frac{dx}{dt}} + \underbrace{2y}_{\frac{\partial u}{\partial y}} \cdot \underbrace{\cos t}_{\frac{dy}{dt}} + \underbrace{(-2z)}_{\frac{\partial u}{\partial z}} \cdot \underbrace{1}_{\frac{dz}{dt}} \quad (\text{calculate each derivative}) \\ &= -2 \cos t \sin t + 2 \sin t \cos t - 2t \quad (\text{write } x, y \text{ and } z \text{ in terms of } t) \\ &= -2t.\end{aligned}$$

As illustrated in the above example, once the chain rule formula is correctly written according to the tree diagram, the remaining computations are straight-forward. From now on, we will investigate how to write down the chain rule under various configuration of variables. The computations will usually be skipped and are left as exercises for readers.

Examples with more diverse variable configurations

Suppose now that the temperature distribution is changing over time as well, i.e. the temperature $u(x, y, z, t)$ is a function of four variables. Again, a particle travels along a path $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$. Then, the tree diagram of variables in this case can be drawn as in Figure 2.8.

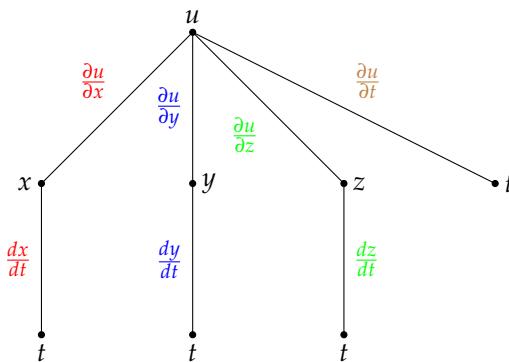


Figure 2.8: tree diagram of variables for $u(x, y, z, t)$.

There are four paths from u to t , so we expect the chain rule formula for $\frac{du}{dt}$ consists of four terms:

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt} + \frac{\partial u}{\partial z} \frac{dz}{dt} + \frac{\partial u}{\partial t}.$$

- (i) Note that $\frac{du}{dt}$ is different from $\frac{\partial u}{\partial t}$. As the particle travels along its path, the temperature the particle "feels" is a ultimately a function of t , and so the rate of change of temperature is represented by $\frac{du}{dt}$ (using d instead of the partial ∂).

On the other hand, the partial $\frac{\partial u}{\partial t}$ is the time derivative of u regarding x, y and z constant! Therefore, it is the rate of change of temperature when the position is **fixed!** It is not the rate of change of temperature for the **moving** particle!

Now consider a slightly more complicated example. Let w be a function of x, y and z , and each of x, y and z is a function of s and t , as illustrated in Figure 2.9. The function w is ultimately a function of s and t .

There are three paths from w to s , so the chain rule for $\frac{\partial w}{\partial s}$ is given by the following:

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial s} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial s}.$$

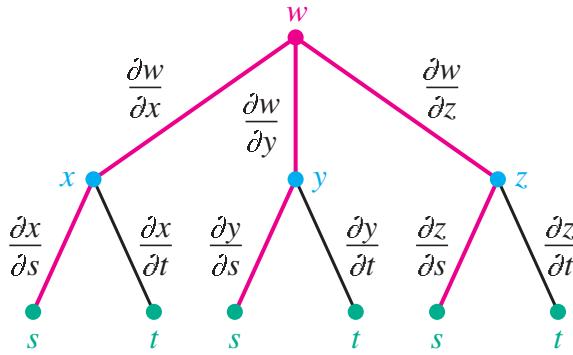


Figure 2.9: tree diagram

A “multi-level” example

Now suppose w is a function of z only, z is a function of x and y , and both x and y are functions of t , as illustrated in Figure 2.10.

Ultimately, w is a function t . There are two paths from w to t in the tree diagram. Each path consists of three segments. The chain rule for $\frac{dw}{dt}$ is given by:

$$\frac{dw}{dt} = \frac{dw}{dz} \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{dw}{dz} \frac{\partial z}{\partial y} \frac{dy}{dt}.$$

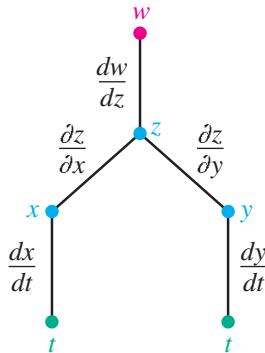


Figure 2.10: tree diagram

2.3.2 Implicit Differentiation

Given an implicit equation such as:

$$x^2 + y^3 + \sin^2 y = 1,$$

it is very difficult (possibly impossible) to express y in terms of x . In single-variable calculus, we learned how to find $\frac{dy}{dx}$ using implicit differentiation – regard y as a function of x , and differentiate both sides by x then solve for $\frac{dy}{dx}$.

The multivariable chain rule offers an alternative approach to implicit differentiation.

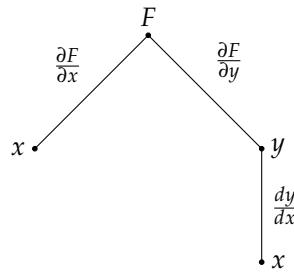


Figure 2.11: tree diagram for implicit differentiation

Define $F(x, y) = x^2 + y^3 + \sin^2 y$, then the above implicit equation can be written as $F(x, y) = 1$. Regarding y as a function of x , $F(x, y)$ is ultimately a function of x . Figure 2.11 shows the tree diagram for the variables.

Therefore, by the chain rule, we have

$$\frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{dy}{dx}.$$

Recall that $F(x, y) = 1$ is a constant, so $\frac{dF}{dx} = 0$, which yields:

$$0 = F_x + F_y \frac{dy}{dx}, \quad \text{and so:} \quad \frac{dy}{dx} = -\frac{F_x}{F_y}.$$

It is a much straight-forward formula than the implicit differentiation method learned in single-variable calculus. When $F(x, y) = x^2 + y^3 + \sin^2 y$, we have:

$$\frac{dy}{dx} = -\frac{2x}{3y^2 + 2\sin y \cos y}.$$

2.4 Directional Derivatives

Recall that the physical meaning of the partial derivative $\frac{\partial f}{\partial x}$ is the rate of change of f in the direction of x . The geometric interpretation can be found in Figure 2.5. Similarly, $\frac{\partial f}{\partial y}$ is the rate of change of f in the direction of y .

In this section, we introduce the way to find the rate of change of f in any other directions, i.e. directional derivative defined as follows.

Definition 2.3 — Directional Derivative. Given a **unit** direction $\mathbf{u} = u_1 \mathbf{i} + u_2 \mathbf{j}$ and a two-variable function $f(x, y)$, the directional derivative of f in the direction of \mathbf{u} at point (x, y) is denoted and defined to be:

$$D_{\mathbf{u}}f(x, y) = \left. \frac{d}{dt} f(x + tu_1, y + tu_2) \right|_{t=0}.$$

(i) When $\mathbf{u} = \mathbf{i}$, then $u_1 = 1$ and $u_2 = 0$ and so

$$D_{\mathbf{i}}f(x, y) = \left. \frac{d}{dt} f(x + t, y) \right|_{t=0} = \lim_{t \rightarrow 0} \frac{f(x + t, y) - f(x, y)}{t} = \frac{\partial f}{\partial x}(x, y).$$

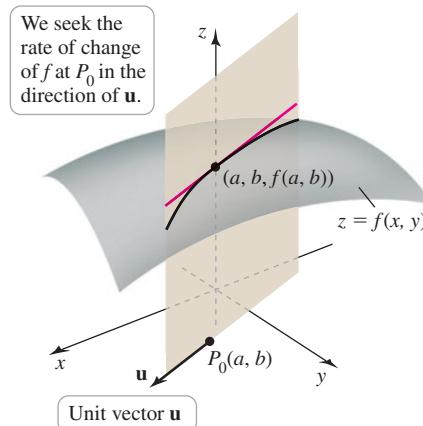


Figure 2.12: directional derivative

In practice, we do not need to compute $D_{\mathbf{u}}f$ from the definition, since we have Theorem 2.2 below to help us. In order to introduce this theorem, we first define:

Definition 2.4 — Gradient Vector. Given a two-variable function $f(x, y)$, the gradient vector of f at (x, y) is denoted and defined as:

$$\nabla f(x, y) = \frac{\partial f}{\partial x}(x, y) \mathbf{i} + \frac{\partial f}{\partial y}(x, y) \mathbf{j}.$$

As an example, let $f(x, y) = x^2y + x^3$, then

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 2xy + 3x^2 \\ \frac{\partial f}{\partial y}(x, y) &= x^2 \end{aligned}$$

Therefore,

$$\nabla f(x, y) = (2xy + 3x^2) \mathbf{i} + x^2 \mathbf{j}.$$

The vector $\nabla f(x, y)$ depends on (x, y) . By taking different values of (x, y) , a different gradient

vector is produced. For instance,

$$\begin{aligned}\nabla f(1, 1) &= 5\mathbf{i} + \mathbf{j}, \\ \nabla f(1, 0) &= 3\mathbf{i} + \mathbf{j}.\end{aligned}$$

Theorem 2.2 Given a two-variable function $f(x, y)$, the directional derivative of f at (x, y) in the **unit direction** $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ is given by:

$$D_{\mathbf{u}}f(x, y) = \nabla f(x, y) \cdot \mathbf{u}.$$

By applying this theorem in the special case $\mathbf{u} = \mathbf{i}$, we can see

$$\nabla f(x, y) \cdot \mathbf{i} = \left(\frac{\partial f}{\partial x}(x, y)\mathbf{i} + \frac{\partial f}{\partial y}(x, y)\mathbf{j} \right) \cdot \mathbf{i} = \frac{\partial f}{\partial x}(x, y) = D_{\mathbf{i}}f(x, y)$$

as expected. Similarly, we can see $\nabla f(x, y) \cdot \mathbf{j} = \frac{\partial f}{\partial y}(x, y) = D_{\mathbf{j}}f(x, y)$ as expected. Let's see the proof of the general case:

Proof of Theorem 2.2. The key idea is to use the chain rule. The directional derivative $D_{\mathbf{u}}f(x_0, y_0)$ at the point (x_0, y_0) is the rate of change of f along the path $\mathbf{r}(t) = \langle x_0, y_0 \rangle + t\langle u_1, u_2 \rangle$, i.e. $x = x_0 + u_1 t$ and $y = y_0 + u_2 t$. By definition of directional derivative, $D_{\mathbf{u}}f(x_0, y_0)$ is the derivative $\frac{d}{dt}f(x_0 + u_1 t, y_0 + u_2 t)$ at $t = 0$. Therefore, f is a function of (x, y) , and (x, y) are functions of t . By the chain rule, we have:

$$\begin{aligned}D_{\mathbf{u}}f &= \frac{df}{dt} \\ &= \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} \\ &= \frac{\partial f}{\partial x} \frac{d}{dt}(x_0 + tu_1) + \frac{\partial f}{\partial y} \frac{d}{dt}(y_0 + tu_2) \\ &= \frac{\partial f}{\partial x} \cdot u_1 + \frac{\partial f}{\partial y} \cdot u_2.\end{aligned}$$

On the other hand,

$$\begin{aligned}\langle \nabla f, \mathbf{u} \rangle &= \left\langle \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}, u_1\mathbf{i} + u_2\mathbf{j} \right\rangle \\ &= \frac{\partial f}{\partial x} \cdot u_1 + \frac{\partial f}{\partial y} \cdot u_2.\end{aligned}$$

Therefore, $D_{\mathbf{u}}f = \langle \nabla f, \mathbf{u} \rangle$. ■

This theorem tells us that the computation of directional derivatives amounts to computing the gradient vector and a dot product. Easy enough?

As an example, given $f(x, y) = x^2y + x^3$. We worked out in the previous example that $\nabla f(1, 1) = 5\mathbf{i} + \mathbf{j}$, and so the directional derivative of f at $(1, 1)$ along $\mathbf{u} = \frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j}$ is given by:

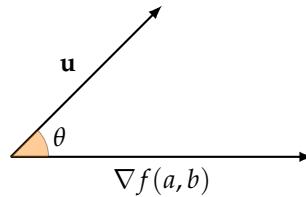
$$D_{\frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j}}f(1, 1) = \nabla f(1, 1) \cdot \left(\frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j} \right) = (5\mathbf{i} + \mathbf{j}) \cdot \left(\frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j} \right) = \frac{6}{\sqrt{2}}.$$

2.4.1 Geometric Interpretation of Gradient Vectors

Theorem 2.2 not only tells us how to compute the directional derivative $D_{\mathbf{u}}f(x, y)$, but also tells us in what direction \mathbf{u} the derivative $D_{\mathbf{u}}f(x, y)$ is the greatest and the smallest.

Given a function $f(x, y)$, fix a point (a, b) . From the dot product formula 1.1, we know:

$$D_{\mathbf{u}}f(a, b) = \nabla f(a, b) \cdot \mathbf{u} = |\nabla f(a, b)| \underbrace{|\mathbf{u}|}_{=1} \cos \theta = |\nabla f(a, b)| \cos \theta$$

Figure 2.13: geometric interpretation of $D_{\mathbf{u}}f$

where θ is the angle between $\nabla f(a, b)$ and \mathbf{u} .

Since $\cos \theta$ is the largest when $\theta = 0$ at which $\cos \theta = 1$, the directional derivative $D_{\mathbf{u}}f(a, b)$ is **maximized** when $\nabla f(a, b)$ is parallel to \mathbf{u} . Therefore, $\nabla f(a, b)$ is pointing in the direction at which f increases most rapidly from (a, b) .

It is quite intuitive that in order to increase the value of f most rapidly, one should go along the direction **perpendicular to the level curve**. It is indeed true. Let's state it as a theorem:

Theorem 2.3 Let $f(x, y)$ be a two-variable function, and (a, b) be a point on the level curve $f(x, y) = c$. Then the gradient vector $\nabla f(a, b)$ is orthogonal to the level curve $f(x, y) = c$ at the point (a, b) .

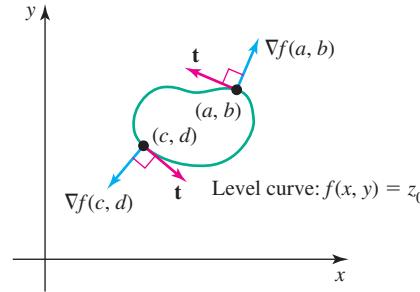
Proof. Let $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ be a parametrization of the level curve $f(x, y) = c$. In other words, we have $f(x(t), y(t)) = c$ for any t , and so

$$\frac{d}{dt}f(x(t), y(t)) = 0.$$

By the chain rule, we get:

$$\begin{aligned} \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} &= 0 \\ \left(\frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} \right) \cdot \left(\frac{dx}{dt} \mathbf{i} + \frac{dy}{dt} \mathbf{j} \right) &= 0 \\ \nabla f \cdot \mathbf{r}'(t) &= 0. \end{aligned}$$

Therefore, the gradient vector ∇f is orthogonal to $\mathbf{r}'(t)$ which is the tangent vector of the level curve $\mathbf{r}(t)$. It completes the proof. ■

Figure 2.14: $\nabla f(a, b)$ is orthogonal to the level curve of f at (a, b)

2.4.2 Directional Derivative of Three-Variable Functions

The directional derivative and the gradient vector for functions $f(x, y, z)$ of three variables is defined in an analogous way as for two-variable functions. Precisely, we have:

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}$$

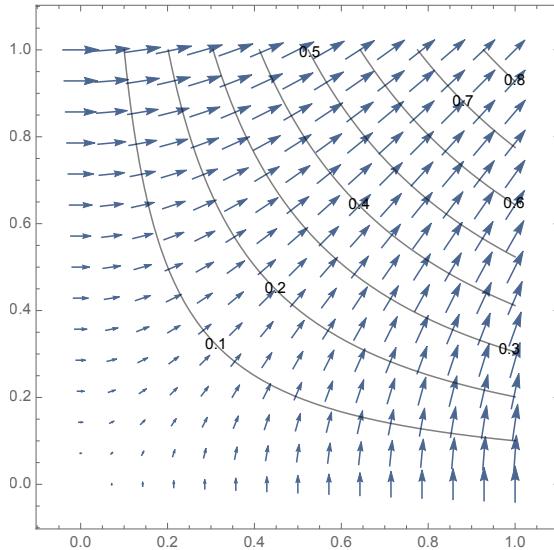


Figure 2.15: a plot of the field $\nabla (\sin xy)$ and the level curves of $\sin xy$.

Given a unit vector \mathbf{u} , the directional derivative of $f(x, y, z)$ in the direction of \mathbf{u} is given by:

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}.$$

Since the level set of a three-variable function is typically a surface, the gradient vector ∇f at any given point is orthogonal to the level surface $f(x, y, z) = c$ at that point. See Figure 2.16.

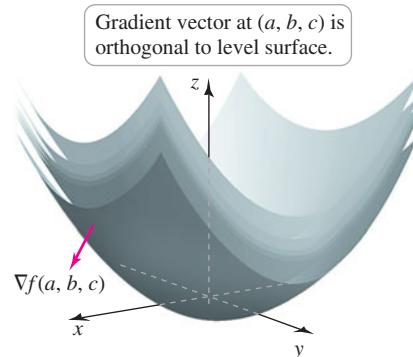


Figure 2.16: $\nabla f(a, b, c)$ is orthogonal to the level surface of f at (a, b, c)

In the next section, we will use this fact to find the equation of the tangent plane to a surface.

2.5 Tangent Planes

In single-variable calculus, the tangent line to the curve $y = f(x)$ at a point $(x_0, f(x_0))$ has slope equal to $f'(x_0)$. Using the slope, one can easily write down the equation of the tangent line as:

$$y = f(x_0) + f'(x_0)(x - x_0).$$

In multivariable calculus, the graph of a two-variable function $z = f(x, y)$ is no longer a curve but a surface. Therefore, there are infinitely many tangent lines passing through any given point on a surface. However, there is only *one* tangent plane at any given point. In this section, we want to find an equation for the tangent plane.

Recall that the two ingredients of finding an equation of a plane are:

- a normal vector to the plane; and
- any given point on the plane.

Naturally, the point can be taken to be the contact point of the surface (x_0, y_0, z_0) . To find the normal vector, we will use the **gradient vector**.

In the previous section, we see that given a three-variable function $g(x, y, z)$, the gradient vector $\nabla g(x, y, z)$ is perpendicular to the level surface $g(x, y, z) = c$. In other words, ∇g is a **normal vector** of the level surface $g(x, y, z) = c$.

■ **Example 2.6** Find the tangent plane to the surface

$$x^2 + y^2 = z^2 + 3$$

at the point $(x, y, z) = (2, 0, -1)$.

■ **Solution** First we need to write the equation of the surface in a level set form, i.e.

$$x^2 + y^2 - z^2 = 3$$

Define $g(x, y, z) = x^2 + y^2 - z^2$, then the given surface is a level set $g(x, y, z) = 3$. By direct computations,

$$\begin{aligned}\nabla g(x, y, z) &= 2x\mathbf{i} + 2y\mathbf{j} + (-2z)\mathbf{k} \\ \nabla g(2, 0, -1) &= 4\mathbf{i} + 2\mathbf{k}.\end{aligned}$$

Then, $\mathbf{n} := 4\mathbf{i} + 2\mathbf{k}$ is a normal vector of the surface at $(2, 0, -1)$. The equation of the tangent plane at $(2, 0, -1)$ is given by:

$$\begin{aligned}4x + 0y + 2z &= 4(2) + 0(0) + 2(-1), \\ 4x + 2z &= 6.\end{aligned}$$

Given a two-variable function $f(x, y)$, the graph $z = f(x, y)$ is a surface. In order to find the tangent plane at a given point, one can rewrite the graph equation $z = f(x, y)$ as:

$$z - f(x, y) = 0.$$

Then, one can define $g(x, y, z) = z - f(x, y)$ so that the **graph** of the two-variable function $f(x, y)$ becomes a **level set** of a three-variable function $g(x, y, z)$. Let's look at an example:

■ **Example 2.7** Given the function $f(x, y) = x \cos y - ye^x$, find the tangent plane at $(0, 0, 0)$ to the graph $z = x \cos y - ye^x$.

■ **Solution** First rearrange the terms so that the graph equation becomes a level set:

$$z - x \cos y + ye^x = 0.$$

Define $g(x, y, z) = z - x \cos y + ye^x$, then the surface under consideration is the level set $g(x, y, z) = 0$.

$$\nabla g(x, y, z) = (-\cos y + ye^x)\mathbf{i} + (x \sin y + e^x)\mathbf{j} + \mathbf{k}.$$

The normal vector at $(0, 0, 0)$ is given by:

$$\mathbf{n} = \nabla g(0, 0, 0) = -\mathbf{i} + \mathbf{j} + \mathbf{k}.$$

The equation of the tangent plane at $(0, 0, 0)$ to the graph is:

$$-x + y + z = 0.$$

Generally, one can derive a formula for finding the tangent plane of any graph of a two-variable function $f(x, y)$:

Theorem 2.4 The equation of the tangent plane for the graph $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$ is given by:

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0).$$

Proof. Write the graph equation $z = f(x, y)$ as:

$$z - f(x, y) = 0$$

and define $g(x, y, z) = z - f(x, y)$, then the graph can be regarded as a level set $g(x, y, z) = 0$ of the three-variable function g .

$$\nabla g(x, y, z) = -\frac{\partial f}{\partial x}\mathbf{i} - \frac{\partial f}{\partial y}\mathbf{j} + \mathbf{k}.$$

At the point $(x_0, y_0, f(x_0, y_0))$, the normal vector to the surface is therefore given by:

$$\mathbf{n} = \nabla g(x_0, y_0, f(x_0, y_0)) = -\frac{\partial f}{\partial x}(x_0, y_0)\mathbf{i} - \frac{\partial f}{\partial y}(x_0, y_0)\mathbf{j} + \mathbf{k}.$$

The equation of the tangent plane at $(x_0, y_0, f(x_0, y_0))$ is:

$$\left(-\frac{\partial f}{\partial x}(x_0, y_0)\right)x + \left(-\frac{\partial f}{\partial y}(x_0, y_0)\right)y + z = \left(-\frac{\partial f}{\partial x}(x_0, y_0)\right)x_0 + \left(-\frac{\partial f}{\partial y}(x_0, y_0)\right)y_0 + f(x_0, y_0)$$

By rearrangement, we get:

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0),$$

as desired. ■

2.6 Local Extrema

In single-variable calculus, the role that tangent lines play in optimization problems is that having a horizontal tangent line at a point indicates the point is critical. This critical point is a *candidate* for local maximum or minimum. The second derivative may be used to determine whether the critical point is a local maximum or a local minimum.

In this section, we will extend the concept of critical points to two-variable functions, and introduce the second derivative test for these functions.

2.6.1 Critical Points

We focus on the case where both $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist so that we don't need to worry whether the function is differentiable or not. Recall that the equation of the tangent plane to the graph $z = f(x, y)$ at a point (a, b) is given by:

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

This plane is horizontal if and only if $f_x(a, b) = f_y(a, b) = 0$. This motivates the definition of critical points for two-variable functions:

Definition 2.5 — Critical Points. Given a function $f(x, y)$ whose partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist everywhere. A point (a, b) is said to be a critical point if the tangent plane at (a, b) to the graph $z = f(x, y)$ is **horizontal**.

Therefore, (a, b) is a critical point $\Leftrightarrow \frac{\partial f}{\partial x}(a, b) = \frac{\partial f}{\partial y}(a, b) = 0 \Leftrightarrow \nabla f(a, b) = \mathbf{0}$.

 As in single-variable calculus, the critical points are just *candidates* of maximum/minimum.

■ **Example 2.8** Find all critical point(s) of the function $f(x, y) = xy - x^2 - y^2 - 2x - 2y + 4$.

■ **Solution** We compute $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, and then set them to zero and solve for (x, y) :

$$\begin{aligned} 0 &= \frac{\partial f}{\partial x} = y - 2x - 2 \\ 0 &= \frac{\partial f}{\partial y} = x - 2y - 2. \end{aligned}$$

It is a system of equations with unknowns x and y . The first equation gives $y = 2x + 2$, and substitute it into the second equation, we get:

$$x - 2(2x + 2) - 2 = 0 \Rightarrow x = -2.$$

When $x = -2$, we have $y = -2$, and so $(x, y) = (-2, -2)$ is a critical point of $f(x, y)$. See Figure 2.17a for its graph.

■ **Example 2.9** Find all critical point(s) of the function $f(x, y) = \sin x \sin y$

■ **Solution** Consider:

$$\begin{aligned} \frac{\partial f}{\partial x} = 0 &\quad \text{and} \quad \frac{\partial f}{\partial y} = 0 \\ \cos x \sin y = 0 &\quad \text{and} \quad \sin x \cos y = 0 \\ (\cos x = 0 \text{ or } \sin y = 0) &\quad \text{and} \quad (\sin x = 0 \text{ or } \cos y = 0) \\ (x = \frac{\pi}{2} + k\pi \text{ or } y = m\pi) &\quad \text{and} \quad (x = n\pi \text{ or } y = \frac{\pi}{2} + p\pi). \end{aligned}$$

Here m, n, k, p are any integers. Some logical deductions show that these imply the following:

$$\begin{aligned} x = \frac{\pi}{2} + k\pi &\quad \text{and} \quad y = \frac{\pi}{2} + p\pi \\ \text{or: } y = m\pi &\quad \text{and} \quad x = n\pi. \end{aligned}$$

Therefore, there are infinitely many critical points:

$$\left(\frac{\pi}{2} + k\pi, \frac{\pi}{2} + p\pi \right), \quad (m\pi, n\pi)$$

where m, n, k, p are any integers. See Figure 2.17b for its graph.

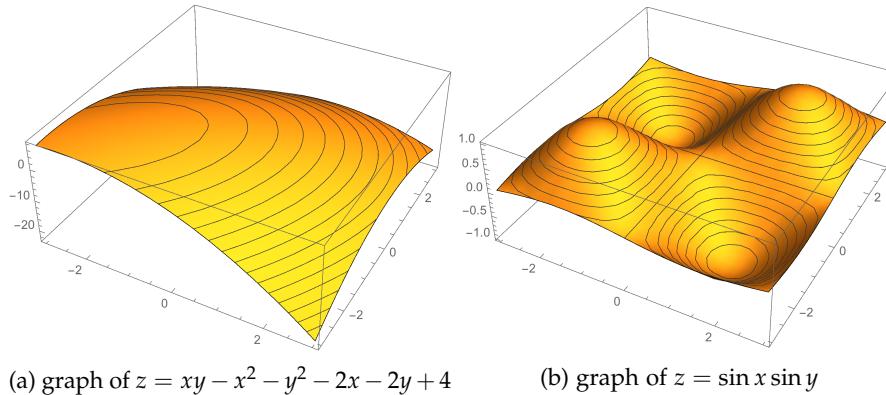


Figure 2.17: graphs of functions in Examples 2.8 and 2.9

2.6.2 Second Derivative Test of Multivariable Functions

In single-variable calculus, to determine the nature of a critical point x_0 of a function $f(x)$, we look at its second derivative at x_0 . If $f''(x_0) > 0$, the graph $y = f(x)$ is concave up around x_0 and so $(x_0, f(x_0))$ is a local minimum point. On the other hand, if $f''(x_0) < 0$, the graph is concave down near x_0 and so $(x_0, f(x_0))$ is a local maximum point.

In multivariable calculus, however, to determine whether a critical point (x_0, y_0) of a two-variable function $f(x, y)$ is not as simple as in single-variable calculus. Take the following function as an example:

$$f(x, y) = x^2 + 4xy + y^2.$$

One can easily verify that $\nabla f(0, 0) = \mathbf{0}$ and so $(0, 0)$ is a critical point. For the second derivatives, we find that:

$$\begin{aligned} f_{xx}(0, 0) &= 2 \\ f_{yy}(0, 0) &= 2 \end{aligned}$$

for every (x, y) on the \mathbb{R}^2 plane. Both are positive numbers. You may be tempted to conclude that $(0, 0)$ is a local maximum point. However, if one plots the graph of this function (see Figure 2.18), one can see easily that $(0, 0)$ is neither a local maximum or a local minimum.

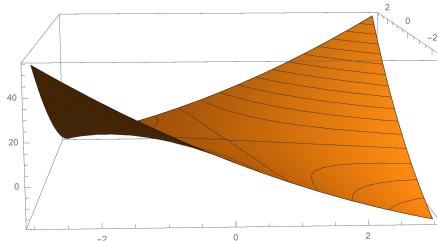


Figure 2.18: $(0, 0)$ is neither a maximum or minimum

Around $(0, 0)$, the graph is a concave up in some directions but concave down in other directions. We call this $(0, 0)$ a **saddle**.

This example shows the signs of f_{xx} and f_{yy} alone could not conclude the nature of the critical point. In fact, the second derivative test for two-variable functions is slightly more complicated than that in single-variable calculus:

Theorem 2.5 — Second Derivative Test for Two-Variable Functions. Let $f(x, y)$ be a twice differentiable function and (x_0, y_0) is a critical point of f , i.e. $\nabla f(x_0, y_0) = \mathbf{0}$. Then the nature of this critical point (x_0, y_0) is determined by the following table:

| $(f_{xx}f_{yy} - f_{xy}^2) \Big _{(x_0, y_0)}$ | $f_{xx}(x_0, y_0)$ | (x_0, y_0) is a: |
|--|--------------------|--------------------|
| > 0 | > 0 | local minimum |
| > 0 | < 0 | local maximum |
| < 0 | anything | saddle |

Any other cases are inconclusive.

For the function $f(x, y) = x^2 + 4xy + y^2$ in the above example, to determine the nature of $(0, 0)$ we also need $f_{xy}(0, 0)$, which can be found as equal to 4.

Therefore, we have:

$$\begin{aligned} (f_{xx}f_{yy} - f_{xy}^2) \Big|_{(0,0)} &= 2 \times 2 - 4^2 < 0, \\ f_{xx}(0,0) &= 2 > 0. \end{aligned}$$

From the table in Theorem 2.5, we conclude $(0, 0)$ is a saddle, as expected from the plot of the its graph.

Let's look at one more example before we learn the proof of the Second Derivative Test.

■ **Example 2.10** Let $f(x, y) = 3y^2 - 2y^3 - 3x^2 + 6xy$. Find all critical points and determine the nature of each of them.

■ **Solution** To find all critical points, we set:

$$\begin{aligned} \frac{\partial f}{\partial x} &= -6x + 6y = 0, \\ \frac{\partial f}{\partial y} &= 6y - 6y^2 + 6x = 0. \end{aligned}$$

From the first equation, we get $y = x$. Substitute this into the second equation, we yield:

$$6x - 6x^2 + 6x = 0, \text{ or equivalently } 2x - x^2 = 0.$$

By factorization, we get $x(2-x) = 0$. Therefore

$$x = 0 \text{ or } x = 2.$$

By noting that $y = x$, we have two critical points: $(0, 0)$ and $(2, 2)$.

Next we compute the second derivatives of f :

$$\begin{array}{ll} f_{xx} = -6 & f_{xy} = 6 \\ f_{yx} = 6 & f_{yy} = 6 - 12y \end{array}$$

| Critical point P | $f_{xx}(P)$ | $f_{yy}(P)$ | $f_{xy}(P)$ | $\left(f_{xx}f_{yy} - f_{xy}^2\right)(P)$ | Nature of P |
|--------------------|-------------|-------------|-------------|---|---------------|
| $(0, 0)$ | -6 | 6 | 6 | -72 | saddle |
| $(2, 2)$ | -6 | -18 | 6 | 72 | local maximum |

Explanation of the Second Derivative Test

In single-variable, the second derivative test can be explained using convexity of the graph $y = f(x)$. However, this approach can hardly be generalized to higher dimensions.

Before we explain why the above second derivative test works for two-variable functions $f(x, y)$, we first seek an alternative explanation of the single-variable second derivative test using Taylor's series.

Recall that the Taylor's series of a given function $f(x)$ about $x = a$ is given by:

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \frac{f'''(a)}{3!}(x - a)^3 + \dots$$

If $f(x)$ has a critical point at $x = a$, then $f'(a) = 0$. Also, when x is very close to a , the higher-order terms $(x - a)^3, (x - a)^4$, etc. are significantly smaller than the quadratic term $(x - a)^2$. Therefore, the function $f(x)$ is approximately given by:

$$f(x) \simeq f(a) + \frac{f''(a)}{2!}(x - a)^2 \quad \text{when } x \text{ is near } a.$$

The right-hand side $f(a) + \frac{f''(a)}{2!}(x - a)^2$ is a quadratic function. If $f''(a) > 0$, then the graph $y = f(a) + \frac{f''(a)}{2!}(x - a)^2$ is a concave up parabola and so $f(a) + \frac{f''(a)}{2!}(x - a)^2 \geq f(a)$. Therefore, $f(x)$, which is approximately $f(a) + \frac{f''(a)}{2!}(x - a)^2$, is also $\geq f(a)$ when x is near a . This explains $f(x)$ has a local minimum at $x = a$.

On the other hand, if $f''(a) < 0$, then the graph $y = f(a) + \frac{f''(a)}{2!}(x - a)^2$ is a concave down parabola. Similar argument as above shows $f(x)$ has a local maximum at $x = a$.

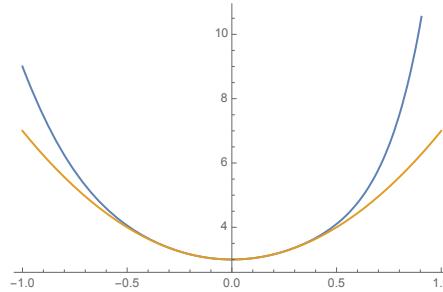


Figure 2.19: blue graph shows $y = f(x)$ where $f'(0) = 0$; yellow graph shows $y = f(0) + \frac{f''(0)}{2!}x^2$ where $f''(0) > 0$

Back to multivariable calculus, we now explain the second derivative test using the Taylor's series approach. Given a function $f(x, y)$, the multivariable Taylor's series about $(x, y) =$

(x_0, y_0) is given by:

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &\quad + \frac{f_{xx}(x_0, y_0)}{2!}(x - x_0)^2 + \frac{2f_{xy}(x_0, y_0)}{2!}(x - x_0)(y - y_0) + \frac{f_{yy}(x_0, y_0)}{2!}(y - y_0)^2 \\ &\quad + \text{higher-order terms} \end{aligned}$$

The proof is beyond the scope of the course. If (x_0, y_0) is a critical point of $f(x, y)$, then $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$. For simplicity, denote $P = (x_0, y_0)$, then when (x, y) is near P , we have:

$$f(x, y) \simeq f(x_0, y_0) + \frac{1}{2} \left(f_{xx}(P)(x - x_0)^2 + 2f_{xy}(P)(x - x_0)(y - y_0) + f_{yy}(P)(y - y_0)^2 \right).$$

Therefore, to determine whether (x_0, y_0) is a local maximum/minimum or a saddle of $f(x, y)$, one should determine whether the quadratic

$$f_{xx}(P)(x - x_0)^2 + 2f_{xy}(P)(x - x_0)(y - y_0) + f_{yy}(P)(y - y_0)^2$$

is positive/negative or neither.

For simplicity, denote

$$A = f_{xx}(P), \quad B = f_{xy}(P), \quad C = f_{yy}(P),$$

$$X = x - x_0, \quad Y = y - y_0.$$

Then, the quadratic expression can be simplified as:

$$f_{xx}(P)(x - x_0)^2 + 2f_{xy}(P)(x - x_0)(y - y_0) + f_{yy}(P)(y - y_0)^2 = AX^2 + 2BXY + CY^2.$$

To determine whether $AX^2 + 2BXY + CY^2$ is always positive/negative or neither, one looks the discriminant $\Delta = (2B)^2 - 4AC = 4(B^2 - AC)$:

| $AC - B^2$ | A | $AX^2 + 2BXY + CY^2$ | near P , $f(x, y)$ is |
|------------|----------|----------------------|--|
| > 0 | > 0 | ≥ 0 | $\geq f(P)$ |
| > 0 | < 0 | ≤ 0 | $\leq f(P)$ |
| < 0 | anything | can be either + or - | can be either $\geq f(P)$ or $\leq f(P)$ |

Translate back to previous notations, we can conclude:

| $(f_{xx}f_{yy} - f_{xy}^2) _{(x_0, y_0)}$ | $f_{xx}(x_0, y_0)$ | (x_0, y_0) is a: |
|---|--------------------|--------------------|
| > 0 | > 0 | local minimum |
| > 0 | < 0 | local maximum |
| < 0 | anything | saddle |

This explains the second derivative test for two-variable functions!

2.7 Lagrange's Multiplier

In the previous section, we learned how to find critical points in the **interior** of a domain, namely by solving $\nabla f = \mathbf{0}$. These critical points are candidates of the maximum or minimum of the function. We also learn how to determine the local nature of the critical points.

However, to determine the maximum/minimum on the boundary of a domain, the gradient method does not work as the tangent plane at the maximum/minimum needs not be horizontal. When the domain of a function $f(x, y)$ is restricted on a level set such as $x^2 + y^2 = 1$, which is a unit circle, we use a method called the **Lagrange's Multiplier**.

We first state the method, then look at a few examples, and finally explain why it works.

Given a function $f(x, y)$ which we want to maximize or minimize, and the variables (x, y) are restricted by the constraint $g(x, y) = c$. Then, to determine all possible candidates of maximum/minimum point on the constraint, we:

1. Solve the system of equations

$$\begin{aligned}\nabla f(x, y) &= \lambda \nabla g(x, y) \\ g(x, y) &= c\end{aligned}$$

Here the unknowns are x, y and λ .

2. The solutions of (x, y) are the possible candidates of the maximum or minimum points of the function $f(x, y)$. We call these **boundary critical points**.
3. Finally, evaluate $f(x, y)$ at each boundary critical point found. The point giving the largest value of $f(x, y)$ is the maximum point on the boundary, and that giving the smallest value of $f(x, y)$ is the minimum.

I We call this the Lagrange's Multiplier method because the scalar λ is called the Lagrange's Multiplier.

Example 2.11 Let $f(x, y) = x^2 + y^2 + 2x + 2y$, find the maximum and minimum values of f when (x, y) is restricted on the constraint $x^2 + y^2 = 1$.

Solution $f(x, y)$ is the function we want to maximize and minimize. Let $g(x, y) = x^2 + y^2$ so that the level set $g(x, y) = 1$ is our constraint. First we compute:

$$\begin{aligned}\nabla f &= \langle 2x + 2, 2y + 2 \rangle \\ \nabla g &= \langle 2x, 2y \rangle.\end{aligned}$$

Therefore, the vector equation $\nabla f(x, y) = \lambda \nabla g(x, y)$ is equivalent to the two equations $2x + 2 = 2\lambda x$ and $2y + 2 = 2\lambda y$. Combining with the constraint equation $x^2 + y^2 = 1$, we get a system of three equations with three unknowns x, y and λ :

$$2x + 2 = 2\lambda x \quad (1)$$

$$2y + 2 = 2\lambda y \quad (2)$$

$$x^2 + y^2 = 1 \quad (3)$$

Since we are interested in solving for (x, y) and finding λ is optional, we divide (1) by (2) so that the λ can be canceled. However, we may worry that whether (2) is zero! Therefore, we split into two cases.

Case 1: $2y + 2 \neq 0$ (and so $2\lambda y \neq 0$ too)

(1) \div (2) gives:

$$\frac{2x + 2}{2y + 2} = \frac{2\lambda x}{2\lambda y}.$$

After cancellations, we get:

$$\frac{x+1}{y+1} = \frac{x}{y}.$$

By cross multiplication:

$$y(x+1) = x(y+1) \Rightarrow xy + y = xy + x \Rightarrow y = x.$$

Substitute $y = x$ into ③, we have $2x^2 = 1$, and so $x = \frac{1}{\sqrt{2}}$ or $-\frac{1}{\sqrt{2}}$. Since $y = x$, the solutions for (x, y) in this case are:

$$(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \quad \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right).$$

Case 2: $2y + 2 = 0$

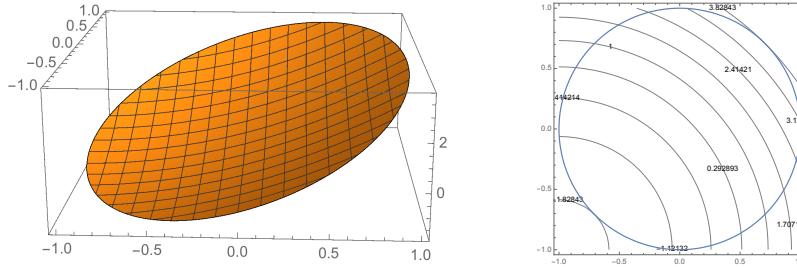
In this case, $y = -1$. Substitute this into ③, we get $x = 0$. However, putting $x = 0$ into ① yields $2 = 0$ which is absurd! Therefore, there is no solution in this case.

To sum up, the boundary critical points are: $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$, $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$. Evaluate $f(x, y)$ at each point gives:

$$f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = 1 + 2\sqrt{2}$$

$$f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = 1 - 2\sqrt{2}$$

Therefore, subject to the constraint $x^2 + y^2 = 1$, $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$ is the maximum point of f with value $1 + 2\sqrt{2}$, and $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$ is the minimum point of f with value $1 - 2\sqrt{2}$. See Figure 2.20 (the blue circle is the constraint).



(a) the graph

(b) the level set diagram

Figure 2.20: $f(x, y) = x^2 + y^2 + 2x + 2y$ in Example 2.11

■ **Example 2.12** Let $f(x, y) = x^2 - 4x + y^2 + 9$. Find the maximum and minimum points and values of $f(x, y)$ subject to the constraint $4x^2 + 9y^2 = 36$.

■ **Solution** Define $g(x, y) = 4x^2 + 9y^2$, then the constraint is the level set $g(x, y) = 36$. Set-up the Lagrange's Multiplier system:

$$\begin{aligned}\nabla f(x, y) &= \lambda \nabla g(x, y) \\ g(x, y) &= 36\end{aligned}$$

By computing ∇f and ∇g , the above is equivalent to a system of three equations:

$$\begin{aligned}2x - 4 &= 8\lambda x & (1) \\ 2y &= 18\lambda y & (2) \\ 4x^2 + 9y^2 &= 36 & (3)\end{aligned}$$

Case 1: $(2) \neq 0$

By $(1) \div (2)$, we get:

$$\begin{aligned}\frac{2x - 4}{2y} &= \frac{8\lambda x}{18\lambda y} \\ \frac{x - 2}{y} &= \frac{4x}{9y} & (\text{cancel } \lambda) \\ 9y(x - 2) &= 4xy & (\text{cross multiplication}) \\ 9(x - 2) &= 4x & (\text{cancel } y \neq 0) \\ x &= \frac{18}{5}\end{aligned}$$

However, substitute $x = \frac{18}{5}$ into (3), we get:

$$9y^2 = 36 - 4 \left(\frac{18}{5} \right)^2$$

which is a negative number, but $9y^2$ must be positive (or zero)! Therefore, there is no solution in this case.

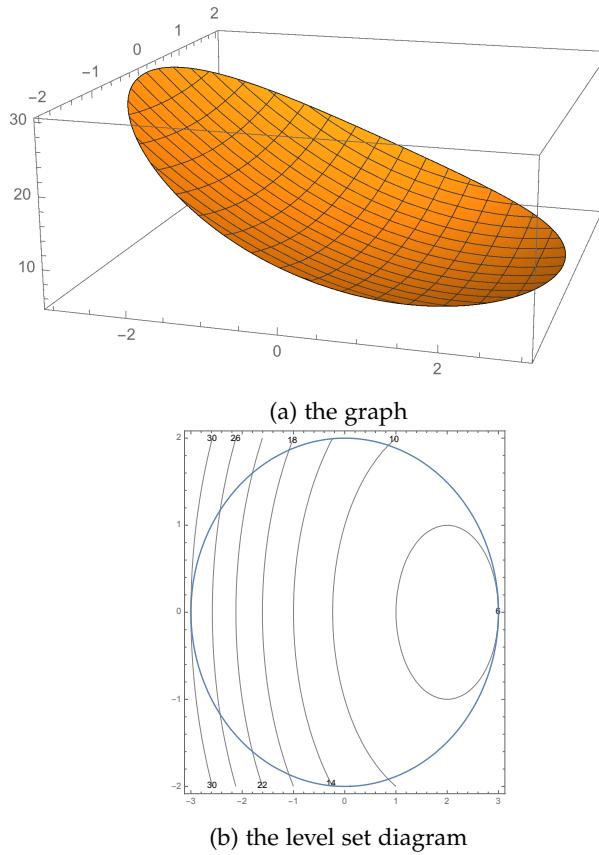
Case 2: $(2) = 0$

In this case, we have $2y = 18\lambda y = 0$, so $y = 0$. From (3), we have $4x^2 = 36$ and so $x = 3$ or $x = -3$. Therefore, $(x, y) = (3, 0)$ and $(x, y) = (-3, 0)$ are the solutions in this case.

Finally, we evaluate f at each boundary critical point:

$$\begin{aligned}f(3, 0) &= 6 \\ f(-3, 0) &= 30\end{aligned}$$

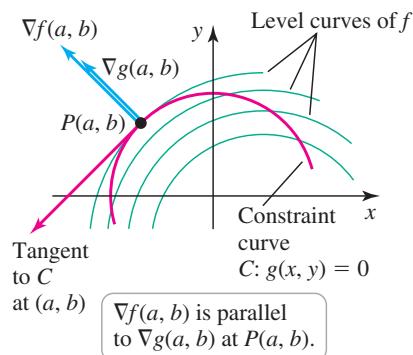
Therefore, minimum point is $(3, 0)$ with value 6; maximum point is $(-3, 0)$ with value 30. See Figure 2.21

Figure 2.21: $f(x, y) = x^2 - 4x + y^2 + 9$ in Example 2.12

The equation $\nabla f = \lambda \nabla g$ concerns about the geometry of the level sets and the constraint at the boundary critical point. Given a function $f(x, y)$ subject to the constraint $g(x, y) = c$. At the point (a, b) on the constraint where the maximum or minimum of $f(x, y)$ is achieved, **the level set of $f(x, y)$ at (a, b) is tangent to the constraint $g(x, y) = c$** . Consequently, the gradient vector $\nabla f(a, b)$, which is perpendicular to the level set of f at (a, b) , must be parallel to the gradient vector $\nabla g(a, b)$, which is perpendicular to the constraint $g = c$. See Figure 2.22 for an illustration. Therefore, at such a point, we must have:

$$\nabla f(a, b) = \lambda \nabla g(a, b)$$

where λ is a scalar.

Figure 2.22: ∇f and ∇g are parallel at the boundary critical point (a, b)

The Lagrange's Multiplier method also works for three-variable functions, yet the system of equations may be more complicated. Let's look at the example:

- **Example 2.13** Find the distance from $(0,0,0)$ to the plane $2x + 3y + 4z = 29$ using Lagrange's Multiplier.

■ **Solution** The distance from a point P to a plane is defined to be the shortest possible distance between the given point P and any point Q on the plane. Let's first formulate this problem in a mathematical way. We want to:

$$\begin{array}{ll} \text{minimize} & \sqrt{x^2 + y^2 + z^2} \\ \text{subject to constraint} & 2x + 3y + 4z = 29 \end{array}$$

However, to minimize $\sqrt{x^2 + y^2 + z^2}$ amounts to calculating $\nabla(\sqrt{x^2 + y^2 + z^2})$. As you can imagine, it would be messy. It is useful to observe that minimizing $\sqrt{x^2 + y^2 + z^2}$ is equivalent to minimizing $x^2 + y^2 + z^2$, i.e. the square of the distance from the origin. The latter is much easier to handle. Let:

$$\begin{aligned} f(x, y, z) &= x^2 + y^2 + z^2 \\ g(x, y, z) &= 2x + 3y + 4z, \end{aligned}$$

then the constraint is the level set $g = 29$. Set-up the Lagrange's Multiplier system $\nabla f = \lambda \nabla g$ as in previous examples:

$$\begin{aligned} 2x &= 2\lambda \\ 2y &= 3\lambda \\ 2z &= 4\lambda \\ 2x + 3y + 4z &= 29 \end{aligned}$$

Then $x = \lambda$, $y = \frac{3\lambda}{2}$ and $z = 2\lambda$. Substitute them into the constraint equation, we get:

$$2\lambda + \frac{9\lambda}{2} + 8\lambda = 29.$$

It is easy to see that $\lambda = 2$, and therefore $(x, y, z) = (2, 3, 4)$. It gives the unique critical point. It is intuitive that the minimum point must exist in this problem (and there is no maximum point), so this unique critical point must give the minimum point. Since $f(2, 3, 4) = 29$, the distance from $(0,0,0)$ to the plane is $\sqrt{29}$.

2.8 Optimizations

In this section we see more optimization examples of using the gradient and/or Lagrange's Multiplier method.

■ **Example 2.14** Many airlines require that the sum of length, width and height of a checked baggage cannot exceed 62 inches. Find the dimensions of the rectangular baggage that has the greatest possible volume under this regulation.

■ **Solution** Denote l, w, h to be the length, width and height respectively. We need to maximize the volume of the baggage, which is given by:

$$V(l, w, h) = lwh \text{ (cubic inches).}$$

The constraint is $l + w + h \leq 62$ (inches), but it is intuitively clear that in order to maximize the volume, the sum $l + w + h$ has better be at maximum possible. Define $g(l, w, h) = l + w + h$, then the constraint can be regarded as the level set $g = 62$. Set up the Lagrange's Multiplier system:

$$\begin{aligned}\nabla V &= \lambda \nabla g \\ g(l, w, h) &= 62\end{aligned}$$

which is equivalent to

$$\begin{aligned}wh &= \lambda \\ lh &= \lambda \\ lw &= \lambda \\ l + w + h &= 62\end{aligned}$$

Although it is not too difficult to solve them by hand, let's type the following command on Mathematica to solve them:

```
Solve[{w h == L, l h == L, l w == L, l + w + h == 62}, {l, w, h, L}]
```

These are all critical points:

$$(l, w, h) = \left(\frac{62}{3}, \frac{62}{3}, \frac{62}{3} \right), \quad (0, 0, 62), \quad (0, 62, 0), \quad (62, 0, 0).$$

Only the first one is physically relevant. Therefore, the rectangular baggage with the largest volume under this restriction is the square cube!

■ **Example 2.15** Three cities A , B and C are located at $(5, 2)$, $(-4, 4)$ and $(-1, -3)$ respectively on the (x, y) -plane. There is a railtrack whose equation is $y = x^3 + 1$, and a station is going to be built on the track so that the sum of squares of the distances from each city to the station is minimized. Find the coordinates of the station.

■ **Solution** Quantity to be minimized is

$$f(x, y, z) = \underbrace{(x - 5)^2 + (y - 2)^2}_{\text{distance}^2 \text{ from station to city A}} + \underbrace{(x + 4)^2 + (y - 4)^2}_{\text{distance}^2 \text{ from station to city B}} + \underbrace{(x + 1)^2 + (y + 3)^2}_{\text{distance}^2 \text{ from station to city C}}.$$

The constraint is that the station has to be on the track, i.e. $y = x^3 + 1$. Define $g(x, y) = y - x^3$, then the constraint can be written as $g(x, y) = 1$. Set up the Lagrange's Multiplier system

$\nabla f = \lambda \nabla g$ and $g(x, y) = 1$:

$$\begin{aligned} 2(x - 5) + 2(x + 4) + 2(x + 1) &= -3\lambda x^2 \\ 2(y - 2) + 2(y - 4) + 2(y + 3) &= \lambda \\ y - x^3 &= 1 \end{aligned}$$

Solving the system, we get $(x, y) = (0, 1)$. Therefore, the station should be located at $(0, 1)$ in order to minimize the sum of squares of the distances.

■ **Example 2.16 — Least Square Approximation.** Given a set of data points:

$$(x_1, y_1), \dots, (x_N, y_N)$$

on the xy -plane. Find the straight-line $y = mx + c$ such that the sum of squares of distances between each (x_i, y_i) and $(x_i, mx_i + c)$ is minimized.

■ **Solution** The quantity to be minimized is:

$$f(m, c) = \sum_{i=1}^N (y_i - mx_i - c)^2.$$

Note that (x_i, y_i) 's are given so they should be regarded as constants. The variables are m and c . Note that there is no constraint for m and c , so we can simply solve $\nabla f(m, c) = \mathbf{0}$ for critical points.

$$\begin{aligned} \frac{\partial f}{\partial m} &= -2 \sum_{i=1}^N (y_i - mx_i - c)x_i = -2 \left(\sum_{i=1}^N x_i y_i - m \sum_{i=1}^N x_i^2 - c \sum_{i=1}^N x_i \right) \\ \frac{\partial f}{\partial c} &= -2 \sum_{i=1}^N (y_i - mx_i - c) = -2 \left(\sum_{i=1}^N y_i - m \sum_{i=1}^N x_i - c N \right) \end{aligned}$$

Set $\frac{\partial f}{\partial m} = \frac{\partial f}{\partial c} = 0$, regarding all x_i 's and y_i 's to be constants, then:

$$\begin{aligned} Am + Bc &= E \\ Bm + Nc &= F \end{aligned}$$

where $A = \sum_{i=1}^N x_i^2$, $B = \sum_{i=1}^N x_i$, $E = \sum_{i=1}^N x_i y_i$ and $F = \sum_{i=1}^N y_i$. By solving the system carefully, one should get:

$$\begin{aligned} m &= \frac{BF - EN}{B^2 - AN} = \frac{(\sum x_i)(\sum y_i) - N(\sum x_i y_i)}{(\sum x_i)^2 - N(\sum x_i^2)} \\ c &= \frac{BE - AF}{B^2 - AN} = \frac{(\sum x_i)(\sum x_i y_i) - (\sum x_i^2)(\sum y_i)}{(\sum x_i)^2 - N(\sum x_i^2)} \end{aligned}$$

It is quite intuitive that this pair of m and c should minimize f since $f \geq 0$ and so a minimum must exist.

■ **Example 2.17** Let $f(x, y) = x^2 - 4x + y^2 + 9$ (which was considered in Example 2.12 in the previous section). Find the absolute maximum and absolute minimum of f restricted to the domain $4x^2 + 9y^2 \leq 36$.

■ **Solution** The Lagrange's Multiplier method finds us the boundary critical points on $4x^2 + 9y^2 = 36$. For the interior $4x^2 + 9y^2 < 36$, the critical points are simply solutions to $\nabla f = \mathbf{0}$. The general procedure of an optimization problem with a solid domain is that:

1. Find all interior critical points by solving $\nabla f = \mathbf{0}$;
2. Find all boundary critical points using Lagrange's Multiplier;
3. Evaluate f at each critical points found, and look for the point that gives greatest/lowest value of f .

Interior: Set $\nabla f = \mathbf{0}$, we get:

$$\begin{aligned} 2x - 4 &= 0 \\ 2y &= 0 \end{aligned}$$

Therefore, the only interior critical point is $(2, 0)$, which can be checked easily that it is in the given region.

Boundary: We have already done in Example 2.12 that the boundary critical points are $(3, 0)$ and $(-3, 0)$.

Finally, evaluate f at each critical point found:

$$\begin{aligned} f(2, 0) &= 5 \\ f(3, 0) &= 6 \\ f(-3, 0) &= 30 \end{aligned}$$

Therefore, the absolute minimum is 5 (attained at $(2, 0)$), and the absolute maximum is 30 (attained at $(-3, 0)$).



3 — Multiple Integrations

3.1 Double Integrals in Rectangular Coordinates

In single-variable calculus, integration is used to find the area of the graph of a function $f(x)$. In this chapter, we will generalize the concept of integrations to multivariable functions. There are many applications of multiple integrals in sciences, including deriving moments of inertia, probability, and in later part of the course: finding surface area and surface flux.

Computations of multivariable integrals are not much different from those in single-variable integrals, but *setting up* a multivariable integral involves a lot more geometric intuitions. Let's first look at some computations first before we explain the geometry of these integrals.

■ **Example 3.1** Compute the following double integral:

$$\int_{y=1}^{y=2} \int_{x=0}^{x=1} (4 - x - y^2 x) dx dy.$$

■ **Solution** A double integral consists of an **inner** integral and an **outer** integral:

$$\overbrace{\int_{y=1}^{y=2} \underbrace{\int_{x=0}^{x=1} (4 - x - y^2 x) dx}_{\text{inner}} dy}^{\text{outer}}.$$

When computing the inner integral (which is respect to x in this example), we regard all other variable(s) (i.e. y) to be constant(s):

$$\begin{aligned} \int_{y=1}^{y=2} \int_{x=0}^{x=1} (4 - x - y^2 x) dx dy &= \int_{y=1}^{y=2} \left[4x - \frac{x^2}{2} - \frac{y^2 x^2}{2} \right]_{x=0}^{x=1} dy \\ &= \int_{y=1}^{y=2} \left(\left(4 - \frac{1}{2} - \frac{y^2}{2} \right) - 0 \right) dy \\ &= \int_{y=1}^{y=2} \left(\frac{7}{2} - \frac{y^2}{2} \right) dy = \left[\frac{7y}{2} - \frac{y^3}{6} \right]_{y=1}^{y=2} = \frac{7}{3}. \end{aligned}$$

It is worthwhile to note that if we switch the inner and outer integrals, the final answer is the

same!

$$\begin{aligned} \int_{x=0}^{x=1} \int_{y=1}^{y=2} (4 - x - y^2 x) dy dx &= \int_{y=1}^{y=2} \left[4y - xy - \frac{y^3 x}{3} \right]_{y=1}^{y=2} dx \\ &= \int_{x=0}^{x=1} \left(8 - 2x - \frac{8x}{3} \right) - \left(4 - x - \frac{x}{3} \right) dx \\ &= \int_{x=0}^{x=1} \left(4 - \frac{10x}{3} \right) dx = \frac{7}{3}. \end{aligned}$$

It is not a coincident! Let's explain why it is true by learning the geometric meaning of double integrals. Consider the integral:

$$\int_{x=a}^{x=b} \int_{y=c}^{y=d} f(x, y) dy dx.$$

The inner integral:

$$A(x) := \int_{y=c}^{y=d} f(x, y) dy$$

is an integral with respect to y keeping x fixed. This quantity represents the **area under the curve obtained by moving along the y -direction from $y = c$ to $y = d$ on the surface $z = f(x, y)$, while keeping x unchanged**. See Figure 3.1.

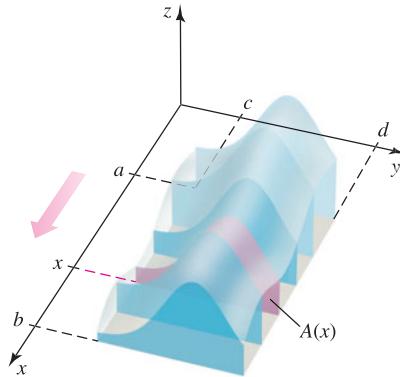


Figure 3.1: geometric meaning of a double integral

The outer integral integrates the inner integral $A(x)$ from $x = a$ to $x = b$, i.e.

$$\int_{x=a}^{x=b} \int_{y=c}^{y=d} f(x, y) dy dx = \int_{x=a}^{x=b} A(x) dx.$$

Since $A(x) dx$ can be thought as the volume of a solid slice with width dx and cross-section area $A(x)$, by integrating $A(x) dx$ it means adding up the volume of these thin slices and so the double integral

$$\int_{x=a}^{x=b} A(x) dx$$

is the **volume under the graph $z = f(x, y)$ over the base rectangle bounded by $x = a, x = b, y = c$ and $y = d$** . It is important to understand the geometric meanings of the inner and outer integrals in order to set-up a double integral correctly.

As a double integral represents the volume of a solid, one should not expect there is any difference if we slice the solid in a different way. For instance, to find the volume under the

graph $z = 6 - 2x - y$ over the rectangular region $0 \leq x \leq 1$ and $0 \leq y \leq 2$, one can set up the double integral in either way:

$$\int_{x=0}^{x=1} \underbrace{\int_{y=0}^{y=2} (6 - 2x - y) dy dx}_{A(x)} \quad \text{see Figure 3.2a}$$

$$\int_{y=0}^{y=2} \underbrace{\int_{x=0}^{x=1} (6 - 2x - y) dx dy}_{A(y)} \quad \text{see Figure 3.2b}$$

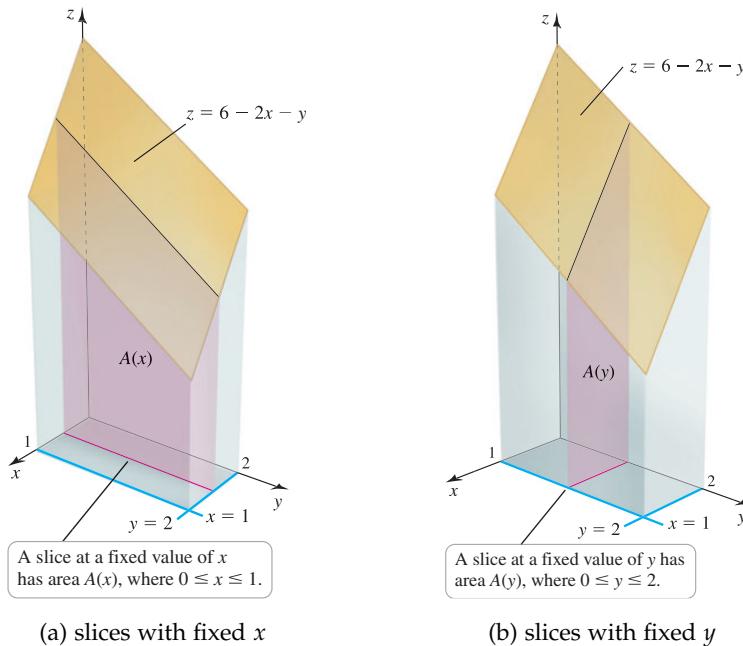


Figure 3.2: volume under the same graph

Readers should verify that the above integrals indeed give the same value (the answer is 8). In general, the following Fubini's Theorem asserts that switching dx and dy (and the corresponding integral signs) give the same double integral. Although the statement of the theorem is geometrically intuitive, the proof is not easy and is beyond the scope of this course.

Theorem 3.1 — Fubini's Theorem for Rectangular Regions. Let $f(x, y)$ be a continuous function over a rectangular region $a \leq x \leq b$ and $c \leq y \leq d$, then:

$$\int_{y=c}^{y=d} \int_{x=a}^{x=b} f(x, y) dx dy = \int_{x=a}^{x=b} \int_{y=c}^{y=d} f(x, y) dy dx.$$

Since the order of integration (i.e. $dxdy$ or $dydx$) determines the order of the integral signs, the above two double integrals can simply be written as:

$$\int_c^d \int_a^b f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx.$$

Even simpler, one may denote $dA := dxdy$ or $dydx$ and the rectangular region by R . Then, we can write the integral as:

$$\iint_R f(x, y) dA.$$

When setting up a double integral to find the volume of the solid under a graph $z = f(x, y)$, it is worthwhile to observe that the lower and upper limits of the integral are not affected by the function $f(x, y)$. Therefore, in order to interpret a double integral in a geometric way, one may simply draw the base region (or in other words, the top-down view) instead of drawing the solid in the three-dimensional space.

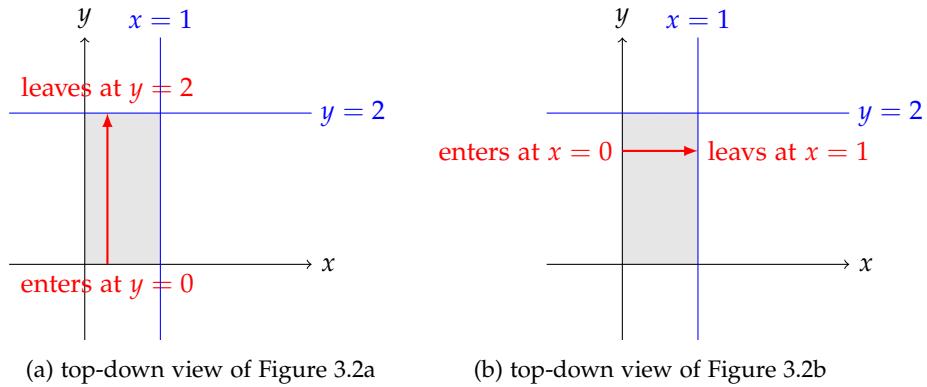


Figure 3.3: the red arrows represent the cross-section slices in Figures 3.2a and 3.2b.

3.2 Fubini's Theorem for General Regions

In this section, we demonstrate some examples of double integrals whose base regions may not be rectangles but some general regions. Setting up these double integrals often require a certain degree of geometric intuition. Therefore it is important to see more examples to acquire the necessary skills.

- **Example 3.2** Find the volume of the solid under the plane $z = 3 - x - y$ over the triangle region R bounded by the x -axis, $x = 1$ and $y = x$.

■ **Solution** First we choose an order of integration, say $dydx$. The inner integral should calculate the area of slices with y varies and x fixed. Since the height $3 - x - y$ of the solid does not affect how we set-up the upper/lower limits, we consider the top-down view of the solid (see Figure 3.4).

The red strip in Figure 3.4b represents a sample slice with fixed x . The strip **enters** at $y = 0$ and **leaves** at $y = x$. Hence, the area of this slice is:

$$\int_{y=0}^{y=x} \underbrace{(3 - x - y)}_{\text{height}} dy.$$

"Summing up" the area of these slices, we integrate by dx over the range of x : $0 \leq x \leq 1$ as shown in Figure 3.4b, i.e.

$$\int_{x=0}^{x=1} \underbrace{\int_{y=0}^{y=x} (3 - x - y) dy}_{\text{inner integral}} dx.$$

It will give the volume of the solid as required in this problem. The rest of the task is to compute the integral:

$$\begin{aligned} \int_{x=0}^{x=1} \int_{y=0}^{y=x} (3 - x - y) dy dx &= \int_{x=0}^{x=1} \left[3y - xy - \frac{y^2}{2} \right]_{y=0}^{y=x} dx \\ &= \int_{x=0}^{x=1} \left(3x - x^2 - \frac{x^2}{2} \right) dx \\ &= \int_0^1 \left(3x - \frac{3x^2}{2} \right) dx \\ &= \left[\frac{3x^2}{2} - \frac{x^3}{2} \right]_0^1 \\ &= 1. \end{aligned}$$

Alternatively, we can also integrate first by dx then by dy . Then, the inner integral is represented by the red strip in Figure 3.4c. It **enters** at $x = y$ and **leaves** at $x = 1$. The double integral is therefore:

$$\int_{y=0}^{y=1} \int_{x=y}^{x=1} (3 - x - y) dx dy.$$

Readers should compute as an exercise that the answer is again 1.

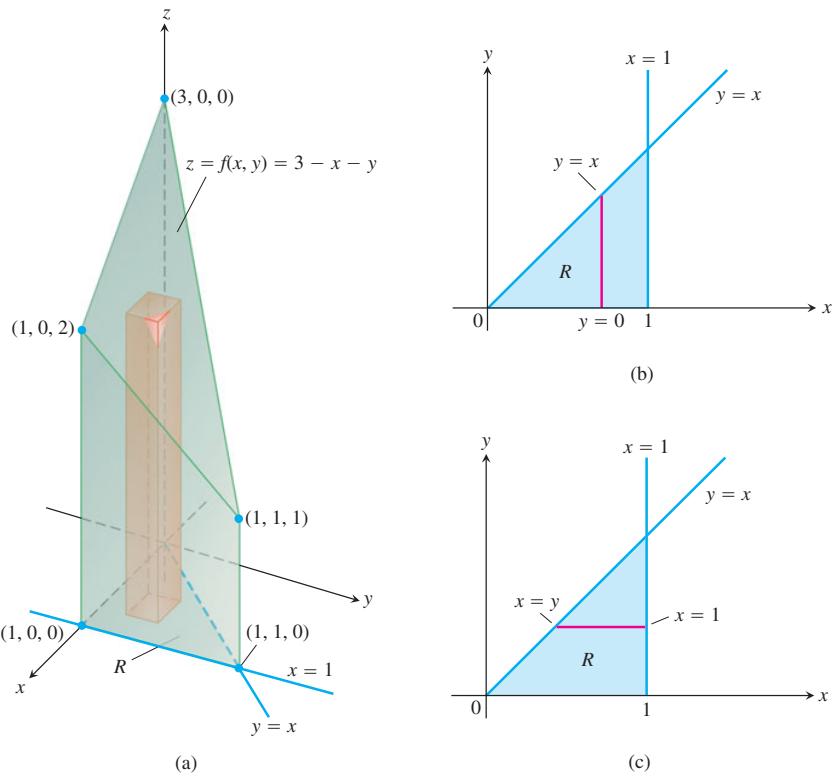


Figure 3.4: the graph and the top-down views of the function in Example 3.2

The fact that we have the freedom to choose our order of integration is guaranteed by the Fubini's Theorem, whose proof is again beyond the scope of this course.

Theorem 3.2 — Fubini's Theorem for General Regions. Let R be a region on the xy -plane and $f(x, y)$ is a continuous function on R , then

$$\iint_R f(x, y) dxdy = \iint_R f(x, y) dydx$$

where the lower/upper limits of each integral are set up according to the shape of the region R .

Notation As there is no difference between $dxdy$ and $dydx$ as far as the upper and lower limits are set according to the same region, we may simply write:

$$dA = dxdy \text{ or } dydx$$

Although choosing the $dxdy$ -order will yield the same result as the $dydx$ -order, it happens often that one order is easier while the other one is harder. Let's look at the following example:

■ **Example 3.3** Let R be the region in the first quadrant of the xy -plane bounded by the unit circle $x^2 + y^2 = 1$ and the straight-line $x + y = 1$. Evaluate the integral

$$\iint_R \sqrt{1 - x^2} dA.$$

■ **Solution** First choose the order of integration. However, it seems like integrating $\sqrt{1-x^2}$ by dx involves trig substitutions that we want to avoid if possible. Let's try integrating first by dy then by dx (to see if there is any luck).

Set-up the double integral according the top-down view of the solid (Figure 3.5):

$$\int_{x=0}^{x=1} \int_{y=1-x}^{y=\sqrt{1-x^2}} \sqrt{1-x^2} dy dx.$$

As x is regarded as a constant when dealing with the inner integral, we can easily see that:

$$\begin{aligned} \int_{x=0}^{x=1} \int_{y=1-x}^{y=\sqrt{1-x^2}} \sqrt{1-x^2} dy dx &= \int_{x=0}^{x=1} \left[y\sqrt{1-x^2} \right]_{y=1-x}^{y=\sqrt{1-x^2}} dx \\ &= \int_{x=0}^{x=1} (1-x^2) - (1-x)\sqrt{1-x^2} dx \\ &= \int_0^1 (1-x^2 - \sqrt{1-x^2} + x\sqrt{1-x^2}) dx \end{aligned}$$

The only two difficult parts are

$$\int_0^1 \sqrt{1-x^2} dx \text{ and } \int_0^1 x\sqrt{1-x^2} dx.$$

The former can be evaluated by substitution $x = \sin \theta$, while the latter can be done by substituting $u = 1 - x^2$. Readers should complete the rest of computations as an exercise. The final answer should be $1 - \frac{\pi}{4}$.

We need a trig substitution anyway, but it is easier than doing a substitution for the inner integral.

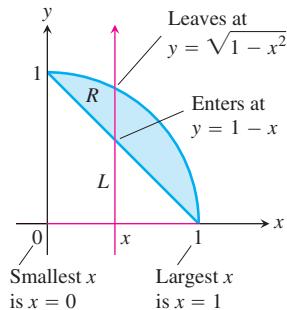


Figure 3.5: top-down view of the region in Example 3.3

In the previous example, we see that although the Fubini's Theorem tells that *in theory* we can choose our favorite the order of integration, *in practice* we sometimes have to make a smart choice. In the next example, let's demonstrate an example that one order gives an integral which is ridiculously hard to compute, while another is extremely easy.

■ **Example 3.4** Evaluate the following double integral:

$$\int_0^1 \int_y^1 \frac{\sin x}{x} dx dy.$$

■ **Solution** The integrand $\frac{\sin x}{x}$ does not have a simple antiderivative when integrating by dx ! Let's switch the order of integration first.

The region corresponds to the double integral is formed by strips entering at $x = y$ and leaving at $x = 1$. The range of y is from 0 to 1. A sketch of the diagram can be found in Figure 3.6.

Switching the order of integration, the strip for each x enters at $y = 0$ and leaves at $y = x$. Therefore, Fubini's Theorem says:

$$\int_0^1 \int_y^1 \frac{\sin x}{x} dx dy = \int_0^1 \int_0^x \frac{\sin x}{x} dy dx.$$

The RHS is much easier to compute:

$$\begin{aligned} \int_0^1 \int_0^x \frac{\sin x}{x} dy dx &= \int_0^1 \left[y \frac{\sin x}{x} \right]_{y=0}^{y=x} dx \\ &= \int_0^1 \left(x \frac{\sin x}{x} - 0 \right) dx \\ &= \int_0^1 \sin x dx = -\cos 1. \end{aligned}$$

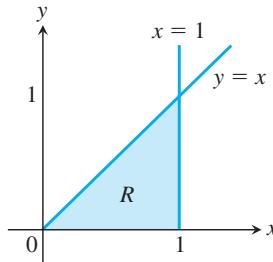
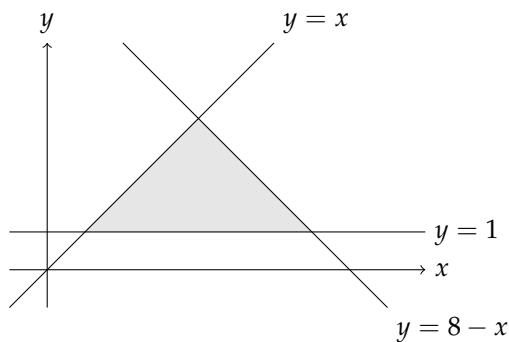


Figure 3.6: the region of integration in Example 3.4

If the region of integration is the shaded triangle below, which order of integration is better?

$$\iint_R f(x, y) dx dy \quad \text{or} \quad \iint_R f(x, y) dy dx,$$

where $f(x, y)$ is not very complicated, say $f(x, y) = x^2 y$.



3.3 Double Integrals in Polar Coordinates

When the region of integration is *circular* in shape, or the integrand is rotationally symmetric, it is often more convenient to use *polar coordinates* to set-up the integral instead of the using the rectangular coordinates as in the previous section. We will see in some examples in this section that some tedious trig substitution can be avoided if polar coordinates are used.

Recall that the polar coordinates (r, θ) and the rectangular coordinates (x, y) are related by the following rules:

$$\begin{aligned} x &= r \cos \theta \\ y &= r \sin \theta \end{aligned}$$

Here r is the distance from the point to the origin, and θ is the angle made with the positive x -axis.

In rectangular coordinates, the region defined by inequalities like $a \leq x \leq b$ and $c \leq y \leq d$, where a, b, c and d are constants, describe a rectangle. We have seen that it is very easy to set-up a double integral when the region is a rectangle.

In polar coordinates, regions defined by inequalities like $a \leq r \leq b$ and $\alpha \leq \theta \leq \beta$, where a, b, α and β are constants, describe a fan shape or a circular sector (see Figure 3.7). It is wise to use polar coordinates instead of rectangular coordinates when the region of integration is given by one of these circular shapes.

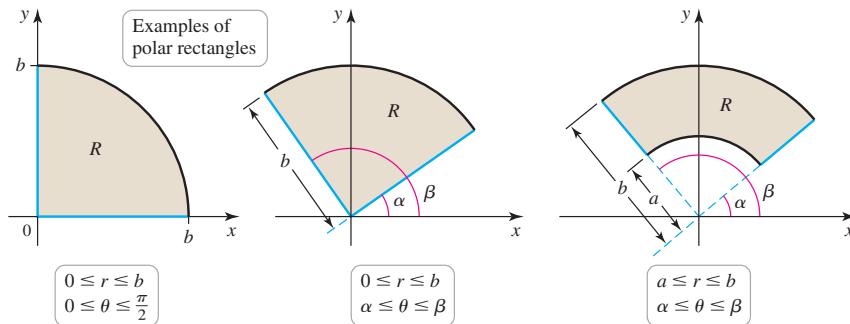


Figure 3.7: examples of regions good for polar coordinates

To set-up a double integral of a region $a \leq r \leq b$ and $\alpha \leq \theta \leq \beta$ using polar coordinates, the upper/lower limits are simply:

$$\int_{\theta=\alpha}^{\theta=\beta} \int_{r=a}^{r=b} \quad \text{or} \quad \int_{r=a}^{r=b} \int_{\theta=\alpha}^{\theta=\beta}$$

depending on the order of integration $drd\theta$ or $d\theta dr$. However, one should be very cautious that while $dA = dx dy$ in rectangular coordinates, it is instead:

$$dA = r dr d\theta \quad \text{or} \quad r d\theta dr$$

in polar coordinates. We will explain why it is so after learning a few examples.

■ **Example 3.5** Evaluate the integral:

$$\iint_R (x^2 + y^2) dA$$

where R is the semicircular region bounded by the x -axis and the curve $y = \sqrt{1 - x^2}$. See Figure 3.8.

■ **Solution** The problem is extremely difficult to do in rectangular coordinates. If one attempts to set it up using xy -coordinates, one will get the following double integral:

$$\int_{-1}^1 \int_0^{\sqrt{1-x^2}} (x^2 + y^2) dy dx.$$

After computing the inner integral, one should get:

$$\int_{-1}^1 \left(x^2 \sqrt{1-x^2} + \frac{(1-x^2)^{3/2}}{3} \right) dx.$$

However, things go much better if we switch to polar coordinates, since the region is in the form $a \leq r \leq b$ and $\alpha \leq \theta \leq \beta$. From Figure 3.8, the region is defined by:

$$0 \leq r \leq 1, \quad 0 \leq \theta \leq \pi.$$

Keeping in mind that $dA = r dr d\theta$, the required integral is:

$$\begin{aligned} \iint_R (x^2 + y^2) dA &= \int_{\theta=0}^{\theta=\pi} \int_{r=0}^{r=1} \underbrace{\left((r \cos \theta)^2 + (r \sin \theta)^2 \right)}_{x^2+y^2} r dr d\theta \\ &= \int_{\theta=0}^{\theta=\pi} \int_{r=0}^{r=1} r^2 (\cos^2 \theta + \sin^2 \theta) r dr d\theta \\ &= \int_{\theta=0}^{\theta=\pi} \int_{r=0}^{r=1} r^3 dr d\theta \\ &= \int_{\theta=0}^{\theta=\pi} \left[\frac{r^4}{4} \right]_{r=0}^{r=1} d\theta \\ &= \int_{\theta=0}^{\theta=\pi} \frac{1}{4} d\theta = \frac{\pi}{4}. \end{aligned}$$

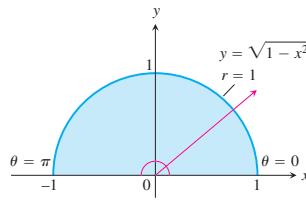


Figure 3.8: region of integration for Example 3.5

Annular regions, to be discussed in the next example, are extremely clumsy using rectangular coordinates but relatively easy using polar coordinates.

■ **Example 3.6** Evaluate the following integral:

$$\iint_R x \, dA$$

where R is an annular region shown in Figure 3.9.

■ **Solution** The annular region R is defined by inequalities:

$$2 \leq r \leq 4, \quad 0 \leq \theta \leq 2\pi.$$

Therefore,

$$\begin{aligned} \iint_R x \, dA &= \int_0^{2\pi} \int_2^4 r \cos \theta \cdot r \, dr \, d\theta \\ &= \int_0^{2\pi} \int_2^4 r^2 \cos \theta \, dr \, d\theta \\ &= \int_0^{2\pi} \left[\frac{r^3}{3} \right]_{r=2}^{r=4} \cdot \cos \theta \, d\theta \\ &= \int_0^{2\pi} \frac{56}{3} \cos \theta \, d\theta \\ &= \left[\frac{56}{3} \sin \theta \right]_0^{2\pi} \\ &= 0 \end{aligned}$$

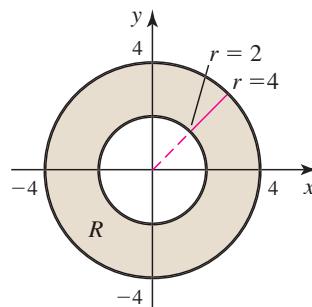


Figure 3.9: region of integration for Example 3.6

A Notoriously Difficult Single-Variable Integral

Next we present an application of evaluating a notoriously difficult single-variable integral using a double integral in polar coordinates.

It is famous (or infamous) that the integral:

$$\int e^{-x^2} \, dx$$

does not have an easy explicit expression. Nonetheless, this integral is extremely important in probability, heat flow and many physics and engineering subjects. While this integral is generally hard to compute, it can be *magically* done, with the help of a double integral, when the upper/lower limits are:

$$\int_0^\infty e^{-x^2} \, dx.$$

Consider the double integral:

$$\int_0^\infty \int_0^\infty e^{-x^2-y^2} dx dy.$$

Observing that $e^{-x^2-y^2} = e^{-x^2}e^{-y^2}$, we can regard e^{-y^2} as a constant in the inner dx -integral:

$$\int_0^\infty \int_0^\infty e^{-x^2-y^2} dx dy = \int_0^\infty \int_0^\infty e^{-x^2} e^{-y^2} dx dy = \int_0^\infty e^{-y^2} \left(\int_0^\infty e^{-x^2} dx \right) dy.$$

Now that $\int_0^\infty e^{-x^2} dx$ is a constant, and in particular, independent of y . Therefore, with respect to the outer dy -integral, one can factor it out and yield:

$$\int_0^\infty e^{-y^2} \left(\int_0^\infty e^{-x^2} dx \right) dy = \left(\int_0^\infty e^{-x^2} dx \right) \left(\int_0^\infty e^{-y^2} dy \right).$$

Now the double integral are split as a product of two single-variable integrals. Note that they two single-variable integrals are the same as x and y are merely dummy variables! Therefore, combining everything we have shown above, we have:

$$\int_0^\infty \int_0^\infty e^{-x^2-y^2} dx dy = \left(\int_0^\infty e^{-x^2} dx \right) \left(\int_0^\infty e^{-y^2} dy \right) = \left(\int_0^\infty e^{-x^2} dx \right)^2.$$

Consequently, in order to evaluate the single-variable integral $\int_0^\infty e^{-x^2} dx$, one can first evaluate the double integral

$$\int_0^\infty \int_0^\infty e^{-x^2-y^2} dx dy$$

and then the square root of the double integral will give the value of $\int_0^\infty e^{-x^2} dx$.

In contrast to the single-variable integral, the double integral is relatively easy to find using polar coordinates. The region of integration is the entire first quadrant which, in polar coordinates, can be described as:

$$0 \leq r \leq \infty, \quad 0 \leq \theta \leq \frac{\pi}{2}.$$

Therefore,

$$\begin{aligned} \int_0^\infty \int_0^\infty e^{-x^2-y^2} dx dy &= \int_0^{\pi/2} \int_0^\infty e^{-r^2} r dr d\theta \\ &= \int_0^{\pi/2} \left[-\frac{1}{2} e^{-r^2} \right]_{r=0}^{r=\infty} d\theta \quad \text{note that } \frac{d}{dr} e^{-r^2} = e^{-r^2} \cdot (-2r) \\ &= \int_0^{\pi/2} \left(-0 + \frac{1}{2} \right) d\theta \\ &= \frac{\pi}{4}. \end{aligned}$$

Therefore,

$$\int_0^\infty e^{-x^2} dx = \sqrt{\frac{\pi}{4}} = \frac{\sqrt{\pi}}{2}.$$

Furthermore, e^{-x^2} is an even function, we also have:

$$\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}.$$

However, this trick does not work well for any similar definite integral such as:

$$\int_0^1 e^{-x^2} dx.$$

Although it is still true (from a similar argument as above) that we have

$$\left(\int_0^1 e^{-x^2} dx \right)^2 = \int_0^1 \int_0^1 e^{-x^2-y^2} dxdy,$$

the double integral is not “polar-friendly” since the region of integration is a square!

Explanation of the Polar Method

We end this section by explaining why $dA = r dr d\theta$ but not simply $dA = dr d\theta$.

In rectangular coordinates, the area of a region is calculated by chopping off the region into tiny rectangular pieces with length and width denoted by the changes Δx and Δy . The area of each rectangular piece is given by $\Delta A = \Delta x \cdot \Delta y$. When these rectangular pieces get smaller and smaller, Δx and Δy become dx and dy , and so the area of these rectangular pieces becomes $dA = dx dy$.

However, things are a bit different in polar coordinates. Instead of chopping off a given region into tiny rectangles, the region is chopped into “fan-shaped” pieces (see Figure 3.10). Given a pair of Δr and $\Delta\theta$, the area ΔA of the piece is *not* simply $\Delta r \cdot \Delta\theta$, but is getting proportionally larger when the piece is further from the origin!

Therefore, one needs to multiply r to $\Delta r \cdot \Delta\theta$ in order to accurately reflect the size of ΔA . Precisely, ΔA can be calculated as the difference between the area of an outer sector (with radius $r + \Delta r$) and an inner sector (with radius r):

$$\begin{aligned} \Delta A &= \underbrace{\frac{1}{2} (r + \Delta r)^2 \Delta\theta}_{\text{area of outer sector}} - \underbrace{\frac{1}{2} r^2 \Delta\theta}_{\text{area of inner sector}} \\ &= \frac{1}{2} (r^2 + 2r\Delta r + (\Delta r)^2 - r^2) \Delta\theta \\ &= r\Delta r \cdot \Delta\theta + \frac{1}{2} (\Delta r)^2 \Delta\theta. \end{aligned}$$

When Δr and $\Delta\theta$ get smaller and smaller, the last term $(\Delta r)^2 \Delta\theta$ is negligible when compared to the other term $r\Delta r \Delta\theta$. Therefore, as the region is chopped into infinitesimal pieces, we have:

$$dA = r dr d\theta.$$

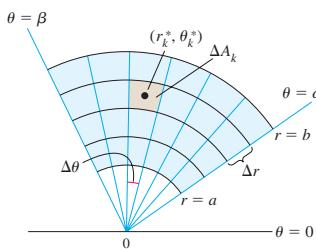


Figure 3.10: area of ΔA gets larger when r gets larger.

3.4 Triple Integrals in Rectangular Coordinates

We now move one dimension up and talk about triple integrals, whose integrands are functions of three variables such as $f(x, y, z)$. An example of such an integral is:

$$\int_0^1 \int_0^y \int_z^y f(x, y, z) dx dz dy.$$

Some triple integrals have important physical meanings. For instance, if $f(x, y, z)$ is the density function, then the triple integral

$$\int_0^1 \int_0^y \int_z^y f(x, y, z) dx dz dy$$

represents the mass of the solid (whose shape is determined by the upper/lower limits of the integral). It is because $dx dz dy$ represents (infinitesimal) volume, and

$$\text{density} \times \text{volume} = \text{mass}.$$

In physics, some calculations such as the center of mass of an object and the moment of inertia about an axis also involve evaluations of triple integrals.

Pillar-Base Approach

Before we see some practical applications of triple integrals, let's learn how a triple integral can be set-up. One common approach is so-called the **pillar-base** approach or **pillar-shadow** approach. Let's explain this through an example:

- **Example 3.7** Let D be the tetrahedral solid bounded by the plane $z = y - x$, the plane $y = 1$, the xy -plane and the yz -plane (see Figure 3.11). Evaluate the triple integral:

$$\iiint_D x^2 dz dy dx.$$

- **Solution** We demonstrate the pillar-base approach in the solution. The so-called *pillar* is the orange ray labeled M in Figure 3.11, and it determines how the inner-most integral is set-up:

$$\int_{z=?}^{z=?} x^2 dz.$$

Its lower and upper limits of z are determined by, respectively, where the pillar **enters** the solid and where the pillar **leaves** the solid. According to the diagram, it enters at $z = 0$, i.e. the xy -plane; and it leaves at $z = y - x$. Therefore, the inner-most integral should be:

$$\int_{z=0}^{z=y-x} x^2 dz.$$

Note that again the integrand x^2 does not affect how we set-up this integral.

Next we call the other two variables x and y to be *base* variables or *shadow* variables. Cast some light on the solid in the direction parallel to the pillar (i.e. z -direction), a shadow appears on the base xy -plane as a triangle. To way to set-up the middle and outer-most integrals are just the same as what we did for double integrals.

Since we picked the $dy dx$ -order, we draw a sample strip L on the base in the direction of y . This strip enters at $y = x$ and leaves at $y = 1$, and therefore the middle and outer-most integrals should be set-up as:

$$\underbrace{\int_{x=0}^{x=1} \int_{y=x}^{y=1} \int_{z=0}^{z=y-x} x^2 dz dy dx}_{\text{base}}_{\text{pillar}}$$

Finally, the easiest step is to evaluate the integral. This is as straight-forward as in double integrals:

$$\begin{aligned}
 \int_0^1 \int_x^1 \int_0^{y-x} x^2 dz dy dx &= \int_0^1 \int_x^1 [x^2 z]_{z=0}^{z=y-x} dy dx \\
 &= \int_0^1 \int_x^1 x^2(y-x) dy dx \\
 &= \int_0^1 \left[\frac{x^2 y^2}{2} - x^3 y \right]_{y=x}^{y=1} dx \\
 &= \int_0^1 \left(\frac{x^2}{2} - x^3 - \frac{x^4}{2} + x^4 \right) dx \\
 &= \int_0^1 \left(\frac{x^2}{2} - x^3 + \frac{x^4}{2} \right) dx \\
 &= \left[\frac{x^3}{6} - \frac{x^4}{4} + \frac{x^5}{10} \right]_0^1 \\
 &= \frac{1}{6} - \frac{1}{4} + \frac{1}{10} = \frac{1}{60}.
 \end{aligned}$$

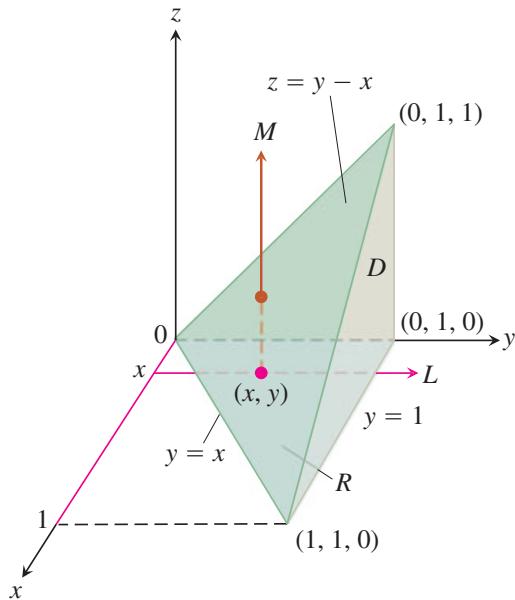


Figure 3.11: solid D in Example 3.7.

The Fubini's Theorem also holds for triple integrals, i.e. we can evaluate the integral by different order of integration, say $dxdydz$ or $dydxdz$ (there are 6 possible orders), we should get the same value provided that the upper and lower limits are adjusted to present the same solid.

■ **Example 3.8** Let D be the tetrahedral solid in Example 3.7. Evaluate the triple integral

$$\iiint_D x^2 dy dz dx$$

using the $dy dz dx$ -order.

■ **Solution** Now the inner-most integral is with respect to dy , meaning the *pillar* is pointing along the positive y -axis. It is the ray labeled as M in Figure 3.12. It enters the solid through the plane $z = y - x$, or equivalently $y = x + z$, and it leaves at $y = 1$. Therefore, the inner-most integral should be:

$$\int_{y=x+z}^{y=1} x^2 dy.$$

The middle and the outer-most variables are z and x , so the *base* is the shadow of the solid on the xz -plane, which is the triangle labeled R in Figure 3.12.

Here the order of integration is $dxdz$, so we draw a sample strip (labeled L) along the z -axis direction. It enters the region through $z = 0$ and leaves at the line $x + z = 1$, or equivalently, $z = 1 - x$. Therefore, the whole triple integral should be set as:

$$\int_{x=0}^{x=1} \int_{z=0}^{z=1-x} \int_{y=x+z}^{y=1} x^2 dy dz dx.$$

It can be verified by straight-forward computations that the answer should be the same as we got in Example 3.7:

$$\begin{aligned} \int_{x=0}^{x=1} \int_{z=0}^{z=1-x} \int_{y=x+z}^{y=1} x^2 dy dz dx &= \int_0^1 \int_0^{1-x} \left[x^2 y \right]_{y=x+z}^{y=1} dz dx \\ &= \int_0^1 \int_0^{1-x} (x^2 - x^2(x+z)) dz dx \\ &= \int_0^1 \left[(x^2 - x^3)z - \frac{x^2 z^2}{2} \right]_{z=0}^{z=1-x} dx \\ &= \int_0^1 \left((x^2 - x^3)(1-x) - \frac{x^2(1-x)^2}{2} \right) dx \\ &= \int_0^1 \left(\frac{x^2}{2} - x^3 + \frac{x^4}{2} \right) dx = \frac{1}{60} \quad (\text{same!}) \end{aligned}$$

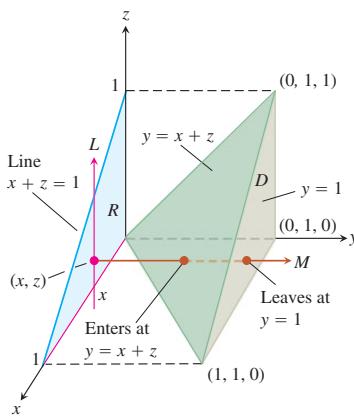


Figure 3.12: the pillar-base diagram for the triple integral in Example 3.8

The Fubini's Theorem tells us that no matter what order of integration we choose, we always get the same answer. Therefore, we can sometimes use the notation dV to denote the volume element $dxdydz$ (or any other order). A generic triple integral can be written as:

$$\iiint_D f(x, y, z) dV.$$

Although Fubini's Theorem allows us to switch the order of integral, owever, to ease our computation, we sometimes need to choose a *smart* choice of pillar and base variables. Let's look at the next example:

■ **Example 3.9** Let D be the solid bounded by the paraboloid $y = x^2 + z^2$ and $y = 16 - 3x^2 - z^2$ (see Figure 3.13). Find the volume of the solid, i.e. evaluate the integral:

$$\iiint_D 1 dV.$$

■ **Solution** After taking a careful look at the diagram, one should see that it would be a bad choice if we chose either x or z to be the pillar variable. The solid can be decomposed into two parts as shown in the diagram. If z were chosen to be the pillar direction, then the pillar would enter the yellow part in a way different from it does in the blue part. The yellow part would have $z = \pm\sqrt{16 - 3x^2 - y}$ as the lower/upper limits, while the blue part have $z = \pm\sqrt{y - x^2}$. Even worse, the part near the intersection of the blue and the yellow surfaces is geometrically complicated – it is not easy to set up the inner integral for that part.

However, life is much easier if we choose y as the pillar variable. Since then the y -pillar will enter through the blue surface and leave through the yellow surface. The shadow is an ellipse on the xz -plane.

To set-up the inner integral, we note that the y -pillar enters at $y = x^2 + z^2$ and leaves at $y = 16 - 3x^2 - z^2$:

$$\int_{y=x^2+z^2}^{y=16-3x^2-z^2} dy.$$

The shadow (or base) is an ellipse, whose equation can be obtained by setting $y = x^2 + z^2 = 16 - 3x^2 - z^2$, which gives

$$2x^2 + z^2 = 8.$$

For the base integral, we can either choose $dzdx$ or $dxdz$. For the former, the sample z -strip enters the ellipse at $-\sqrt{8 - 2x^2}$ and leaves at $\sqrt{8 - 2x^2}$. The min/max values of x are -2 and 2 for the ellipse. Therefore, the middle and outer integrals should:

$$\int_{x=-2}^{x=2} \int_{z=-\sqrt{8-2x^2}}^{z=\sqrt{8-2x^2}} \int_{y=x^2+z^2}^{y=16-3x^2-z^2} dy dz dx.$$

Evaluation of this integral is straight-forward (but be careful). It is left as an exercise for readers. The answer should be $32\pi\sqrt{2}$.

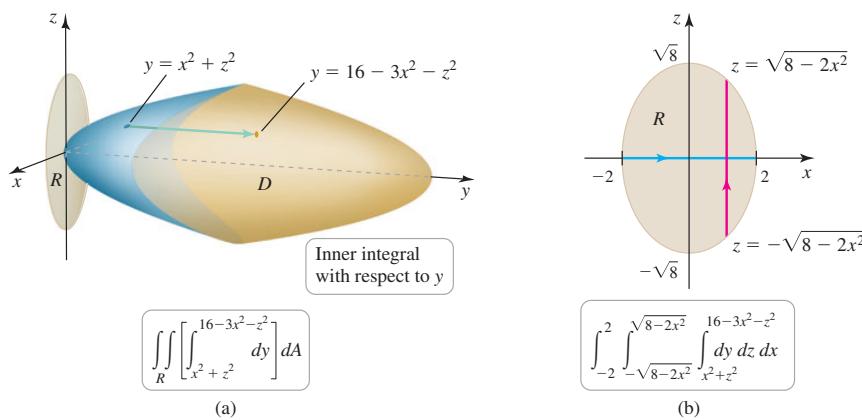


Figure 3.13: the pillar-base diagram for the triple integral in Example 3.9

Suppose D be the tetrahedron solid in the first octant bounded by the plane $2x + 3y + 6z = 12$ (see Figure 3.14). Now you are asked to evaluate

$$\iiint_D \frac{2}{12 - 3y - 6z} dV.$$

Which is the best order of integration? Why?

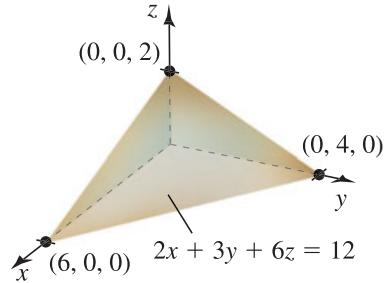


Figure 3.14: the tetrahedron D in the above question.

3.5 Triple Integrals in Cylindrical Coordinates

In order to set-up or compute a double integral of a *circular* region, it is often more convenient to convert the problem into polar coordinates. For triple integrals, there are also some solids that are not so “compatible” with rectangular coordinates but are easily to handle if one convert the problem into cylindrical or spherical coordinates. These solids include cylinders, cones, spheres, etc.

3.5.1 Cylindrical Coordinates

Cylindrical coordinates in \mathbb{R}^3 is simply combining the polar coordinates for the (x, y) -directions, and keeping the z coordinates. The conversion rule is given by:

$$\begin{aligned}x &= r \cos \theta \\y &= r \sin \theta \\z &= z\end{aligned}$$

where $r \geq 0$, $0 \leq \theta \leq 2\pi$ and z can be any real number. Just like polar coordinates, it is good to keep in mind that $x^2 + y^2 = r^2$, which will sometimes simplify your calculations. Figure 3.15 explains the geometry of these conversion rules.

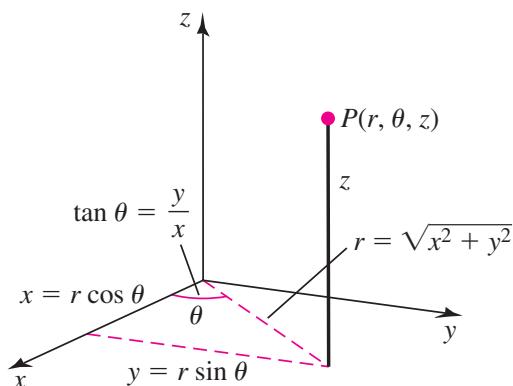


Figure 3.15: cylindrical coordinates

If one sets $r = \text{constant}$, then it describes an infinite cylinder in \mathbb{R}^3 with z -axis as the central axis. Therefore, if the solid is cylindrical in shape, it is usually easier to set-up a triple integral using cylindrical coordinates. Analogous to polar coordinates, the volume element dV is given by:

Theorem 3.3 Under cylindrical coordinates (r, θ, z) , we have:

$$dV = rdzdrd\theta.$$

Let's look at an example:

■ **Example 3.10** The volume of the solid bounded by two surfaces $z = 4 - 4(x^2 + y^2)$ and $z = (x^2 + y^2)^2 - 1$, as shown in Figure 3.16.

■ **Solution** When using cylindrical coordinates, one often (though not always) set z as the pillar variable, then r and θ become the base variable.

The z -pillar enters the solid at $z = (x^2 + y^2)^2 - 1$, which is $z = r^4 - 1$ in cylindrical coordinates, and leaves the solid at $z = 4 - 4(x^2 + y^2)$, which is $z = 4 - 4r^2$ in cylindrical coordinates. These determine the lower and upper limit of the inner integral.

The shadow is a circle centered at the origin. To find its radius, we solve:

$$r^4 - 1 = z = 4 - 4r^2$$

which gives $r = 1$. Therefore, the outer and middle integrals have upper and lower limits given by:

$$\int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=1}.$$

Combining all of the above, the volume of the solid is given by:

$$\begin{aligned} \iiint_D dV &= \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=1} \int_{z=r^4-1}^{z=4-4r^2} 1 \, r dz dr d\theta \\ &= \int_0^{2\pi} \int_0^1 [z]_{r^4-1}^{z=4-4r^2} r dr d\theta \\ &= \int_0^{2\pi} \int_0^1 (4 - 4r^2 - r^4 + 1)r \, dr d\theta \\ &= \int_0^{2\pi} \int_0^1 (5r - 4r^3 - r^5) \, dr d\theta \\ &= \int_0^{2\pi} \left[-\frac{5r^2}{2} - r^4 - \frac{r^6}{6} \right]_{r=0}^{r=1} d\theta \\ &= \int_0^{2\pi} \left(\frac{5}{2} - 1 - \frac{1}{6} \right) d\theta = \frac{8\pi}{3}. \end{aligned}$$

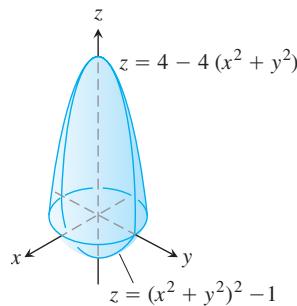
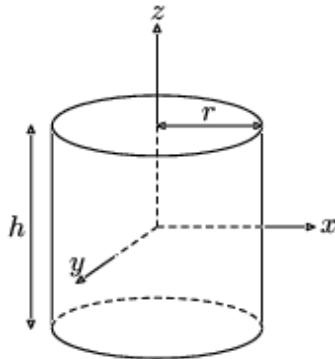


Figure 3.16: the solid in Example 3.10.

■ **Example 3.11** Let D be a solid cylinder of radius a with z -axis as the central axis, is bounded by planes $z = z_0$ and $z = z_0 + h$, where z_0 and h are constants. Therefore, the cylinder has height h . Suppose the solid has uniform density δ . Evaluate the following triple integral:

$$I_z := \iiint_D \delta(x^2 + y^2) dV$$

which is the **moment of inertia** of the solid about the z -axis.



■ **Solution** We choose the order of integration $dzdrd\theta$. The z -pillar enters the solid at $z = z_0$ and leaves at $z = z_0 + h$. The shadow is the circle with radius a centered at the origin. Therefore,

$$\begin{aligned}
 I_z &= \int_0^{2\pi} \int_0^a \int_{z_0}^{z_0+h} \delta(x^2 + y^2) \cdot r dz dr d\theta \\
 &= \int_0^{2\pi} \int_0^a \int_{z_0}^{z_0+h} \delta r^3 dz dr d\theta \\
 &= \int_0^{2\pi} \int_0^a \delta r^3 h dr d\theta \\
 &= \int_0^{2\pi} \left[\frac{\delta hr^4}{4} \right]_{r=0}^{r=a} d\theta \\
 &= \int_0^{2\pi} \frac{\delta ha^4}{4} d\theta \\
 &= \frac{\delta h \pi a^4}{2}.
 \end{aligned}$$

Most physics/engineering textbook expresses the moment of inertia in terms of the total mass m rather than the density δ . To rewrite the above answer in terms of m , we note that:

$$\delta = \frac{m}{V}$$

where V is the total volume of the solid, which is $\pi a^2 h$. Therefore, combining with the above calculation, we have:

$$I_z = \frac{m}{\pi a^2 h} \cdot \frac{h \pi a^4}{2} = \frac{ma^2}{2},$$

which is exactly what you can find in physics or engineering textbooks.

3.6 Triple Integrals in Spherical Coordinates

Another common coordinate system in physics and engineering is the spherical coordinates. A point in \mathbb{R}^3 can be represented by three numbers (ρ, θ, φ) which have the following geometric meaning:

ρ = distance from the point to the origin

θ = the projected angle on the xy -plane counting from the positive x -axis

φ = the angle counting from the positive z -axis

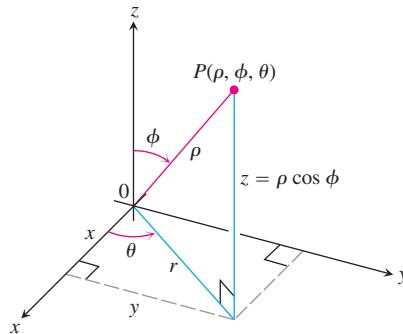


Figure 3.17: spherical coordinates

The ranges of (ρ, θ, φ) are:

$$\rho \geq 0, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \varphi \leq \pi.$$

From standard trigonometry, one can figure out the following conversion rules:

$$\begin{aligned} x &= \rho \sin \varphi \cos \theta \\ y &= \rho \sin \varphi \sin \theta \\ z &= \rho \cos \varphi \end{aligned}$$

Similar to polar and cylindrical coordinates, it is good to keep in mind that

$$\rho^2 = x^2 + y^2 + z^2.$$

The equation $\rho = \rho_0$, where ρ_0 is a positive constant, represents the sphere of radius ρ_0 centered at the origin. The equation $\varphi = \varphi_0$ represents a cone with angle φ_0 with the origin as its vertex. Therefore, the spherical coordinates are particularly useful when handling spherical or conical objects (see Figure 3.18).

To integrate using spherical coordinates, one should note that:

Theorem 3.4 Under the spherical coordinates (ρ, θ, φ) , the volume element is given by:

$$dV = \rho^2 \sin \varphi \, d\rho \, d\varphi \, d\theta.$$

Infinitesimally, $\rho^2 \sin \varphi \, d\rho \, d\varphi \, d\theta$ is the volume of the little cube in red in Figure 3.19. The volume of this little piece decreases as it approaches toward the north/south pole. It accounts for the factor $\sin \varphi$, which is smaller when φ is close to 0 or π , but large when φ is close to $\frac{\pi}{2}$.

Although a more precise explanation is available, we omit it in this note because this $\rho^2 \sin \varphi$ can be derived systematically using *Jacobian matrix*. Meanwhile let's look at some examples on how to use the spherical coordinates to do integration.

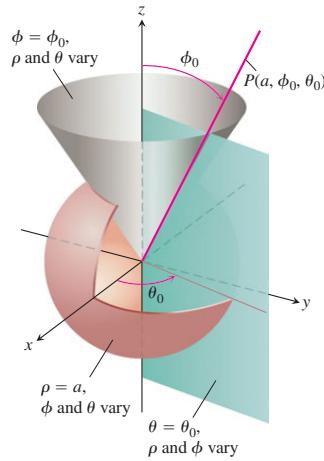
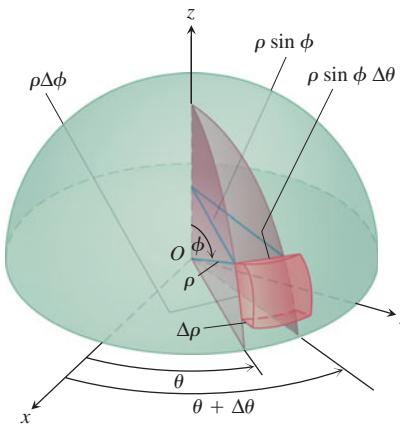
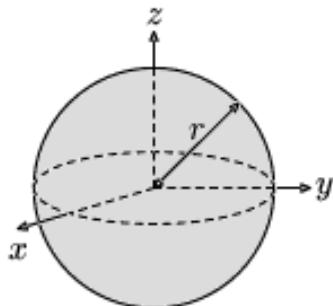


Figure 3.18: coordinate planes

Figure 3.19: geometric meaning of $dV = \rho^2 \sin \phi \, d\rho \, d\phi \, d\theta$.

■ **Example 3.12** Consider a solid sphere S of radius r centered at the origin. Suppose it has uniform density δ . Derive the moment of inertia about z -axis, which is given by:

$$I_z := \iiint_S \delta(x^2 + y^2) dV$$



■ **Solution** The solid sphere S can be written as $0 \leq \rho \leq r$ in spherical coordinates. As a full sphere, the ranges for the angles are:

$$0 \leq \theta \leq 2\pi, \quad 0 \leq \varphi \leq \pi.$$

Therefore, the set-up of the integral is:

$$I_z = \int_{\theta=0}^{\theta=2\pi} \int_{\varphi=0}^{\varphi=\pi} \int_{\rho=0}^{\rho=r} \delta(x^2 + y^2) \rho^2 \sin \varphi \, d\rho d\varphi d\theta.$$

To compute the integral, we need to express $x^2 + y^2$ using spherical coordinates:

$$x^2 + y^2 = (\rho \sin \varphi \cos \theta)^2 + (\rho \sin \varphi \sin \theta)^2 = \rho^2 \sin^2 \varphi.$$

Therefore,

$$\begin{aligned} I_z &= \int_0^{2\pi} \int_0^\pi \int_0^r \delta \cdot \rho^2 \sin^2 \varphi \cdot \rho^2 \sin \varphi \, d\rho d\varphi d\theta \\ &= \int_0^{2\pi} \int_0^\pi \int_0^r \delta \rho^4 \sin^3 \varphi \, d\rho d\varphi d\theta \\ &= \left(\int_0^{2\pi} d\theta \right) \left(\int_0^\pi \sin^3 \varphi \, d\varphi \right) \left(\int_0^r \delta \rho^4 \, d\rho \right) \\ &= 2\pi \cdot \left(\int_0^\pi (1 - \cos^2 \varphi) \sin \varphi \, d\varphi \right) \cdot \frac{\delta r^5}{5} \\ &= 2\pi \cdot \left[-\cos \varphi + \frac{\cos^3 \varphi}{3} \right]_0^\pi \cdot \frac{\delta r^5}{5} \\ &= 2\pi \cdot \left(1 - \frac{1}{3} + 1 - \frac{1}{3} \right) \cdot \frac{\delta r^5}{5}. \\ &= \frac{8\pi \delta r^5}{15}. \end{aligned}$$

Although it is a perfectly acceptable answer, the moment of inertia in many physics and engineering books is expressed in terms of the total mass m rather than of the density. Since the density in this problem is uniform, one can easily derive the moment of inertia formula in terms of m :

$$I_z = \frac{8\pi r^5}{15} \cdot \frac{m}{\frac{4\pi r^3}{3}} = \frac{2mr^2}{5}$$

which is what one can find in physics or engineering books.

Conical objects are another type of solids which are compatible with spherical coordinates. The inequality $0 \leq \varphi \leq \frac{\pi}{3}$ represents an infinite cone above the xy -plane with cone angle $\frac{\pi}{3}$ counting from the positive z -axis. If one further impose the inequality $0 \leq \rho \leq 1$ which represents the sphere with radius 1 centered at the origin, then the combined inequalities:

$$0 \leq \rho \leq 1, \quad 0 \leq \varphi \leq \frac{\pi}{3}$$

represents the common part of the sphere and the cone, which is in an ice-cream shape. Let's consider the following example:

■ **Example 3.13** Find the volume of the “ice-cream cone” D cut from the solid sphere $\rho \leq 1$ by the cone $\varphi = \frac{\pi}{3}$, as shown in Figure 3.20.

■ **Solution** The volume is given by $\iiint_D \rho^2 \sin \varphi d\rho d\varphi d\theta$. The limits of the inner-most integral are determined by where the ρ -ray, labeled M in the diagram, enters the solid and where it leaves the solid. Evidently, it enters from the origin $\rho = 0$. When it leaves, it always leaves from the sphere part and never the cone part, so the upper limit should be $\rho = 1$.

The φ -angle runs from 0 to $\frac{\pi}{3}$, and the θ -ray, labeled L in the diagram, sweeps over the shadow from 0 to 2π .

Combining all these limits, the volume is given by the integral:

$$\begin{aligned} \int_{\theta=0}^{\theta=2\pi} \int_{\varphi=0}^{\varphi=\pi/3} \int_{\rho=0}^{\rho=1} \rho^2 \sin \varphi d\rho d\varphi d\theta &= \left(\int_0^{2\pi} d\theta \right) \left(\int_0^{\pi/3} \sin \varphi d\varphi \right) \left(\int_0^1 \rho^2 d\rho \right) \\ &= 2\pi \cdot [-\cos \varphi]_0^{\pi/3} \cdot \frac{1}{3} \\ &= 2\pi \cdot \left(-\frac{1}{2} + 1 \right) \cdot \frac{1}{3} = \frac{\pi}{3}. \end{aligned}$$

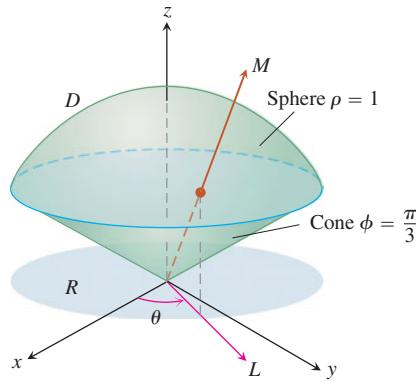
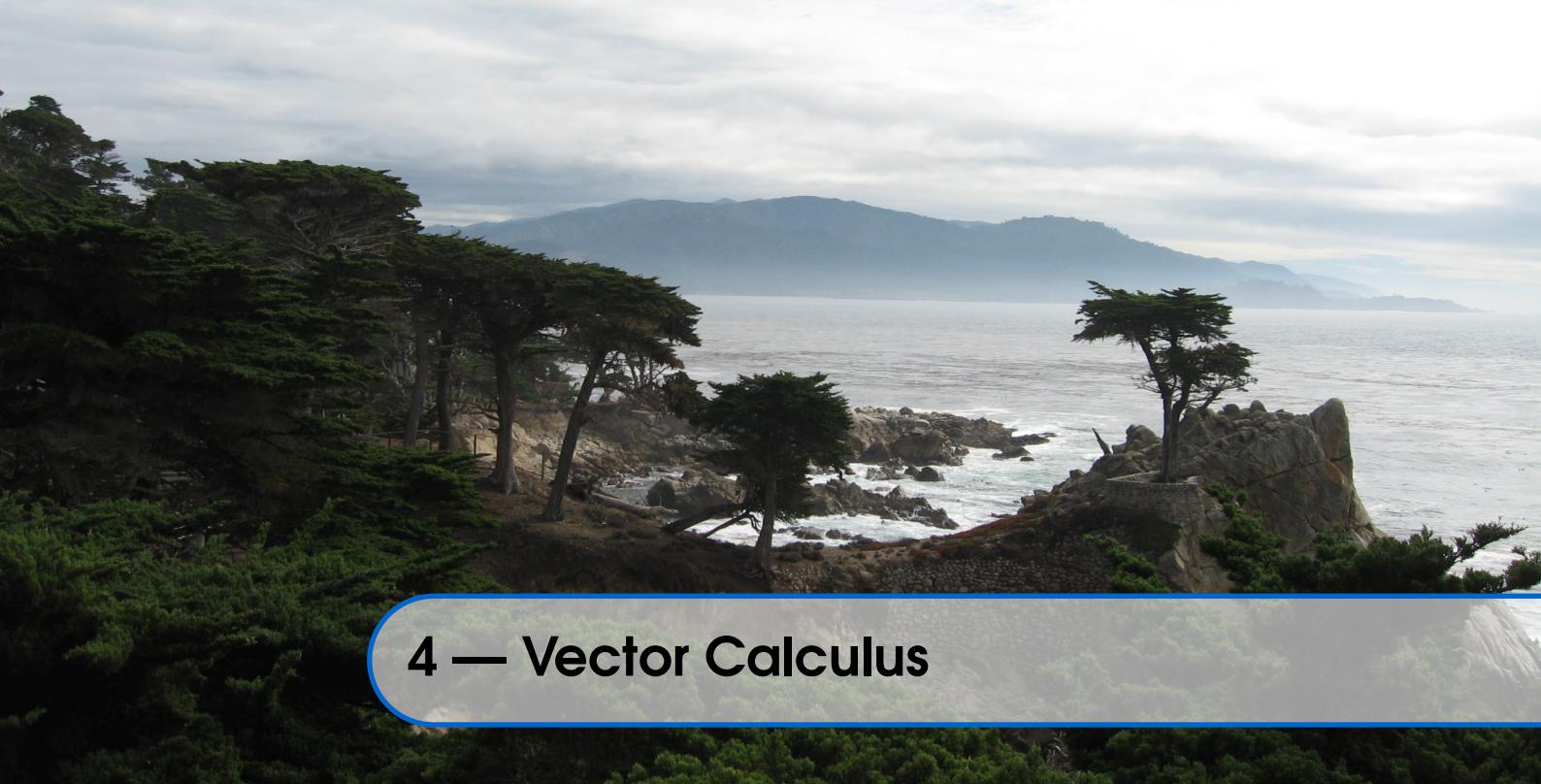


Figure 3.20: ice cream cone



4 — Vector Calculus

4.1 Vector Fields on \mathbb{R}^2 and \mathbb{R}^3

4.1.1 Examples of Vector Fields

Vector Calculus is an important tool in physics and engineering. Many concepts in physics, such as gravitational and electrostatic forces, fluid flow, heat flow, etc. are described using a mathematical concept called vector fields. Gravitational, magnetic and electric forces are not just one single vector. Both directions and magnitudes may vary from place to place and even changing over time. A vector field consists of a collection of vectors that are denoted by a vector-valued function $\mathbf{F}(x, y, z) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, which assigns to each point (x, y, z) a vector labeled $\mathbf{F}(x, y, z)$. For instance, the gravitational force field due to the Sun (whose center is at the origin) is given by:

$$\mathbf{F}(x, y, z) = -GMm \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{(x^2 + y^2 + z^2)^{3/2}},$$

where G , M and m are constants. That is to say, at the point $(1, 0, 0)$, the gravitational force is $-GMm \mathbf{i}$, whereas at the point $(1, 0, 1)$, the gravitation force is $-GMm \left(\frac{1}{2^{3/2}} \mathbf{i} + \frac{1}{2^{3/2}} \mathbf{k} \right)$.

With the help of computer software (such as Mathematica), one can visualize a vector field easily. A plot of the above gravitational field can be found in Figure 4.1.

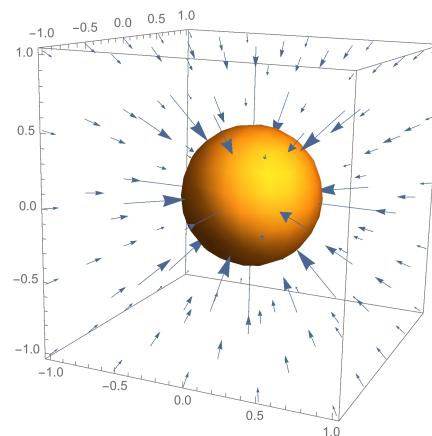


Figure 4.1: plot of the gravitational force \mathbf{F} , the central sphere is the Sun (not a part of the vector field)

The general form of a vector field in \mathbb{R}^3 is given by:

$$\mathbf{F}(x, y, z) = F_x(x, y, z)\mathbf{i} + F_y(x, y, z)\mathbf{j} + F_z(x, y, z)\mathbf{k}$$

where each of F_x , F_y and F_z is a scalar-valued function.

- i** Do NOT confuse F_x with the partial derivative $\frac{\partial F}{\partial x}$! Although we used the subscript notation F_x to denote a partial derivative before, we should avoid using it in this chapter. Here F_x means the x -component of the vector field \mathbf{F} .

Many vector fields in physics are three-dimensional. To start with, we will also study two-dimensional vector fields. A two-dimensional vector field is a collection of vectors in \mathbb{R}^2 that are denoted by a vector-valued function $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, which assigns to each point (x, y) a vector $\mathbf{F}(x, y)$. The general form of a two-dimensional vector field is given by:

$$\mathbf{F}(x, y) = F_x(x, y)\mathbf{i} + F_y(x, y)\mathbf{j}.$$

Figure 4.2 shows the plot of two examples of two-dimensional vector fields: $\mathbf{F}(x, y) = (1 - y^2)\mathbf{i}$ and $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$.

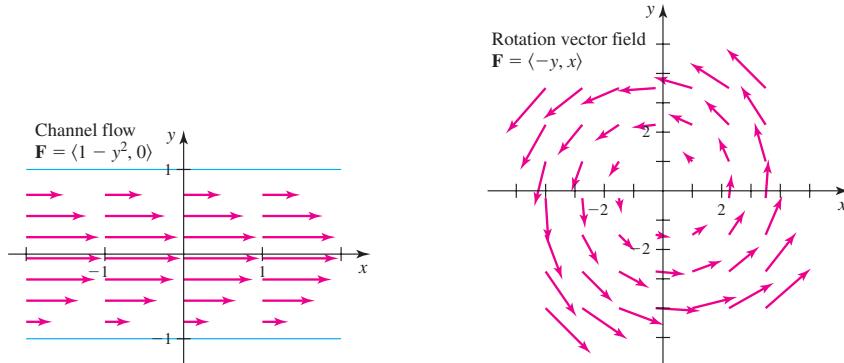


Figure 4.2: two examples of vector fields in \mathbb{R}^2

4.1.2 Vector Fields in Polar Coordinates

Sometimes it is more convenient to use other coordinate systems to represent a vector field, especially those which are radially symmetric (such as gravitation and electrostatic forces).

Think of \mathbf{i} and \mathbf{j} are the unit vectors in the increasing x and y directions in \mathbb{R}^2 respectively. One can also define a pair of *basis* vectors \mathbf{e}_r and \mathbf{e}_θ for polar coordinates, which are the unit vectors in the increasing r and θ directions respectively. See Figure 4.3.

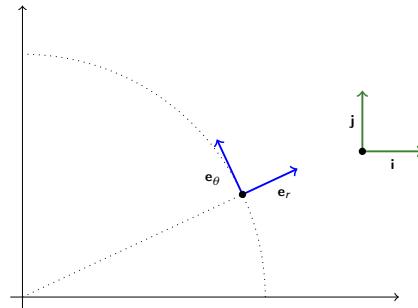


Figure 4.3: basis vectors \mathbf{e}_r and \mathbf{e}_θ in polar coordinates

This is worthwhile to note that unlike \mathbf{i} and \mathbf{j} in rectangular coordinates, the pair \mathbf{e}_r and \mathbf{e}_θ depends on the base point! To express \mathbf{e}_r and \mathbf{e}_θ in terms of \mathbf{i} and \mathbf{j} , one can do the following:

Take \mathbf{e}_θ as an example. It is the unit tangent vector along the θ -coordinate curve (with r fixed). Such a curve can be parametrized by:

$$\mathbf{r}(\theta) = x\mathbf{i} + y\mathbf{j} = (r \cos \theta)\mathbf{i} + (r \sin \theta)\mathbf{j}$$

where r is regarded to be a constant. The tangent vector is given by:

$$\mathbf{r}'(\theta) = (-r \sin \theta)\mathbf{i} + (r \cos \theta)\mathbf{j}.$$

Therefore, the unit tangent is given by:

$$\mathbf{e}_\theta = \frac{\mathbf{r}'(\theta)}{|\mathbf{r}'(\theta)|} = (-\sin \theta)\mathbf{i} + (\cos \theta)\mathbf{j}.$$

In essence, one can derive \mathbf{e}_θ by:

1. differentiate the position vector $x\mathbf{i} + y\mathbf{j}$ with respect to θ ;
2. then divide the resulting vector by its magnitude to get a unit vector.

Let's perform these procedures to derive the unit vector \mathbf{e}_r :

$$\begin{aligned}\frac{\partial}{\partial r}(x\mathbf{i} + y\mathbf{j}) &= \left(\frac{\partial(r \cos \theta)}{\partial r}\right)\mathbf{i} + \left(\frac{\partial(r \sin \theta)}{\partial r}\right)\mathbf{j} \\ &= (\cos \theta)\mathbf{i} + (\sin \theta)\mathbf{j}.\end{aligned}$$

Since it is a unit vector already, we conclude:

$$\mathbf{e}_r = (\cos \theta)\mathbf{i} + (\sin \theta)\mathbf{j}.$$

It is worthwhile to note that \mathbf{e}_r and \mathbf{e}_θ are orthogonal to each other, since $\mathbf{e}_r \cdot \mathbf{e}_\theta = 0$.

A general vector field $\mathbf{F}(r, \theta)$ in \mathbb{R}^2 in polar coordinates is given by:

$$\mathbf{F}(r, \theta) = F_r(r, \theta)\mathbf{e}_r + F_\theta(r, \theta)\mathbf{e}_\theta$$

where F_r and F_θ are scalar-valued functions, which represent the radial and rotational components of \mathbf{F} respectively.

It is sometimes more convenient to use polar coordinates to express a radial or a rotational vector field. For instance, if $F_\theta = 0$, then \mathbf{F} is a radial vector field, i.e.

$$\mathbf{F} = F_r(r, \theta)\mathbf{e}_r \quad (\text{radial})$$

On the other hand, if $F_r = 0$ then \mathbf{F} is a rotational vector field, i.e.

$$\mathbf{F} = F_\theta(r, \theta)\mathbf{e}_\theta \quad (\text{rotational})$$

A general vector field $\mathbf{F} = F_r\mathbf{e}_r + F_\theta\mathbf{e}_\theta$ can be rewritten in terms of \mathbf{i} and \mathbf{j} using the relations $\mathbf{e}_r = (\cos \theta)\mathbf{i} + (\sin \theta)\mathbf{j}$ and $\mathbf{e}_\theta = (-\sin \theta)\mathbf{i} + (\cos \theta)\mathbf{j}$:

$$\begin{aligned}\mathbf{F} &= F_r((\cos \theta)\mathbf{i} + (\sin \theta)\mathbf{j}) + F_\theta((- \sin \theta)\mathbf{i} + (\cos \theta)\mathbf{j}) \\ &= (F_r \cos \theta - F_\theta \sin \theta)\mathbf{i} + (F_r \sin \theta + F_\theta \cos \theta)\mathbf{j}.\end{aligned}$$

Since we also have $\mathbf{F} = F_x\mathbf{i} + F_y\mathbf{j}$, the conversion rule of vectors fields from polar to rectangular coordinates can be stated as:

$$\begin{aligned}F_x &= F_r \cos \theta - F_\theta \sin \theta = \frac{F_r x - F_\theta y}{\sqrt{x^2 + y^2}}, \\ F_y &= F_r \sin \theta + F_\theta \cos \theta = \frac{F_r y + F_\theta x}{\sqrt{x^2 + y^2}}.\end{aligned}$$

As an example, for a vector field $\mathbf{F} = r\mathbf{e}_\theta$, we have $F_r = 0$ and $F_\theta = r$. Therefore, from the above conversion rules, we have $F_x = -y$ and $F_y = x$. In rectangular coordinates, this vector field is expressed as $\mathbf{F} = -y\mathbf{i} + x\mathbf{j}$.

We next derive some conversion rules of vector field from rectangular to polar coordinates. First we need to write \mathbf{i} and \mathbf{j} in terms of \mathbf{e}_r and \mathbf{e}_θ . While it can be worked out by solving the equations $\mathbf{e}_r = (\cos \theta)\mathbf{i} + (\sin \theta)\mathbf{j}$ and $\mathbf{e}_\theta = (-\sin \theta)\mathbf{i} + (\cos \theta)\mathbf{j}$ for \mathbf{i} and \mathbf{j} , we adopt a more elegant approach.

Let $\mathbf{i} = A\mathbf{e}_r + B\mathbf{e}_\theta$ where A and B are to be determined. Observing that \mathbf{e}_r and \mathbf{e}_θ are orthogonal and they are unit vectors, we consider:

$$\begin{aligned}\mathbf{i} \cdot \mathbf{e}_r &= (A\mathbf{e}_r + B\mathbf{e}_\theta) \cdot \mathbf{e}_r \\ &= A\mathbf{e}_r \cdot \mathbf{e}_r + B\mathbf{e}_r \cdot \mathbf{e}_\theta \\ &= A + 0.\end{aligned}$$

Therefore, $A = \mathbf{i} \cdot \mathbf{e}_r$. Similarly, one can work out that $B = \mathbf{i} \cdot \mathbf{e}_\theta$. Hence, we have:

$$\mathbf{i} = (\mathbf{i} \cdot \mathbf{e}_r)\mathbf{e}_r + (\mathbf{i} \cdot \mathbf{e}_\theta)\mathbf{e}_\theta = (\cos \theta)\mathbf{e}_r - (\sin \theta)\mathbf{e}_\theta.$$

Analogously, one can also derive:

$$\mathbf{j} = (\mathbf{j} \cdot \mathbf{e}_r)\mathbf{e}_r + (\mathbf{j} \cdot \mathbf{e}_\theta)\mathbf{e}_\theta = (\sin \theta)\mathbf{e}_r + (\cos \theta)\mathbf{e}_\theta.$$

With these conversion rules for basis vectors, one can derive conversion rules for components of a vector field:

$$\begin{aligned}F_x \mathbf{i} + F_y \mathbf{j} &= F_r \mathbf{e}_r + F_\theta \mathbf{e}_\theta \\ F_x ((\cos \theta)\mathbf{e}_r - (\sin \theta)\mathbf{e}_\theta) + F_y ((\sin \theta)\mathbf{e}_r + (\cos \theta)\mathbf{e}_\theta) &= F_r \mathbf{e}_r + F_\theta \mathbf{e}_\theta \\ (F_x \cos \theta + F_y \sin \theta)\mathbf{e}_r + (-F_x \sin \theta + F_y \cos \theta)\mathbf{e}_\theta &= F_r \mathbf{e}_r + F_\theta \mathbf{e}_\theta.\end{aligned}$$

Equating the components, we have:

$$\begin{aligned}F_r &= F_x \cos \theta + F_y \sin \theta \\ F_\theta &= -F_x \sin \theta + F_y \cos \theta\end{aligned}$$

As an example, given a vector field $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$ in rectangular coordinates, then $F_x = x$ and $F_y = y$. Therefore,

$$\begin{aligned}F_r &= x \cos \theta + y \sin \theta = r \cos^2 \theta + r \sin^2 \theta = r \\ F_\theta &= -x \sin \theta + y \cos \theta = -r \cos \theta \sin \theta + r \sin \theta \cos \theta = 0\end{aligned}$$

In polar coordinates, the vector field can be expressed as $\mathbf{F} = r\mathbf{e}_r$, which is a radial vector field.

To summarize all conversion rules derived so far, we have:

Polar and Rectangular Conversions

Conversion rules of coordinates:

$$\begin{array}{ll}x = r \cos \theta & r = \sqrt{x^2 + y^2} \\ y = r \sin \theta & \theta = \tan^{-1} \frac{y}{x}\end{array}$$

Conversion rules of basis vectors:

$$\begin{array}{ll}\mathbf{i} = (\cos \theta)\mathbf{e}_r - (\sin \theta)\mathbf{e}_\theta & \mathbf{e}_r = (\cos \theta)\mathbf{i} + (\sin \theta)\mathbf{j} \\ \mathbf{j} = (\sin \theta)\mathbf{e}_r + (\cos \theta)\mathbf{e}_\theta & \mathbf{e}_\theta = (-\sin \theta)\mathbf{i} + (\cos \theta)\mathbf{j}\end{array}$$

Conversion rules of components of vector fields:

$$\begin{aligned} F_x &= \frac{F_r x}{\sqrt{x^2 + y^2}} - \frac{F_\theta y}{\sqrt{x^2 + y^2}} & F_r &= F_x \cos \theta + F_y \sin \theta \\ F_y &= \frac{F_r y}{\sqrt{x^2 + y^2}} + \frac{F_\theta x}{\sqrt{x^2 + y^2}} & F_\theta &= -F_x \sin \theta + F_y \cos \theta \end{aligned}$$

Generally, we do not usually express a vector field in polar coordinates unless it is radial or rotational. Take $\mathbf{F} = (1 - y^2)\mathbf{i}$ as an example, which we have $F_x = 1 - y^2$ and $F_y = 0$. In polar coordinates, the components are given by:

$$\begin{aligned} F_r &= (1 - y^2) \cos \theta = (1 - r^2 \sin^2 \theta) \cos \theta \\ F_\theta &= -(1 - y^2) \cos \theta = -(1 - r^2 \sin^2 \theta) \cos \theta \end{aligned}$$

Therefore, $\mathbf{F} = \{(1 - r^2 \sin^2 \theta) \cos \theta\}\mathbf{e}_r - \{(1 - r^2 \sin^2 \theta) \cos \theta\}\mathbf{e}_\theta$, which is way more complicated than it was in rectangular coordinates.

4.1.3 Vector Fields in Cylindrical Coordinates

In \mathbb{R}^3 , the cylindrical coordinates are essentially “polar” on the xy -plane, together with the vertical z variable. The basis vectors are denoted as \mathbf{e}_r , \mathbf{e}_θ and \mathbf{k} , which are unit vectors in the direction of increasing r , θ and z respectively. Note that these basis vectors are mutually perpendicular, meaning:

$$\mathbf{e}_r \cdot \mathbf{e}_\theta = 0, \quad \mathbf{e}_r \cdot \mathbf{k} = 0, \quad \mathbf{e}_\theta \cdot \mathbf{k} = 0.$$

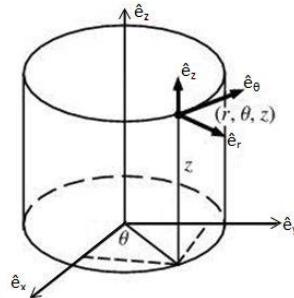


Figure 4.4: basis unit vectors in cylindrical coordinates

The general form of a vector field in \mathbb{R}^3 in cylindrical coordinates is:

$$\mathbf{F}(r, \theta, z) = F_r(r, \theta, z)\mathbf{e}_r + F_\theta(r, \theta, z)\mathbf{e}_\theta + F_z(r, \theta, z)\mathbf{k}$$

where F_r , F_θ and F_z are scalar-valued functions on \mathbb{R}^3 .

The conversion rules for cylindrical coordinates are inherited from those for polar coordinates in \mathbb{R}^2 , and the derivation of these conversion rules are almost the same. We omit the derivations here.

Cylindrical and Rectangular Conversions *Conversion rules of coordinates:*

$$\begin{aligned} x &= r \cos \theta & r &= \sqrt{x^2 + y^2} \\ y &= r \sin \theta & \theta &= \tan^{-1} \frac{y}{x} \\ z &= z & z &= z \end{aligned}$$

Conversion rules of basis vectors:

$$\begin{aligned}\mathbf{i} &= (\cos \theta) \mathbf{e}_r - (\sin \theta) \mathbf{e}_\theta \\ \mathbf{j} &= (\sin \theta) \mathbf{e}_r + (\cos \theta) \mathbf{e}_\theta \\ \mathbf{k} &= \mathbf{k}\end{aligned}$$

$$\begin{aligned}\mathbf{e}_r &= (\cos \theta) \mathbf{i} + (\sin \theta) \mathbf{j} \\ \mathbf{e}_\theta &= (-\sin \theta) \mathbf{i} + (\cos \theta) \mathbf{j} \\ \mathbf{k} &= \mathbf{k}\end{aligned}$$

Conversion rules of components of vector fields:

$$\begin{aligned}F_x &= \frac{F_r x}{\sqrt{x^2 + y^2}} - \frac{F_\theta y}{\sqrt{x^2 + y^2}} & F_r &= F_x \cos \theta + F_y \sin \theta \\ F_y &= \frac{F_r y}{\sqrt{x^2 + y^2}} + \frac{F_\theta x}{\sqrt{x^2 + y^2}} & F_\theta &= -F_x \sin \theta + F_y \cos \theta \\ F_z &= F_z & F_z &= F_z\end{aligned}$$

4.1.4 Vector Fields in Spherical Coordinates

In spherical coordinates (ρ, θ, φ) in \mathbb{R}^3 , we label the basis vectors by \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ . Let's demonstrate how to write \mathbf{e}_ρ in terms of \mathbf{i} , \mathbf{j} and \mathbf{k} explicitly, and leave the other two as exercises.

To find \mathbf{e}_ρ , just like in polar coordinates, we first differentiate the position vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ with respect to ρ :

$$\begin{aligned}\mathbf{r} &= x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = (\rho \sin \varphi \cos \theta) \mathbf{i} + (\rho \sin \varphi \sin \theta) \mathbf{j} + (\rho \cos \varphi) \mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial \rho} &= (\sin \varphi \cos \theta) \mathbf{i} + (\sin \varphi \sin \theta) \mathbf{j} + (\cos \varphi) \mathbf{k}\end{aligned}$$

Incidentally, $\frac{\partial \mathbf{r}}{\partial \rho}$ is unit, and so:

$$\mathbf{e}_\rho = (\sin \varphi \cos \theta) \mathbf{i} + (\sin \varphi \sin \theta) \mathbf{j} + (\cos \varphi) \mathbf{k}.$$

Similarly, differentiate \mathbf{r} with respect to θ or φ , and divide the resulting vector by its length to get a unit vector, one can get the expression of \mathbf{e}_θ and \mathbf{e}_φ as:

$$\begin{aligned}\mathbf{e}_\theta &= -(\sin \theta) \mathbf{i} + (\cos \theta) \mathbf{j} \\ \mathbf{e}_\varphi &= (\cos \varphi \cos \theta) \mathbf{i} + (\cos \varphi \sin \theta) \mathbf{j} - (\sin \varphi) \mathbf{k}\end{aligned}$$

Note that also \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ are mutually perpendicular:

$$\mathbf{e}_\rho \cdot \mathbf{e}_\theta = 0, \quad \mathbf{e}_\rho \cdot \mathbf{e}_\varphi = 0, \quad \mathbf{e}_\theta \cdot \mathbf{e}_\varphi = 0.$$

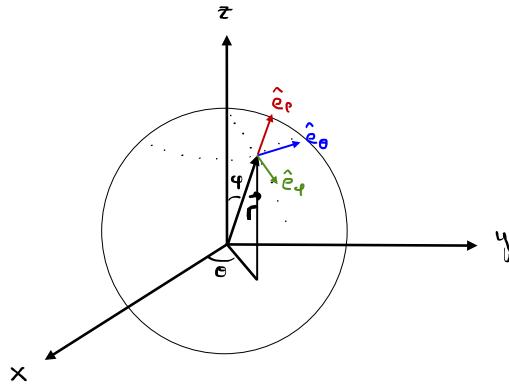


Figure 4.5: basis unit vectors in spherical coordinates

The general form of a vector field in spherical coordinates is given by:

$$\mathbf{F}(\rho, \theta, \varphi) = F_\rho(\rho, \theta, \varphi)\mathbf{e}_\rho + F_\theta(\rho, \theta, \varphi)\mathbf{e}_\theta + F_\varphi(\rho, \theta, \varphi)\mathbf{e}_\varphi.$$

A vector field of the form $\mathbf{F} = F_\rho \mathbf{e}_\rho$, i.e. $F_\theta = F_\varphi = 0$ at any point, is called a radial vector field. Using spherical coordinates, the gravitational force can be conveniently written as:

$$\mathbf{F} = -\frac{GMm}{\rho^2}\mathbf{e}_\rho,$$

which is much more elegant and concise than it were using rectangular coordinates.

Since \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ are orthonormal (i.e. mutually perpendicular and unit), one can mimic what we did for polar coordinates to derive conversion rules between spherical and rectangular coordinates.

For instance, the vector \mathbf{i} can be expressed in terms of \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ by considering:

$$\mathbf{i} = (\mathbf{i} \cdot \mathbf{e}_\rho)\mathbf{e}_\rho + (\mathbf{i} \cdot \mathbf{e}_\theta)\mathbf{e}_\theta + (\mathbf{i} \cdot \mathbf{e}_\varphi)\mathbf{e}_\varphi,$$

and so it only amounts to computing the dot products between \mathbf{i} and each of \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ .

\mathbf{j} and \mathbf{k} can be written in terms of \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ in a similar way.

After establishing the conversion rules between the basis vectors $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ and $\{\mathbf{e}_\rho, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$, one can derive the conversion rules between the components F_x, F_y, F_z and $F_\rho, F_\theta, F_\varphi$ (see the derivation for the polar-rectangular conversion rules).

We skip all the detail of derivation here since it is tedious and not enlightening.

Spherical and Rectangular Conversions *Conversion rules of coordinates:*

$$\begin{aligned} x &= \rho \sin \varphi \cos \theta & \rho &= \sqrt{x^2 + y^2 + z^2} \\ y &= \rho \sin \varphi \sin \theta & \theta &= \tan^{-1} \frac{y}{x} \\ z &= \rho \cos \varphi & \varphi &= \tan^{-1} \frac{\sqrt{x^2 + y^2}}{z} \end{aligned}$$

Conversion rules of basis vectors:

$$\begin{aligned} \mathbf{i} &= (\sin \varphi \cos \theta)\mathbf{e}_\rho + (\cos \varphi \cos \theta)\mathbf{e}_\varphi - (\sin \theta)\mathbf{e}_\theta \\ \mathbf{j} &= (\sin \varphi \cos \theta)\mathbf{e}_\rho + (\cos \varphi \sin \theta)\mathbf{e}_\varphi + (\cos \theta)\mathbf{e}_\theta \\ \mathbf{k} &= (\cos \varphi)\mathbf{e}_\rho - (\sin \varphi)\mathbf{e}_\varphi \end{aligned}$$

$$\begin{aligned} \mathbf{e}_\rho &= (\sin \varphi \cos \theta)\mathbf{i} + (\sin \varphi \sin \theta)\mathbf{j} + (\cos \varphi)\mathbf{k} \\ \mathbf{e}_\theta &= -(\sin \theta)\mathbf{i} + (\cos \theta)\mathbf{j} \\ \mathbf{e}_\varphi &= (\cos \varphi \cos \theta)\mathbf{i} + (\cos \varphi \sin \theta)\mathbf{j} - (\sin \varphi)\mathbf{k} \end{aligned}$$

Conversion rules of components of vector fields:

$$\begin{aligned} F_x &= (\sin \varphi \cos \theta)F_\rho + (\cos \varphi \cos \theta)F_\varphi - (\sin \theta)F_\theta \\ F_y &= (\sin \varphi \sin \theta)F_\rho + (\cos \varphi \sin \theta)F_\varphi + (\cos \theta)F_\theta \\ F_z &= (\cos \varphi)F_\rho - (\sin \varphi)F_\varphi \end{aligned}$$

$$\begin{aligned} F_\rho &= (\sin \varphi \cos \theta)F_x + (\sin \varphi \sin \theta)F_y + (\cos \varphi)F_z \\ F_\theta &= -(\sin \theta)F_x + (\cos \theta)F_y \\ F_\varphi &= (\cos \varphi \cos \theta)F_x + (\cos \varphi \sin \theta)F_y - (\sin \varphi)F_z \end{aligned}$$

We convert the gravitational force

$$\mathbf{F} = -\frac{GMm}{\rho^2}\mathbf{e}_\rho$$

into rectangular coordinates as an example. From the above expression, we have

$$F_\rho = -\frac{GMm}{\rho^2}, \quad F_\theta = 0, \quad F_\varphi = 0.$$

Therefore,

$$\begin{aligned} F_x &= -\frac{GMm}{\rho^2}(\sin \varphi \cos \theta) = -\frac{GMm}{\rho^2} \cdot \frac{x}{\rho} = -\frac{GMm}{\rho^3}x \\ F_y &= -\frac{GMm}{\rho^2}(\sin \varphi \sin \theta) = -\frac{GMm}{\rho^2} \cdot \frac{y}{\rho} = -\frac{GMm}{\rho^3}y \\ F_z &= -\frac{GMm}{\rho^2}(\cos \varphi) = -\frac{GMm}{\rho^2} \cdot \frac{z}{\rho} = -\frac{GMm}{\rho^3}z, \end{aligned}$$

and so:

$$\mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k} = -\frac{GMm}{\rho^3}(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = -GMm \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{(x^2 + y^2 + z^2)^{3/2}},$$

which is exactly the same as we stated in the previous subsection.

4.2 Line Integrals of Vector Fields

4.2.1 Definition of Line Integrals

Work is defined to be: force \times displacement. Precisely, if \mathbf{F} is a **constant** force, and $\Delta\mathbf{r} := \mathbf{r}_2 - \mathbf{r}_1$ is the displacement vector (which denotes the change of position from \mathbf{r}_1 to \mathbf{r}_2 , then:

$$\text{Work} = \mathbf{F} \cdot \Delta\mathbf{r}.$$

However, if the force is not a constant, meaning that either the direction or the magnitude is not uniform, the work done by the force is not as simple as stated above. Likewise, if the path is not a straight-path but a curved one so that $\Delta\mathbf{r}$ is changing over time, the work done by the force is again a bit more complicated. **Line integrals** of vector fields are introduced to handle these more complicated scenarios.

Definition 4.1 — Line Integrals of Vector Fields. Given a vector field $\mathbf{F}(x, y, z)$ and a path C which is parametrized by $\mathbf{r}(t)$, $a \leq t \leq b$, the line integral of \mathbf{F} over C is defined to be:

$$\int_a^b \mathbf{F} \cdot \mathbf{r}'(t) dt.$$

Notation In the spirit of $d\mathbf{r} = \mathbf{r}'(t) dt$, we may denote the above line integral in a concise way as:

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

where C is the given path.

Let's first look at some computational examples before we explain the physical and geometric meanings of line integrals.

■ **Example 4.1** Let $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$ and C be the counter-clockwise path along the circular arc from $(1, 0)$ to $(0, 1)$. Find the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$.

■ **Solution** First we parametrize C . A quick sketch of the path should tell us that C can be parametrized by:

$$\mathbf{r}(t) = (\cos t)\mathbf{i} + (\sin t)\mathbf{j}, \quad 0 \leq t \leq \frac{\pi}{2}.$$

By the definition of line integrals,

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_{t=0}^{t=\pi/2} \mathbf{F} \cdot \mathbf{r}'(t) dt \\ &= \int_0^{\pi/2} (-y\mathbf{i} + x\mathbf{j}) \cdot ((-\sin t)\mathbf{i} + (\cos t)\mathbf{j}) dt \\ &= \int_0^{\pi/2} (y \sin t + x \cos t) dt. \end{aligned}$$

Along the path C , we have $x = \cos t$ and $y = \sin t$ according to the parametrization, so we have:

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^{\pi/2} (\sin^2 t + \cos^2 t) dt \\ &= \int_0^{\pi/2} 1 dt = \frac{\pi}{2}. \end{aligned}$$

A line integral in three dimension can be computed in an exactly the same way as in two dimension. Let's see one example:

■ **Example 4.2** Let $\mathbf{F}(x, y, z) = (y - x^2)\mathbf{i} + (z - y^2)\mathbf{j} + (x - z^2)\mathbf{k}$ and C be a path parametrized by: $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$ where $0 \leq t \leq 1$. Compute the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$.

■ **Solution** We are given the parametrization in the problem so we can proceed to the computation of the line integral:

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^1 \langle y - x^2, z - y^2, x - z^2 \rangle \cdot \mathbf{r}'(t) dt \\ &= \int_0^1 \langle y - x^2, z - y^2, x - z^2 \rangle \cdot \langle 1, 2t, 3t^2 \rangle dt \\ &= \int_0^1 (y - x^2 + 2t(z - y^2) + 3t^2(x - z^2)) dt.\end{aligned}$$

Along the path C , we have $x = t$, $y = t^2$ and $z = t^3$, so:

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^1 (t^2 - t^2 + 2t(t^3 - t^4) + 3t^2(t - t^6)) dt \\ &= \int_0^1 (3t^3 + 2t^4 - 2t^5 - 3t^8) dt = \frac{29}{60}.\end{aligned}$$

4.2.2 Physical Meaning of Line Integrals

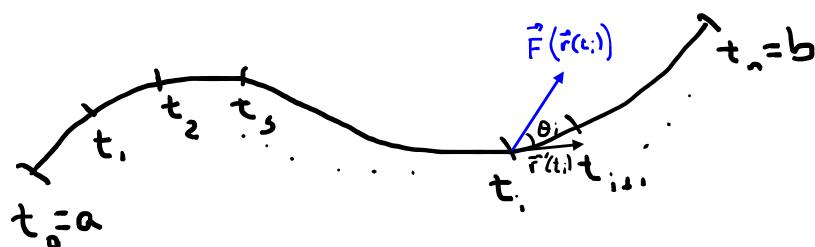
We introduce line integrals here because $\int_C \mathbf{F} \cdot d\mathbf{r}$ is exactly equal to the **work done by the force \mathbf{F} to move a particle from the starting point of C to the ending point of C** .

By the virtue of Riemann sums, the line integral can be thought as:

$$\int_a^b \mathbf{F} \cdot \mathbf{r}'(t) dt \simeq \sum_i \mathbf{F} \cdot \mathbf{r}'(t_i) \Delta t_i$$

where we divide the time interval $[a, b]$ into small subdivisions:

$$a =: t_0 < t_1 < \dots < t_i < \dots < t_n := b.$$



Assume each subdivision is very small that \mathbf{F} and \mathbf{r}' are roughly constants in each subdivision. By definition of derivatives, we have:

$$\mathbf{r}'(t_i) \simeq \frac{\Delta \mathbf{r}(t_i)}{\Delta t_i}.$$

Therefore,

$$\int_a^b \mathbf{F} \cdot \mathbf{r}'(t) dt \simeq \sum_i \mathbf{F} \cdot \Delta \mathbf{r}(t_i)$$

Since $\mathbf{F} \cdot \Delta \mathbf{r}(t_i)$ is the work done by \mathbf{F} with displacement $\Delta \mathbf{r}(t_i)$ (as they are roughly constants), summing them up gives the approximated total work done by the force over the whole path C . As $n \rightarrow \infty$, the subdivisions become infinitesimal and $\sum_i \mathbf{F} \cdot \Delta \mathbf{r}(t_i)$ becomes more accurate and approaches to the total work done by the force.

4.2.3 Geometric meaning of line integrals

The line integral of a given vector field \mathbf{F} over a path C can indicate whether the path C is overall *flowing along* the vector field. Recall that the line integral is given by:

$$\int_C \mathbf{F} \cdot \mathbf{r}'(t) dt.$$

The value of the integral is determined by many factors, including the length of C , the magnitude of \mathbf{F} and the velocity of $\mathbf{r}(t)$. However, the sign of $\mathbf{F} \cdot \mathbf{r}'(t)$ is solely determined by the angle θ between \mathbf{F} and the tangent vector $\mathbf{r}'(t)$ of the curve C , since:

$$\mathbf{F} \cdot \mathbf{r}'(t) = |\mathbf{F}| |\mathbf{r}'(t)| \cos \theta.$$

It is *positive* when \mathbf{F} and $\mathbf{r}'(t)$ make an acute angle, and is *negative* when they make an obtuse angle. Therefore, the sign of the line integral can *roughly* reveal whether the path C is *along* or *against* the direction of the vector field \mathbf{F} . The more positive is the value, the more often the path is traveling along the vector field.

This geometric meaning is very vague when compared to the physical meaning, but is a good one for us to develop intuition towards line integrals.

4.2.4 Independence of parametrization

Recall that when computing a line integral, we first pick a parametrization of the path. There is one logical gap we need to fill in, namely if we choose two different parametrizations of the same path, will we get the same answer for the line integral?

The answer is positive, as we can show it is true using the chain rule:

Suppose $\mathbf{r}(\tau)$, $a \leq \tau \leq b$, and $\mathbf{r}(t)$, $c \leq t \leq d$, are two parametrizations of a path C . Therefore, their endpoints must match, i.e. $\tau = a$ if and only if $t = c$, and $\tau = b$ if and only if $t = d$.

Then, using the τ -parametrization, the line integral is given by:

$$\int_{\tau=a}^{\tau=b} \mathbf{F} \cdot \mathbf{r}'(\tau) d\tau.$$

By chain rule, we have:

$$\mathbf{r}'(\tau) = \frac{d\mathbf{r}}{d\tau} = \frac{d\mathbf{r}}{dt} \frac{dt}{d\tau} = \mathbf{r}'(t) \frac{dt}{d\tau},$$

and so

$$\int_{\tau=a}^{\tau=b} \mathbf{F} \cdot \mathbf{r}'(\tau) d\tau = \int_{\tau=a}^{\tau=b} \mathbf{F} \cdot \mathbf{r}'(t) \frac{dt}{d\tau} d\tau.$$

By change of variables, we have $dt = \frac{dt}{d\tau} d\tau$ and so:

$$\int_{\tau=a}^{\tau=b} \mathbf{F} \cdot \mathbf{r}'(\tau) d\tau = \int_{t=c}^{t=d} \mathbf{F} \cdot \mathbf{r}'(t) dt$$

which is exactly how we defined line integral using t as the parameter.

The above shows that no matter how we parametrize the path, as far as the endpoints of the path are kept unchanged, the line integral must be the same. We call this **independence of parametrization**. Physically speaking, this tells us as far as a particle travels along a fixed path C , the work done by the force does not depend on how fast the particle travels.

4.2.5 Alternative notations for line integrals

In some textbooks, a line integral is denoted using the arc-length parametrization. Recall that a path $\mathbf{r}(s)$ is said to be arc-length parametrized if $|\mathbf{r}'(s)| = 1$ for any s , i.e. unit speed. For such a parametrization, it is a convention to use s as a parameter.

Since the line integral of a given vector field \mathbf{F} over a path C is independent of parametrization, we can also denote the line integral using the s parameter:

$$\int_C \mathbf{F} \cdot \mathbf{r}'(s) ds.$$

The value is the same using any other parametrizations.

Since $\mathbf{r}'(s)$ is a unit vector, conventionally it is often denoted by $\hat{\mathbf{T}}$ inside a line integral, meaning that it is a unit tangent vector of the path C . Therefore, you may see that occasionally the line integral is denoted by:

$$\int_C \mathbf{F} \cdot \hat{\mathbf{T}} ds.$$

It simply means $\int_C \mathbf{F} \cdot d\mathbf{r}$.

Another common notation for line integrals is so-called the *differential form* notation. Recall that \mathbf{r} is the position vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, and so symbolically $d\mathbf{r}$ can be regarded as:

$$d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}.$$

Let $\mathbf{F} = F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k}$, then (again symbolically)

$$\mathbf{F} \cdot d\mathbf{r} = F_x dx + F_y dy + F_z dz.$$

Therefore, another common notation for line integrals is:

$$\int_C F_x dx + F_y dy + F_z dz.$$

Again, it simply means $\int_C \mathbf{F} \cdot d\mathbf{r}$.

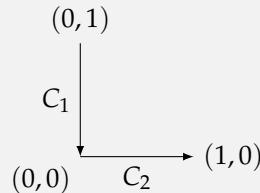
While the ds -notation does not have much practical use, there are some practical advantages for using the differential form notations if the path C is parallel to one of the coordinate axes. Let's illustrate this through an example:

■ **Example 4.3** Compute the line integral

$$\int_C -y dx + x dy$$

where C is the path from $(0, 1)$ down to $(0, 0)$ along the y -axis, then to $(1, 0)$ along the x -axis.

■ **Solution** Note that the path C has two segments. Break C into two segments C_1 and C_2 where C_1 is the path from $(0, 1)$ to $(0, 0)$ along the y -axis, and C_2 is the path from $(0, 0)$ to $(1, 0)$ along the x -axis.



The line integral is broken down into two:

$$\int_C -y dx + x dy = \int_{C_1} -y dx + x dy + \int_{C_2} -y dx + x dy.$$

Along C_1 , we have $x = 0$, and so $dx = 0$ too. From this we can immediately tell that

$$\int_{C_1} -y \, dx + x \, dy = \int_{C_1} -y \, d(0) + 0 \, dy = 0.$$

Similarly, along C_2 , we have $y = 0$, and so:

$$\int_{C_2} -y \, dx + x \, dy = \int_{C_2} -0 \, dx + x \, d(0) = 0.$$

Combining the two results, we know:

$$\int_C -y \, dx + x \, dy = 0.$$

To conclude, given a path C parametrized by $\mathbf{r}(t)$ and a vector field $\mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k}$, then the following are all equivalent notations for line integrals:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{r}'(t) \, dt = \int_C \mathbf{F} \cdot \hat{\mathbf{T}} \, ds = \int_C F_x \, dx + F_y \, dy + F_z \, dz.$$

4.3 Conservative Vector Fields

4.3.1 Definition and Consequences

This section discusses a special type of fields called *conservative vector fields*. The term *conservative* is rooted from physics, not politics! We will first study its definition, and then investigate the features that make conservative vector fields distinguished from generic ones.

Definition 4.2 — Conservative Vector Field. A vector field \mathbf{F} is called a conservative vector field if and only if it is in the form of $\mathbf{F} = \nabla f$ where f is a scalar function. The scalar function f is called a **potential function** of the vector field \mathbf{F} .

Recall that ∇f denotes the gradient vector of f , defined by:

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}.$$

The most preliminary method to determine whether a given vector field is conservative is to solve for the scalar potential f , as illustrated by the following example:

■ **Example 4.4** Consider the vector field

$$\mathbf{F}(x, y, z) = (2x + y)\mathbf{i} + (x + z^3)\mathbf{j} + (3yz^2 + 1)\mathbf{k}.$$

Determine whether or not \mathbf{F} is a conservative vector field. If so, find its potential function f such that $\mathbf{F} = \nabla f$.

■ **Solution** \mathbf{F} is conservative if and only if $\mathbf{F} = \nabla f$ for some scalar function f , or equivalently, the following equations hold simultaneously:

$$\begin{aligned} \textcircled{1} \quad \frac{\partial f}{\partial x} &= 2x + y \\ \textcircled{2} \quad \frac{\partial f}{\partial y} &= x + z^3 \\ \textcircled{3} \quad \frac{\partial f}{\partial z} &= 3yz^2 + 1 \end{aligned}$$

From $\textcircled{1}$, one can find $f(x, y, z)$ by integrating $2x + y$ by x , regarding y to be a constant. In single variable calculus, an integration constant will be added after the integration. However, we are now considering partial derivatives, and not only constants but also y or z will vanish after differentiating by x ! In other words, the “integration constant” is no longer a constant but instead a function of y and z . Precisely, we have:

$$\textcircled{4} \quad f(x, y, z) = x^2 + yx + g(y, z)$$

where $g(y, z)$ is some function of y and z . We will figure out $g(y, z)$ in the remaining steps.

By differentiating both sides of $\textcircled{4}$ with respect to y , we get:

$$\frac{\partial f}{\partial y} = x + \frac{\partial g}{\partial y}.$$

Compare this with $\textcircled{2}$, one need to have

$$\frac{\partial g}{\partial y} = z^3.$$

An integration by y yields:

$$g(y, z) = yz^3 + h(z).$$

Note that from similar principle discussed above for f , the integration “constant” is no longer just a constant but a function not depending on y and hence a function of z only.

By differentiating both sides with respect to z , we get:

$$\frac{\partial g}{\partial z} = 3yz^2 + h'(z).$$

Finally, by comparing this result with ③, one must have $h'(z) = 1$, and so clearly $h(z) = z + C$, where C is genuinely a constant this time!

Combining all results above, we have

$$f(x, y, z) = x^2 + yx + yz^3 + z + C$$

where C is any real constant.

It can be easily checked that $\mathbf{F} = \nabla f$. Therefore, \mathbf{F} is a conservative vector field with potential functions given by the above f 's.

It is important to keep in mind that not all vector fields are conservative! Here is one which such an f does not exist:

■ **Example 4.5** Let $\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$. Determine whether or not \mathbf{F} is a conservative vector field. If so, find its potential function f such that $\mathbf{F} = \nabla f$.

■ **Solution** If f is a scalar function such that $\mathbf{F} = \nabla f$, then we have:

$$\begin{aligned} \textcircled{1} \quad \frac{\partial f}{\partial x} &= -y \\ \textcircled{2} \quad \frac{\partial f}{\partial y} &= x \end{aligned}$$

Solving ① for f by integration, we get $f(x, y) = -xy + g(y)$ for some function $g(y)$. However, by differentiating this result with respect to y , we get:

$$\frac{\partial f}{\partial y} = -x + g'(y).$$

In order to be consistent with ②, we would require

$$-x + g'(y) = x, \text{ or equivalently, } g'(y) = 2x.$$

However, $g'(y)$ is a function of y , not of x ! Since it leads to inconsistency, such an f cannot exist and so \mathbf{F} is not conservative.

One important feature of conservative vector fields is the **path independence** of line integral, meaning that the line integral depends only on the end-points of the path. Precisely, we have:

Theorem 4.1 Given a conservative vector field $\mathbf{F} = \nabla f$, where f is a potential function, then along any path C connecting from point $P_0(x_0, y_0, z_0)$ to point $P_1(x_1, y_1, z_1)$, then the line integral is given by:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = f(x_1, y_1, z_1) - f(x_0, y_0, z_0).$$

Proof. The proof is a consequence of the multivariable chain rule and the Fundamental Theorem of Calculus.

Suppose $\mathbf{F} = \nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle$ and the path C is parametrized by:

$$\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle, \quad a \leq t \leq b.$$

Then,

$$\begin{aligned}
 \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_C \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} dt \\
 &= \int_C \nabla f \cdot \frac{d\mathbf{r}}{dt} dt \\
 &= \int_a^b \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle \cdot \left\langle \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right\rangle dt \\
 &= \int_a^b \left(\frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt} \right) dt \\
 &= \int_a^b \frac{d}{dt} f(\mathbf{r}(t)) dt \quad (\text{chain rule}) \\
 &= f(\mathbf{r}(b)) - f(\mathbf{r}(a)) \quad (\text{Fundamental Theorem of Calculus})
 \end{aligned}$$

If (x_0, y_0, z_0) and (x_1, y_1, z_1) are the initial and final positions respectively, then $\mathbf{r}(a) = \langle x_0, y_0, z_0 \rangle$ and $\mathbf{r}(b) = \langle x_1, y_1, z_1 \rangle$. Therefore,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = f(x_1, y_1, z_1) - f(x_0, y_0, z_0).$$

■

The significance of this theorem is that the RHS depends only on the initial and final points of the curve C , but not the intermediate path. In other words, if C_1 and C_2 are two paths with the same initial and final points, then we have $\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}$. Moreover, if C is closed path whose initial and final positions are the same, then $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$.

Notation If C is a closed path, meaning that the two endpoints are the same, it is a convention to use denote the line integral as:

$$\oint_C \mathbf{F} \cdot d\mathbf{r}.$$

To summarize, we have:

Corollary 4.2 For a conservative vector field \mathbf{F} , if C_1 and C_2 are two paths with the same initial and final positions, then

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

Moreover, if C is a closed path, then

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

Theorem 4.1 can be applied to find the line integral of a conservative vector field if path is too complicated.

■ **Example 4.6** Consider the vector field $\mathbf{F}(x, y, z) = (2x + y)\mathbf{i} + (x + z^3)\mathbf{j} + (3yz^2 + 1)\mathbf{k}$ which appeared in Example 4.4, and the path C given by the parametric equation

$$\mathbf{r}(t) = (e^t \sin^{2012} t)\mathbf{i} + (\cos^{2013} t)\mathbf{j} + \frac{t}{\pi}\mathbf{k}, \quad 0 \leq t \leq 2\pi.$$

Find the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$.

■ **Solution** It is clear that direct computation of this line integral is extremely laborious (and may be impossible). Fortunately, it was shown in Example 4.4 that \mathbf{F} is a conservative vector field with potential function

$$f(x, y, z) = x^2 + yx + yz^3 + z + C.$$

The initial and final positions of the given path are respectively:

$$\begin{aligned}\mathbf{r}(0) &= \langle 0, 1, 0 \rangle \\ \mathbf{r}(2\pi) &= \langle 0, 1, 2 \rangle.\end{aligned}$$

By Theorem 4.1, the line integral in question is simply given by:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = f(0, 1, 2) - f(0, 1, 0) = (10 + C) - C = 10.$$

Alternatively, one can also find this line integral using the path independence property of conservative vector fields. Let L be the straight path connecting from $(0, 1, 0)$ to $(0, 1, 2)$ which are the initial and final positions respectively. Since \mathbf{F} is conservative, we must have:

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

The latter is much easier to figure out: along the line L , we have

$$x \equiv 0, \quad y \equiv 1, \quad 0 \leq z \leq 2.$$

Therefore

$$\begin{aligned}\int_L \mathbf{F} \cdot d\mathbf{r} &= \int_L (2x + y)dx + (x + z^3)dy + (3yz^2 + 1)dz \\ &= \int_{z=0}^{z=2} (0 + 1)d(0) + (0 + 2^3)d(1) + (3 \cdot 1 \cdot z^2 + 1)dz \\ &= \int_0^2 (3z^2 + 1)dz \\ &= z^3 + z \Big|_0^2 = 10.\end{aligned}$$

By path independence, $\int_C \mathbf{F} \cdot d\mathbf{r} = \int_L \mathbf{F} \cdot d\mathbf{r} = 10$.

4.3.2 Conservation of Total Energy

In physics, it is a convention to define the potential function V of a conservative vector fields $\mathbf{F}(x, y, z)$ by:

$$\mathbf{F}(x, y, z) = -\nabla V(x, y, z).$$

The scalar potential function V is often called by physicists the **potential energy**.

Let $\mathbf{r}(t)$ represent the path of a particle with mass m . Its **kinetic energy** is defined as:

$$\text{KE} = \frac{1}{2}m|\mathbf{r}'(t)|^2.$$

It is a nice exercise to show that the total energy is *conserved*:

■ **Exercise 4.1** Let $\mathbf{F} = -\nabla V$ and $\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle$, show that the total energy:

$$E(t) := \frac{1}{2}m|\mathbf{r}'(t)|^2 + V(\mathbf{r}(t))$$

is a constant. [Hint: multivariable chain rule and Newton's Second Law $\mathbf{F} = m\mathbf{r}''(t)$.]

4.3.3 Gradient Vector in Various Coordinate Systems

The gravitational force field between the Earth (with mass M) and a point particle (with mass m) is given by:

$$\mathbf{F}(x, y, z) = -GmM \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{(x^2 + y^2 + z^2)^{3/2}}$$

where G is the gravitational constant, and the (x, y, z) coordinates are chosen so that $(0, 0, 0)$ is the center of the Earth.

It is a very clumsy expression, but we have seen in the previous section, it becomes more succinct if one converts it into spherical coordinates:

$$\mathbf{F} = -\frac{GMm}{\rho^2} \mathbf{e}_\rho.$$

To determine whether this kind of radially symmetric vector field is conservative and find its potential function, it is more efficient to use the gradient ∇f in spherical coordinates form:

$$\nabla f = \frac{\partial f}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \mathbf{e}_\varphi + \frac{1}{\rho \sin \varphi} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta.$$

The derivation of the above conversion is straight-forward but quite tedious. One can first write

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k},$$

then convert each of \mathbf{i} , \mathbf{j} and \mathbf{k} into \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ . Secondly, use the chain rule to express each of $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$ in terms of $\frac{\partial f}{\partial \rho}$, $\frac{\partial f}{\partial \theta}$ and $\frac{\partial f}{\partial \varphi}$.

We omit this tedious derivation here (readers may verify this as an exercise if you wish). Here we are going to use the spherical form of ∇ to find a potential function for \mathbf{F} .

Firstly, set $\mathbf{F} = \nabla f$ and try to solve for f :

$$-\frac{GMm}{\rho^2} \mathbf{e}_\rho = \nabla f = \frac{\partial f}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \mathbf{e}_\varphi + \frac{1}{\rho \sin \varphi} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta.$$

Equating the components, we get a system of equations:

$$\begin{aligned} -\frac{GMm}{\rho^2} &= \frac{\partial f}{\partial \rho} \\ 0 &= \frac{\partial f}{\partial \varphi} \\ 0 &= \frac{\partial f}{\partial \theta} \end{aligned}$$

The first equation gives

$$f(\rho, \theta, \varphi) = \frac{GMm}{\rho} + g(\theta, \varphi)$$

where $g(\theta, \varphi)$ is an arbitrary function of θ and φ . However, the second and the third equations demand that

$$\frac{\partial g}{\partial \theta} = \frac{\partial g}{\partial \varphi} = 0.$$

Therefore, g must be a constant.

To conclude, the gravitational force field \mathbf{F} is conservative with potential function $f = \frac{GMm}{\rho}$. Many physicists prefer using the convention $\mathbf{F} = -\nabla V$ when defining the potential energy V . Hence, the gravitational potential energy is given by:

$$V = -\frac{GMm}{\rho}.$$

Likewise, there is an analogous expression for ∇f in cylindrical coordinates:

$$\nabla f = \frac{\partial f}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta + \frac{\partial f}{\partial z} \mathbf{k}$$

Using this expression, the vector field $\mathbf{F} = r\mathbf{e}_\theta$ (which is equal to $-y\mathbf{i} + x\mathbf{j}$ in rectangular coordinates) can be seen to be not conservative. Set $\mathbf{F} = \nabla f$, then:

$$r\mathbf{e}_\theta = \frac{\partial f}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta + \frac{\partial f}{\partial z} \mathbf{k},$$

and so:

$$\begin{aligned} 0 &= \frac{\partial f}{\partial r} \\ r &= \frac{1}{r} \frac{\partial f}{\partial \theta} \\ 0 &= \frac{\partial f}{\partial z} \end{aligned}$$

The second equation shows $f(r, \theta, z) = r^2\theta + g(r, z)$ for some arbitrary function g . Then, we will have $\frac{\partial f}{\partial z} = \frac{\partial g}{\partial z}$, and so the third equation demands that $\frac{\partial g}{\partial z} = 0$. Therefore, g depends only on r , and we can simply write $g = g(r)$. Finally, we have $f(r, \theta, z) = r^2\theta + g(r)$. Differentiation with respect to r gives $\frac{\partial f}{\partial r} = 2r\theta + g'(r)$. Then the first equation demands that:

$$0 = 2r\theta + g'(r), \text{ or equivalently: } \frac{g'(r)}{2r} = -\theta.$$

However, the LHS would be a function of r but the RHS is a function of θ . This cannot be true. This shows the vector field \mathbf{F} is not conservative.

4.3.4 Curl Test

It is possible to test whether a given vector field is conservative or not without attempting to find its potential function. This test is commonly called the **curl test**. The use of term *curl* comes from the fact that it involves calculations of the curl of a vector field.

Definition 4.3 — Curl of a Vector Field. Given a vector field

$$\mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k},$$

the curl of the vector field \mathbf{F} , denoted either by $\text{curl}(\mathbf{F})$ or $\nabla \times \mathbf{F}$, is defined as:

$$\begin{aligned} \nabla \times \mathbf{F} &= \left(\frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \right) \times (F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k}) \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \\ &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathbf{k}. \end{aligned}$$

- i** The symbol $\nabla := \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}$ should be considered as an operator rather than a vector. It carries no physical or geometric meaning. "Multiplying" $\frac{\partial}{\partial x}$ by a function P gives the partial derivative $\frac{\partial P}{\partial x}$ as the product.

- i** If $\mathbf{F} = F_x\mathbf{i} + F_y\mathbf{j}$ is a two dimensional vector field, the curl $\nabla \times \mathbf{F}$ can also be defined by regarding the \mathbf{k} -component to be zero, i.e.

$$\mathbf{F} = F_x\mathbf{i} + F_y\mathbf{j} + 0\mathbf{k}.$$

It can be easily verified that

$$\nabla \times \mathbf{F} = \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathbf{k}.$$

- i** $\nabla \times \mathbf{F}$ is called the *curl* of \mathbf{F} because it measures how circular the vector field \mathbf{F} is, as a consequence of the Green's and Stokes' Theorems which we will learn very soon.

■ **Example 4.7** Compute the curl of the vector field:

$$\mathbf{F} = (2x + y)\mathbf{i} + (x + z^3)\mathbf{j} + (3yz^2 + 1)\mathbf{k}.$$

■ **Solution**

$$\begin{aligned} \nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2x + y & x + z^3 & 3yz^2 + 1 \end{vmatrix} \\ &= \left(\frac{\partial}{\partial y}(3yz^2 + 1) - \frac{\partial}{\partial z}(x + z^3) \right) \mathbf{i} + \left(\frac{\partial}{\partial z}(2x + y) - \frac{\partial}{\partial x}(3yz^2 + 1) \right) \mathbf{j} \\ &\quad + \left(\frac{\partial}{\partial x}(x + z^3) - \frac{\partial}{\partial y}(2x + y) \right) \mathbf{k} \\ &= (3z^2 - 3z^2)\mathbf{i} + (0 - 0)\mathbf{j} + (1 - 1)\mathbf{k} = \mathbf{0}. \end{aligned}$$

The curl test is a very straight-forward test to check whether a given vector field \mathbf{F} is conservative, without solving for the potential function f .

Theorem 4.3 — Curl Test. Given a vector field \mathbf{F} is defined and continuously differentiable everywhere in \mathbb{R}^3 (or everywhere in \mathbb{R}^2 for vector fields in \mathbb{R}^2), then \mathbf{F} is conservative if and only if $\nabla \times \mathbf{F} = \mathbf{0}$.

Proof. The theorem has two directions (if and only if). One direction is easy, while the other direction is more technical. We only present the (\Rightarrow)-part here

The (\Rightarrow)-part of the proof is a consequence of the Mixed Partial Theorem. Suppose $\mathbf{F} = \nabla f$, then $\mathbf{F} = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}$, and so

$$\begin{aligned} \nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} \\ &= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y} \right) \mathbf{i} + \left(\frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z} \right) \mathbf{j} + \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) \mathbf{k} \\ &= 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k} = \mathbf{0}. \end{aligned}$$

■

Therefore, to check whether or not \mathbf{F} is conservative assuming it is defined and smooth everywhere, it is not necessary to solve for the potential function f . All is needed is to find $\nabla \times \mathbf{F}$ which only involves *differentiation* but not *integration*. However, the curl test only tells you whether or not the vector field is conservative. It fails to tell you what the potential function is.

In Example 4.7, the vector field \mathbf{F} is defined (and is smooth) everywhere in \mathbb{R}^3 . Therefore, the curl test can be applied on this \mathbf{F} . Since $\nabla \times \mathbf{F} = \mathbf{0}$ as shown in the example, the curl test asserts that \mathbf{F} is conservative. Note that the curl test only asserts *whether or not* \mathbf{F} is conservative. It provides no information about the potential function f . However, knowing that \mathbf{F} is conservative without knowing the potential f is still enlightening as demonstrated in the alternative solution of Problem 4.6: one can calculate the line integral by choosing an easier path.

On the other hand, if $\mathbf{G} = -y\mathbf{i} + x\mathbf{j}$ which is defined and is smooth everywhere in \mathbb{R}^2 , the curl test can be applied to \mathbf{G} . One can verify that $\nabla \times \mathbf{G} = 2\mathbf{k} \neq \mathbf{0}$, so the curl test concludes that \mathbf{G} is *not* conservative. There is no need to attempt solving the potential function, as it cannot never exist.

It is worthwhile to note that the condition \mathbf{F} is defined and continuously differentiable everywhere is crucial when applying the curl test. The following example tells us why:

Let

$$\mathbf{F}(x, y) = -\frac{y}{x^2 + y^2}\mathbf{i} + \frac{x}{x^2 + y^2}\mathbf{j}.$$

This vector field is not defined when $(x, y) \neq (0, 0)$. It can be easily verified (left as an exercise) that $\nabla \times \mathbf{F} = \mathbf{0}$ for any $(x, y) \neq (0, 0)$. However, when integrating along the unit circle $C : \mathbf{r}(t) = (\cos t)\mathbf{i} + (\sin t)\mathbf{j}$ where $0 \leq t \leq 2\pi$:

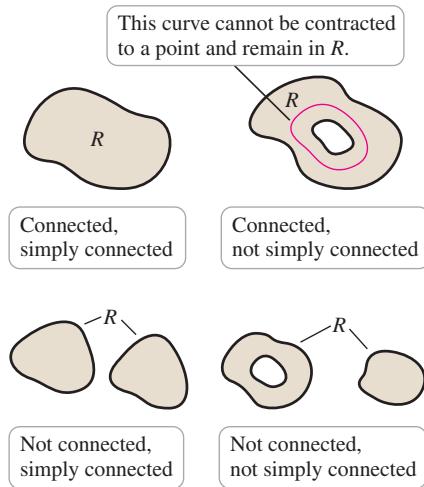
$$\begin{aligned} \oint_C \mathbf{F} \cdot d\mathbf{r} &= \oint_C \frac{-ydx + xdy}{x^2 + y^2} \\ &= \int_0^{2\pi} \frac{-\sin t d(\cos t) + \cos t d(\sin t)}{\sin^2 t + \cos^2 t} \\ &= \int_0^{2\pi} \frac{(\sin^2 t + \cos^2 t)dt}{\sin^2 t + \cos^2 t} \\ &= \int_0^{2\pi} 1dt = 2\pi \neq 0. \end{aligned}$$

C is a closed path, so \mathbf{F} cannot be conservative even though the curl of \mathbf{F} is zero! The curl test fails here because the vector field is not defined everywhere.

4.3.5 Simply-Connected Regions

You have seen that if a two-dimensional vector field \mathbf{F} is not defined at $(0, 0)$, one cannot always use the curl test to conclude whether it is conservative. However, one can relax the everywhere-defined condition a little bit to allow us to use the curl test on vector fields which can be undefined at some points. First we introduce:

Definition 4.4 — Simply Connected Regions. A region Ω is simply connected if every closed loop in Ω can be contracted to a point continuously without leaving the region Ω .



The set \mathbb{R}^2 with the origin removed, is not simply-connected, as the loops that go around the origin cannot be contracted to a point without “touching” the origin. However, the set \mathbb{R}^3 with the origin removed is simply-connected – draw a picture to convince yourself on that!

Simply connected regions are related to the curl test because one can also apply the curl test even though \mathbf{F} is not defined everywhere but is defined on a simply connected region.

For instance, the gravitational vector field

$$\mathbf{F}(x, y, z) = -GMm \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{(x^2 + y^2 + z^2)^{3/2}}$$

is not defined on $(0, 0, 0)$ but is defined on everywhere else in \mathbb{R}^3 . Since the set \mathbb{R}^3 with origin removed is simply connected, the curl test applies to this vector field!

With some straight-forward (but lengthy) computations, one can verify that $\nabla \times \mathbf{F} = \mathbf{0}$ for any $(x, y, z) \neq (0, 0, 0)$. Since the curl test applies here, it concludes that \mathbf{F} is conservative.

4.3.6 Curl Operator in Various Coordinate Systems

Note the definition of $\nabla \times \mathbf{F}$, i.e.

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathbf{k}$$

applies to rectangular coordinates only! For spherical coordinates (and analogously for cylindrical coordinates), $\nabla \times \mathbf{F}$ is **not** defined as:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{e}_\rho & \mathbf{e}_\theta & \mathbf{e}_\varphi \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \varphi} \\ F_\rho & F_\theta & F_\varphi \end{vmatrix} \quad \text{WRONG!}$$

Instead, $\nabla \times \mathbf{F}$ in spherical coordinates is simply converting the curl of \mathbf{F} in rectangular coordinates into spherical coordinates using the conversion rules established in the previous section. Namely, we need to:

1. convert all \mathbf{i} , \mathbf{j} and \mathbf{k} into \mathbf{e}_ρ , \mathbf{e}_θ and \mathbf{e}_φ ;
2. convert all F_x , F_y and F_z into F_ρ , F_θ and F_φ ;
3. use the chain rule (repeatedly) to convert all partial derivatives with respect to x , y and z to partial derivatives with respect to ρ , θ and φ . For instance,

$$\frac{\partial F_\rho}{\partial y} = \frac{\partial F_\rho}{\partial \rho} \frac{\partial \rho}{\partial y} + \frac{\partial F_\rho}{\partial \theta} \frac{\partial \theta}{\partial y} + \frac{\partial F_\rho}{\partial \varphi} \frac{\partial \varphi}{\partial y}.$$

It is extremely tedious to derive such a conversion rule! We omit the detail of the derivation here. In spherical coordinates, the curl of \mathbf{F} can be shown to have a nice determinant formula:

$$\nabla \times \mathbf{F} = \frac{1}{\rho^2 \sin \varphi} \begin{vmatrix} \mathbf{e}_\rho & \rho \mathbf{e}_\varphi & (\rho \sin \varphi) \mathbf{e}_\theta \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \varphi} & \frac{\partial}{\partial \theta} \\ F_\rho & \rho F_\varphi & (\rho \sin \varphi) F_\theta \end{vmatrix}.$$

Although the derivation of the above formula is extremely tedious, one of its neat consequence is that if \mathbf{F} is radially symmetric, i.e. $F_\varphi = F_\theta = 0$ and F_ρ depends only on ρ , then:

$$\nabla \times \mathbf{F} = \frac{1}{\rho^2 \sin \varphi} \begin{vmatrix} \mathbf{e}_\rho & \rho \mathbf{e}_\varphi & (\rho \sin \varphi) \mathbf{e}_\theta \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \varphi} & \frac{\partial}{\partial \theta} \\ F_\rho & 0 & 0 \end{vmatrix} = \mathbf{0}.$$

For instance, the gravitational force field is radially symmetric, so its curl is a zero vector. Although the spherical coordinates are undefined at the origin, it is not an issue when applying the curl test using spherical coordinates since \mathbb{R}^3 with the origin removed is a simply connected region. Therefore, any radially symmetric vector field \mathbb{R}^3 is conservative.

There is an analogous formula for cylindrical coordinates which is given by the following. Again, we omit its derivation since it is also tedious to do so.

$$\nabla \times \mathbf{F} = \frac{1}{r} \begin{vmatrix} \mathbf{e}_r & r \mathbf{e}_\theta & \mathbf{k} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial z} \\ F_r & r F_\theta & F_z \end{vmatrix}.$$

Similarly, it can also be checked that for vector fields with $F_\theta = F_z = 0$ and F_r depends only on r , then $\nabla \times \mathbf{F} = \mathbf{0}$. However, cylindrical coordinates cannot be defined on the z -axis. Since \mathbb{R}^3 with the z -axis removed is not simply connected, one needs to be very careful when applying the curl test under cylindrical coordinates!

4.4 Green's Theorem

4.4.1 The Theorem and its Uses

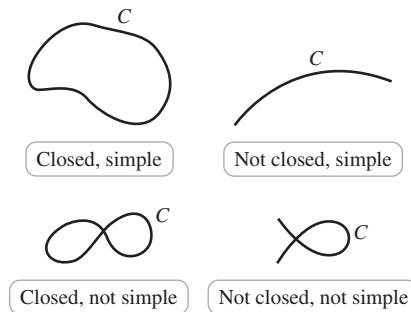
We will exclusively deal with two-dimensional vector fields in this section. In the previous section, we see that if \mathbf{F} is a two-dimensional vector field which is defined on a simply-connected region Ω in \mathbb{R}^2 (such as the entire \mathbb{R}^2 plane), then the curl test says $\nabla \times \mathbf{F} = \mathbf{0}$ if and only if \mathbf{F} is conservative, and so for such a vector field, we have

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$$

for any closed curve C in \mathbb{R}^2 .

It is natural to ask if there is any hidden relation between $\nabla \times \mathbf{F}$ and $\oint_C \mathbf{F} \cdot d\mathbf{r}$, given that the former being a zero vector implies the latter is zero. The Green's Theorem gives the relationship between them.

Definition 4.5 — Simple Closed Curves. A curve C is called a simple closed curve if the two endpoints coincide and it does not intersect itself at any point (other than the endpoints).



Theorem 4.4 — Green's Theorem. Let C be a simple closed curve in \mathbb{R}^2 which is counter-clockwise oriented. Suppose the curve C encloses region R . Let $\mathbf{F}(x, y)$ be a vector field which is defined and is continuously differentiable at every point in R , then:

$$\underbrace{\oint_C \mathbf{F} \cdot d\mathbf{r}}_{\text{line integral}} = \underbrace{\iint_R (\nabla \times \mathbf{F}) \cdot \mathbf{k} dA}_{\text{double integral}}$$

- (i) In terms of components of \mathbf{F} in rectangular coordinates, i.e. $\mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j}$, one can easily verify that $\nabla \times \mathbf{F} = \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathbf{k}$, so the above Green's Theorem can be stated as:

$$\oint_C F_x dx + F_y dy = \iint_R \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dA.$$

- (i) In particular, if $\nabla \times \mathbf{F} = \mathbf{0}$ and \mathbf{F} is defined on all of \mathbb{R}^2 , then the Green's Theorem tells us that

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R \mathbf{0} \cdot \mathbf{k} dA = 0$$

for any simple closed curve C . This is exactly what the curl test tells us!

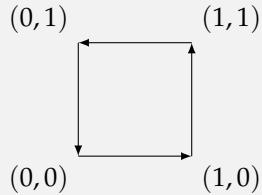
The proof of the Green's Theorem is quite technical if the curve C is complicated. The proof of the theorem in some special cases can be found in the textbook. Let's first look at some examples.

■ **Example 4.8** Use the Green's Theorem to evaluate the line integral:

$$\oint_C -ydx + xdy$$

where C is a rectangular loop $(0,0) \rightarrow (1,0) \rightarrow (1,1) \rightarrow (0,1) \rightarrow (0,0)$.

■ **Solution** The vector field corresponding to this line integral is $\mathbf{F} = -y\mathbf{i} + x\mathbf{j}$, which is defined and is smooth everywhere on \mathbb{R}^2 . The given path C encloses the rectangle R with vertices $(0,0), (1,0), (1,1)$ and $(0,1)$.



By straight-forward calculation, the curl of \mathbf{F} is given by:

$$\nabla \times \mathbf{F} = 2\mathbf{k}.$$

By the Green's Theorem, we have:

$$\begin{aligned} \oint_C -ydx + xdy &= \oint_C \mathbf{F} \cdot d\mathbf{r} \\ &= \iint_R (\nabla \times \mathbf{F}) \cdot \mathbf{k} \, dA \\ &= \iint_R 2\mathbf{k} \cdot \mathbf{k} \, dA \\ &= \iint_R 2 \, dA \\ &= 2 \times \text{area of } R \\ &= 2. \end{aligned}$$

■ **Example 4.9** Let $\mathbf{F} = -y^2\mathbf{i} + xy\mathbf{j}$ and C be a simple closed curve in \mathbb{R}^2 enclosing a region R which is symmetric across the x -axis. Show that:

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

■ **Solution** Since \mathbf{F} is defined and is continuously differentiable everywhere, we can apply the Green's Theorem without any issues. First we compute its curl:

$$\nabla \times \mathbf{F} = 3y\mathbf{k}.$$

By the Green's Theorem

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R (\nabla \times \mathbf{F}) \cdot \mathbf{k} \, dA$$

$$= \iint_R 3y \mathbf{k} \cdot \mathbf{k} \, dA$$

$$= \iint_R 3y \, dA.$$

Since the region R is symmetric across the x -axis and $3y$ is an odd function. If we split R into two: R_+ (the part above the x -axis) and R_- (the part below the x -axis), then

$$\iint_{R_+} 3y \, dA \text{ and } \iint_{R_-} 3y \, dA$$

are numerically the same except that one is the negative of the other. Therefore, by

$$\iint_R 3y \, dA = \iint_{R_+} 3y \, dA + \iint_{R_-} 3y \, dA,$$

we get:

$$\iint_R 3y \, dA = 0$$

since the R_+ -integral gets canceled out by the R_- -integral. Therefore,

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

To summarize, applying Green's Theorem is very straightforward. It often simplifies the computation of a line integral over a closed loop since there is no need to parametrize the curve. If the curve encloses a rectangular region, the Green's Theorem is an effective tool for evaluating the line integral because a double integral over a rectangle is easy to set up.

4.4.2 Significance of the Green's Theorem

There are far-reaching significance of the Green's Theorem apart from making computations of line integrals easier. One important geometric or physical significance is that it gives an interpretation of the curl of a two-dimensional vector field.

Consider a vector field \mathbf{F} in \mathbb{R}^2 defined and being continuously differentiable everywhere. Suppose C is a very tiny simple closed curve enclosing a tiny region R . Then, the quantity $(\nabla \times \mathbf{F}) \cdot \mathbf{k}$ can be regarded as a constant inside the region R .

By the Green's Theorem, we have:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R (\nabla \times \mathbf{F}) \cdot \mathbf{k} \, dA \simeq (\nabla \times \mathbf{F}) \cdot \mathbf{k} \text{ (area of } R).$$

In other words,

$$(\nabla \times \mathbf{F}) \cdot \mathbf{k} \simeq \frac{1}{\text{area of } R} \oint_C \mathbf{F} \cdot d\mathbf{r}.$$

Since $\oint_C \mathbf{F} \cdot d\mathbf{r}$ indicates how *circular* the vector field \mathbf{F} is around the curve C , the above result tells us that the quantity $(\nabla \times \mathbf{F}) \cdot \mathbf{k}$ at a point measures the density of circulation around that point. That's why we call $\nabla \times \mathbf{F}$ to be curl of \mathbf{F} because it is an indicator of curlessness of a vector field!

As an example, you can verify that the curls of the vector fields $\mathbf{F} = -y\mathbf{i} + x\mathbf{j}$ and $\mathbf{G} = x\mathbf{i} + y\mathbf{j}$ are respectively given by:

$$\begin{aligned}(\nabla \times \mathbf{F}) \cdot \mathbf{k} &= 2 \\(\nabla \times \mathbf{G}) \cdot \mathbf{k} &= 0\end{aligned}$$

It suggests that \mathbf{F} is more circular than \mathbf{G} . By plotting them in Mathematica, we can see that vectors in \mathbf{F} are circling around the origin, while vectors in \mathbf{G} are diverging from the origin.

4.4.3 Limitations of the Green's Theorem

We have stated in the Green's Theorem that the vector field \mathbf{F} needs to be defined at every point in the region R enclosed by the simple closed curve C . It is a crucial condition!

Let's consider the vector field

$$\mathbf{F} = -\frac{y}{x^2+y^2}\mathbf{i} + \frac{x}{x^2+y^2}\mathbf{j}.$$

It is not defined at the origin $(0,0)$. If we let C be the (counter-clockwise) unit circle centered at the origin, then C can be parametrized by:

$$\mathbf{r}(t) = (\cos t)\mathbf{i} + (\sin t)\mathbf{j}, \quad 0 \leq t \leq 2\pi.$$

By computing the line integral over C directly, one should get:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_C \mathbf{F} \cdot \mathbf{r}'(t) dt = \oint_C \langle -\sin t, \cos t \rangle \cdot \langle -\sin t, \cos t \rangle dt = \oint_C 1 dt = 2\pi.$$

However, one can check from direct computations that $\nabla \times \mathbf{F} = \mathbf{0}$ at every point except the origin, and is undefined at the origin. Therefore, the double integral

$$\iint_R \underbrace{(\nabla \times \mathbf{F}) \cdot \mathbf{k}}_{\text{undefined at } (0,0)} dA$$

is an *improper* integral! To summarize, one cannot directly apply the Green's Theorem if the given curve encloses a point at which the vector field is not defined. In such cases, we should either compute the line integral directly by parametrization, or use some other tools to compute it. We will compute this integral by so-called the *hole-drilling technique* to be presented in the next subsection.

If the curve does not enclose the origin (at which \mathbf{F} is undefined), then Green's Theorem applies to \mathbf{F} and such a curve without any issues. For instance, if C is a circle centered at $(3,0)$ with radius 2, then it does not enclose the origin which is the only point at which \mathbf{F} is undefined. The Green's Theorem does show that

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R \underbrace{(\nabla \times \mathbf{F}) \cdot \mathbf{k}}_{=0 \text{ at every point in } R} dA = \iint_R 0 dA = 0.$$

4.4.4 Winding Number of a Closed Curve

The following vector field (discussed in the previous part)

$$\mathbf{F} = -\frac{y}{x^2+y^2}\mathbf{i} + \frac{x}{x^2+y^2}\mathbf{j}$$

is a famous one which concerns about the winding number of a closed curve. As discussed before, it is not defined at the origin, and $\nabla \times \mathbf{F} = \mathbf{0}$ at every point except the origin. If C is a closed curve (not necessarily simple closed) in \mathbb{R}^2 not passing through the origin, there is a celebrated result in topology that says:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 2\pi \times \text{number of times } C \text{ travels around the origin counter-clockwisely.}$$

Consequently, the quantity $\frac{1}{2\pi} \oint_C \frac{-y}{r^2} dx + \frac{x}{r^2} dy$ is often called the *winding number* of the curve C . This number is not only important in pure mathematics but also in physics. Furthermore, the computation of surface flux of a vector field satisfying the inverse square law, as we will see later, are also in the same spirit as the computation of the winding number.

Our goal here is to use the Green's Theorem to explain why this line integral gives the winding number for some not-too-complicated curves.

Circles Centered at Origin

We start with the simplest case where the curve is a circle with radius ε centered at the origin, which from now on denoted by Γ_ε . The path is parametrized by:

$$\Gamma_\varepsilon : \mathbf{r}(t) = (\varepsilon \cos t) \mathbf{i} + (\varepsilon \sin t) \mathbf{j}, \quad 0 \leq t \leq 2\pi.$$

Then, by straight-forward calculations, we get:

$$\begin{aligned} \oint_{\Gamma_\varepsilon} -\frac{y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy &= \int_{t=0}^{t=2\pi} -\frac{\varepsilon \sin t}{\varepsilon^2} d(\varepsilon \cos t) + \frac{\varepsilon \cos t}{\varepsilon^2} d(\varepsilon \sin t) \\ &= \int_{t=0}^{t=2\pi} (\sin^2 t + \cos^2 t) dt \\ &= \int_{t=0}^{t=2\pi} 1 dt = 2\pi. \end{aligned}$$

Simple Closed Curves Enclosing the Origin

Next we see how to apply the Green's Theorem on a more arbitrary curve, say the curve C in Figure 4.6a.

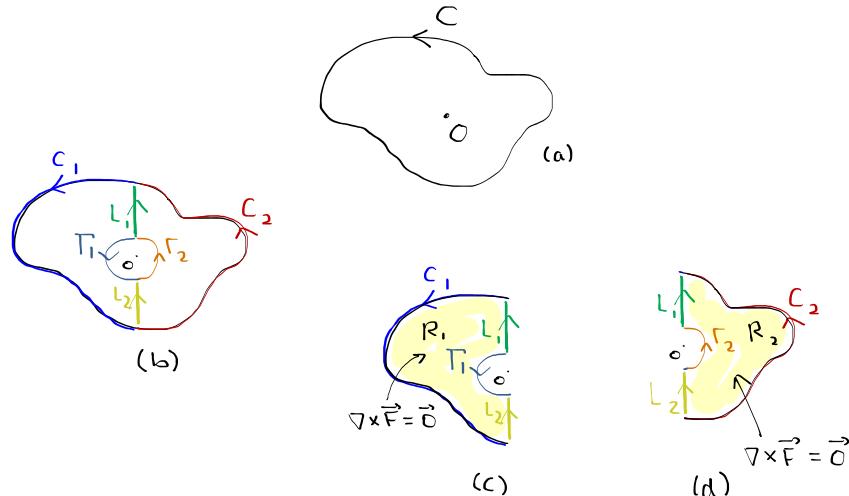


Figure 4.6: a curve circling around the origin once

We are going to show that the line integral $\oint_C \mathbf{F} \cdot d\mathbf{r}$, where \mathbf{F} is defined as before, is equal to 2π . First note that \mathbf{F} is not defined at the origin which is enclosed by the curve C , so we can't directly apply the Green's Theorem on this curve C .

To handle this issue, we drill a hole near the origin, i.e. we remove a tiny ball with radius ε centered at the origin from the region (see Figure 4.6b), and further split the punctured region into two parts R_1 and R_2 by cutting it through line segments L_1 and L_2 (see Figure 4.6cd). We label each segment of the boundaries by C_1 , C_2 , L_1 , L_2 , Γ_1 and Γ_2 with directions indicated in the diagram. Then, according to the directions of C_1 , C_2 and C , the line integral in question can be expressed as:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_{C_1+C_2} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_1} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

Likewise, according to the directions of Γ_1 , Γ_2 and Γ_ε (recall that Γ_ε is the counter-clockwise circle with radius ε centered at the origin), we have:

$$\oint_{\Gamma_\varepsilon} \mathbf{F} \cdot d\mathbf{r} = \oint_{\Gamma_1 + \Gamma_2} \mathbf{F} \cdot d\mathbf{r} = \oint_{\Gamma_1} \mathbf{F} \cdot d\mathbf{r} + \oint_{\Gamma_2} \mathbf{F} \cdot d\mathbf{r}.$$

Now our goal is to show that

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_{\Gamma_\varepsilon} \mathbf{F} \cdot d\mathbf{r}.$$

Note that we have already computed the RHS, which is 2π . The above result will show that the line integral over C is also 2π .

In order to show this, we consider each of the regions R_1 and R_2 shown in Figure 4.6cd. Since R_1 does not enclose the origin and \mathbf{F} is defined everywhere in R_1 , the Green's Theorem can be applied to the region R_1 . According to the indicated directions of the boundary curves, we see that:

$$\text{boundary of } R_1 = C_1 + L_2 - \Gamma_1 + L_1.$$

Therefore, by the Green's Theorem, we have:

$$\underbrace{\int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{L_2} \mathbf{F} \cdot d\mathbf{r} - \int_{\Gamma_1} \mathbf{F} \cdot d\mathbf{r} + \int_{L_1} \mathbf{F} \cdot d\mathbf{r}}_{\oint_{\text{boundary of } R_1} \mathbf{F} \cdot d\mathbf{r}} = \iint_{R_1} (\nabla \times \mathbf{F}) \cdot \mathbf{k} dA = 0. \quad (4.1)$$

Similarly, R_2 does not enclose the origin so the Green's Theorem can be applied to R_2 :

$$\underbrace{\int_{C_2} \mathbf{F} \cdot d\mathbf{r} - \int_{L_1} \mathbf{F} \cdot d\mathbf{r} - \int_{\Gamma_2} \mathbf{F} \cdot d\mathbf{r} - \int_{L_2} \mathbf{F} \cdot d\mathbf{r}}_{\oint_{\text{boundary of } R_2} \mathbf{F} \cdot d\mathbf{r}} = \iint_{R_2} (\nabla \times \mathbf{F}) \cdot \mathbf{k} dA = 0. \quad (4.2)$$

By summing up (4.1) and (4.2) and canceling out the line integrals over L_1 and L_2 , we get:

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r} - \int_{\Gamma_1} \mathbf{F} \cdot d\mathbf{r} - \int_{\Gamma_2} \mathbf{F} \cdot d\mathbf{r} = 0.$$

Since $C = C_1 + C_2$ and $\Gamma_\varepsilon = \Gamma_1 + \Gamma_2$ according to their directions in the diagram, we have:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_{\Gamma_\varepsilon} \mathbf{F} \cdot d\mathbf{r} = 2\pi.$$

Self-Intersecting Curves

Next, let's use the above result to deal with some intersecting curves. Consider the curve C in Figure 4.7.

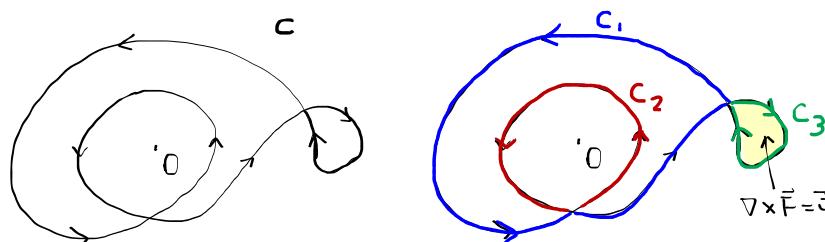


Figure 4.7: a self-intersecting curve circling around the origin twice

Since we have already know how to deal with any simple closed curve enclosing the origin, we are going to split the curve C into simple closed curves C_1 , C_2 and C_3 according to Figure 4.7.

As both C_1 and C_2 are simple closed curves enclosing the origin, by our previous discussion we know:

$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = 2\pi.$$

For C_3 , it is a simple closed curve not enclosing the origin. Therefore, the Green's Theorem can be applied on C_3 without any issues. Hence, we have

$$\oint_{-C_3} \mathbf{F} \cdot d\mathbf{r} = \iint_R (\nabla \times \mathbf{F}) \cdot \mathbf{k} \, dA = 0$$

where R is the region enclosed by C_3 . Here we have again used the fact that $\nabla \times \mathbf{F} = \mathbf{0}$ inside the region R .

According the directions indicated on the diagram, we have:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_{C_1} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_3} \mathbf{F} \cdot d\mathbf{r} = 2\pi + 2\pi + 0 = 4\pi.$$

Therefore, the winding number of this curve C is equal to 2. Similar technique can be applied to more complicated curves which go around the origin a lot of times.

4.5 Parametric Surfaces

The goal of this and the next sections is to generalize the Green's Theorem for vector fields on the xy -plane to the Stokes' Theorem for vector fields on the xyz -space. While the Green's Theorem relates the line integral of a closed curve to the double integral over the enclosed region, the Stokes' Theorem relates the line integral to the an integral over a *surface* whose boundary is the a closed curve. For that we need to define and make sense of *surface integrals*.

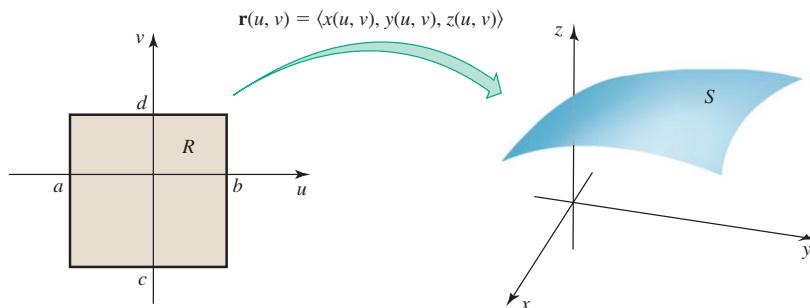
4.5.1 Surface Parametrizations

To begin with, we need to know how to describe a surface in the xyz -space. Recall that a curve in space is represented in parametric form $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$, and is thought as the path of a particle. The values of $x(t)$, $y(t)$ and $z(t)$ represent the coordinates of the particle at time t .

To present a surface, we need *two* parameters, say u and v . The general form of a parametric equation of a surface is:

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k}.$$

Instead of regarding u and v as time variables, we regard them as the coordinates on a uv -plane, and the vector $\mathbf{r}(u, v)$ is a function that associates each point (u, v) on the uv -plane to a point $(x(u, v), y(u, v), z(u, v))$ in the xyz -space. Since there are two parameters, the image of the function is a surface in the xyz -space. In other words, the function $\mathbf{r}(u, v)$ can be thought as a transformation that "wraps" the uv -paper onto the surface.



The function $\mathbf{r}(u, v)$ is called a **parametrization** of the surface, and by what we mean "parametrizing a surface" is to give a parametrization of the surface. As we will see later, parametrizing a surface is often the first step of computing a surface integral.

- (i) Although we use (u, v) to denote the parameters, you can use whatever pair of variables for the parameters, provided that there is no confusion.

Parametrization via Coordinate Systems

Let's look at several elementary examples such as cylinders, spheres and cones.

- **Example 4.10** Find a parametrization for the cylinder with radius r_0 and with z -axis as the central axis.

- **Solution** If, under a certain coordinate system, the surface has *one* of the coordinates being constant, then giving a parametrization to that surface is fairly easy: simply take the other *two* coordinates as parameters, and define \mathbf{r} according to the conversion rules between this coordinate system and the rectangular coordinate system.

The cylinder described in the problem can be presented by equation $r = r_0$ under cylindrical coordinates (r, θ, z) . The conversion rule between cylindrical and rectangular

coordinates is given by:

$$\begin{aligned}x &= r \cos \theta \\y &= r \sin \theta \\z &= z\end{aligned}$$

Fix $r = r_0$, then x , y and z are functions of (θ, z) . Simply set:

$$x(\theta, z) = r_0 \cos \theta, \quad y(\theta, z) = r_0 \sin \theta, \quad z(\theta, z) = z,$$

then the parametrization is given by:

$$\mathbf{r}(\theta, z) = (r_0 \cos \theta)\mathbf{i} + (r_0 \sin \theta)\mathbf{j} + z\mathbf{k}, \quad 0 \leq \theta \leq 2\pi, \quad -\infty < z < \infty.$$

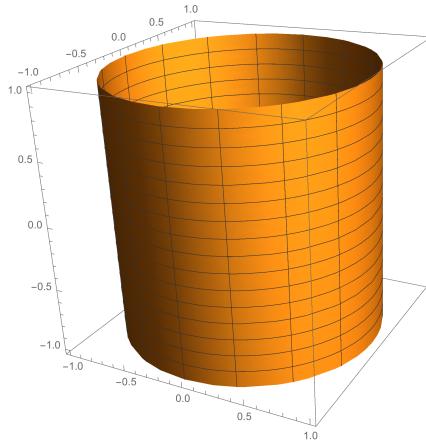


Figure 4.8: parametric plot of a cylinder

One can also specify the range of z so that the parametrization gives a finite cylinder. For instance,

$$\mathbf{r}(\theta, z) = (r_0 \cos \theta)\mathbf{i} + (r_0 \sin \theta)\mathbf{j} + z\mathbf{k}, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq z \leq 1$$

gives the finite cylinder with height 1 unit (from $z = 0$ to $z = 1$).

Similarly, a cone making $\frac{\pi}{4}$ angle with the z -axis can be represented by $z = r$ in cylindrical coordinates, or $\varphi = \frac{\pi}{4}$ in spherical coordinates. Therefore, it is can be parametrized by two different ways:

$$\begin{aligned}\mathbf{r}_1(r, \theta) &= (r \cos \theta)\mathbf{i} + (r \sin \theta)\mathbf{j} + r\mathbf{k}, \quad 0 \leq r < \infty, \quad 0 \leq \theta \leq 2\pi \\\mathbf{r}_2(\rho, \theta) &= \left(\rho \sin \frac{\pi}{4} \cos \theta\right)\mathbf{i} + \left(\rho \cos \frac{\pi}{4} \sin \theta\right)\mathbf{j} + \left(\rho \cos \frac{\pi}{4}\right)\mathbf{k} \\&= \left(\frac{\rho \cos \theta}{\sqrt{2}}\right)\mathbf{i} + \left(\frac{\rho \sin \theta}{\sqrt{2}}\right)\mathbf{j} + \frac{\rho}{\sqrt{2}}\mathbf{k}, \quad 0 \leq \rho < \infty, \quad 0 \leq \theta \leq 2\pi\end{aligned}$$

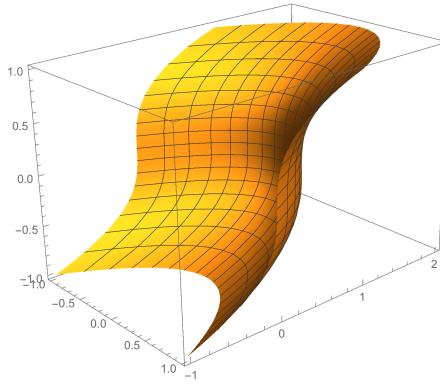
A sphere with radius 3 can be presented by $\rho = 3$ in spherical coordinates, and so it can be parametrized by:

$$\mathbf{r}_3(\theta, \varphi) = (3 \sin \varphi \cos \theta)\mathbf{i} + (3 \sin \varphi \sin \theta)\mathbf{j} + (3 \sin \varphi)\mathbf{k}, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \varphi \leq \pi.$$

Parametrization of Graphs

If a surface can be represented by a Cartesian equation (i.e. level-set form) such as $x^2 + y - z^3 = 1$ and that you can write one of the variables as a function of the other two variables (such as $y = z^3 - x^2 + 1$ in this example), then we can use the other two variables as parameters. For instance, the surface $y = z^3 - x^2 + 1$ can be parametrized as:

$$\mathbf{r}(x, z) = x\mathbf{i} + \underbrace{(z^3 - x^2 + 1)}_y\mathbf{j} + z\mathbf{k}.$$



However, x cannot be written as a function of y and z for the surface $y = z^3 - x^2 + 1$, since $x = \pm\sqrt{z^3 - y + 1}$ and the \pm makes it not a function of y and z . On the other hand, z can be written as a function of x and y , since $z^3 = x^2 + y - 1$ and there is exactly one cubic root for $x^2 + y - 1$. However, the resulting parametrization

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + (x^2 + y - 1)^{1/3}\mathbf{k}$$

is not easy to work with.

Some Interesting Surfaces

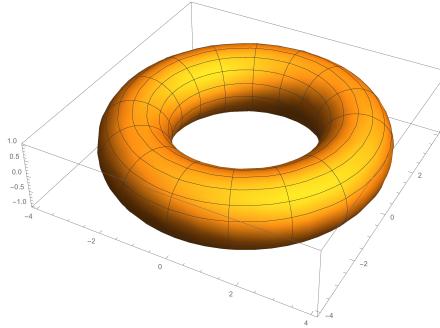
Although surfaces such as cones, spheres and cylinders can be easily parametrized, it is not the case for many interesting surfaces.

Below are some interesting examples of surfaces. It is not easy to write down (or even to explain) the parametrization since it demands quite a lot of geometric intuitions.

A torus (i.e. donut), for example, has a complicated parametrization given by:

$$\mathbf{r}(u, v) = ((3 + \cos u) \cos v)\mathbf{i} + ((3 + \cos u) \sin v)\mathbf{j} + (\sin u)\mathbf{k},$$

with $0 \leq u \leq 2\pi$, $0 \leq v \leq 2\pi$.



The Möbius strip, a famous object in topology, has the following parametrization

$$\mathbf{r}(u, v) = \left(\left(1 + \frac{v}{2} \cos \frac{u}{2} \right) \cos u \right) \mathbf{i} + \left(\left(1 + \frac{v}{2} \cos \frac{u}{2} \right) \sin u \right) \mathbf{j} + \left(\frac{v}{2} \sin \frac{u}{2} \right) \mathbf{k}$$

where $0 \leq u \leq 2\pi$ and $-1 \leq v \leq 1$.

Since \mathbf{r} denotes the position vector $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, we can also write down a parametrization in an equation form (especially when the expression of the parametrization is too long to fit in one line). For instance, the Möbius strip parametrization can be equivalently written as:

$$\begin{aligned} x &= \left(1 + \frac{v}{2} \cos \frac{u}{2} \right) \cos u \\ y &= \left(1 + \frac{v}{2} \cos \frac{u}{2} \right) \sin u \\ z &= \frac{v}{2} \sin \frac{u}{2} \end{aligned}$$

where $0 \leq u \leq 2\pi$ and $-1 \leq v \leq 1$.

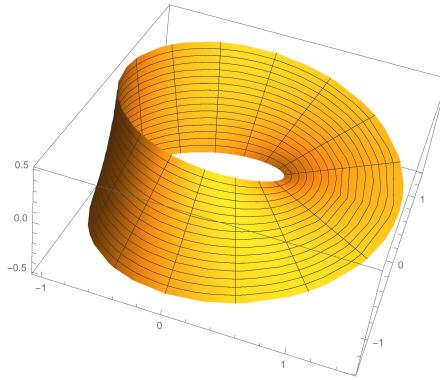


Figure 4.9: Möbius Strip

The following parametrization, which looks intimidating, describes a very beautiful surface called the Klein bottle (see Figure 4.10):

$$\begin{aligned}x &= -\frac{2}{15} \cos u \left(3 \cos v - 30 \sin u + 90 \cos^4 u \sin u - 60 \cos^6 u \sin u + 5 \cos u \sin u \cos v \right) \\y &= -\frac{1}{15} \sin u \left(3 \cos v - 3 \cos^2 u \cos v - 48 \cos^4 u \cos v + 48 \cos^6 u \cos v \right. \\&\quad \left. - 60 \sin u + 5 \cos u \sin u \cos v - 5 \cos^3 u \cos v \sin u \right. \\&\quad \left. - 80 \cos^5 u \sin u \cos v + 80 \cos^7 u \sin u \cos v \right) \\z &= \frac{2}{15} (3 + 5 \cos u \sin u) \sin v\end{aligned}$$

for $0 \leq u \leq \pi$ and $0 \leq v \leq 2\pi$.

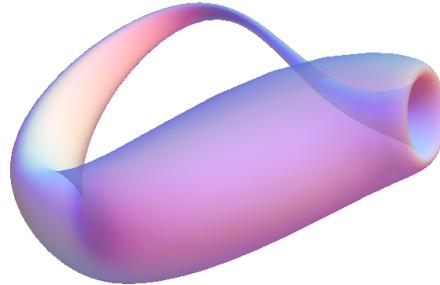


Figure 4.10: the Klein Bottle

Note that the Klein bottle is self-intersecting as depicted in the diagram.

Geometry of Parametric Surfaces

Consider a parametrization $\mathbf{r}(u, v)$ of a surface. Keeping $v = v_0$ fixed and letting u vary, the parametrization function $\mathbf{r}(u, v_0)$ depends only on u and it will give a curve on the surface. This curve is often called a u -curve for u is varying. The tangent vector to any u -curve can be computed by taking the u -derivative of \mathbf{r} , i.e.

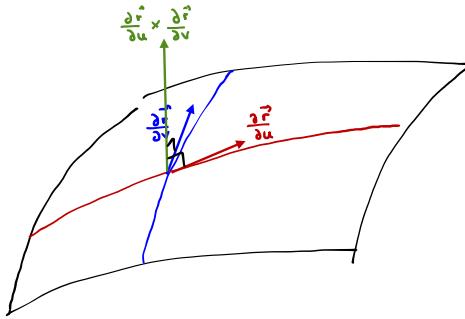
$$\frac{\partial \mathbf{r}}{\partial u} = \text{tangent vector to the } u\text{-curve}$$

Likewise, when $u = u_0$ is fixed while v varies, the function $\mathbf{r}(u_0, v)$ depends only on v . The curve traced out by this function when v varies is called a v -curve, and

$$\frac{\partial \mathbf{r}}{\partial v} = \text{tangent vector to the } v\text{-curve}$$

Both $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ are *tangent* to the surface, hence their cross product is *normal* to the surface (precisely, normal to the tangent plane). A unit normal at each point is therefore given by:

$$\hat{\mathbf{n}} = \frac{\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}}{\left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right|}.$$



4.5.2 Surface Integrals

We are going to introduce surface integrals in this subsection. Double integrals such as

$$\iint_R f(x, y) dA$$

is a special type of surface integral where the region of integration R is on the flat xy -plane. A *surface integral* is one that the region of integration can be a curved surface. Many geometric and physical quantities, such as surface area, surface flux, and moment of inertia for some shell objects, can be computed using surface integrals. The Stokes' Theorem to be introduced in the next section also involves surface integrals.

We first state the definition of surface integrals, compute some examples and then explain its geometric and physical meaning.

Definition 4.6 — Surface Integrals. Given a surface S parametrized by $\mathbf{r}(u, v)$ with $a \leq u \leq b$ and $c \leq v \leq d$, and a continuous, scaled-valued function $f(x, y, z)$, the surface integral of f over the surface S is denoted and defined to be:

$$\iint_S f dS = \int_{v=c}^{v=d} \int_{u=a}^{u=b} f(\mathbf{r}(u, v)) \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| du dv.$$

- (1) The line integral of a vector field over a curve C is independent of how we parametrize C . It is also true that the surface integral over a surface S is also independent of the surface parametrization $\mathbf{r}(u, v)$. The proof involves change of variables technique in multivariable calculus and is omitted here.

Notation If a surface is closed (meaning it has no boundaries), it is conventional to denote the integral sign by:

$$\iint_S$$

Examples of closed surfaces include spheres and torus, while a hemisphere (only the spherical part, the flat part is not included) is not closed since it has a circle as its boundary.

■ **Example 4.11** Let S be the sphere of radius a centered at the origin. Evaluate the surface integral

$$\iint_S (x^2 + y^2) \, dS$$

■ **Solution** We first parametrize the surface. Using spherical coordinates, the sphere is presented by $\rho = a$. Take (θ, φ) as parameters, then a parametrization of the sphere is given by:

$$\mathbf{r}(\theta, \varphi) = (a \sin \varphi \cos \theta) \mathbf{i} + (a \sin \varphi \sin \theta) \mathbf{j} + (a \cos \varphi) \mathbf{k}$$

where $0 \leq \theta \leq 2\pi$ and $0 \leq \varphi \leq \pi$.

According to the definition of surface integrals, we need to compute $\left| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} \right|$. It is straight-forward, although quite tedious:

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial \theta} &= (-a \sin \varphi \sin \theta) \mathbf{i} + (a \sin \varphi \cos \theta) \mathbf{j} + 0 \mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial \varphi} &= (a \cos \varphi \cos \theta) \mathbf{i} + (a \cos \varphi \sin \theta) \mathbf{j} + (-a \sin \varphi) \mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -a \sin \varphi \sin \theta & a \sin \varphi \cos \theta & 0 \\ a \cos \varphi \cos \theta & a \cos \varphi \sin \theta & -a \sin \varphi \end{vmatrix} \\ &= (-a^2 \sin^2 \varphi \cos \theta) \mathbf{i} + (-a^2 \sin^2 \varphi \sin \theta) \mathbf{j} + (-a^2 \sin \varphi \cos \varphi) \mathbf{k} \\ \left| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} \right| &= \sqrt{a^4 \sin^4 \varphi (\cos^2 \theta + \sin^2 \theta) + a^4 \sin^2 \varphi \cos^2 \varphi} \\ &= \sqrt{a^4 \sin^2 \varphi (\sin^2 \varphi + \cos^2 \varphi)} \\ &= a^2 \sin \varphi. \end{aligned}$$

Next we compute the surface integral. When (x, y, z) is on the sphere S , we have: $x = a \sin \varphi \cos \theta$ and $y = a \sin \varphi \sin \theta$. Therefore, the integrand is given by:

$$x^2 + y^2 = a^2 \sin^2 \varphi (\cos^2 \theta + \sin^2 \theta) = a^2 \sin^2 \varphi.$$

According to the bounds of θ and φ in the parametrization, the surface integral of $x^2 + y^2$ over S is given by:

$$\begin{aligned} \iint_S (x^2 + y^2) \, dS &= \int_{\varphi=0}^{\varphi=\pi} \int_{\theta=0}^{\theta=2\pi} (x^2 + y^2) \left| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} \right| d\theta d\varphi \\ &= \int_{\varphi=0}^{\varphi=\pi} \int_{\theta=0}^{\theta=2\pi} \underbrace{a^2 \sin^2 \varphi}_{x^2+y^2} \cdot \underbrace{a^2 \sin \varphi}_{\left| \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} \right|} d\theta d\varphi \\ &= \int_{\varphi=0}^{\varphi=\pi} \int_{\theta=0}^{\theta=2\pi} a^4 \sin^3 \varphi d\theta d\varphi \\ &= \int_{\varphi=0}^{\varphi=\pi} 2\pi a^4 \sin^3 \varphi d\varphi \\ &= 2\pi a^4 \cdot \frac{4}{3} = \frac{8\pi a^4}{3}. \end{aligned}$$

Here we used the fact from single-variable calculus that:

$$\int_0^\pi \sin^3 \varphi \, d\varphi = \frac{4}{3}.$$

We leave this part as an exercise.

■ **Example 4.12** Let S be the plane $3x + 2y + z = 1$ defined over the region $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Compute the surface integral:

$$\iint_S (x + y + z) \, dS.$$

■ **Solution** The equation of the given plane can be written as a graph $z = 1 - 3x - 2y$ over the given region $0 \leq x \leq 1$ and $0 \leq y \leq 1$. We can take x and y as parameters and so a parametrization of the plane is given by:

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + (1 - 3x - 2y)\mathbf{k}$$

where $0 \leq x \leq 1$ and $0 \leq y \leq 1$.

Next we compute all the ingredients of the surface integral:

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial x} &= \mathbf{i} - 3\mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial y} &= \mathbf{j} - 2\mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} &= 3\mathbf{i} + 2\mathbf{j} + \mathbf{k} \\ \left| \frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} \right| &= \sqrt{3^2 + 2^2 + 1^2} = \sqrt{14} \\ x + y + z &= x + y + (1 - 3x - 2y) = 1 - 2x - y \end{aligned}$$

Finally we compute the surface integral:

$$\begin{aligned} \iint_S (x + y + z) \, dS &= \int_0^1 \int_0^1 (1 - 2x - y) \cdot \sqrt{14} \, dx \, dy \\ &= \sqrt{14} \int_0^1 \left[x - x^2 - xy \right]_{x=0}^{x=1} \, dy \\ &= \sqrt{14} \int_0^1 -y \, dy \\ &= \left[-\frac{y^2 \sqrt{14}}{2} \right]_0^1 \\ &= -\frac{\sqrt{14}}{2}. \end{aligned}$$

Surface Element

Now we explain the geometric and physical meaning of surface integrals. Given a parametric surface S with parametrization $\mathbf{r}(u, v)$, $a \leq u \leq b$ and $c \leq v \leq d$, from now on we denote:

Notation $dS = \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| du dv$

It is called the *surface element* of the integral. If we subdivide the domain in the uv -plane into small rectangular pieces with area $\Delta u \Delta v$, the parametrization $\mathbf{r}(u, v)$ transform them into small pieces ΔS on the surface (see Figure 4.11).

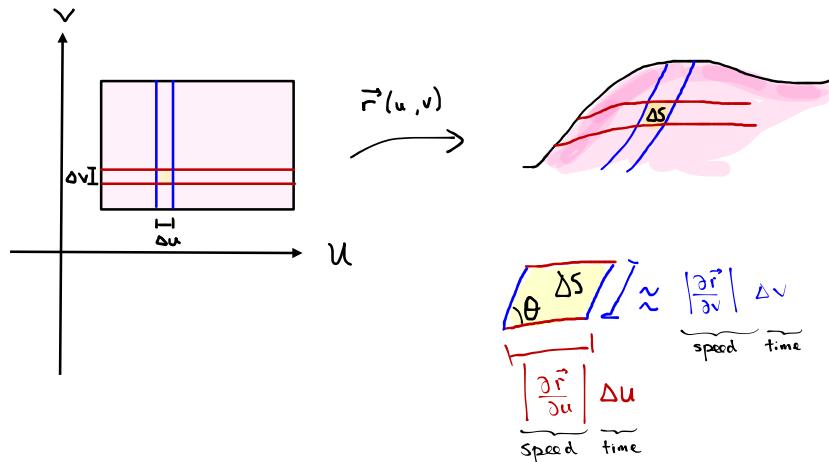


Figure 4.11: geometric meaning of the surface element

If the number of subdivisions is very large so that each ΔS is very small, then ΔS is approximately a parallelogram. The red side of the parallelogram has length approximately equal to $\left| \frac{\partial \mathbf{r}}{\partial u} \right| \Delta u$ (to see this, regard u is the time then the distance traveled after Δu unit time is approximately the speed \times time). Similarly, the blue side of the parallelogram has length approximately equal to $\left| \frac{\partial \mathbf{r}}{\partial v} \right| \Delta v$. Suppose the angle of the parallelogram is θ , then the area of ΔS is approximately:

$$\left| \frac{\partial \mathbf{r}}{\partial u} \right| \Delta u \cdot \left| \frac{\partial \mathbf{r}}{\partial v} \right| \Delta v \cdot \sin \theta = \left| \frac{\partial \mathbf{r}}{\partial u} \right| \left| \frac{\partial \mathbf{r}}{\partial v} \right| \sin \theta \cdot \Delta u \Delta v = \underbrace{\left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right|}_{\left| \frac{\partial \mathbf{r}}{\partial u} \right| \left| \frac{\partial \mathbf{r}}{\partial v} \right| \sin \theta} \Delta u \Delta v.$$

Infinitesimally, the ΔS becomes the surface element dS , and $\Delta u \Delta v$ becomes $du dv$. In other words, the surface element dS presents the area of a very tiny piece of subdivision on the surface. Summing up the area of all tiny subdivisions, we get the **surface area** of S , i.e.

$$\iint_S dS = \text{surface area of } S.$$

Different geometric or physical meanings of the integrand f give different meanings to the surface integral of f over S . For instance,

| If f is | each element $f dS$ means | $\iint_S f dS$ means |
|---|---|---|
| 1 surface density $\text{density} \times (x^2 + y^2)$ | area of dS mass of dS moment of inertia of dS | total surface area total mass of the surface moment of inertia of S |

Let S be the sphere of radius a centered at the origin. We have computed in Example 4.11 that

$$\iint_S (x^2 + y^2) dS = \frac{8\pi a^4}{3}.$$

Suppose this spherical shell has a uniform surface density δ , then the integral:

$$\iint_S \delta (x^2 + y^2) dS$$

is the moment of inertia I_z about z -axis. Since its formula in many physics textbooks is written in terms of the total mass m of the sphere, let's rewrite it in terms of m .

$$\begin{aligned} I_z &= \iint_S \delta(x^2 + y^2) \, dS \\ &= \delta \left(\frac{8\pi a^4}{3} \right) \\ &= \frac{m}{4\pi a^2} \cdot \frac{8\pi a^4}{3} \\ &= \frac{2}{3} m a^2. \end{aligned}$$

4.5.3 Surface Flux

Surface flux is an important type of surface integrals in both mathematics and physics. It will appear in the statement of the Stokes' Theorem, and also plays an important role in electricity and magnetism. In a nutshell, the surface flux of a vector field \mathbf{F} through a surface S measures the "amount" of vectors passing through S .

Let's begin our discussion with the simplest case where the vector field \mathbf{F} is constant and the surface is simply a flat plane.

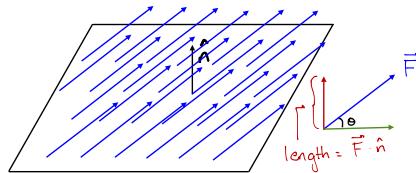


Figure 4.12: surface flux for uniform vector field through a plane

The force \mathbf{F} can be decomposed into two components, one perpendicular to the plane, another parallel to the plane. As the flux counts only the amount of vectors *through* the plane, only those perpendicular to the plane should be counted. Suppose the force \mathbf{F} makes an angle θ with the unit normal vector $\hat{\mathbf{n}}$, then the perpendicular component of the force has length $\mathbf{F} \cdot \hat{\mathbf{n}} = |\mathbf{F}| \cos \theta$.

Furthermore, there are vectors at every point on the plane. The larger the plane, the more vectors passing through it. Therefore, the flux of \mathbf{F} through the plane should be defined as:

$$(\mathbf{F} \cdot \hat{\mathbf{n}})(\text{area of the plane}).$$

Now consider a curved surface (so that $\hat{\mathbf{n}}$ is no longer constant) and \mathbf{F} is no longer a constant vector field. Infinitesimally, each area element can be regarded as a tiny flat parallelogram, and both \mathbf{F} and $\hat{\mathbf{n}}$ can be regarded as constant vectors over each of tiny parallelogram. Recall that dS is the area of this element. Therefore, the flux of \mathbf{F} through this tiny bit of surface is given by:

$$\mathbf{F} \cdot \hat{\mathbf{n}} \, dS$$

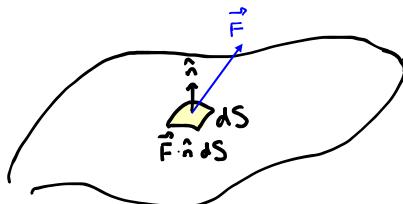


Figure 4.13: flux through a tiny surface element

Summing up, the total flux of \mathbf{F} through the entire surface is given by a surface integral as stated in the definition below:

Definition 4.7 — Surface Flux. Given a vector field \mathbf{F} and a surface S , the surface flux of \mathbf{F} through S is defined to be the surface integral:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS$$

where $\hat{\mathbf{n}}$ denotes unit normal vector to S at each point.

- (i) Note that there are often two choice of $\hat{\mathbf{n}}$, so the sign of the surface flux depends on which direction of $\hat{\mathbf{n}}$ is chosen. For a closed surface such as a sphere, it is a convention to choose the **outward** unit normal.
- (ii) There are some surfaces which you cannot “choose” a unit normal convention. For instance, if you pick a normal vector on the Möbius strip and let it vary continuously over the strip, the normal vector may end up pointing at opposite direction when it returns to the original position! We call these **non-orientable** surfaces. We do not define surface flux on non-orientable surfaces.

Generally speaking, we need to parametrize the surface in order to compute the flux. However, there are some exceptions that can be done without parametrization.

■ **Example 4.13** Consider the gravitational vector field given in spherical coordinates by:

$$\mathbf{F} = -\frac{GMm}{\rho^2} \mathbf{e}_\rho.$$

Compute the *outward* flux of \mathbf{F} through the sphere S with radius ρ_0 centered at the origin, i.e.

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS.$$

■ **Solution** Outward flux means the unit normal vector $\hat{\mathbf{n}}$ is chosen to be the outward normal. For a round sphere centered at the origin, clearly we have $\hat{\mathbf{n}} = \mathbf{e}_\rho$.

Therefore,

$$\begin{aligned} \mathbf{F} \cdot \hat{\mathbf{n}} &= -\frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \mathbf{e}_\rho \\ &= -\frac{GMm}{\rho^2} |\mathbf{e}_\rho|^2 \\ &= -\frac{GMm}{\rho^2}. \end{aligned}$$

Hence, the outward flux is given by:

$$\begin{aligned}
 \iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS &= \iint_S -\frac{GMm}{\rho^2} \, dS \\
 &= \iint_S -\frac{GMm}{\rho_0^2} \, dS && (\text{since } \rho = \rho_0 \text{ on } S) \\
 &= -\frac{GMm}{\rho_0^2} \underbrace{\iint_S \, dS}_{\text{surface area of } S} \\
 &= -\frac{GMm}{\rho_0^2} \cdot 4\pi\rho_0^2 = -4\pi GMm.
 \end{aligned}$$

- (i)** Using the Divergence Theorem, we will later see that the above flux is always $-4\pi GMm$ for any closed orientable surface S provided that it encloses the origin. This fact is commonly called the *Gauss's Law for Gravity*.

Most surface flux cannot be compute as neatly as in the previous example. Generally speaking, the surface flux can be computed by parametrizing the surface. Given a parametrization $\mathbf{r}(u, v)$, with $a \leq u \leq b$ and $c \leq v \leq d$, of a surface S , the unit normal vector $\hat{\mathbf{n}}$ is given by:

$$\hat{\mathbf{n}} = \pm \frac{\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}}{\left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right|}$$

where \pm depends on the convention chosen.

It looks complicated to compute the normal vector and the surface flux. However, as a coincidence, the area element dS is given by:

$$dS = \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \, dudv.$$

Therefore, when one multiplies $\mathbf{F} \cdot \hat{\mathbf{n}}$ by dS , the term $\left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right|$ got canceled! Precisely, we have:

Theorem 4.5 Let $\mathbf{r}(u, v)$, with $a \leq u \leq b$ and $c \leq v \leq d$, be a parametrization of a surface S . The surface flux of a vector field \mathbf{F} through S can be computed by:

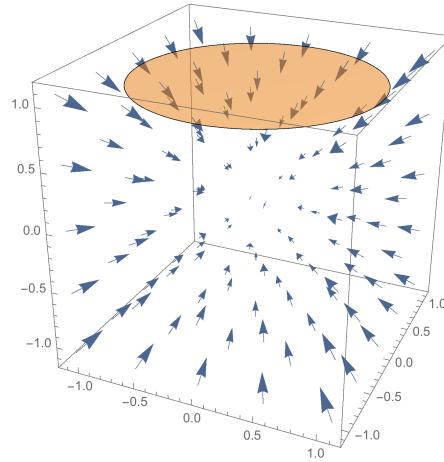
$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS = \pm \int_c^d \int_a^b \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \, dudv$$

where \pm depends on the chosen convention of $\hat{\mathbf{n}}$.

Proof. From the given parametrization $\mathbf{r}(u, v)$, we have

$$\begin{aligned}
 \iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS &= \int_c^d \int_a^b \mathbf{F} \cdot \hat{\mathbf{n}} \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \, dudv \\
 &= \pm \int_c^d \int_a^b \mathbf{F} \cdot \frac{\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}}{\left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right|} \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \, dudv \\
 &= \pm \int_c^d \int_a^b \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \, dudv.
 \end{aligned}$$

■

Figure 4.14: \mathbf{F} and S in Example 4.14

■ **Example 4.14** Consider the vector field:

$$\mathbf{F} = -GMm \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{(x^2 + y^2 + z^2)^{3/2}}.$$

Let S be part of the horizontal plane $z = 1$ over the region $x^2 + y^2 \leq 1$. Compute the surface flux of \mathbf{F} through S , i.e.

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS.$$

Here $\hat{\mathbf{n}}$ is chosen to be the upward normal.

■ **Solution** First we parametrize the surface S . Since the plane $z = 1$ is a graph over the xy -plane, it seems the easiest way is to use x and y as parameters. However, the base region $x^2 + y^2 \leq 1$ is not a rectangle so it may be difficult to set up the ranges of x and y for the parametrization. Since the base region is a solid circle, we can use cylindrical coordinates too. Let

$$\mathbf{r}(r, \theta) = (r \cos \theta)\mathbf{i} + (r \sin \theta)\mathbf{j} + \mathbf{k} \quad (\text{since } z = 1),$$

where $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. Next we compute all the ingredients:

$$\begin{aligned}
 \frac{\partial \mathbf{r}}{\partial r} &= (\cos \theta) \mathbf{i} + (\sin \theta) \mathbf{j} \\
 \frac{\partial \mathbf{r}}{\partial \theta} &= (-r \sin \theta) \mathbf{i} + (r \cos \theta) \mathbf{j} \\
 \frac{\partial \mathbf{r}}{\partial r} \times \frac{\partial \mathbf{r}}{\partial \theta} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos \theta & \sin \theta & 0 \\ -r \sin \theta & r \cos \theta & 0 \end{vmatrix} \\
 &= (r \cos^2 \theta + r \sin^2 \theta) \mathbf{k} \\
 &= r \mathbf{k} \quad (\text{which is upward}) \\
 \mathbf{F} &= -GMm \frac{(r \cos \theta) \mathbf{i} + (r \sin \theta) \mathbf{j} + z \mathbf{k}}{(r^2 \cos^2 \theta + r^2 \sin^2 \theta + z^2)^{3/2}} \\
 &= -GMm \frac{(r \cos \theta) \mathbf{i} + (r \sin \theta) \mathbf{j} + \mathbf{k}}{(r^2 + 1)^{3/2}} \quad (\text{since } z = 1) \\
 \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial r} \times \frac{\partial \mathbf{r}}{\partial \theta} \right) &= -\frac{GMmr}{(r^2 + 1)^{3/2}}.
 \end{aligned}$$

Therefore, by Theorem 4.5, the surface flux is given by

$$\begin{aligned}
 \iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS &= \int_0^{2\pi} \int_0^1 \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial r} \times \frac{\partial \mathbf{r}}{\partial \theta} \right) dr d\theta \\
 &= - \int_0^{2\pi} \int_0^1 \frac{GMmr}{(r^2 + 1)^{3/2}} dr d\theta \\
 &= \int_0^{2\pi} \left[\frac{GMm}{\sqrt{1+r^2}} \right]_{r=0}^{r=1} d\theta \quad (\text{by inspection}) \\
 &= GMm \int_0^{2\pi} \left(\frac{1}{\sqrt{2}} - 1 \right) d\theta \\
 &= 2\pi GMm \left(\frac{1}{\sqrt{2}} - 1 \right).
 \end{aligned}$$

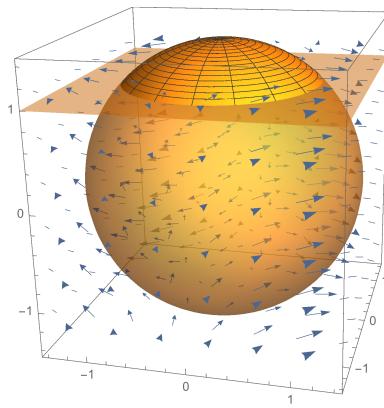


Figure 4.15: \mathbf{F} and S in Example 4.15

■ **Example 4.15** Let $\mathbf{F} = xi - yj$. Find the upward flux of \mathbf{F} over S which is the upper part of the sphere with radius $\sqrt{2}$ centered at the origin cut out by the plane $z = 1$.

■ **Solution** Since the surface is spherical, it is usually the best to use spherical coordinates to parametrize it. Under spherical coordinates, the sphere is represented by $\rho = \sqrt{2}$, so we use θ and φ for the parameters. Let

$$\mathbf{r}(\theta, \varphi) = \left\langle \sqrt{2} \sin \varphi \cos \theta, \sqrt{2} \sin \varphi \sin \theta, \sqrt{2} \cos \varphi \right\rangle.$$

The domain of the parameters are $0 \leq \theta \leq 2\pi$ and $0 \leq \varphi \leq \frac{\pi}{4}$.

As in the previous example, we first compute all the ingredients. Since they are all straight-forward computations, some detail will be omitted here.

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial \theta} &= \left\langle -\sqrt{2} \sin \varphi \sin \theta, \sqrt{2} \sin \varphi \cos \theta, 0 \right\rangle \\ \frac{\partial \mathbf{r}}{\partial \varphi} &= \left\langle \sqrt{2} \cos \varphi \cos \theta, \sqrt{2} \cos \varphi \sin \theta, -\sqrt{2} \sin \varphi \right\rangle \\ \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} &= \left\langle -2 \cos \varphi \sin^2 \theta, -2 \sin \varphi \sin^2 \theta, -2 \cos \varphi \sin \varphi \right\rangle \\ \mathbf{F} &= x\mathbf{i} - y\mathbf{j} = \left\langle \sqrt{2} \sin \varphi \cos \theta, -\sqrt{2} \sin \varphi \sin \theta, 0 \right\rangle \\ \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} \right) &= -2\sqrt{2} \cos 2\theta \sin^3 \varphi.\end{aligned}$$

Note that $\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi}$ obtained above is a downward normal since the \mathbf{k} -component is negative. The upward flux is given by:

$$\begin{aligned}\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS &= \int_0^{2\pi} \int_0^{\pi/4} \mathbf{F} \cdot \left(-\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} \right) d\varphi d\theta \\ &= \int_0^{2\pi} \int_0^{\pi/4} 2\sqrt{2} \cos 2\theta \sin^3 \varphi d\varphi d\theta \\ &= 2\sqrt{2} \underbrace{\left(\int_0^{2\pi} \cos 2\theta d\theta \right)}_{=0} \left(\int_0^{\pi/4} \sin^3 \varphi d\varphi \right) = 0.\end{aligned}$$

Physical Interpretations of Surface Flux

Generally speaking, the flux of a vector field \mathbf{F} through a surface S measures the net “amount” of vectors \mathbf{F} passing through S along a chosen direction of normal vector $\hat{\mathbf{n}}$. The unit of the “amount” depends on the physical meaning of the vector field \mathbf{F} .

For instance, if \mathbf{u} is the velocity vector field of fluid (in ms^{-1}), and the surface area of S has unit m^2 , then the surface flux

$$\iint_S \mathbf{u} \cdot \hat{\mathbf{n}} dS$$

has unit $m^3 s^{-1}$ and so it measures the net volume of fluid through the surface S in the direction of $\hat{\mathbf{n}}$. If the flux is positive, then there is more fluid flowing along the direction $\hat{\mathbf{n}}$ than against it. On the other hand, if the flux is negative, there is more fluid flowing against $\hat{\mathbf{n}}$ than along it.

If S is a closed surface (such as a sphere, a cube, a torus), it is a convention to take $\hat{\mathbf{n}}$ as the outward unit normal. The surface flux

$$\iint_S \mathbf{u} \cdot \hat{\mathbf{n}} dS$$

over this closed surface measures the net volume of fluid flowing in the direction of $\hat{\mathbf{n}}$ through the surface, or in other words, the net volume of fluid flowing **out** from region D enclosed by S .

If one denotes ρ as the uniform density of the fluid, then

$$\iint_S \rho \mathbf{u} \cdot \hat{\mathbf{n}} \, dS$$

measures the net mass of the fluid flowing out from the enclosed region D through its boundary S per unit time. Assuming there is no sink or source inside D , by the conservation of mass, the rate of change of the total mass of fluid enclosed by S is related to the surface flux by:

$$\frac{\partial}{\partial t} \underbrace{\iiint_D \rho \, dV}_{\text{total mass in } D} = - \iint_S \rho \mathbf{u} \cdot \hat{\mathbf{n}} \, dS.$$

Later we can apply the Divergence Theorem on the above relation to derive an important equation to the fluid flow.

Now suppose the vector field \mathbf{J} represents the transfer of heat energy in unit Joule per second. Note that while energy is a scalar, the transfer of heat at different point may have a different direction and so it is a vector quantity. Again, take S to be a closed surface enclosing a solid region D , then the surface flux of \mathbf{j} through S :

$$\iint_S \mathbf{J} \cdot \hat{\mathbf{n}} \, dS$$

measures the amount of heat energy flowing out from the region D through S . This flux integral is commonly called the *heat flux* through S by physicists.

If \mathbf{E} is a electric field and, again, S is a closed surface, then the flux integral

$$\iint_S \mathbf{E} \cdot \hat{\mathbf{n}} \, dS$$

is commonly called the *electric flux* through S . A result by Gauss claims that this flux integral is proportional to the total amount of charges enclosed by the surface. If \mathbf{B} is a magnetic field and, again, S is a closed surface, then the flux integral

$$\iint_S \mathbf{B} \cdot \hat{\mathbf{n}} \, dS$$

is commonly called the *magnetic flux* through S . Gauss said that it must be zero.

4.6 Stokes' Theorem

This section was written during the author's first flight of Boeing 787 from Boston to Tokyo. The author is grateful that Japan Airlines offer in-seat power for his laptop throughout the entire 14-hour journey.

4.6.1 Stokes' Theorem for Simply-Connected Surfaces

Recall that the Green's Theorem relates the line integral of a closed curve to a certain double integral over the region enclosed by the curve. The Stokes' Theorem is its generalization to the three dimensions, which allows the region to be a curved surface in \mathbb{R}^3 .

Theorem 4.6 — Stokes' Theorem. Let S be an orientable, simply-connected surface in \mathbb{R}^3 , and C be the boundary curve of the surface S . Suppose \mathbf{F} is a vector field which is defined and is continuously differentiable on and near the surface S , then we have:

$$\underbrace{\oint_C \mathbf{F} \cdot d\mathbf{r}}_{\text{line integral}} = \underbrace{\iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS}_{\text{surface integral}}$$

where $\hat{\mathbf{n}}$ is the unit normal vector to S , with direction determined by the right-hand rule (see Figure 4.16).

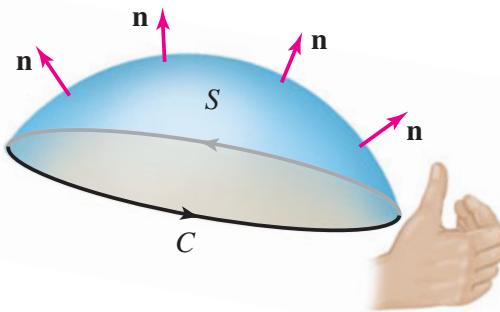


Figure 4.16: right-hand rule

- (i) By comparing the statements of the Green's and Stokes' Theorems, one can easily see that the Green's Theorem is a special case of the Stokes' Theorem, in a sense that the former applies to plane curve and the flat region enclosed by the curve on the xy -plane. The unit normal vector $\hat{\mathbf{n}}$ for the plane region is obviously given by \mathbf{k} if the plane curve is traveling in the counter-clockwise orientation.
- (i) For closed curves in the three-dimensional space, one cannot say whether they are counter-clockwise or clockwise as it depends on the direction of observations. Therefore, the counter-clockwise convention of the Green's Theorem is generalized to the right-hand rule condition in the statement of the Stokes' Theorem.
- (i) The Stokes' Theorem applies only on orientable surfaces. That says, it may not hold for surfaces such as the Möbius strip. Also, the condition where the vector field \mathbf{F} needs to be defined and is continuously differentiable on and near the surface S is crucial. However, we will mostly deal with vector fields that satisfy this condition.
- (i) The condition that S has to be simply-connected is also crucial, but we will later learn how to modify the Stokes' Theorem so as to allow non-simply-connected surface S .

The proof of the Stokes' Theorem is omitted here. Interested readers may consult the textbook for a proof of one special case. Let's look at some examples.

■ **Example 4.16** Let S be the hemisphere $x^2 + y^2 + z^2 = 4$, $z \geq 0$ above the the xy -plane, and C be its boundary curve oriented counter-clockwise on the xy -plane. Given $\mathbf{F} = (z - y)\mathbf{i} + x\mathbf{j} - x\mathbf{k}$. Determine the line integral

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

using the Stokes' Theorem.

■ **Solution** The Stokes' Theorem asserts that:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, dS.$$

Since the RHS is a surface integral, we need to parametrize it first in order to compute it.

Using spherical coordinates, the parametrization of S is given by:

$$\mathbf{r}(\theta, \varphi) = (2 \sin \varphi \cos \theta)\mathbf{i} + (2 \sin \varphi \sin \theta)\mathbf{j} + (2 \cos \varphi)\mathbf{k}$$

where $0 \leq \theta \leq 2\pi$ and $0 \leq \varphi \leq \frac{\pi}{2}$. Then,

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial \theta} &= (-2 \sin \varphi \cos \theta)\mathbf{i} + (2 \sin \varphi \cos \theta)\mathbf{j} + 0\mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial \varphi} &= (2 \cos \varphi \cos \theta)\mathbf{i} + 2 \cos \varphi \sin \theta)\mathbf{j} + (-2 \sin \varphi)\mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} &= (-4 \sin^2 \varphi \cos \theta)\mathbf{i} + (-4 \sin^2 \varphi \sin \theta)\mathbf{j} + (-4 \sin \varphi \cos \varphi)\mathbf{k} \\ \frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta} &= (4 \sin^2 \varphi \cos \theta)\mathbf{i} + (4 \sin^2 \varphi \sin \theta)\mathbf{j} + (4 \sin \varphi \cos \varphi)\mathbf{k}\end{aligned}$$

Note that $\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi}$ is pointing downward, so we use $\frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta}$ instead.

Next we need to compute $\nabla \times \mathbf{F}$:

$$\begin{aligned}\nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z - y & x & -x \end{vmatrix} \\ &= 2\mathbf{j} + 2\mathbf{k}.\end{aligned}$$

Therefore, using the Stokes' Theorem, we have

$$\begin{aligned}
 \oint_C \mathbf{F} \cdot d\mathbf{r} &= \iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \, dS \\
 &= \iint_S (2\mathbf{j} + 2\mathbf{k}) \cdot \hat{\mathbf{n}} \, dS \\
 &= \int_0^{\pi/2} \int_0^{2\pi} (2\mathbf{j} + 2\mathbf{k}) \cdot \left(\frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta} \right) \, d\theta d\varphi \quad (\text{using Theorem 4.5}) \\
 &= \int_0^{\pi/2} \int_0^{2\pi} (8 \sin^2 \varphi \sin \theta + 8 \sin \varphi \cos \varphi) \, d\theta d\varphi \\
 &= 8 \left(\int_0^{\pi/2} \sin^2 \varphi \, d\varphi \right) \underbrace{\left(\int_0^{2\pi} \sin \theta \, d\theta \right)}_{=0} + 2\pi \int_0^{\pi/2} 4 \sin 2\varphi \, d\varphi \\
 &= 8\pi \left[-\frac{1}{2} \cos 2\varphi \right]_0^{\pi/2} = 8\pi.
 \end{aligned}$$

i Although the Stokes' Theorem was used in the previous example (as required by the problem), it is actually easier to compute the line integral directly by parametrizing the curve:

Let $\mathbf{r}(t) = (2 \cos t)\mathbf{i} + (2 \sin t)\mathbf{j} + 0\mathbf{k}$ where $0 \leq t \leq 2\pi$. Then

$$\begin{aligned}
 \oint_C \mathbf{F} \cdot d\mathbf{r} &= \oint_C \mathbf{F} \cdot \mathbf{r}'(t) \, dt \\
 &= \int_0^{2\pi} ((z-y)\mathbf{i} + x\mathbf{j} - x\mathbf{k}) \cdot ((-2 \sin t)\mathbf{i} + (2 \cos t)\mathbf{j} + 0\mathbf{k}) \, dt \\
 &= \int_0^{2\pi} (-2(z-y) \sin t + 2x \cos t) \, dt \\
 &= \int_0^{2\pi} -2(-2 \sin t) \sin t + 2(2 \cos t) \cos t \, dt \\
 &= \int_0^{2\pi} 4(\sin^2 t + \cos^2 t) \, dt = 8\pi.
 \end{aligned}$$

The purpose of the previous example is simply to illustrate how to use the Stokes' Theorem, although it is not necessary to use it. The line integral in the next example, however, would be extremely difficult to compute without the Stokes' Theorem.

■ **Example 4.17** Let C be the curve of intersection of the plane $Ax + By + Cz = 0$ and the sphere $x^2 + y^2 + z^2 = a^2$. Show that

$$\oint_C ydx + zdy + xdz = \pm \frac{\pi a^2(A + B + C)}{\sqrt{A^2 + B^2 + C^2}}$$

where \pm is determined by the orientation of C .

■ **Solution** The plane $Ax + By + Cz = 0$ passes through the origin. Therefore, the curve C is a great circle on the sphere. However, this great circle is neither horizontal or vertical, so it is difficult to parametrize C to compute the line integral.

Let's use the Stokes' Theorem to see if there is any luck! When using the Stokes' Theorem, one needs to pick a surface S whose boundary curve is C . There are three choices:

1. region of the plane enclosed by C
2. the hemisphere above the plane
3. the hemisphere below the plane

All of the above choice should give the same conclusion. However, let's pick the region of the plane enclosed by C to be the surface S and we will explain why it is the smartest choice among all three.

Now the given line integral is associated to the vector field $\mathbf{F} = y\mathbf{i} + z\mathbf{j} + x\mathbf{k}$, i.e.

$$\oint_C ydx + zd\mathbf{y} + xdz = \oint_C \mathbf{F} \cdot d\mathbf{r}.$$

In order to apply the Stokes' Theorem, we need to compute the curl:

$$\begin{aligned}\nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & z & x \end{vmatrix} \\ &= -\mathbf{i} - \mathbf{j} - \mathbf{k}.\end{aligned}$$

We also need the unit normal vector $\hat{\mathbf{n}}$, but since the region S is a plane whose equation is $Ax + By + Cz = 0$. The unit normal is a constant vector given by:

$$\hat{\mathbf{n}} = \pm \frac{A\mathbf{i} + B\mathbf{j} + C\mathbf{k}}{\sqrt{A^2 + B^2 + C^2}}$$

where \pm is determined by the orientation of C .

Therefore,

$$(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} = \pm (-\mathbf{i} - \mathbf{j} - \mathbf{k}) \cdot \left(\frac{A\mathbf{i} + B\mathbf{j} + C\mathbf{k}}{\sqrt{A^2 + B^2 + C^2}} \right) = \pm \frac{A + B + C}{\sqrt{A^2 + B^2 + C^2}}.$$

Next, we apply the Stokes' Theorem on these C and S :

$$\begin{aligned}\oint_C \mathbf{F} \cdot d\mathbf{r} &= \iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS \\ &= \pm \iint_S \frac{A + B + C}{\sqrt{A^2 + B^2 + C^2}} dS \\ &= \pm \frac{A + B + C}{\sqrt{A^2 + B^2 + C^2}} \underbrace{\iint_S dS}_{\text{surface area}} \\ &= \pm \frac{\pi a^2 (A + B + C)}{\sqrt{A^2 + B^2 + C^2}}.\end{aligned}$$

Note that the surface area of the region S on the plane is πa^2 , since its boundary is a circle with radius a .

There are two major reasons why the part of the plane enclosed by C is a smarter choice for S than the hemispheres. For one thing, both $\nabla \times \mathbf{F}$ and $\hat{\mathbf{n}}$ are constant vector field if S is chosen to be a planar region, so that computing its surface integral is very easy – no parametrization! For another, if any of the hemispheres were chosen to be S , then the surface integral needs to be computed by parametrization – which can be tedious. It is also very difficult to determine the range of values of φ and θ since the plane cutting the sphere is not a horizontal one.

Occasionally, the Stokes' Theorem can be applied to evaluate a surface integral over an arbitrary or complicated surface, which is not easy to be parametrized. If the given vector field \mathbf{G} can be expressed in the form of $\mathbf{F} = \nabla \times \mathbf{G}$ for another vector field \mathbf{G} , by the (backward) Stokes' Theorem asserts that:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS = \iint_S (\nabla \times \mathbf{G}) \cdot \hat{\mathbf{n}} dS = \oint_C \mathbf{G} \cdot d\mathbf{r}$$

where C is the boundary curve of the surface S . Very often, the line integral is easier to compute if one can parametrize it.

- (1)** Note that the above discussion holds only when the given vector field \mathbf{F} is of the form $\mathbf{F} = \nabla \times \mathbf{G}$. If such an \mathbf{F} is not in this form, there is no easy way to apply the Stokes' Theorem backward.

■ **Example 4.18** Let C be an arbitrary simple closed curve on the xy -plane in the xyz -space, and S be an arbitrary surface above the xy -plane with boundary curve C .

1. Verify that $\mathbf{i} = \nabla \times \left(-\frac{z}{2}\mathbf{j} + \frac{y}{2}\mathbf{k}\right)$.
2. Show that:

$$\iint_S \mathbf{i} \cdot \hat{\mathbf{n}} dS = 0.$$

■ **Solution** Part (1) is straight-forward:

$$\begin{aligned} \nabla \times \left(-\frac{z}{2}\mathbf{j} + \frac{y}{2}\mathbf{k}\right) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & -\frac{z}{2} & \frac{y}{2} \end{vmatrix} \\ &= \left(\frac{\partial}{\partial y}\left(\frac{y}{2}\right) - \frac{\partial}{\partial z}\left(-\frac{z}{2}\right)\right)\mathbf{i} - 0\mathbf{j} + 0\mathbf{k} \\ &= \mathbf{i} \end{aligned}$$

For part (2), we denote $\mathbf{G} = -\frac{z}{2}\mathbf{j} + \frac{y}{2}\mathbf{k}$ for simplicity. Then $\mathbf{i} = \nabla \times \mathbf{G}$. Applying the Stokes' Theorem backward, we get:

$$\begin{aligned} \iint_S \mathbf{i} \cdot \hat{\mathbf{n}} dS &= \iint_S (\nabla \times \mathbf{G}) \cdot \hat{\mathbf{n}} dS \\ &= \oint_C \mathbf{G} \cdot d\mathbf{r}. \end{aligned}$$

Since the curve C is arbitrary in nature, there is no way to parametrize C . However, it is given that C is on the xy -plane! Therefore, one can use the Green's Theorem to evaluate the above line integral. Denote R to be the region on the xy -plane enclosed by the curve C , then the Green's Theorem asserts that:

$$\begin{aligned} \oint_C \mathbf{G} \cdot d\mathbf{r} &= \iint_R (\nabla \times \mathbf{G}) \cdot \mathbf{k} dA \\ &= \iint_R \mathbf{i} \cdot \mathbf{k} dA = \iint_R 0 dA = 0. \end{aligned}$$

4.6.2 Significance of the Stokes' Theorem

Interpretation of Curl

Using the Stokes' Theorem, one can give a geometric interpretation of $\nabla \times \mathbf{F}$. Consider a tiny surface S with boundary curve C . Denote $\hat{\mathbf{n}}$ to be the unit normal vector of S with direction determined by the right-hand rule. Since the surface S is very small, one can regard the quantity $(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}}$ is nearly a constant over the surface S . By the Stokes' Theorem, we have:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS \simeq (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \left(\iint_S dS \right).$$

Therefore,

$$(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \simeq \frac{\oint_C \mathbf{F} \cdot d\mathbf{r}}{\text{Surface area of } S}.$$

This quantity is large when the vector field \mathbf{F} is circular about the normal vector $\hat{\mathbf{n}}$. In other words, the quantity $(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}}$ measures the circulation density around any given point. That's why $\nabla \times \mathbf{F}$ is often called the *curl* of \mathbf{F} .

Conservative Vector Fields

Recall that a vector field \mathbf{F} is conservative if there exists a scalar function f such that $\mathbf{F} = \nabla f$. If \mathbf{F} is defined and continuously differentiable everywhere (or on a simply-connected domain), then the curl test asserts that \mathbf{F} is conservative if and only if $\nabla \times \mathbf{F} = \mathbf{0}$. Using the Stokes' Theorem, if \mathbf{F} is conservative, then:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_S \underbrace{(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}}}_{=0} dS = 0,$$

which recovers the result we proved before in Theorem 4.1. Of course, the Stokes' Theorem asserts more than Theorem 4.1 does because it applies to any vector field, not just the conservative ones.

4.6.3 Surfaces with Multiple Boundaries

When applying the Stokes' Theorem on surfaces with multiple boundaries (i.e. not simply-connected), such as the one in Figure 4.17, one needs to be very careful when dealing with the inner boundary.

The form of the Stokes' Theorem, as stated in Theorem 4.6, applies only to simply-connected regions. However, using a similar technique illustrated in the subsection about the winding number, one can extend the Stokes' Theorem so that it can be applied to surfaces with *holes* as well. Take the region in Figure 4.17 as an example. One can subdivide the surface by cutting along arcs that connect the outer boundary and the inner boundaries.

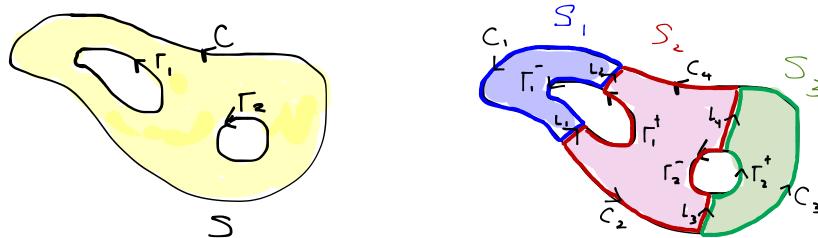


Figure 4.17: apply Stokes' Theorem for surfaces with holes

Each sub-surface of S_1 , S_2 and S_3 now becomes simply-connected and so the Stokes' Theorem applies to each sub-surface:

$$\underbrace{\left(\int_{C_1} + \int_{L_1} - \int_{\Gamma_1^-} + \int_{L_2} \right) \mathbf{F} \cdot d\mathbf{r}}_{\text{form the boundary of } S_1} = \iint_{S_1} (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS$$

$$\underbrace{\left(\int_{C_2} + \int_{L_3} - \int_{\Gamma_2^-} + \int_{L_4} + \int_{C_4} - \int_{L_2} - \int_{\Gamma_1^+} - \int_{L_1} \right) \mathbf{F} \cdot d\mathbf{r}}_{\text{form the boundary of } S_2} = \iint_{S_2} (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS$$

$$\underbrace{\left(\int_{C_3} - \int_{L_4} - \int_{\Gamma_2^+} - \int_{L_3} \right) \mathbf{F} \cdot d\mathbf{r}}_{\text{form the boundary of } S_3} = \iint_{S_3} (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS.$$

Summing up all three equations and cancelling out the L_i 's terms, we get:

$$\begin{aligned} & \left(\int_{C_1} + \int_{C_2} + \int_{C_3} + \int_{C_4} - \int_{\Gamma_1^-} - \int_{\Gamma_1^+} - \int_{\Gamma_2^-} - \int_{\Gamma_2^+} \right) \mathbf{F} \cdot d\mathbf{r} \\ &= \left(\iint_{S_1} + \iint_{S_2} + \iint_{S_3} \right) (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS \end{aligned}$$

Combining all C_i 's, Γ_i 's and S_i 's, we yield:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} - \oint_{\Gamma_1} \mathbf{F} \cdot d\mathbf{r} - \oint_{\Gamma_2} \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS.$$

While the surface integral (RHS) is in the same form as in the usual Stokes' Theorem, the LHS is not *summing* up the boundary line integrals, but rather the outer boundary has a plus sign in front and the inner boundaries each has a minus sign in front.

For more complicated surfaces (with many, but finitely many, holes), one can apply the technique illustrated above to establish:

Theorem 4.7 — Stokes' Theorem for Higher Genus^a Surfaces. Let S be an orientable surface in \mathbb{R}^3 with n holes. Denote C to be its outer boundary, and $\Gamma_1, \Gamma_2, \dots, \Gamma_n$ to be its inner boundaries. Suppose \mathbf{F} is a vector field is defined and continuously differentiable on and near the surface S , then:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} - \sum_{i=1}^n \oint_{\Gamma_i} \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS.$$

Here $\hat{\mathbf{n}}$ is the unit normal vector to S with orientation determined by the right-hand rule applied to the outer boundary C .

^aThe word *genus* is the mathematical term for number of “holes” inside the surface.

4.7 Divergence Theorem

The Green's and Stokes' Theorems relate the line integral of a vector field over a simple closed curve with a double/surface integral of the curl of the vector field. In this section, we are going to learn the Divergence Theorem, which relates the surface integral over a closed surface with a triple integral over the solid region enclosed by the surface.

4.7.1 Divergence Operator

In order to state the Divergence Theorem, we need to define:

Definition 4.8 — Divergence Operator. Given a continuously differentiable vector field \mathbf{F} in \mathbb{R}^3 whose components in rectangular coordinates are:

$$\mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k},$$

the divergence of \mathbf{F} , denoted by $\nabla \cdot \mathbf{F}$, is defined as:

$$\nabla \cdot \mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}.$$

Note that if the vector field \mathbf{F} is given in spherical coordinates, the divergence of \mathbf{F} is **not** defined as:

$$\nabla \cdot \mathbf{F} = \frac{\partial F_\rho}{\partial \rho} + \frac{\partial F_\theta}{\partial \theta} + \frac{\partial F_\varphi}{\partial \varphi} \quad \text{WRONG!}$$

Using the chain rule and the conversion rules of components, one can convert $\nabla \times \mathbf{F}$ from the rectangular coordinate form to both spherical and cylindrical forms:

$$\begin{aligned} \nabla \cdot \mathbf{F} &= \frac{1}{\rho^2} \frac{\partial}{\partial \rho} (\rho^2 F_\rho) + \frac{1}{\rho \sin \varphi} \frac{\partial}{\partial \varphi} (F_\theta \sin \varphi) + \frac{1}{\rho \sin \varphi} \frac{\partial F_\varphi}{\partial \theta} \\ &= \frac{1}{r} \frac{\partial}{\partial r} (r F_r) + \frac{1}{r} \frac{\partial F_\theta}{\partial \theta} + \frac{\partial F_z}{\partial z}. \end{aligned}$$

The derivation, although straight-forward, is very tedious and is therefore omitted here.

We will see the geometric interpretation of $\nabla \times \mathbf{F}$ after we state the Divergence Theorem. Essentially, it measures how *diverging* the vector field is.

4.7.2 Divergence Theorem for Solids without Holes

Theorem 4.8 — Divergence Theorem. Let S be a closed orientable surface enclosing a simply-connected solid region D . Suppose \mathbf{F} is a vector field defined and being continuously differentiable in and near the region D , then we have:

$$\underbrace{\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS}_{\text{surface integral}} = \underbrace{\iiint_D \nabla \cdot \mathbf{F} dV}_{\text{triple integral}}.$$

Here $\hat{\mathbf{n}}$ is the outward normal of S .

The Divergence Theorem is particularly useful for computing the flux over a closed surface, as the theorem says we do not need to parametrize the surface and compute the normal vector, and all is needed is to compute a triple integral instead, which is often easier than computing surface integrals. Let's look at some examples.

■ **Example 4.19** Consider the vector field $\mathbf{F} = 3x\mathbf{i} + 4y\mathbf{j} - 5z\mathbf{k}$. Let S be the sphere with radius a centered at the origin. Evaluate the flux integral:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS$$

where $\hat{\mathbf{n}}$ is the outward normal.

■ **Solution** You can imagine the computation would be quite tedious if we computed this flux integral directly by parametrizing the sphere. However, since this is a flux integral over a closed surface, one can try to use the Divergence Theorem to compute it!

By easy computations, we have:

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(3x) + \frac{\partial}{\partial y}(4y) + \frac{\partial}{\partial z}(-5z) = 3 + 4 - 5 = 2.$$

Denote D to be the solid sphere with radius a centered at the origin, i.e. the solid region enclosed by S . By the Divergence Theorem, we have:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS = \iiint_D \nabla \cdot \mathbf{F} \, dV = \iiint_D 2 \, dV = 2 \cdot \frac{4}{3}\pi a^3 = \frac{8}{3}\pi a^3.$$

■ **Example 4.20** Let $\mathbf{F} = x^2\mathbf{i} + 4xyz\mathbf{j} + ze^x\mathbf{k}$, D be the rectangular box defined by $0 \leq x \leq 3$, $0 \leq y \leq 2$ and $0 \leq z \leq 1$ and S be the boundary surface of the D (i.e. S is the shell of the box). Evaluate the flux integral:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS$$

where $\hat{\mathbf{n}}$ is the outward normal.

■ **Solution** The closed surface S has six faces! If one attempts to compute the flux integral directly, one needs to split it into six integrals, corresponding to each of its six faces.

However, if one applies the Divergence Theorem, the difficult surface integral becomes a triple integral over a rectangular region which is very easy to set-up.

We first compute:

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(x^2) + \frac{\partial}{\partial y}(4xyz) + \frac{\partial}{\partial z}(ze^x) = 2x + 4xz + e^x.$$

By the Divergence Theorem, we get:

$$\begin{aligned} \iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, dS &= \iiint_D \nabla \cdot \mathbf{F} \, dV \\ &= \int_{z=0}^{z=1} \int_{y=0}^{y=2} \int_{x=0}^{x=3} (2x + 4xz + e^x) \, dx \, dy \, dz \\ &= 34 + 2e^3. \end{aligned}$$

We omit the computational detail of the triple integral above, which is a very straight-forward.

Let's demonstrate one example using the cylindrical form of the divergence.

■ **Example 4.21** Let $\mathbf{F} = r(2 + \sin^2 \theta)\mathbf{e}_r + (r \sin \theta \cos \theta) \mathbf{e}_\theta + 3z\mathbf{k}$, and S be the quarter-cylinder $r = 2$, $0 \leq \theta \leq \pi$ and $0 \leq z \leq 5$. Compute the surface flux:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS.$$

■ **Solution** The cylinder has 5 faces. It would be time-consuming to compute the total flux face-by-face directly. Let's try to use the Divergence Theorem (since S is a closed surface).

The vector field \mathbf{F} is given in cylindrical form, and its components are:

$$F_r = r(2 + \sin^2 \theta), \quad F_\theta = r \sin \theta \cos \theta, \quad F_z = 3z.$$

Therefore, we have:

$$\begin{aligned} \nabla \cdot \mathbf{F} &= \frac{1}{r} \frac{\partial}{\partial r} (r^2(2 + \sin^2 \theta)) + \frac{1}{r} \frac{\partial}{\partial \theta} (r \sin \theta \cos \theta) + \frac{\partial(3z)}{\partial z} \\ &= \frac{1}{r} \cdot 2r(2 + \sin^2 \theta) + (\cos^2 \theta - \sin^2 \theta) + 3 \\ &= 4 + 2 \sin^2 \theta + \cos^2 \theta - \sin^2 \theta \\ &= 4 + \cos^2 \theta + \sin^2 \theta = 5. \end{aligned}$$

Therefore, by the Divergence Theorem,

$$\iint_S \nabla \cdot \mathbf{F} dS = \iiint_D \nabla \cdot \mathbf{F} dV = \iiint_D 5 dV = 5 \cdot \underbrace{\frac{\pi(2)^2(5)}{4}}_{\text{volume of the cylinder}} = 25\pi.$$

Here D denotes the solid enclosed by S .

One should note that the surface integral stated in the Divergence Theorem is

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS$$

but not of $(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}}$. In fact, using the Divergence Theorem, one can show that

$$\iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS = 0$$

for any continuously differentiable vector field \mathbf{F} .

It is because:

$$\begin{aligned} \iint_S \underbrace{(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}}}_{=: \mathbf{G}} dS &= \iiint_D \nabla \cdot \mathbf{G} dV \\ &= \iiint_D \nabla \cdot (\nabla \times \mathbf{F}) dV. \end{aligned}$$

We next show that $\nabla \cdot (\nabla \times \mathbf{F}) = 0$ for any continuously differentiable vector field \mathbf{F} .

$$\begin{aligned} \nabla \times \mathbf{F} &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathbf{k} \\ \nabla \cdot (\nabla \times \mathbf{F}) &= \frac{\partial}{\partial x} \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) + \frac{\partial}{\partial y} \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \\ &= \frac{\partial^2 F_z}{\partial x \partial y} - \frac{\partial^2 F_y}{\partial x \partial z} + \frac{\partial^2 F_x}{\partial y \partial z} - \frac{\partial^2 F_z}{\partial y \partial x} + \frac{\partial^2 F_y}{\partial z \partial x} - \frac{\partial^2 F_x}{\partial z \partial y}. \end{aligned}$$

Using the Mixed Partial Derivatives Theorem, all of the above second derivatives are canceled out, and so $\nabla \cdot (\nabla \times \mathbf{F}) = 0$.

Therefore, we get:

$$\iint_S (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dS = 0.$$

4.7.3 Interpretation of Divergence Operator

By taking a tiny solid region D with boundary surface S , then the Divergence Theorem applied to a vector field \mathbf{F} asserts that:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS = \iiint_D \nabla \cdot \mathbf{F} dV.$$

When the region D is very small, one can regard $\nabla \times \mathbf{F}$ is nearly a constant, and so we have:

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \simeq (\nabla \cdot \mathbf{F}) \times \text{volume of } D.$$

This result gives a geometric interpretation of $\nabla \cdot \mathbf{F}$:

$$\nabla \cdot \mathbf{F} \simeq \frac{1}{\text{volume of } D} \iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS.$$

In other words, $\nabla \times \mathbf{F}$ measures the *flux density* near a point. The more diverging \mathbf{F} is around a point, the higher the flux over a tiny closed surface around that point, resulting in greater value of $\nabla \cdot \mathbf{F}$. This justifies the use of name *divergence* for $\nabla \cdot \mathbf{F}$.

4.7.4 Limitations of Divergence Theorem

The condition that \mathbf{F} has to be defined in the region D is crucial. Consider the gravitational force field:

$$\mathbf{F} = -\frac{GMm}{\rho^2} \mathbf{e}_\rho.$$

If S is the sphere with radius a centered at the origin, then its outward unit normal is simply \mathbf{e}_ρ and so

$$\iint_S \mathbf{F} \cdot \hat{\mathbf{n}} dS = -GMm \iint_S \frac{1}{\rho^2} \mathbf{e}_\rho \cdot \mathbf{e}_\rho dS = -\frac{GMm}{a^2} \cdot 4\pi a^2 = -4\pi GMm.$$

If one attempts to apply the Divergence Theorem on this vector field, then by the spherical form of the divergence operator, we can easily see that:

$$\nabla \cdot \mathbf{F} = -GMm \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho^2 \cdot \frac{1}{\rho^2} \right) = -GMm \frac{1}{\rho} \frac{\partial}{\partial \rho} (1) = 0$$

whenever $\rho \neq 0$, i.e. everywhere except the origin.

The triple integral

$$\iiint_D \nabla \cdot \mathbf{F} dV$$

is an *improper* integral since D (which is the solid sphere enclosed by S) contains the origin at which \mathbf{F} is undefined!

To conclude, one needs to be very careful when applying the Divergence Theorem if the region D contains some points at which the vector field is not defined.

4.7.5 Gauss's Law for Gravity

The purpose of this subsection is to give a proof of the *Gauss's Law for Gravity* (assuming the inverse-square law), which says that the gravitational flux:

$$\iint_S \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}} dS$$

is given by $4\pi GMm$ for any closed surface S enclosing the origin.

We have shown that it is so when S is a sphere centered at the origin, and we are going to use the Divergence Theorem to show that it is always true for any closed surface S enclosing the origin. However, we need to be very careful when applying the Divergence Theorem since the gravitational field is undefined at the origin.

We will adopt the "hole-drilling" technique which was previously used in computing the winding number integral. Given a solid D containing the origin, we first construct a small sphere B with radius a centered at the origin. Then, the solid $D \setminus B$ (i.e. the solid D with B removed) is a solid not enclosing the "bad" point origin.

Next, we cut this solid into two parts by the horizontal plane $z = 0$. Label each side of the resulting solids by S_i , Π and Σ_i as shown in the Figure 4.18. Note that Π is the common side.

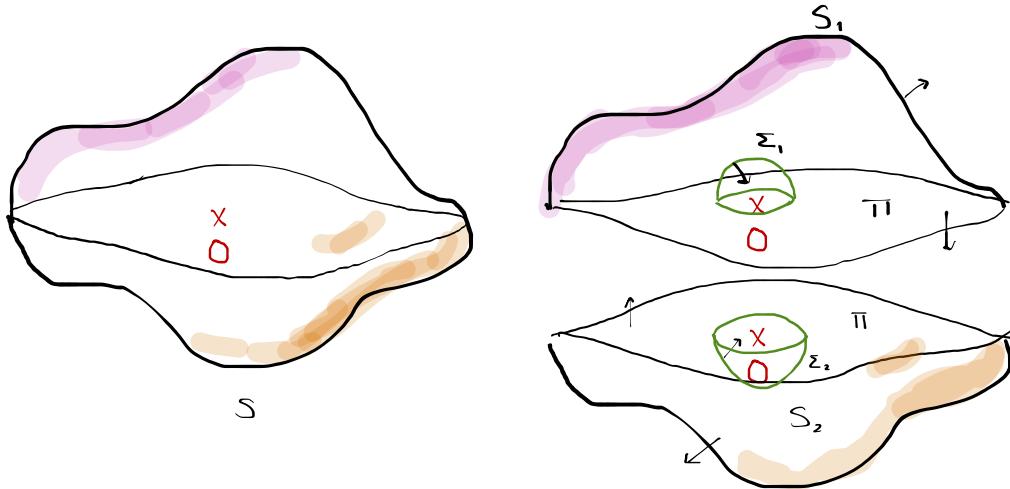


Figure 4.18: applying Divergence Theorem on the gravitational force field

Gluing S_1 , Π and Σ_1 together gives a closed surface not enclosing the origin. Denote D_1 to be the solid enclosed by this closed surface. Hence, one can apply the Divergence Theorem without any issue:

$$\left(\iint_{S_1} + \iint_{\Pi} + \iint_{\Sigma_1} \right) \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}} dS = \iiint_{D_1} \underbrace{\nabla \cdot \left(\frac{GMm}{\rho^2} \mathbf{e}_\rho \right)}_{=0} dV = 0$$

where $\hat{\mathbf{n}}$ is the **outward** unit normal of the boundary surface of D_1 . Denote $\hat{\mathbf{n}}_{\text{up}}$ and $\hat{\mathbf{n}}_{\text{down}}$ to be the upward and downward normal vector respectively. The above integrals can be expressed as:

$$\iint_{S_1} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{up}} dS + \iint_{\Pi} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{down}} dS + \iint_{\Sigma_1} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{down}} dS = 0. \quad (4.3)$$

Similarly, gluing S_2 , Π and Σ_2 together gives a closed surface not enclosing the origin. By the Divergence Theorem applied to this surface, we get:

$$\iint_{S_2} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{down}} dS + \iint_{\Pi} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{up}} dS + \iint_{\Sigma_2} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{up}} dS = 0. \quad (4.4)$$

We then add up the above two equations. First note that S_1 and S_2 can glue together to form the closed surface S . Both $\hat{\mathbf{n}}_{\text{up}}$ of S_1 , and $\hat{\mathbf{n}}_{\text{down}}$ of S_2 become the **outward** normal of S . Therefore,

$$\iint_{S_1} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{down}} dS + \iint_{S_2} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{up}} dS = \iint_S \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{outward}} dS.$$

For the planar surface Π , the downward normal $\hat{\mathbf{n}}_{\text{down}}$ is in the opposite direction of the upward normal $\hat{\mathbf{n}}_{\text{up}}$, i.e. $\hat{\mathbf{n}}_{\text{down}} = -\hat{\mathbf{n}}_{\text{up}}$. Therefore,

$$\iint_{\Pi} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{down}} dS + \iint_{\Pi} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{up}} dS = 0.$$

Finally, the surfaces Σ_1 and Σ_2 glue together to form the closed sphere Σ . The normal vectors $\hat{\mathbf{n}}_{\text{down}}$ of Σ_1 , and $\hat{\mathbf{n}}_{\text{up}}$ of Σ_2 , are the **inward** unit normal of Σ . Therefore,

$$\iint_{\Sigma_1} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{down}} dS + \iint_{\Sigma_2} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{up}} dS = \iint_{\Sigma} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{inward}} dS.$$

Summing up (4.3) and (4.4), we get:

$$\oint\!\oint_S \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{outward}} dS + 0 + \oint\!\oint_{\Sigma} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{inward}} dS = 0.$$

Therefore,

$$\begin{aligned} \oint\!\oint_S \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{outward}} dS &= - \oint\!\oint_{\Sigma} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{inward}} dS \\ &= \oint\!\oint_{\Sigma} \frac{GMm}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}}_{\text{outward}} dS \\ &= 4\pi GMm \quad (\text{computed before}) \end{aligned}$$

This holds true for **any closed surface** S enclosing the origin. This proves the Gauss's Law for Gravity (assuming the inverse-square law).

However, if S **does not** enclose the origin, then one can apply the Divergence Theorem on the gravitational vector field without any issue.

To conclude, for any closed surface S not passing through the origin, we have:

$$\oint\!\oint_S \frac{1}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}} dS = \begin{cases} 4\pi & \text{if } S \text{ encloses the origin;} \\ 0 & \text{otherwise.} \end{cases}$$

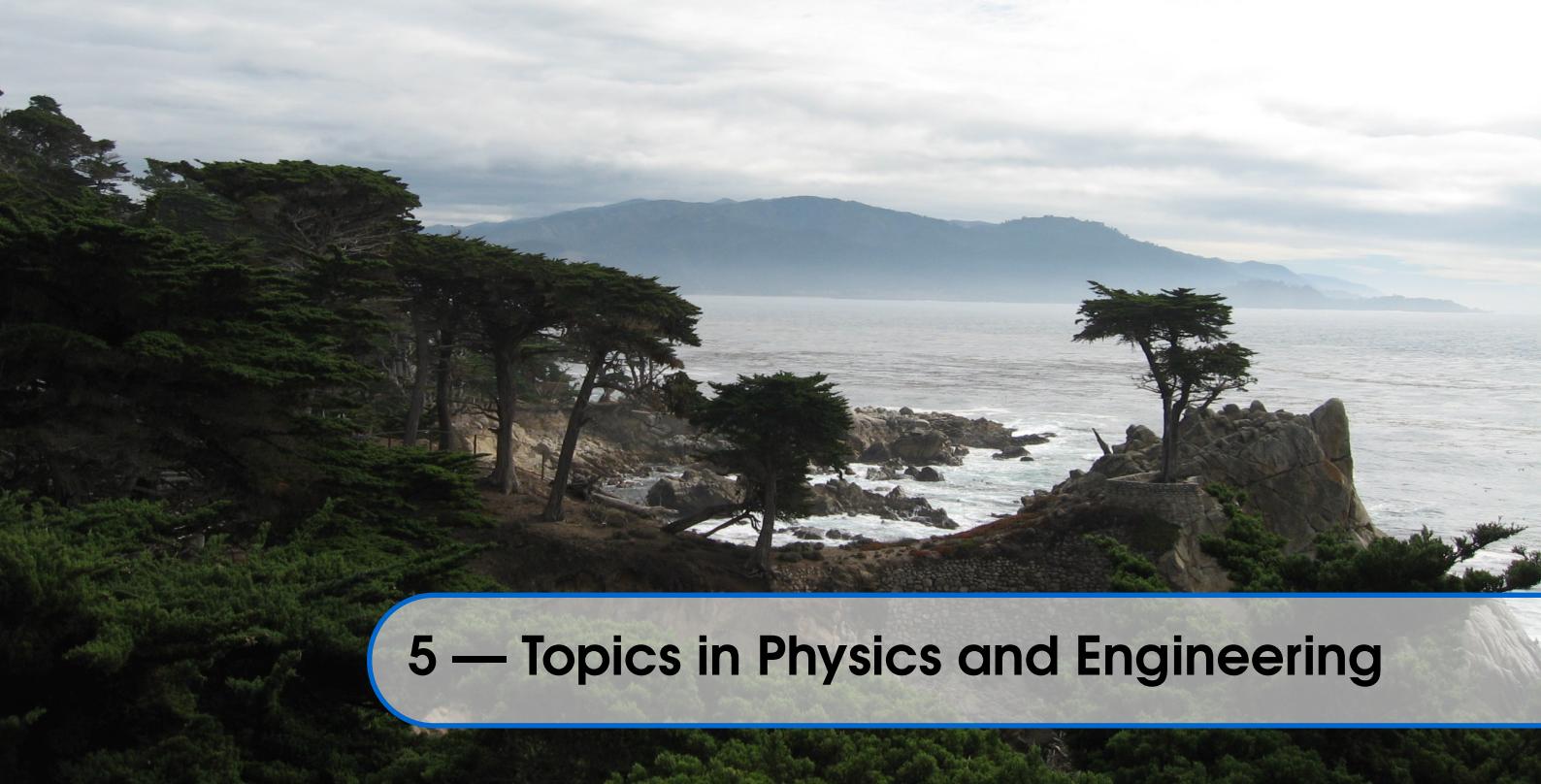
In the spirit of the Divergence Theorem, i.e.

$$\oint\!\oint_S \frac{1}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}} dS = \iiint_D \nabla \cdot \left(\frac{1}{\rho^2} \mathbf{e}_\rho \right) dV,$$

we can *decree* that

$$\iiint_D \nabla \cdot \left(\frac{1}{\rho^2} \mathbf{e}_\rho \right) dV = \begin{cases} 4\pi & \text{if } D \text{ contains the origin;} \\ 0 & \text{otherwise} \end{cases}$$

even though the integrand $\nabla \cdot \left(\frac{1}{\rho^2} \mathbf{e}_\rho \right)$ is undefined at the origin. This fact motivates the introduction of *Dirac delta functions* to be discussed in the next chapter.



5 — Topics in Physics and Engineering

5.1 Coulomb's Law

An electric field, commonly denoted by \mathbf{E} , is a vector field in \mathbb{R}^3 . At every point (x, y, z) , the vector $\mathbf{E}(x, y, z)$ represents the force that would be exerted on a **stationary test particle of unit positive charge**. An electric field can be generated by the presence of electric charges, or time-varying magnetic forces. We are going to discuss the former in this section.

A magnetic field, commonly denoted by \mathbf{B} , is a vector field in \mathbb{R}^3 that describes the direction and strength of magnetic force exerted at each point. Both electric current and presence of magnetic material can generate a magnetic field.

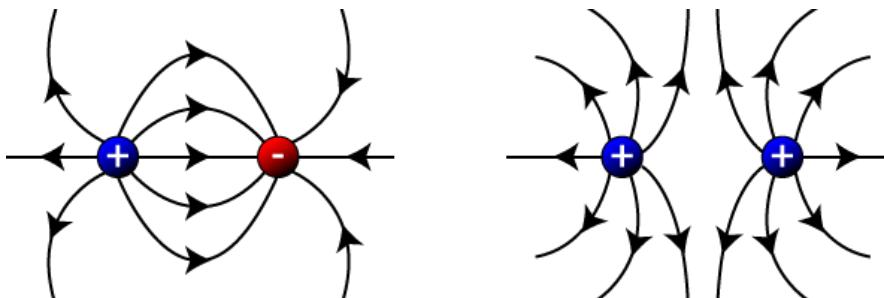


Figure 5.1: electric fields generated by point charges

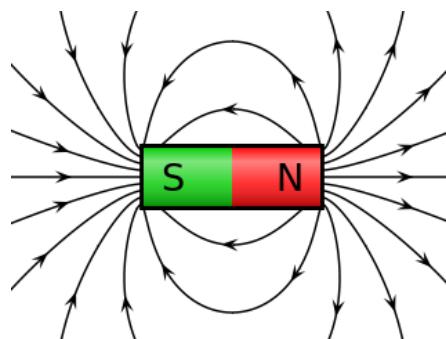


Figure 5.2: a magnetic field generated by a bar magnet

5.1.1 Electric Field of Point Charges

A point charge located at the origin generates an E-field according to the Coulomb's Inverse-Square Law:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0\rho^2} \mathbf{e}_\rho,$$

where q is the charge of the point particle, and ϵ_0 is a positive constant depending on the background material (i.e. permittivity). Clearly, this E-field is radially symmetric, and whether it points inward or outward depends on the sign of q .

If there is another point particle with charge Q , then the electric force exerted by the q -particle on the Q -particle is given by:

$$\mathbf{F} = Q\mathbf{E} = \frac{qQ}{4\pi\epsilon_0\rho^2} \mathbf{e}_\rho$$

where ρ and \mathbf{e}_ρ are evaluated at the point of the Q -particle. Therefore, if both q and Q are of the same sign (i.e. like charges), the force \mathbf{F} exerted on the Q -particle is pointing outward from the origin. This indicates the two charges repulse each other. On the other hand, if q and Q are of different sign, the force \mathbf{F} exerted on the Q -particle is pointing inward, indicating an attraction.

Note that the above E-field is not defined at the origin. Physically speaking, it can be interpreted as having highly concentrated charge density at the origin. We leave it as an exercise for readers to verify that:

$$\begin{aligned}\nabla \cdot \mathbf{E} &= 0 \\ \nabla \times \mathbf{E} &= \mathbf{0}\end{aligned}$$

everywhere except the origin.

Since \mathbb{R}^3 with the origin removed is a simply-connected region, we can still apply the curl test to claim that \mathbf{E} is conservative. In spherical coordinates, the gradient operator is given by:

$$\nabla f = \frac{\partial f}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \mathbf{e}_\varphi + \frac{1}{\rho \sin \varphi} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta.$$

In physics, we usually define potential energy according to the negative convention: $\mathbf{E} = -\nabla V$. Set:

$$\frac{q}{4\pi\epsilon_0\rho^2} \mathbf{e}_\rho = -\nabla V = -\frac{\partial V}{\partial \rho} \mathbf{e}_\rho - \frac{1}{\rho} \frac{\partial V}{\partial \varphi} \mathbf{e}_\varphi - \frac{1}{\rho \sin \varphi} \frac{\partial V}{\partial \theta} \mathbf{e}_\theta.$$

By equating the components, we have

$$\begin{aligned}\frac{q}{4\pi\epsilon_0\rho^2} &= -\frac{\partial V}{\partial \rho} \\ 0 &= -\frac{\partial V}{\partial \varphi} \\ 0 &= -\frac{\partial V}{\partial \theta}\end{aligned}$$

Clearly, V depends only on ρ , and by the first equation, we have:

$$V = - \int \frac{q}{4\pi\epsilon_0\rho^2} d\rho = \frac{q}{4\pi\epsilon_0\rho} + \text{constant.}$$

By picking the constant to be zero (so that the potential energy is 0 at infinity), the electric potential energy due to the q -particle is therefore given by:

$$V = \frac{q}{4\pi\epsilon_0\rho}.$$

The above discussion assumes the point charge is located at the origin, so that the E-field generated can be conveniently expressed using spherical coordinates. Using rectangular coordinates, the above E-field and the electric potential V are given by:

$$\mathbf{E} = \frac{q(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})}{4\pi\epsilon_0(x^2 + y^2 + z^2)^{3/2}}$$

$$V = \frac{q}{4\pi\epsilon_0\sqrt{x^2 + y^2 + z^2}}$$

If the point charge is located at (x_0, y_0, z_0) , then the E-field generated and its electric potential V are simply translations of the above, i.e.

$$\mathbf{E} = \frac{q((x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k})}{4\pi\epsilon_0((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2)^{3/2}} = \frac{q}{4\pi\epsilon_0} \cdot \frac{\mathbf{r} - \mathbf{r}_0}{|\mathbf{r} - \mathbf{r}_0|^3}$$

$$V = \frac{q}{4\pi\epsilon_0\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}} = \frac{q}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}_0|}$$

where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and $\mathbf{r}_0 = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k}$.

If there are more than one point charges, the resulting E-field generated by these point charges is the vector sum of all E-fields generated by these charges. Precisely, if these particles are located at $\mathbf{r}_1, \mathbf{r}_2, \dots$ with charges q_1, q_2, \dots correspondingly, then the overall E-field is given by:

$$\mathbf{E} = \frac{q_1}{4\pi\epsilon_0} \cdot \frac{\mathbf{r} - \mathbf{r}_1}{|\mathbf{r} - \mathbf{r}_1|^3} + \frac{q_2}{4\pi\epsilon_0} \cdot \frac{\mathbf{r} - \mathbf{r}_2}{|\mathbf{r} - \mathbf{r}_2|^3} + \dots = \sum_i \frac{q_i}{4\pi\epsilon_0} \cdot \frac{\mathbf{r} - \mathbf{r}_i}{|\mathbf{r} - \mathbf{r}_i|^3}.$$

Physicists often call this rule the Principle of Superposition.

5.1.2 Electric Flux

For a uniform E-field across a flat surface with unit normal $\hat{\mathbf{n}}$ and surface area A , the electric flux through this surface is defined to be:

$$(\mathbf{E} \cdot \hat{\mathbf{n}}) A$$

For a non-uniform E-field across a curved surface S , the electric flux through the surface is defined using a surface integral:

$$\Phi_E = \iint_S \mathbf{E} \cdot \hat{\mathbf{n}} dS.$$

Recall that the E-field generated by a point particle with charge q located at the origin is given by:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0\rho^2} \mathbf{e}_\rho.$$

Given a closed surface S , we have computed in the previous chapter that:

$$\iint_S \frac{1}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}} dS = \begin{cases} 4\pi & \text{if } S \text{ encloses the origin;} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the electric flux through S due to a particle located at the origin is given by:

$$\Phi_E = \iint_S \frac{q}{4\pi\epsilon_0} \frac{\mathbf{e}_\rho}{\rho^2} \cdot \hat{\mathbf{n}} dS = \begin{cases} \frac{q}{4\pi\epsilon_0} 4\pi = \frac{q}{\epsilon_0} & \text{if } S \text{ encloses the origin;} \\ 0 & \text{otherwise.} \end{cases}$$

By translation invariance, the electric flux through a closed surface S due to a particle located at *any* position is given by:

$$\Phi_E = \begin{cases} \frac{q}{\epsilon_0} & \text{if } S \text{ encloses the particle;} \\ 0 & \text{otherwise.} \end{cases}$$

For more than one point charges (say with charges q_1, q_2, \dots), the overall E-field is given by

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \dots$$

where \mathbf{E}_i is the electric field generated by the q_i -particle. Therefore, the resulting electric flux through a closed surface S is given by:

$$\Phi_{\mathbf{E}} = \iint_S (\mathbf{E}_1 + \mathbf{E}_2 + \dots) \cdot \hat{\mathbf{n}} dS = \sum_i \iint_S \mathbf{E}_i \cdot \hat{\mathbf{n}} dS$$

For each i , the flux integral $\iint_S \mathbf{E}_i \cdot \hat{\mathbf{n}} dS$ is zero if S does not enclose the q_i -particle, and is equal to $\frac{q_i}{\epsilon_0}$ if S encloses the q_i -particle. Therefore, by summing up each individual electric flux, we conclude that:

$$\Phi_{\mathbf{E}} = \frac{Q_{\text{enc}}}{\epsilon_0}$$

where Q_{enc} is the sum of charges enclosed by S . It is commonly called the **Gauss's Law for Electricity** (in the case of discrete charges).

From practical viewpoint, a continuous charge distribution (i.e. charge cloud) is composed of densely packed discrete particles. Therefore, physicists assert that the Gauss's Law for Electricity is true for *any* (discrete or continuous) charge distribution:

$$\iint_S \mathbf{E} \cdot \hat{\mathbf{n}} dS = \frac{Q_{\text{enc}}}{\epsilon_0}$$

where Q_{enc} is the amount of charges enclosed by the closed surface S .

5.2 Introduction to Maxwell's Equations

5.2.1 Integral Form

Maxwell's Equations are a set of four partial differential equations (PDE) which govern the interaction of **E**-fields, **B**-fields and currents. Each of the four has both an integral form and a differential form. The two forms are equivalent to each other by the Stokes' and Divergence Theorems discussed in the previous chapter.

Given any closed surface S enclosing a solid region D , and simply-connected surface Σ with a simple closed boundary curve C , the Maxwell's Equations in integral form are:

$$\iint_S \mathbf{E} \cdot \hat{\mathbf{n}} dS = \frac{Q}{\epsilon_0} \quad (\text{Gauss's Law for Electricity})$$

$$\iint_S \mathbf{B} \cdot \hat{\mathbf{n}} dS = 0 \quad (\text{Gauss's Law for Magnetism})$$

$$\oint_C \mathbf{E} \cdot d\mathbf{r} = -\frac{\partial}{\partial t} \iint_{\Sigma} \mathbf{B} \cdot \hat{\mathbf{n}} dS \quad (\text{Faraday's Law})$$

$$\oint_C \mathbf{B} \cdot d\mathbf{r} = \mu_0 \left(\iint_{\Sigma} \mathbf{J} \cdot \hat{\mathbf{n}} dS + \epsilon_0 \frac{\partial}{\partial t} \iint_{\Sigma} \mathbf{E} \cdot \hat{\mathbf{n}} dS \right) \quad (\text{Ampère's Law})$$

Here **E** is the electric field, **B** is the magnetic field and **J** is the current. Q is the total amount of charges enclosed by the surface S . Both μ_0 and ϵ_0 are positive constants.

We have discussed the Gauss's Law for Electricity in the previous section. It can be derived by assuming the Coulomb's Inverse-Square Law. The Gauss's Law for Magnetism asserts that the magnetic flux through any closed surface must be zero! In physical terms, it means the amount of magnetic field *entering* the surface is exactly equal to that *leaving* the surface.

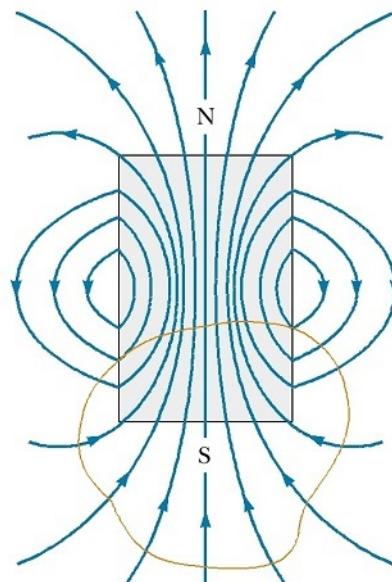


Figure 5.3: Gauss's Law for Magnetism

Under the assumption that the Gauss's Law for Magnetism is true, *magnetic monopoles* (i.e. magnets with only one pole) are ruled out. It is because a magnet with only one pole will generate a **B**-field similar to the **E**-field generated by a point charge, which is radially outward or inward, giving non-zero flux (positive if radially outward, negative if radially inward) on a surface enclosing the magnetic monopole. It would lead to a contradiction to the Gauss's Law for Magnetism.

Easy experiments tell us that breaking an ordinary bar magnet cannot create a magnetic monopole. Instead, the magnet will be broken into two magnets each of which contains both

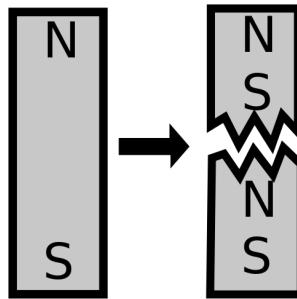


Figure 5.4: breaking a bar magnet does not create magnetic monopoles

south and north poles. If in someday a magnet monopole is discovered (maybe using some exotic materials), the Gauss's Law for Magnetism will need to be modified.

The Faraday's and Ampère's Laws in integral form look a bit more complicated than the two Gauss's Laws. We will discuss them after we derive the differential form of Maxwell's Equations.

5.2.2 Differential Form

Using the Stokes' and Divergence Theorems, one can convert the integral form of Maxwell's Equations to the differential form, which is more elegant to state and easier to work with:

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} && \text{(Gauss's Law for Electricity)} \\ \nabla \cdot \mathbf{B} &= 0 && \text{(Gauss's Law for Magnetism)} \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && \text{(Faraday's Law)} \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} && \text{(Ampère's Law)}\end{aligned}$$

To begin with, let's describe a common technique that is often applied when converting an integral expression to a differential one:

Theorem 5.1 Suppose $f(x, y, z)$ and $g(x, y, z)$ are two continuous functions such that

$$\iiint_D f(x, y, z) dV = \iiint_D g(x, y, z) dV$$

for **any** solid region D in \mathbb{R}^3 , then the functions f and g are identical, i.e. $f(x, y, z) = g(x, y, z)$ for any (x, y, z) in \mathbb{R}^3 .

Proof. Here we present a *heuristic*, rather than a rigorous, proof. Since it is assumed that

$$\iiint_D f(x, y, z) dV = \iiint_D g(x, y, z) dV$$

for **any** solid region D , it is in particular true for tiny ones. Take D to be a very tiny region in \mathbb{R}^3 , then both f and g are approximately constants over the region D . Therefore,

$$f(x, y, z) \iiint_D dV \simeq g(x, y, z) \iiint_D dV.$$

By canceling out $\iiint_D dV$ on both sides, we get $f(x, y, z) = g(x, y, z)$ over the region D . Since this tiny region D is arbitrarily chosen, one can repeat the same argument with other tiny regions to conclude that $f(x, y, z) = g(x, y, z)$ over the whole \mathbb{R}^3 . ■

One can use this theorem and the Divergence Theorem to derive the differential form of Gauss's Laws. The Gauss's Law for Magnetism asserts that:

$$\oint_S \mathbf{B} \cdot \hat{\mathbf{n}} dS = 0$$

for any closed surface S . Denote D to be the solid enclosed by S , then the Divergence Theorem asserts that:

$$\iiint_D \nabla \cdot \mathbf{B} dV = \oint_S \mathbf{B} \cdot \hat{\mathbf{n}} dS = 0 = \iiint_D 0 dV.$$

Note that S , and hence D , are arbitrary. By Theorem 5.1, we have:

$$\nabla \cdot \mathbf{B} = 0.$$

This equation is called the differential form of the Gauss's Law for Magnetism.

One can also convert the Gauss's Law for Electricity into a differential form by introducing the charge density function ρ , which infinitesimally measures the amount of charges per unit volume. The amount of charge Q in a solid region D is related to the charge density by:

$$Q = \iiint_D \rho dV.$$

Recall that the Gauss's Law for Electricity asserts that:

$$\oint_S \mathbf{E} \cdot \hat{\mathbf{n}} dS = \frac{Q}{\epsilon_0} = \iiint_D \frac{\rho}{\epsilon_0} dV$$

for any closed surface S . Here D is the solid region enclosed by S . Using the Divergence Theorem, we get:

$$\iiint_D \nabla \cdot \mathbf{E} dV = \iiint_D \frac{\rho}{\epsilon_0} dV.$$

Since S and hence D are arbitrary, Theorem 5.1 shows:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}.$$

It is called the differential form of the Gauss's Law for Electricity. In physics term, this equation tells us the *denser* the charges, the more *diverging* the \mathbf{E} -field.

5.2.3 Faraday's and Ampère's Laws

Similar technique can be applied to derive the differential form of Faraday's and Ampère's Laws. The integral form of the Faraday's Law says that:

$$\oint_C \mathbf{E} \cdot d\mathbf{r} = -\frac{\partial}{\partial t} \iint_{\Sigma} \mathbf{B} \cdot \hat{\mathbf{n}} dS$$

for any simply-connected surface Σ with boundary C .

Using Stokes' Theorem, we have:

$$\iint_{\Sigma} (\nabla \times \mathbf{E}) \cdot \hat{\mathbf{n}} dS = \oint_C \mathbf{E} \cdot d\mathbf{r} = -\frac{\partial}{\partial t} \iint_{\Sigma} \mathbf{B} \cdot \hat{\mathbf{n}} dS = \iint_{\Sigma} -\frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}} dS.$$

Since it is true for any surface Σ , one can pick a tiny surface Σ so that $(\nabla \times \mathbf{E}) \cdot \hat{\mathbf{n}}$ and $-\frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}}$ are approximately constants, and so:

$$(\nabla \times \mathbf{E}) \cdot \hat{\mathbf{n}} \iint_{\Sigma} dS = -\frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}} \iint_{\Sigma} dS.$$

By canceling out $\iint_{\Sigma} dS$ on both sides, we have

$$(\nabla \times \mathbf{E}) \cdot \hat{\mathbf{n}} = -\frac{\partial B}{\partial t} \cdot \hat{\mathbf{n}}.$$

Since the surface Σ is arbitrary, the unit normal $\hat{\mathbf{n}}$ is also arbitrary. In order for the above to be true, we get the differential form of the Faraday's Law:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

The Faraday's Law in differential form tells us that a *changing* \mathbf{B} -field (so that $\frac{\partial \mathbf{B}}{\partial t}$ is a non-zero vector) will induce a \mathbf{E} -field. Moreover, the more rapid is the change of \mathbf{B} -field, the more *circular* is the induced \mathbf{E} -field.

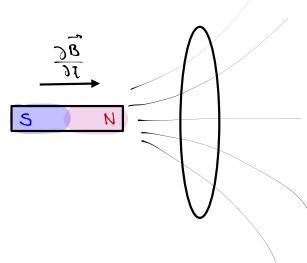


Figure 5.5: a moving magnet induces a \mathbf{E} -field on the wire

Similarly, the Ampère's Law in integral form can be converted into differential form by Stokes' Theorem:

$$\oint_C \mathbf{B} \cdot d\mathbf{r} = \mu_0 \left(\iint_{\Sigma} \mathbf{J} \cdot \hat{\mathbf{n}} dS + \epsilon_0 \frac{\partial}{\partial t} \iint_{\Sigma} \mathbf{E} \cdot \hat{\mathbf{n}} dS \right)$$

$$\iint_{\Sigma} (\nabla \times \mathbf{B}) \cdot \hat{\mathbf{n}} dS = \iint_{\Sigma} \left(\mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot \hat{\mathbf{n}} dS$$

Since C and Σ are arbitrary, we must have

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

The physical interpretation of the Ampère's Law is a bit more complicated than that of Faraday's Law. For the sake of simplicity, let's consider only the case where the \mathbf{E} -field is unchanging:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad (\text{when } \mathbf{E} \text{ is unchanging})$$

Recall that \mathbf{J} is the current. This equation tells us that when the \mathbf{E} -field is unchanging, the presence of an electric current in a wire will induce a magnetic field circling around the wire.

In fact, the equation $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$ is the original form of Ampère's Law. Later Maxwell found that it is incomplete and correct it by adding a $\frac{\partial \mathbf{E}}{\partial t}$ term.

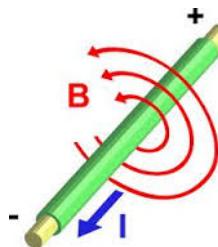


Figure 5.6: Ampère's Law

5.3 Heat Diffusion

5.3.1 Derivation of Heat Equation

Let $u(x, y, z, t)$ be the temperature at point (x, y, z) at time t . The heat equation:

$$\frac{\partial u}{\partial t} = k \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right)$$

governs the diffusion of heat. In this section, we will use several fundamental laws in physics and the Divergence Theorem to derive the heat equation.

Heat diffusion is caused by displacement of heat energy. Fourier's Law in physics asserts that heat energy transfers according to the following rule:

$$\mathbf{J} = -a \nabla u$$

where \mathbf{J} is a vector field (in energy per second) representing the flow of heat energy, and a is a positive constant depending on the medium. In other words, heat energy diffuses from higher temperature regions to lower ones, and the rate of diffusion is proportional to the magnitude of ∇u .

Let D be an arbitrary solid region with boundary surface S . Denote ϱ to be the energy density function (in energy per volume), which equals to bu for some positive constant b whose value depends on the medium. Then the triple integral:

$$\iiint_D \varrho \, dV$$

is the **total** amount of heat energy contained in the region D .

On the other hand, the outward flux

$$\oint\!\oint_S \mathbf{J} \cdot \hat{\mathbf{n}} \, dS$$

measures the amount of heat loss through the closed surface S . By the *conservation of heat energy*, heat energy must escape through the surface S . In mathematical terms, it is stated as:

$$\frac{\partial}{\partial t} \iiint_D \varrho \, dV = - \oint\!\oint_S \mathbf{J} \cdot \hat{\mathbf{n}} \, dS.$$

The negative sign appears on the RHS because of the outward convention of $\hat{\mathbf{n}}$.

Applying the above physical laws, we get:

$$\begin{aligned} \iiint_D b \frac{\partial u}{\partial t} \, dV &= - \oint\!\oint_S (-a \nabla u) \cdot \hat{\mathbf{n}} \, dS \\ \iiint_D b \frac{\partial u}{\partial t} \, dV &= \oint\!\oint_S a \nabla u \cdot \hat{\mathbf{n}} \, dS \\ \iiint_D b \frac{\partial u}{\partial t} \, dV &= \iiint_D \nabla \cdot (a \nabla u) \, dV \quad (\text{Divergence Theorem}) \end{aligned}$$

Since D is arbitrary, we must have:

$$b \frac{\partial u}{\partial t} = \nabla \cdot (a \nabla u) = a \nabla \cdot \nabla u.$$

We leave it as an exercise for readers to verify that:

$$\nabla \cdot \nabla u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

Therefore, we can conclude that:

$$\frac{\partial u}{\partial t} = \frac{b}{a} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right)$$

which is exactly the heat equation by defining $k = \frac{b}{a}$.

Very often, $\nabla \cdot \nabla u$ is denoted by $\nabla^2 u$ or Δu . As such, the heat equation can be written as:

$$\frac{\partial u}{\partial t} = k\Delta u.$$

5.3.2 Fundamental Solution

It can be verified that the following function satisfies the heat equation:

$$\Phi(x, y, z, t) = \frac{1}{(4\pi kt)^{3/2}} \exp\left(-\frac{x^2 + y^2 + z^2}{4kt}\right).$$

At a point $(x, y, z) = (0, 0, 0)$, we have $\Phi(0, 0, 0, t) = \frac{1}{(4\pi kt)^{3/2}}$, and so:

$$\lim_{t \rightarrow 0} \Phi(0, 0, 0, t) = \lim_{t \rightarrow 0} \frac{1}{(4\pi kt)^{3/2}} = \infty.$$

In contrast, if $(x, y, z) \neq (0, 0, 0)$, both $\exp\left(-\frac{x^2+y^2+z^2}{4kt}\right)$ and $(4\pi kt)^{3/2}$ go to 0 as $t \rightarrow 0$. However, the exponential term goes to 0 faster than the $t^{3/2}$ term, so

$$\lim_{t \rightarrow 0} \Phi(x, y, z, t) = \lim_{t \rightarrow 0} \frac{1}{(4\pi kt)^{3/2}} \exp\left(-\frac{x^2 + y^2 + z^2}{4kt}\right) = 0 \quad \text{when } (x, y, z) \neq (0, 0, 0).$$

Therefore, the function $\Phi(x, y, z, t)$ represents the heat diffusion starting from a highly concentrated heat source at $t = 0$. As time goes, the temperature distribution becomes more and more uniform.

In general, if the initial temperature distribution is given by the function $g(x, y, z)$, it can be shown (proof beyond the scope of the course) that the following function

$$u(x, y, z, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x - u, y - v, z - w, t) g(u, v, w) du dv dw$$

satisfies the heat equation $\frac{\partial u}{\partial t} = k\Delta u$ with initial condition $g(x, y, z)$, meaning that

$$\lim_{t \rightarrow 0} u(x, y, z, t) = g(x, y, z).$$

In other words, the function $u(x, y, z, t)$ predicts how heat diffuses when given an initial temperature profile $g(x, y, z)$. However, the integral defining $u(x, y, z, t)$ is in general difficult to compute explicitly.

5.3.3 Steady State

A temperature distribution $u(x, y, z, t)$ is said to be at steady state if it is independent of the time t , i.e. $\frac{\partial u}{\partial t} = 0$. For such a temperature distribution, the heat equation implies that

$$\Delta u = 0.$$

The above equation is often called the Laplace Equation.

Now given a closed surface S which encloses a solid region D . Using the Divergence Theorem, one can show that at a steady state if the temperature *on* the surface S is constant, then the temperature *inside* the surface S is also a constant. To argue this, we denote $u(x, y, z)$ to be a steady state temperature distribution (i.e. $\Delta u = 0$), and that:

$$u(x, y, z) = C \quad \text{for any } (x, y, z) \text{ on } S.$$

Next we consider the vector field $(u - C)\nabla u$. We leave it as an exercise for readers to verify from the definition that:

$$\nabla \cdot ((u - C)\nabla u) = |\nabla u|^2 + (u - C)\Delta u.$$

At steady state, we have $\Delta u = 0$ and so $\nabla \cdot ((u - C)\nabla u) = |\nabla u|^2$. Next we integrate this result over D :

$$\iiint_D \nabla \cdot ((u - C)\nabla u) \, dV = \iiint_D |\nabla u|^2 \, dV.$$

Applying the Divergence Theorem on the LHS, we get:

$$\iint_S (u - C)\nabla u \cdot \hat{n} \, dS = \iiint_D |\nabla u|^2 \, dV.$$

From our assumption, we have $u = C$ for any point on S . Therefore, the integrand $(u - C)\nabla u \cdot \hat{n}$ of the flux integral on LHS is zero, and so:

$$0 = \iiint_D |\nabla u|^2 \, dV.$$

Since the integrand $|\nabla u|^2$ is non-negative, the only scenario for the above to happen is that

$$\nabla u(x, y, z) = 0 \quad \text{for any } (x, y, z) \text{ in } D.$$

Therefore, u must be a constant in the region D , and by continuity, this constant must be C .

5.4 Dirac Delta Functions

Recall that the following function

$$\Phi(x, y, z, t) = \frac{1}{(4\pi kt)^{3/2}} \exp\left(-\frac{x^2 + y^2 + z^2}{4kt}\right)$$

is the solution to the heat equation starting at $t = 0$ with a highly concentrated heat source at the origin. As $t \rightarrow 0$, we have shown that:

$$\lim_{t \rightarrow 0} \Phi(x, y, z, t) = \begin{cases} 0 & \text{if } (x, y, z) \neq (0, 0, 0); \\ \infty & \text{otherwise.} \end{cases}$$

The *Dirac delta function* in \mathbb{R}^3 , denoted by δ , is a “function” which can be *loosely* defined as follows:

$$\delta(x, y, z) = \begin{cases} 0 & \text{if } (x, y, z) \neq (0, 0, 0); \\ \infty & \text{otherwise.} \end{cases}$$

As such, we then have

$$\lim_{t \rightarrow 0} \Phi(x, y, z, t) = \delta(x, y, z)$$

and so the delta function δ can be thought as the initial condition for Φ .

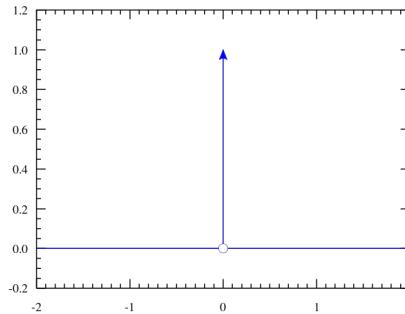


Figure 5.7: the Dirac delta function can be *loosely* regarded as the above graph

While the above definition of δ is very neat, it is *paradoxical*. From mathematical viewpoint, it is a function which is zero *almost everywhere* in \mathbb{R}^3 . By the definition of integral, the above definition of δ will imply:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y, z) dx dy dz = 0.$$

On the other hand in physics, the integral

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y, z) dx dy dz$$

represents the total heat energy. It being 0 means there is no heat energy over the entire \mathbb{R}^3 . As time goes by, the heat distribution becomes Φ since it solves the heat equation and $\Phi \rightarrow \delta$ as $t \rightarrow 0$. However, it can be computed (using spherical coordinates) that:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x, y, z, t) dx dy dz = 1$$

for any $t > 0$, meaning the total heat energy in \mathbb{R}^3 suddenly jumps to 1 as t becomes positive. It contradicts the conservation of energy!

To resolve this paradox, mathematicians introduced the concept of **generalized functions** so that the Dirac delta “function” is regarded as a *distribution* rather than an ordinary *function*. The rigorous theory of generalized functions is very advanced and well beyond the scope of this course. To put it in simpler term, the Dirac delta “function” δ is a *generalized function* that satisfies:

$$\iiint_D \delta \, dV = \begin{cases} 1 & \text{if } D \text{ contains the origin} \\ 0 & \text{otherwise.} \end{cases}$$

There is no ordinary function that satisfies the above. Therefore, the Dirac delta “function” is not regarded by mathematicians as a function despite its name.

When we discuss the Gauss's Law for Gravity, we figured out that:

$$\iint_S \frac{1}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}} \, dS = \begin{cases} 4\pi & \text{if } S \text{ encloses the origin;} \\ 0 & \text{otherwise.} \end{cases}$$

“Applying” the Divergence Theorem to this vector field, we get:

$$\iint_S \frac{1}{\rho^2} \mathbf{e}_\rho \cdot \hat{\mathbf{n}} \, dS = \iiint_D \nabla \cdot \left(\frac{1}{\rho^2} \mathbf{e}_\rho \right) \, dV.$$

We put “applying” in quote because the vector field $\frac{1}{\rho^2} \mathbf{e}_\rho$ is not defined at the origin, so *rigorously* speaking one cannot applying the Divergence Theorem here. By *decreeing* the Divergence Theorem holds for this vector field, we obtain:

$$\iiint_D \nabla \cdot \left(\frac{1}{\rho^2} \mathbf{e}_\rho \right) \, dV = \begin{cases} 4\pi & \text{if } D \text{ contains the origin;} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$\frac{1}{4\pi} \nabla \cdot \left(\frac{1}{\rho^2} \mathbf{e}_\rho \right)$$

satisfies the definition of the Dirac delta function, and so one can write:

$$\nabla \cdot \left(\frac{1}{\rho^2} \mathbf{e}_\rho \right) = 4\pi\delta.$$

According to the Coulomb's Law, a point particle with charge q at the origin generates a E-field given by:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0\rho^2} \mathbf{e}_\rho.$$

From the above discussion, we get:

$$\nabla \cdot \mathbf{E} = \frac{q}{4\pi\epsilon_0} \cdot 4\pi\delta = \frac{q}{\epsilon_0} \delta.$$

The Gauss's Law for Electricity in differential form states that $\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$ where ρ is the charge density. Therefore, using the Dirac delta function (again *decree* that the Divergence Theorem is true for the above E-field), one can derive:

$$\rho = q\delta.$$

In practice, one can regard $q\delta$ as the charge density distribution for a point particle (with charge q).