PHYS 2155 - Methods in Physics II

Lee Kai Ming
CYMP Rm 415A
kmlee@lily.physics.hku.hk
2859 2370

March 2020[1]

[1]Most of the material here comes from notes of previous years, with examples from various books. Thank Dr. Judy Chow.

Contents

Chapter 1

Matrices

Matrices are rectangular array of quantities. They arise naturally in many fields including mathematics, engineering, and physics. For example, the current flowing in each branch of an electrical circuit can be found by solving the matrix equation arising from this circuit which are obtained using Kirchhoff's laws. In this chapter, we will review basic concepts about matrices, and discuss how to find determinants and inverses of square matrices. Then, we study the eigenvalues and eigenvectors as well as diagonalization of square matrices which are useful in the analysis of dynamical systems.

1.1 Definitions and Notations

Definition 1.1. *A* **matrix** *of size $m \times n$ is a rectangular array of quantities such as numbers, symbols, or functions arranged in m rows and n columns which has the form*

$$A = (a_{ij})_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{1.1}$$

The entries $a_{11}, a_{12}, \dots, a_{mn}$ in the matrix $A = (a_{ij})_{m\times n}$ are called the **elements** of the matrix. With i and j given, the element in the i-th row and j-th column is said to have **row index** i and **column index** j. Elements with the same row and column indices, $a_{11}, a_{22}, a_{33}, \dots$, are called **diagonal elements** (or **diagonal entries**).

A matrix with m rows and n columns is called an $m \times n$ matrix or m-by-n matrix, while m and n are called its **dimensions**. Below are examples of a 2×3, a 2×1 and a 3×3 matrix, respectively

$$\begin{pmatrix} \sqrt{5} & -1 & 0 \\ 2 & \pi & 1/2 \end{pmatrix}, \qquad \begin{pmatrix} 11 \\ 9 \end{pmatrix}, \qquad \begin{pmatrix} 2 & -1 & 3 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.2}$$

Two matrices are equal if they have the same size and if their corresponding entries are equal. Therefore, if $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{p\times q}$, then $A = B$ if and only if $m = p$ and $n = q$ as well as $a_{ij} = b_{ij}$ for all $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

A **row matrix** (or **row vector**) is a $1 \times n$ matrix (a matrix consisting of a single row of n elements)

$$\mathbf{u} = \begin{pmatrix} u_1 & u_2 & \ldots & u_n \end{pmatrix} . \tag{1.3}$$

Similarly, a **column matrix** (or **column vector**) is a $m \times 1$ matrix (a matrix consisting of a single column of m elements)

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} . \tag{1.4}$$

If the row matrices $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ and the column matrices $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are the rows and columns of an $m \times n$ matrix A, then A can be written as

$$A = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_m \end{pmatrix} \qquad \text{or} \qquad A = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix} . \tag{1.5}$$

An $n \times n$ matrix has the same numbers of rows as columns and is called a **square matrix of order** n. A matrix, whether square or not, whose entries are all zero is called a **zero matrix** (or **null matrix**), denoted by O.

Definition 1.2. *If A and B are are both $m \times n$ matrices, then their* **sum** *$A + B$ is the $m \times n$ matrix obtained by adding corresponding elements of A and B. That is to say, if $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$, then $A + B = (a_{ij} + b_{ij})_{m\times n}$. Note that the sum $A + B$ is defined only when A and B have the same size.*

Definition 1.3. *If A is a $m \times n$ matrix and α is a scalar, then the* **scalar multiple***, αA, is the $m \times n$ matrix obtained by multiplying each element of A by α. That is to say, if $A = (a_{ij})_{m\times n}$, then $\alpha A = (\alpha a_{ij})_{m\times n}$.*

By using familiar properties of numbers, it is easy to verify that matrices have the same algebraic properties of addition and scalar multiplication as vectors. These properties are summarized in the following theorem.

Theorem 1.4. Let A, B, and C be matrices of the same size, and let α and β be scalars. Then we have

(a) $A + B = B + A$;

(b) $(A + B) + C = A + (B + C)$;

(c) $A + O = A$;

(d) $A + (-A) = O$;

(e) $\alpha(A + B) = \alpha A + \alpha B$;

(f) $(\alpha + \beta)A = \alpha A + \beta A$;

(g) $(\alpha\beta)A = \alpha(\beta A)$.

■

Just like vectors, the matrix $(-1)A$ is denoted as $-A$ which is called the **negative** of A. Then we can define the **difference** of two matrices with the same size by

$$A - B = A + (-1)B \ . \tag{1.6}$$

Therefore, if $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$, then $A - B = (a_{ij} - b_{ij})_{m\times n}$. In other words, the difference $A - B$ is obtained by subtracting the corresponding elements of A and B.

Definition 1.5. *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are both* $n \times 1$ *column matrices with elements* $u_1, u_2, \ldots, u_n$ *and* $v_1, v_2, \ldots, v_n$*, respectively, then their* **scalar product** *(or* **dot product***)* $\mathbf{u} \cdot \mathbf{v}$ *is the number defined by*

$$\mathbf{u} \cdot \mathbf{v} = \overline{u_1}v_1 + \overline{u_2}v_2 + \cdots + \overline{u_n}v_n \ , \tag{1.7}$$

where $\overline{u_i}$ *is the complex conjugate of* u_i*. Two column matrices* $\mathbf{u}$ *and* $\mathbf{v}$ *are said to be* **orthogonal** *if and only if* $\mathbf{u} \cdot \mathbf{v} = 0$*. We will say much more about orthogonality later.*

Note that if $\mathbf{u}$ and $\mathbf{v}$ are column matrices of the same size whose elements are real, then their dot product is equal to

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n \ , \tag{1.8}$$

which is equivalent to the dot product of vectors with real components.

Definition 1.6. *If* $A = (a_{ij})_{m\times n}$ *is an* $m \times n$ *matrix and* $B = (b_{ij})_{n\times p}$ *is an* $n \times p$ *matrix, then their* **matrix product** $C = AB$ *is the* $m \times p$ *matrix in which the element* c_{ij} *in the* i*th row and* j*th column is defined by*

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj} \ . \tag{1.9}$$

Note that the matrix product AB *is defined only when the number of columns of* A *is same as the number of rows of* B*. And even if both* AB *and* BA *are defined, in general* $AB \neq BA$*. In general, matrix multiplication is* **noncommutative**.

The following theorem summarizes the main properties of matrix multiplication. It tells us that matrix multiplication is associative and distributive over addition.

Theorem 1.7. Let A, B, C, and D be matrices have sizes for which the indicated sums and/or products are defined, and let α be a scalar. We have

(a) $A(BC) = (AB)C$,

(b) $\alpha(AB) = (\alpha A)B = A(\alpha B)$,

(c) $A(B + C) = AB + AC$,

(d) $(B + C)D = BD + CD$.

■

Definition 1.8. *If A is an $n \times n$ square matrix and k is a positive integer, then the product of k copies of A is also an $n \times n$ matrix denoted as A^k,*

$$A^k = \underbrace{AA \cdots A}_{k \text{ times}} \ . \tag{1.10}$$

Obviously, if A is a square matrix and r and s are positive integers, then $A^r A^s = A^{r+s}$ and $(A^r)^s = A^{rs}$.

Definition 1.9. *If A is an $m \times n$ matrix, then the* **transpose** *of A is the $n \times m$ matrix A^t obtained by interchanging the rows and columns of A. That is to say, if $A = (a_{ij})_{m \times n}$, then $A^t = (a_{ji})_{n \times m}$.*

It is obvious that if A is a row matrix, then its transpose is a column matrix, and vice versa.

Theorem 1.10. Let A, B, and C be matrices have sizes for which the indicated sums and products are defined, and let α be a scalar. Then we have

(a) $(A^t)^t = A$,

(b) $(\alpha A)^t = \alpha(A^t)$,

(c) $(A + C)^t = A^t + C^t$,

(d) $(AB)^t = B^t A^t$.

■

We can generalize Theorem 1.10(d) to products of more than two factors. It can be shown that the transpose of a product of matrices is equal to the product of their transposes in *reverse* order.

Definition 1.11. *For a square matrix $A = (a_{ij})_{n \times n}$, the elements a_{11}, a_{22}, $\ldots, a_{nn}$ form the* **leading diagonal** *(or* **main diagonal***) of A. The sum of the leading diagonal elements of a square matrix A is called the* **trace** *of A and is denoted as $\mathrm{Tr}(A)$. Thus, if A is a square matrix of order n, then*

$$\mathrm{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii} \ . \tag{1.11}$$

Note that if A and B are both $n \times n$ matrices and α is a scalar, then

$$\begin{aligned} \mathrm{Tr}(\alpha A) &= \alpha\,\mathrm{Tr}(A)\ , && (1.12)\\ \mathrm{Tr}(AB) &= \mathrm{Tr}(BA)\ . && (1.13) \end{aligned}$$

Definition 1.12. *A square matrix $D = [d_{ij}]_{n\times n}$ that has all non-diagonal elements equal to zero, $d_{ij} = 0$ if $i \neq j$, is called a* **diagonal matrix**. *A* **scalar matrix** *is a diagonal matrix whose diagonal elements are the same. A scalar matrix in which each diagonal element is 1 is called an* **identity matrix** *or* **unit matrix**. *An $n \times n$ identity matrix is denoted by the symbol I_n or simply I if its size is understood. The elements of I_n can be represented by the* **Kronecker delta**, *defined by*

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \quad . \tag{1.14}$$

Thus,

$$I_n = (\delta_{ij})_{n\times n}\ . \tag{1.15}$$

From the definition of matrix multiplication, we can see that for any square matrix A

$$IA = AI = A\ . \tag{1.16}$$

Besides, it is convenient to define $A^0 = I_n$ for any $n \times n$ matrix A.

Definition 1.13. *(1) A square matrix in which every element below the leading diagonal is zero is called an* **upper triangular matrix**. *An example is*

$$\begin{pmatrix} 1 & 6 & -4 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{pmatrix} . \tag{1.17}$$

(2) A square matrix in which every element above the leading diagonal is zero is called a **lower triangular matrix**. *An example is*

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 0 \\ 3 & -2 & 5 \end{pmatrix} . \tag{1.18}$$

(3) A square matrix A such that $A = A^t$ is called a **symmetric matrix**. *That is to say, if $A = (a_{ij})_{n\times n}$, then $a_{ij} = a_{ji}$ for $1 \leq i, j \leq n$. An example is*

$$\begin{pmatrix} 1 & 5 & -3 \\ 5 & 4 & 2 \\ -3 & 2 & 7 \end{pmatrix} . \tag{1.19}$$

(4) A square matrix A such that $A = -A^t$ is called a **skew-symmetric matrix**. *That is to say, if $A = (a_{ij})_{n\times n}$, then $a_{ij} = -a_{ji}$ for $1 \leq i, j \leq n$. An example is*

$$\begin{pmatrix} 0 & 3 & -5 \\ -3 & 0 & 2 \\ 5 & -2 & 0 \end{pmatrix} . \tag{1.20}$$

(5) *A square matrix A with real numbers as elements such that $AA^t = A^tA = I$ is called an* **orthogonal matrix**. *An example is*

$$\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} . \tag{1.21}$$

Up to now, we have only considered real-valued matrices. Of course, there exist matrices in which the elements are complex numbers.

Definition 1.14. *The* **conjugate** *of a matrix A is the matrix $\overline{A}$ obtained by replacing every element of A by its complex conjugate. That is to say, if $A = (a_{ij})_{m\times n}$, then $\overline{A} = (\overline{a_{ij}})_{m\times n}$.*

A matrix A is real if and only if $A = \overline{A}$, whereas a matrix A is imaginary if and only if $A = -\overline{A}$.

The **hermitian conjugate** *of a matrix A is the matrix $A^\dagger$ defined by*

$$A^\dagger = \overline{A^t} = (\overline{A})^t . \tag{1.22}$$

Obviously, if A is a $m \times n$ matrix, then $A^\dagger$ is a $n \times m$ matrix.

Example 1.15. The hermitian conjugates of the following matrices,

$$A = \begin{pmatrix} 1 & 2i \\ -i & 3 \end{pmatrix} \tag{1.23}$$

and

$$B = \begin{pmatrix} 2-i & 1+3i & -2 \\ 4 & 0 & 3-4i \end{pmatrix} , \tag{1.24}$$

are

$$A^\dagger = \begin{pmatrix} 1 & i \\ -2i & 3 \end{pmatrix} , \tag{1.25}$$

and

$$B^\dagger = \begin{pmatrix} 2+i & 4 \\ 1-3i & 0 \\ -2 & 3+4i \end{pmatrix} . \tag{1.26}$$

■

Definition 1.16. *A matrix A such that $A = A^\dagger$ is said to be* **hermitian** *while a matrix A such that $A = -A^\dagger$ is said to be* **skew-hermitian**. *Obviously, both hermitian and skew-hermitian matrices must be square matrices. In addition, a real symmetric matrix is a hermitian matrix which is real, and a real skew-symmetric matrix is a skew-hermitian matrix which is real. Below are examples of a hermitian and a skew-hermitian matrix, respectively.*

$$\begin{pmatrix} 7 & 1-4i \\ 1+4i & 3 \end{pmatrix} \qquad \begin{pmatrix} 3i & 5+2i \\ -5+2i & 0 \end{pmatrix} . \tag{1.27}$$

(In quantum mechanics, physical quantities are represented by hermitian operators, and they correspond to hermitian matrices in some basis.)

1.2 Determinants

The determinant of a square matrix A is a number that depends upon the elements of A. Just like the trace, the determinant is only defined for square matrices. If A is a $n \times n$ matrix, then the determinant of A is denoted by

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} . \tag{1.28}$$

We define the determinant of a 1×1 matrix $A = (a_{11})$ to be $\det A = a_{11}$.

To calculate the value of a determinant, we first need to introduce the concepts of the minor and cofactor of an element of a matrix A.

Definition 1.17. *Let $A = (a_{ij})_{n\times n}$ be an $n \times n$ matrix. The* **minor** *M_{ij} of the element a_{ij} is the determinant of the $(n-1) \times (n-1)$ submatrix obtained by deleting all the elements in the ith row and jth column of A,*

$$M_{ij} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1\,(j-1)} & a_{1\,(j+1)} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2\,(j-1)} & a_{2\,(j+1)} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{(i-1)\,1} & a_{(i-1)\,2} & \dots & a_{(i-1)\,(j-1)} & a_{(i-1)\,(j+1)} & \dots & a_{(i-1)\,n} \\ a_{(i+1)\,1} & a_{(i+1)\,2} & \dots & a_{(i+1)\,(j-1)} & a_{(i+1)\,(j+1)} & \dots & a_{(i+1)\,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\,(j-1)} & a_{n\,(j+1)} & \dots & a_{nn} \end{vmatrix} . \tag{1.29}$$

The **cofactor** *C_{ij} of the element a_{ij} is the signed minor defined by*

$$C_{ij} = (-1)^{i+j} M_{ij} \quad \textit{for } i, j = 1, 2, \dots, n. \tag{1.30}$$

The plus or minus sign in the cofactor C_{ij} is given by the term $(-1)^{i+j}$. We can easily check the sign of C_{ij} by noting that the signs form a checkerboard pattern

$$\begin{pmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} . \tag{1.31}$$

Example 1.18. Find the minors and cofactors of all the elements in the following matrices

$$A = \begin{pmatrix} 2 & -3 \\ 1 & 4 \end{pmatrix} \tag{1.32}$$

and

$$B = \begin{pmatrix} 12 & 7 & 0 \\ 5 & 8 & 3 \\ 6 & 7 & 0 \end{pmatrix} . \tag{1.33}$$

The minors of all the elements in the matrix A are

$$M_{11} = |4| = 4 \ , \qquad M_{12} = |1| = 1 \ , \tag{1.34}$$
$$M_{21} = |-3| = -3 \ , \qquad M_{22} = |2| = 2 \ . \tag{1.35}$$

(Note that in the above equation, $|-3|$ is the determinant, not the absolute value.) So the cofactors of all the elements in the matrix A are

$$C_{11} = (-1)^{1+1}M_{11} = 4, \qquad C_{12} = (-1)^{1+2}M_{12} = -1, \tag{1.36}$$
$$C_{21} = (-1)^{2+1}M_{21} = 3, \qquad C_{22} = (-1)^{2+2}M_{22} = 2. \tag{1.37}$$

Similarly, the minors of all the elements in the matrix B are

$$\begin{aligned}
M_{11} &= \begin{vmatrix} 8 & 3 \\ 7 & 0 \end{vmatrix} = -21 \ , \qquad M_{12} = \begin{vmatrix} 5 & 3 \\ 6 & 0 \end{vmatrix} = -18 \ , \qquad M_{13} = \begin{vmatrix} 5 & 8 \\ 6 & 7 \end{vmatrix} = -13 \ , \\
M_{21} &= \begin{vmatrix} 7 & 0 \\ 7 & 0 \end{vmatrix} = 0 \ , \qquad M_{22} = \begin{vmatrix} 12 & 0 \\ 6 & 0 \end{vmatrix} = 0 \ , \qquad M_{23} = \begin{vmatrix} 12 & 7 \\ 6 & 7 \end{vmatrix} = 42 \ , \\
M_{31} &= \begin{vmatrix} 7 & 0 \\ 8 & 3 \end{vmatrix} = 21 \ , \qquad M_{32} = \begin{vmatrix} 12 & 0 \\ 5 & 3 \end{vmatrix} = 36 \ , \qquad M_{33} = \begin{vmatrix} 12 & 7 \\ 5 & 8 \end{vmatrix} = 61 \ .
\end{aligned} \tag{1.38}$$

(We have not talked how to calculate determinant yet. Please see below.) So the cofactors of all the elements in the matrix B are

$$\begin{aligned}
& C_{11} = (-1)^{1+1}M_{11} = -21 \ , \qquad C_{12} = (-1)^{1+2}M_{12} = 18 \ , \\
& C_{13} = (-1)^{1+3}M_{13} = -13 \ , \\
& C_{21} = (-1)^{2+1}M_{21} = 0 \ , \qquad C_{22} = (-1)^{2+2}M_{22} = 0 \ , \\
& C_{23} = (-1)^{2+3}M_{23} = -42 \ , \\
& C_{31} = (-1)^{3+1}M_{31} = 21 \ , \qquad C_{32} = (-1)^{3+2}M_{32} = -36 \ , \\
& C_{33} = (-1)^{3+3}M_{33} = 61 \ .
\end{aligned} \tag{1.39}$$

■

We are now ready to give a general definition for the determinant of an $n \times n$ matrix. There are various ways to define an order-n determinant. Here we choose to give a recursive definition of a determinant.

Definition 1.19. *For any square matrix A of order $n \geq 2$, the* **determinant** *of A is defined as*

$$\det A \equiv \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \equiv \sum_{j=1}^{n} a_{1j}C_{1j} \tag{1.40}$$

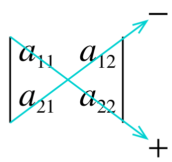
where C_{ij} is the cofactor of the element a_{ij} of A.

Eq. (1.40) is called a **cofactor expansion along the first row of** A. It tells us that the value of the determinant of any $n \times n$ matrix with $n \geq 2$ is equal to the sum of the products of the elements of its first row and their respective cofactors. As a result, in general, a determinant of order n is defined by n determinants of order $n-1$ which in turn is defined by $n-1$ determinants of order $n-2$, and so on, until finally the expansion involves only numbers — determinants of order 1.

Applying Eq. (1.40) to a 2×2 matrix, we obtain

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \ . \tag{1.41}$$

This expansion is just the difference between the product of the main diagonal elements and the product of the non-diagonal elements as shown in the following



Example 1.20. Compute the determinant of the matrix

$$A = \begin{pmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{pmatrix} \ . \tag{1.42}$$

We first compute the cofactors of the elements of the first row of A

$$\begin{aligned} C_{11} &= (-1)^{1+1}M_{11} = \begin{vmatrix} 0 & 2 \\ -1 & 3 \end{vmatrix} = 2 \ , \\ C_{12} &= (-1)^{1+2}M_{12} = -\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = 1 \ , \\ C_{13} &= (-1)^{1+3}M_{13} = \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} = -1 \ . \end{aligned}$$

Then we find from Eq. (1.40) that

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = (5)(2) + (-3)(1) + (2)(-1) = 5 \ . \tag{1.43}$$

∎

For the determinant of a 3×3 matrix, we use Eq. (1.40) together with

Eq. (1.41)

$$
\begin{aligned}
&\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
=\ & a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\
=\ & a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\
& + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
=\ & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} \\
& - a_{12}a_{21}a_{33}\ . \qquad (1.44)
\end{aligned}
$$

This expansion can be also obtained by diagonal multiplication as shown below. We first copy the first two columns to the right of the matrix and compute the products of the elements on the six diagonals. Then the determinant is obtained by adding the downward diagonal products and subtracting the upward diagonal products. Note that the diagonal scheme of expanding determinants shown in above is **not** valid for determinants of order $n > 3$.

$$
\begin{array}{ccc|cc}
 & & & - & - \quad - \\
a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\
a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\
a_{31} & a_{23} & a_{33} & a_{31} & a_{23} \\
 & & & + & + \quad +
\end{array}
$$

Example 1.21. Calculate the determinant of the matrix in Example 1.20 using the diagonal scheme.

We first copy the first two columns to the right of the matrix and compute the six indicated products as shown here.

$$
\begin{array}{ccc|cc}
 & & & 0 & -10 \quad -9 \\
5 & -3 & 2 & 5 & -3 \\
1 & 0 & 2 & 1 & 0 \\
2 & -1 & 3 & 2 & -1 \\
 & & & 0 & -12 \quad -2
\end{array}
$$

Adding the three products at the bottom and subtracting the three products at the top gives

$$
\det A = 0 + (-12) + (-2) - 0 - (-10) - (-9) = 5\ . \qquad (1.45)
$$

■

By definition, the determinant of any square matrix of order $n \geq 2$ can be expanded using a cofactor expansion along the first row. Indeed, we would get exactly the same result by expanding *along any row or any column.* This fact is summarized in the following theorem[1].

Theorem 1.22. Laplace Expansion Theorem
Let $A = (A_{ij})_{n\times n}$ be an $n \times n$ matrix where $n \geq 2$. If we multiply the elements in any row or column of A by the corresponding cofactors, then the sum of the resulting products is the determinant of A. Therefore,

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = \sum_{j=1}^{n} a_{ij}C_{ij} \tag{1.46}$$

for any fixed i with $1 \leq i \leq n$ and

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} = \sum_{i=1}^{n} a_{ij}C_{ij} \tag{1.47}$$

for any fixed j with $1 \leq j \leq n$. They are known as **cofactor expansion along the ith row** and **cofactor expansion along the jth column**, respectively. ■

Example 1.23. Use cofactor expansion along (a) the second column and (b) the third row to find the determinant of the matrix

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 1 \\ 6 & 3 & 0 \end{pmatrix} . \tag{1.48}$$

(a) Using cofactor expansion along the second column, we have

$$\begin{aligned} \det A &= a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} \\ &= -(3)\begin{vmatrix} 1 & 1 \\ 6 & 0 \end{vmatrix} + (-1)\begin{vmatrix} 2 & 4 \\ 6 & 0 \end{vmatrix} - (3)\begin{vmatrix} 2 & 4 \\ 1 & 1 \end{vmatrix} \\ &= -(3)(-6) + (-1)(-24) - (3)(-2) = 48 . \end{aligned} \tag{1.49}$$

Using cofactor expansion along the third row, we have

$$\begin{aligned} \det A &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= (6)\begin{vmatrix} 3 & 4 \\ -1 & 1 \end{vmatrix} - (3)\begin{vmatrix} 2 & 4 \\ 1 & 1 \end{vmatrix} + (0)\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} \\ &= (6)(7) - (3)(-2) + 0 = 48 . \end{aligned} \tag{1.50}$$

■

Note that in Example 1.23, we needed to do fewer calculations in part (b) than in part (a) since we were expanding a column with a zero entry and so we did not need to calculate one cofactor. We can see that the Laplace Expansion Theorem is most useful when the matrix contains a row or column

[1]For the proof of this theorem, see, e.g., C. Ray Wylie, *Advanced Engineering Mathematics*, 4th ed,,pp. 457-459, McGraw-Hill, New York, 1975

with lots of zeros because the number of cofactors needed to be computed can be minimized by expanding along that row or column.

Example 1.24. Compute the determinant of the matrix

$$A = \begin{pmatrix} 2 & -3 & 0 & 1 \\ 5 & 4 & 2 & 0 \\ 1 & -1 & 0 & 3 \\ -2 & 1 & 0 & 0 \end{pmatrix} . \tag{1.51}$$

Notice that we should expand the determinant of A along the third column because this column has only one nonzero element. Moreover, the cofactor C_{23} has a minus sign since $(-1)^{2+3} = -1$. Therefore, we have

$$\begin{aligned} \det A &= a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33} + a_{43}C_{43} \\ &= (0)(C_{13}) + (2)(C_{23}) + (0)(C_{33}) + (0)(C_{34}) \\ &= -2 \begin{vmatrix} 2 & -3 & 1 \\ 1 & -1 & 3 \\ -2 & 1 & 0 \end{vmatrix} . \end{aligned} \tag{1.52}$$

Next we find the determinant of the above matrix by expanding along the third row since it has one zero element. Then we obtain

$$\begin{aligned} \det A &= -2 \left[(-2) \begin{vmatrix} -3 & 1 \\ -1 & 3 \end{vmatrix} - (1) \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} + (0) \begin{vmatrix} 2 & -3 \\ 1 & -1 \end{vmatrix} \right] \\ &= -2[(-2)(-8) - (1)(5) + 0] \\ &= -22 . \end{aligned} \tag{1.53}$$

■

According to Theorem 1.22, the same result would be obtained no matter we expand a determinant in terms of the elements of an arbitrary row or an arbitrary column. As a result, we are led to the following theorem.

Theorem 1.25. For any square matrix A, $\det A^t = \det A$. ■

Example 1.26. Verify that $\det A^t = \det A$ for

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 5 & -2 & 0 \\ 9 & 14 & 3 \end{pmatrix} . \tag{1.54}$$

Expanding successive determinants by elements of their first rows, we find

$$\det A = \begin{vmatrix} 1 & 0 & 0 \\ 5 & -2 & 0 \\ 9 & 14 & 3 \end{vmatrix} = (1) \begin{vmatrix} -2 & 0 \\ 14 & 3 \end{vmatrix} = (1)(-2)(3) = -6 . \tag{1.55}$$

Expanding successive determinants by elements of their first columns, we find

$$\det A^t = \begin{vmatrix} 1 & 5 & 9 \\ 0 & -2 & 14 \\ 0 & 0 & 3 \end{vmatrix} = (1) \begin{vmatrix} -2 & 14 \\ 0 & 3 \end{vmatrix} = (1)(-2)(3) = -6 = \det A . \tag{1.56}$$

■

In Example 1.26, both A and A^t are triangular matrices, and the determinant of either matrix is equal to the product of its diagonal elements. This suggests another theorem for determinants.

Theorem 1.27. The determinant of a triangular matrix is equal to the product of its diagonal elements. Specifically, if $A = (A_{ij})_{n\times n}$ is an $n \times n$ triangular matrix, then

$$\det A = a_{11}a_{22}\cdots a_{nn} \ . \tag{1.57}$$

Proof. To prove the theorem, we let A be an $n \times n$ lower triangular matrix. Then successive expansions of $\det A$ by the elements in the first row of A yields

$$\det A = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \dots & 0 \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} = \dots = a_{11}a_{22}\cdots a_{nn} \ . \tag{1.58}$$

If A is an upper triangular matrix, successive expansions of $\det A$ by the elements in the first column of A prove the theorem. ■

Determinants have a number of special properties that can be used to simplify their expansions. The most important and useful of these properties are summarized in the next theorem.

Theorem 1.28. The determinant of a square matrix A has the following properties.

(a) If A has either a row (or a column) that only contains zero elements, then $\det A = 0$.

(b) If B is obtained by multiplying a row (or a column) of A by a number k, then $\det B = k \det A$.

(c) If B is obtained by interchanging any two rows (or two columns) of A, then $\det B = -\det A$.

(d) If A has two identical rows (or columns), then $\det A = 0$.

(e) If A, B, and C are identical except that the ith row (or the i column) of C is the sum of the ith rows (or the ith columns) of A and B, then $\det C = \det A + \det B$.

(f) If B is obtained by adding a multiple of one row (or column) of A to another row (or column), then $\det B = \det A$.

(g) The sum of the products of the elements in the ith row (or the ith column) of A with the cofactors of the corresponding elements in the jth row (or jth column) of A is equal to zero if $i \neq j$.

(h) If A and B are square matrices of the same order, then $\det(AB) = (\det A)(\det B)$.

■

Example 1.29. Compute the determinant of

$$A = \begin{pmatrix} 0 & 2 & -4 & 5 \\ 3 & 0 & -3 & 6 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{pmatrix} . \tag{1.59}$$

The determinant of A can be computed efficiently if it can be shown to be equal to the determinant of an upper triangular matrix. We use this approach to compute $\det A$ by using Theorem 1.28 as follows.

First, taking a factor 3 out of the second row and then interchanging the first and second rows gives

$$\det A = 3 \begin{vmatrix} 0 & 2 & -4 & 5 \\ 1 & 0 & -1 & 2 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{vmatrix} = -3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & -4 & 5 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{vmatrix} . \tag{1.60}$$

Next, subtracting the third row by two times the first row and the fourth row by five times the first row and then interchanging the second and fourth rows yields

$$\begin{aligned} \det A &= -3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & -4 & 5 \\ 0 & 4 & 7 & 3 \\ 5 & -1 & -3 & 1 \end{vmatrix} \\ &= -3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & -4 & 5 \\ 0 & 4 & 7 & 3 \\ 0 & -1 & 2 & -9 \end{vmatrix} \\ &= 3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & -9 \\ 0 & 4 & 7 & 3 \\ 0 & 2 & -4 & 5 \end{vmatrix} . \end{aligned} \tag{1.61}$$

Finally, adding four times the second row to the third row and two times the second row to the fourth row, and computing the "triangular" determinant by Theorem 1.27, we obtain

$$\begin{aligned} \det A &= -3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & -9 \\ 0 & 0 & 15 & -33 \\ 0 & 2 & -4 & 5 \end{vmatrix} = 3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & -9 \\ 0 & 0 & 15 & -33 \\ 0 & 0 & 0 & -13 \end{vmatrix} \\ &= (3)(1)(-1)(15)(-13) = 585 . \end{aligned} \tag{1.62}$$

■

Example 1.30. Verify that Theorem 1.28(g) is true for

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} . \tag{1.63}$$

If we multiply the elements in the first row of A by the cofactors of the corresponding elements in the third row of A, we obtain

$$a_{11} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{12} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{13} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} . \tag{1.64}$$

Clearly, it is the expansion of the determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} \tag{1.65}$$

along the third row. Thus this determinant is equal to zero since it has two identical rows. ■

Example 1.31. Verify that $(\det A)(\det B) = \det(AB)$ for the matrices

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 2 & -5 & 4 \\ -2 & 3 & -1 \end{pmatrix} \tag{1.66}$$

and

$$B = \begin{pmatrix} -3 & -2 & 12 \\ 4 & 1 & -6 \\ 2 & 3 & -10 \end{pmatrix} . \tag{1.67}$$

Adding three times the first column of $\det A$ to its third column, we obtain

$$\det A = \begin{vmatrix} 1 & 0 & 0 \\ 2 & -5 & 10 \\ -2 & 3 & -7 \end{vmatrix} = (1) \begin{vmatrix} -5 & 10 \\ 3 & -7 \end{vmatrix} = (1)(5) = 5 . \tag{1.68}$$

Adding two times the second row of $\det B$ to its first row, we obtain

$$\det B = \begin{vmatrix} 5 & 0 & 0 \\ 4 & 1 & -6 \\ 2 & 3 & -10 \end{vmatrix} = (5) \begin{vmatrix} 1 & -6 \\ 3 & -10 \end{vmatrix} = (5)(8) = 40 . \tag{1.69}$$

We next compute the matrix product AB

$$AB = \begin{pmatrix} 1 & 0 & -3 \\ 2 & -5 & 4 \\ -2 & 3 & -1 \end{pmatrix} \begin{pmatrix} -3 & -2 & 12 \\ 4 & 1 & -6 \\ 2 & 3 & -10 \end{pmatrix} = \begin{pmatrix} -9 & -11 & 42 \\ -18 & 3 & 14 \\ 16 & 4 & -32 \end{pmatrix} . \tag{1.70}$$

Using properties (b) and (f) of Theorem 1.28, we find that

$$\begin{aligned}
\det(AB) &= 4\begin{vmatrix} -9 & -11 & 42 \\ -18 & 3 & 14 \\ 4 & 1 & -8 \end{vmatrix} \\
&= 4\begin{vmatrix} 35 & -11 & 42 \\ -30 & 3 & 14 \\ 0 & 1 & -8 \end{vmatrix} \\
&= 4\begin{vmatrix} 35 & -11 & -46 \\ -30 & 3 & 38 \\ 0 & 1 & 0 \end{vmatrix} \\
&= -(4)\begin{vmatrix} 35 & -46 \\ -30 & 38 \end{vmatrix} \\
&= -(4)(-50) = 200 = (\det A)(\det B) \ . \qquad (1.71)
\end{aligned}$$

■

1.3 Matrix Inverses

If a, b, and c are ordinary numbers, we know that the solution of the equation $ab = c$ is simply $b = c/a$ provided that $a \neq 0$. However, we cannot find the solution in this way if a, b, and c were matrices since division is not defined for matrices. However, if A is an $n \times n$ matrix for which $\det A \neq 0$, then we can define an $n \times n$ matrix A^{-1} called the inverse of A such that the solution of the matrix equation $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = A^{-1}\mathbf{b}$ where $\mathbf{x}$ and $\mathbf{b}$ are both $n \times 1$ column matrices. Notice that if the matrix A^{-1} exists, then we expect $AA^{-1} = I$. Therefore, we give a formal definition for the inverse of a matrix as follows.

Definition 1.32. *If A is an $n \times n$ matrix, an* **inverse** *of A is an $n \times n$ matrix A^{-1} with the property that*

$$AA^{-1} = A^{-1}A = I \ , \qquad (1.72)$$

where $I \equiv I_n$ is the $n \times n$ identity matrix. A matrix A is said to be **invertible** *if its inverse A^{-1} exists.*

It can be shown that a matrix A is invertible if and only if its determinant is non-zero, $\det A \neq 0$. A matrix that is not invertible is called singular; otherwise it is called non-singular.

Theorem 1.33. If A is an invertible matrix, then its inverse A^{-1} is unique.

Proof. If a square matrix A has two inverses B and C, by definition, $AC = I$ and $BA = I$. So we have

$$B = BI = B(AC) = (BA)C = IC = C \ . \qquad (1.73)$$

Hence, the inverse is unique. ■

Before proceeding further, we consider the following theorem which is a set of equivalent statements telling us what it means for a matrix to be invertible.

Theorem 1.34. Let A be an $n \times n$ matrix. The following statements are equivalent.

(a) A is an invertible matrix.

(b) $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$ for any column matrix $\mathbf{b}$.

(c) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

(d) The columns of A are linearly independent.

Proof. Here we only give proof for the following chain of implications

$$\text{(a)} \Rightarrow \text{(b)} \Rightarrow \text{(c)} \Rightarrow \text{(d)} \; .$$

The proof for the other implications in this theorem is omitted since it requires the concepts of row echelon form and elementary matrices which are not discussed.

In order to prove (a) $\Rightarrow$ (b), we need to demonstrate that $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution of the matrix equation $A\mathbf{x} = \mathbf{b}$ and it is the *only* solution to this equation. To show that the solution $\mathbf{x} = A^{-1}\mathbf{b}$ works, we substitute it back into $A\mathbf{x} = \mathbf{b}$ to obtain

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b} \; . \tag{1.74}$$

So $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution of the equation $A\mathbf{x} = \mathbf{b}$. To show that this solution is unique, suppose $\mathbf{y}$ is another solution such that $A\mathbf{y} = \mathbf{b}$. Multiplying both sides of this equation by A^{-1} from the left, we obtain

$$\begin{aligned} A^{-1}(A\mathbf{y}) &= A^{-1}\mathbf{b} \\ (A^{-1}A)\mathbf{y} &= A^{-1}\mathbf{b} \\ I\mathbf{y} &= A^{-1}\mathbf{b} \\ \mathbf{y} &= A^{-1}\mathbf{b} \; . \end{aligned} \tag{1.75}$$

It means that any solution of $A\mathbf{x} = \mathbf{b}$ must be $A^{-1}\mathbf{b}$, and thus the solution is unique.

Next, we prove that (b) $\Rightarrow$ (c). Statement (b) tells us that the equation $A\mathbf{x} = \mathbf{0}$ has a unique solution. Obviously, $A\mathbf{x} = \mathbf{0}$ always has $\mathbf{x} = \mathbf{0}$ as *one* solution. So $\mathbf{x} = \mathbf{0}$ must be the *only* solution in this case.

Finally, to prove (c) $\Rightarrow$ (d), we denote the matrix $A = (\mathbf{u}_1\, \mathbf{u}_2\, \ldots\, \mathbf{u}_n)$, where $\mathbf{u}_i$ is the ith column of A. Then the matrix equation $A\mathbf{x} = \mathbf{0}$ can be written as $x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbf{0}$. Each linear dependence relationship among the $\mathbf{u}_i$ corresponds to a nontrivial solution of $A\mathbf{x} = \mathbf{0}$. Therefore, the columns of a square matrix A must be linearly independent if $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. ■

Theorem 1.35. Let A and B be invertible matrices of the same size, and let α be a scalar. Then,

(a) A^{-1} is invertible and $(A^{-1})^{-1} = A$;

(b) AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$;

(c) A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$;

(d) $A^\dagger$ is invertible and $(A^\dagger)^{-1} = (A^{-1})^\dagger$ if A is a complex matrix;

(e) if $\alpha \neq 0$, αA is invertible and $(\alpha A)^{-1} = (1/\alpha)A^{-1}$;

(f) A^n is invertible for any non-negative integer n and $(A^n)^{-1} = (A^{-1})^n$

Proof. We will prove (a)-(c), leaving the proof of (d)-(f) as exercises.

(a) Consider a matrix B such that $A^{-1}B = I$ and $BA^{-1} = I$. Of course, these equations are satisfied with B replaced by A. Thus, A^{-1} is invertible and A is its inverse since matrix inverses are unique.

(b) We multiply $B^{-1}A^{-1}$ by AB to obtain

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I \ . \tag{1.76}$$

A similar calculation show that $(B^{-1}A^{-1})AB = I$. Thus AB is invertible and $B^{-1}A^{-1}$ is its inverse.

(c) Recall that $B^tA^t = (AB)^t$ according to Theorem 1.10(d). Replacing B by A^{-1} gives

$$(A^{-1})^tA^t = (AA^{-1})^t = I^t = I \ . \tag{1.77}$$

Thus A^t is invertible and $(A^{-1})^t$ is its inverse.

∎

Property (b) can be generalized to products of finitely many invertible matrices: If $A_1, A_2, \ldots, A_n$ are invertible matrices of the same size, then $A_1A_2\cdots A_n$ is invertible and $(A_1A_2\cdots A_n)^{-1} = A_n^{-1}\cdots A_2^{-1}A_1^{-1}$. Then we can state that the inverse of a product of invertible matrices is the product of their inverses in the reverse order.

Property (f) allows us to define negative integer powers of an invertible matrix: If A is an invertible matrix and n is a positive integer, then A^{-n} is defined by $A^{-n} = (A^n)^{-1} = (A^{-1})^n$.

Example 1.36. Let A, B, and X be square matrices of the same size. Find X in terms of A and B if $A^{-1}(BX)^{-1} = (A^{-1}B^3)^2$.

We have

$$\begin{aligned} A^{-1}(BX)^{-1} &= (A^{-1}B^3)^2 \\ A^{-1}X^{-1}B^{-1} &= A^{-1}B^3A^{-1}B^3 \\ X^{-1}B^{-1} &= B^3A^{-1}B^3 \\ X^{-1} &= B^3A^{-1}B^4 \\ X &= B^{-4}AB^{-3} \ . \end{aligned} \tag{1.78}$$

∎

There are many methods for finding the inverse of an invertible matrix. For example, the inverse of a 2×2 matrix can be found using a simple formula as stated in the next theorem.

Theorem 1.37. For

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} , \tag{1.79}$$

if $\det A = ad - bc \neq 0$, then A is invertible and its inverse is

$$A^{-1} = \frac{1}{(ad - bc)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} . \tag{1.80}$$

If $\det A = 0$, then A is not invertible. ■

Example 1.38. Find the inverses of the following matrices if exist.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \tag{1.81}$$

and

$$B = \begin{pmatrix} 12 & -15 \\ 4 & -5 \end{pmatrix} . \tag{1.82}$$

Since $\det A = (1)(4) - (2)(3) = -2 \neq 0$, A is invertible. Using Eq. (1.80), we obtain

$$A^{-1} = \frac{1}{-2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} . \tag{1.83}$$

B is not invertible because $\det B = (12)(-5) - (-15)(4) = 0$. ■

To calculate the inverse of an invertible matrix of any order, we need the concept of adjoint.

Definition 1.39. *Let C be an $n \times n$ matrix whose elements are the cofactors C_{ij} of the corresponding elements of an $n \times n$ matrix A. The transpose of C, C^t, is called the* **adjoint** *(or* **adjugate***) of A and is denoted by adj A,*

$$adj\, A = \begin{pmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{pmatrix} . \tag{1.84}$$

Theorem 1.40. If A is an invertible $n \times n$ matrix, then its inverse is

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A . \tag{1.85}$$

Proof. We consider the matrix product AB where $B = \operatorname{adj} A$. Denoting the element in the ith row and jth column of a matrix X as x_{ij}, we have

$$(A\,(\mathrm{adj}\,A))_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} a_{ik} C_{jk} = \det A\,\delta_{ij} \tag{1.86}$$

where the last step follows from Theorem 1.22 and 1.28(g). It implies that

$$A\,(\mathrm{adj}\,A) = (\det A)\,I \tag{1.87}$$

where $I \equiv I_n$ is the identity matrix of order n. Similarly, we can show that

$$(\mathrm{adj}\,A)\,A = (\det A)\,I \ . \tag{1.88}$$

Thus, if A is invertible ($\det A \neq 0$), we have

$$A^{-1} = \frac{1}{\det A}\,(\mathrm{adj}\,A) \ . \tag{1.89}$$

■

Example 1.41. Use the adjoint method to find the inverse of the matrix

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{pmatrix} \ . \tag{1.90}$$

The adjoint of A is

$$\mathrm{adj}\,A = \begin{pmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{pmatrix} \ . \tag{1.91}$$

The determinant of A is

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = (2)(-2) + (1)(3) + (3)(5) = 14 \ . \tag{1.92}$$

Hence,

$$A^{-1} = \frac{1}{\det A}\,\mathrm{adj}\,A = \frac{1}{14}\begin{pmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{pmatrix} \ . \tag{1.93}$$

■

Finally, we consider two special types of matrices called, unitary matrices and normal matrices. We will discuss interesting properties of the eigenvalues and eigenvectors of a normal matrix in later section. Besides, we will show that a matrix can be diagonalized using a unitary matrix under specific conditions later.

A square matrix A such that $A^\dagger = A^{-1}$ is called a **unitary matrix**. For a unitary matrix A,

$$A^\dagger A = AA^\dagger = I \ , \tag{1.94}$$

and so

$$\det(A^\dagger A) = (\det A^\dagger)(\det A) = \overline{(\det A)}(\det A) = \det I = 1 \ . \tag{1.95}$$

Thus the determinant of a unitary matrix has unit modulus. Note that the inverse A^{-1} of a unitary matrix A is also unitary. Obviously, $A^\dagger = A^t$ if A is real. It implies that a real orthogonal matrix is a special case of a unitary matrix in which all the elements are real.

A square matrix A that commutes with its Hermitian conjugate, $AA^\dagger = A^\dagger A$, is called a **normal matrix**. Note that both Hermitian, skew-hermitian, and unitary matrices are examples of normal matrices. For a unitary matrix A, $A^\dagger = A^{-1}$ and so

$$AA^\dagger = AA^{-1} = I = A^{-1}A = A^\dagger A \ . \tag{1.96}$$

Similarly, for a Hermitian matrix A, $A = A^\dagger$ and so $AA^\dagger = A^\dagger A$. We can show that a skew-hermitian matrix is normal using similar arguments. Note that the inverse A^{-1} of a normal matrix A is also a normal matrix.

1.4 Eigenvalues and Eigenvectors

In this section, we study whether there exist non-zero column matrix $\mathbf{v}$ such that the matrix product $A\mathbf{v}$ is just a scalar multiple of $\mathbf{v}$ where A is a square matrix for which $A\mathbf{v}$ is defined. This is the **eigenvalue problem** (or **characteristic value problem**), and it is one of the most important problem in linear algebra. It has applications not only in mathematics but also in many other fields as well.

Definition 1.42. *Let A be an $n \times n$ matrix. A scalar λ is called an* **eigenvalue** *(or* **characteristic value***) of A if there exists a non-zero column matrix $\mathbf{v}$ such that*

$$A\mathbf{v} = \lambda\mathbf{v} \ . \tag{1.97}$$

Such a column matrix $\mathbf{v}$ is called an **eigenvector** *(or* **characteristic vector***) of A corresponding to the eigenvalue λ.*

Notice that Eq. (1.97) can be rewritten in the equivalent form

$$(A - \lambda I)\mathbf{v} = \mathbf{0} \tag{1.98}$$

where I is the $n \times n$ identity matrix and $\mathbf{0}$ is the $n \times 1$ zero matrix. Eq. (1.98) has nontrivial solutions if and only if

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \ . \tag{1.99}$$

Hence, the matrix $A - \lambda I$ is not invertible. When we expand $p(\lambda) =$

$\det(A - \lambda I)$, we get a polynomial in the parameter λ called the **characteristic polynomial** of A. Eq. (1.99) is called the **characteristic equation** of A. If A is an $n \times n$ matrix, $p(\lambda)$ will be of degree n and so $p(\lambda) = 0$ has at most n distinct roots. The roots of $p(\lambda) = 0$ are the eigenvalues of A and the corresponding nontrivial solutions are the eigenvectors of A. Thus an $n \times n$ matrix has at most n distinct eigenvalues.

Example 1.43. Find the eigenvalues and the corresponding eigenvectors of the matrix

$$A = \begin{pmatrix} 4 & -5 \\ 1 & -2 \end{pmatrix} . \tag{1.100}$$

The characteristic equation of the matrix A is

$$\det(A - \lambda I) = \begin{vmatrix} 4-\lambda & -5 \\ 1 & -2-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3 = (\lambda + 1)(\lambda - 3) = 0 . \tag{1.101}$$

So the eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = 3$.

To find the eigenvector $\mathbf{v}_1$ of A corresponding to the eigenvalue $\lambda_1 = -1$, we need to solve the equation $(A - \lambda_1)\mathbf{v}_1 = \mathbf{0}$

$$\begin{pmatrix} 4-(-1) & -5 \\ 1 & -2-(-1) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} . \tag{1.102}$$

It is equivalent to the system of linear algebraic equations $5x_1 - 5x_2 = 0$ and $x_1 - x_2 = 0$, which implies $x_1 = x_2$. So the eigenvector $\mathbf{v}_1$ corresponding to the eigenvalue $\lambda_1 = -1$ is given by

$$\mathbf{v}_1 = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{1.103}$$

where c_1 is any non-zero real number. Similarly, the eigenvector $\mathbf{v}_2$ corresponding to the eigenvalue $\lambda_2 = 3$ is given by

$$\mathbf{v}_2 = c_2 \begin{pmatrix} 5 \\ 1 \end{pmatrix} \tag{1.104}$$

where c_2 is any non-zero real number. It can be shown that the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. Notice that here we only consider real eigenvectors for simplicity. ■

Example 1.44. It is straight forward to find the eigenvalues and the corresponding eigenvectors of the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{1.105}$$

They are $\lambda_1 = i$ with $\mathbf{v}_1 = c_1(1\ i)^t$, and $\lambda_2 = -i$ with $\mathbf{v}_2 = c_2(1\ -i)^t$.

Consider the function

$$f_{a,b}(x) \equiv a\cos x + b\sin x \ . \tag{1.106}$$

Its derivative is

$$\frac{\mathrm{d}f_{a,b}}{\mathrm{d}x} = -a\sin x + b\cos x \ . \tag{1.107}$$

If we represent the function $f_{a,b}$ by a column matrix

$$\begin{pmatrix} a \\ b \end{pmatrix} , \tag{1.108}$$

then the action of differentiation is same as the action of the matrix A

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ -a \end{pmatrix} . \tag{1.109}$$

The eigenvectors are somethings that you knew. For example, $\mathbf{v}_1$ is, for example,

$$f_{1,i}(x) = \cos x + i\sin x = e^{ix} \ . \tag{1.110}$$

■

The eigenvalues of some matrices are very easy to find.

Theorem 1.45. The eigenvalues of a triangular matrix are the elements on its main diagonal. ■

The next theorem tells us the relationship between eigenvectors corresponding to distinct eigenvalue.

Theorem 1.46. Let A be an $n\times n$ matrix. If $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the distinct eigenvalues of A with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ where $m \le n$, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly independent,

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0} \tag{1.111}$$

implies that the scalars $c_1 = c_2 = \cdots = c_m = 0$.

Proof. We assume that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly dependent and show that this assumption leads to contradiction. Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent for some $k < m$ and there exists scalars $c_1, c_2, \ldots, c_k$ such that

$$\mathbf{v}_{k+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \ . \tag{1.112}$$

Multiplying both sides of this equation by A and using the fact that $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ for each i yields

$$\lambda_{k+1}\mathbf{v}_{k+1} = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_k\lambda_k\mathbf{v}_k \ . \tag{1.113}$$

However, multiplying both sides of Eq. (1.112) by λ_{k+1} gives

$$\lambda_{k+1}\mathbf{v}_{k+1} = c_1\lambda_{k+1}\mathbf{v}_1 + c_2\lambda_{k+1}\mathbf{v}_2 + \cdots + c_k\lambda_{k+1}\mathbf{v}_k. \tag{1.114}$$

We subtract this equation from Eq. (1.113) to obtain

$$c_1(\lambda_1 - \lambda_{k+1})\mathbf{v}_1 + c_2(\lambda_2 - \lambda_{k+1})\mathbf{v}_2 + \cdots + c_k(\lambda_k - \lambda_{k+1})\mathbf{v}_k = \mathbf{0} \ . \tag{1.115}$$

The linearly independence of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ implies that

$$c_1(\lambda_1 - \lambda_{k+1}) = c_2(\lambda_2 - \lambda_{k+1}) = \cdots = c_k(\lambda_k - \lambda_{k+1}) = 0 \ . \tag{1.116}$$

As a result, $c_1 = c_2 = \cdots = c_k = 0$ since the eigenvalues λ_i are all distinct.

Therefore, $\mathbf{v}_{k+1} = 0$, which is impossible since the eigenvectors $\mathbf{v}_{k+1} \neq \mathbf{0}$. Thus, we have a contradiction and so $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ must be linearly independent. ■

Before ending this section, we will discuss the properties of the eigenvalues and eigenvectors of normal matrices which play an important role in quantum mechanics.

Theorem 1.47. Let A be a normal square matrix.

(a) Eigenvalues of $A^\dagger$ are the complex conjugates of the eigenvalues of A.

(b) The eigenvectors corresponding to different eigenvalues of A are orthogonal.

Proof.

(a) Assume that λ is an eigenvalue of a normal matrix A with corresponding eigenvector $\mathbf{v}$. Since $A\mathbf{v} = \lambda\mathbf{v}$, we have $B\mathbf{v} = 0$ where $B = A - \lambda I$. Taking the Hermitian conjugate on both sides of this equation gives

$$(B\mathbf{v})^\dagger = \mathbf{v}^\dagger B^\dagger = 0 \ . \tag{1.117}$$

Multiplying the above equation by $B\mathbf{v}$, we obtain

$$\mathbf{v}^\dagger B^\dagger B\mathbf{v} = 0 \ . \tag{1.118}$$

Note that

$$\begin{aligned} B^\dagger B &= (A - \lambda I)^\dagger (A - \lambda I) \\ &= (A^\dagger - \overline{\lambda} I)(A - \lambda I) \\ &= A^\dagger A - \overline{\lambda} A - \lambda A^\dagger + \lambda\overline{\lambda} \\ &= AA^\dagger - \overline{\lambda} A - \lambda A^\dagger + \lambda\overline{\lambda} \\ &= (A - \lambda I)(A - \lambda I)^\dagger \\ &= BB^\dagger \ , \end{aligned} \tag{1.119}$$

since A is normal, which implies that B is also normal. Hence,

$$\mathbf{v}^\dagger B^\dagger B\mathbf{v} = \mathbf{v}^\dagger BB^\dagger\mathbf{v} = (B^\dagger\mathbf{v})^\dagger B^\dagger\mathbf{v} = 0 \ , \tag{1.120}$$

and $B^\dagger\mathbf{v} = (A^\dagger - \overline{\lambda} I)\mathbf{v} = 0$. Therefore, if A is a normal matrix for which $A\mathbf{v} = \lambda\mathbf{v}$, then $A^\dagger\mathbf{v} = \overline{\lambda}\mathbf{v}$. That is to say, the eigenvalues of $A^\dagger$ are the complex conjugates of the eigenvalues of A. However, if A is also Hermitian, then $A = A^\dagger$ and it follows immediately that $\lambda = \overline{\lambda}$. Thus all the eigenvalues of a Hermitian matrix are real.

(b) We let λ_i and λ_j be two distinct eigenvalues of a normal matrix A. Obviously, their corresponding eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ are distinct and they must satisfy the equations $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ and $A\mathbf{v}_j = \lambda_j \mathbf{v}_j$. Multiplying the second one by $\mathbf{v}_i^\dagger$, we obtain

$$\mathbf{v}_i^\dagger A\mathbf{v}_j = \lambda_j \mathbf{v}_i^\dagger \mathbf{v}_j \tag{1.121}$$

In addition, since A is normal, $A^\dagger \mathbf{v}_i = \overline{\lambda_i}\mathbf{v}_i$ and so we have

$$\mathbf{v}_i^\dagger A\mathbf{v}_j = (A^\dagger \mathbf{v}_i)^\dagger \mathbf{v}_j = (\overline{\lambda_i}\mathbf{v}_i)^\dagger \mathbf{v}_j = \lambda_i \mathbf{v}_i^\dagger \mathbf{v}_j \ . \tag{1.122}$$

Subtracting yields $(\lambda_i - \lambda_j)\mathbf{v}_i^\dagger \mathbf{v}_j = \mathbf{0}$, and we have

$$\mathbf{v}_i^\dagger \mathbf{v}_j = \mathbf{v}_i \cdot \mathbf{v}_j = 0 \ . \tag{1.123}$$

Thus, if $\lambda_i \neq \lambda_j$, the corresponding eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ are orthogonal.

■

If $\mathbf{u}$ and $\mathbf{v}$ are $n \times 1$ column matrices with complex elements, then

$$\mathbf{u}^\dagger \mathbf{v} = \overline{u_1}v_1 + \overline{u_2}v_2 + \cdots + \overline{u_n}v_n = \begin{pmatrix} \overline{u_1} & \overline{u_2} & \ldots & \overline{u_n} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \ . \tag{1.124}$$

Therefore, the matrix product $\mathbf{u}^\dagger \mathbf{v} = 0$ if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

1.5 Diagonalization of Matrices

As seen in last section, we can find the eigenvalues of triangular and diagonal matrices very easily. It would be nice if a square matrix can be related to a triangular or diagonal one such that they have exactly the same eigenvalues. This can be done if the matrix is **similar** to a diagonal matrix. The process of finding a diagonal matrix similar to a square matrix is known as the **diagonalization** of the square matrix.

Definition 1.48. *Let A and B be square matrices of the same size. We say that A is* **similar** *to B, denoted as $A \sim B$, if there exists an invertible matrix P such that $PAP^{-1} = B$.*

If $A \sim B$, then we can write $A = PBP^{-1}$ or $AP = PB$. Besides, the invertible matrix P such that $B = P^{-1}AP$ depends on the matrices A and B. It is not unique for a given pair of similar matrices A and B. For example, if $A = B = I$, then $PIP^{-1} = I$ for *any* invertible matrix P. A transformation mapping of a matrix A into PAP^{-1} is called a **similarity transformation** of the matrix A.

Theorem 1.49. Similarity is an Equivalence Relation Let A, B, and C be square matrices of the same size.

(a) A is similar to itself, $A \sim A$.

(b) If $A \sim B$, then $B \sim A$.

(c) If $A \sim B$ and $B \sim C$, then $A \sim C$.

■

Theorem 1.50. Let A and B be square matrices of the same size with $A \sim B$.

(a) A and B have the same determinant, $\det A = \det B$.

(b) A and B have the same characteristic polynomial and hence the same eigenvalues.

Proof. Using property (h) of Theorem 1.28. ■

Definition 1.51. *A square matrix A is said to be* **diagonalizable** *if A is similar to a diagonal matrix D, $A = PDP^{-1}$ for some invertible matrix P.*

Theorem 1.52. The Diagonalization Theorem
An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. More precisely, there exist an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$ if and only if the columns of P are n linearly independent eigenvectors of A. In such case, the diagonal entries of D are the eigenvalues of A corresponding to the eigenvectors of P in the same order.

Proof. We let P be an $n \times n$ matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ and D be a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then,

$$AP = A\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix} = \begin{pmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \ldots & A\mathbf{v}_n \end{pmatrix} \tag{1.125}$$

and

$$\begin{aligned} PD &= \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \ldots & \lambda_n\mathbf{v}_n \end{pmatrix} . \end{aligned} \tag{1.126}$$

Suppose a matrix A is diagonalizable, $A = PDP^{-1}$ or $AP = PD$. Then we have

$$\begin{pmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \ldots & A\mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \ldots & \lambda_n\mathbf{v}_n \end{pmatrix} , \tag{1.127}$$

which implies

$$A\mathbf{v}_i = \lambda_i\mathbf{v}_i \tag{1.128}$$

for $i = 1, \ldots, n$. Hence, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are the corresponding eigenvectors. Theorem 1.34 also tells us that since P

is invertible, its columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ must be linearly independent. This argument proves the "only if" parts of the first and second statements along with the third statement of the theorem. The proof of the "if" parts of the first and second statements are omitted as it is beyond our scope. ■

Example 1.53. If possible, diagonalize the matrix

$$A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} . \tag{1.129}$$

In other words, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

There are four steps to implement the description in the above theorem.

Firstly, find the eigenvalues of A. The characteristic equation of A is

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix} = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2 . \tag{1.130}$$

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = -2$.

Secondly, find three linearly independent eigenvectors of A. Three eigenvectors are needed since A is a 3×3 matrix. This is the critical step. We find that the eigenvectors are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} , \tag{1.131}$$

and

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} . \tag{1.132}$$

Since $\lambda_2 = \lambda_3$, there is some freedom in choosing $\mathbf{v}_2$ and $\mathbf{v}_3$. It is easy to check that these vectors are linearly independent.

Thirdly, construct P from the eigenvectors. The order of the eigenvectors is not important. Using the order chosen in step 2,

$$P = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} . \tag{1.133}$$

Last step, construct D from the corresponding eigenvalues. The eigenvalues should come up on the diagonal of D in the same order as their corresponding eigenvectors in P. Therefore, we expect

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} . \tag{1.134}$$

We need to verify that $AP = PD$.

$$\begin{aligned}
AP &= \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ -1 & 0 & -2 \\ 1 & 2 & 2 \end{pmatrix} , \\
PD &= \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ -1 & 0 & -2 \\ 1 & 2 & 2 \end{pmatrix} .
\end{aligned} \tag{1.135}$$

■

Example 1.54. We will study a linear system of first order differential equations with constant coefficients

$$\begin{aligned}
\frac{\mathrm{d}y_1}{\mathrm{d}x} &= y_1 + 3y_2 + 3y_3 + f_1 , \\
\frac{\mathrm{d}y_2}{\mathrm{d}x} &= -3y_1 + -5y_2 - 3y_3 + f_2 , \\
\frac{\mathrm{d}y_n}{\mathrm{d}x} &= 3y_1 + 3y_2 + y_3 + f_3 ,
\end{aligned} \tag{1.136}$$

where f_i are constants. In obvious notations, we can rewrite the system as

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}x} = A\,\mathbf{y} + \mathbf{f} , \tag{1.137}$$

where A is just the matrix in the last example, Example 1.53. Define $\mathbf{w} \equiv P^{-1}\mathbf{y}$, we have

$$\begin{aligned}
P\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}x} &= A\,P\,\mathbf{w} + \mathbf{f} \\
\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}x} &= P^{-1}AP\,\mathbf{w} + P^{-1}\mathbf{f} \\
\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}x} &= D\,\mathbf{w} + P^{-1}\mathbf{f} .
\end{aligned} \tag{1.138}$$

We can find that

$$P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ 1 & 2 & 1 \end{pmatrix} . \tag{1.139}$$

Hence, the system in component form is

$$\begin{aligned}
\frac{\mathrm{d}w_1}{\mathrm{d}x} &= w_1 + (f_1 + f_2 + f_3) \\
\frac{\mathrm{d}w_2}{\mathrm{d}x} &= -2w_2 - (f_2 + f_3) \\
\frac{\mathrm{d}w_3}{\mathrm{d}x} &= -2w_3 + (f_1 + 2f_2 + f_3) .
\end{aligned} \tag{1.140}$$

The solution is

$$\begin{aligned} w_1 &= C_1 e^{x} - (f_1 + f_2 + f_3) \\ w_2 &= C_2 e^{-2x} + (f_2 + f_3) \\ w_3 &= C_3 e^{-2x} - (f_1 + 2f_2 + f_3)\ , \end{aligned} \tag{1.141}$$

where C_i are integration constants. We can express the y_i in terms of the w_i. (Note that this is a special example. Solving differential equations is in general difficult.) ■

Theorem 1.55. If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.

Proof. All we need is to note that by Theorem 1.46, eigenvectors corresponding to distinct eigenvalues are linearly independent. ■

If an $n \times n$ matrix has degenerate eigenvalues (some eigenvalues are the same), then it may not have n linearly independent eigenvectors such that the matrix cannot be diagonalized.

Example 1.56. Find A^k for positive integer k and

$$A = \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix}\ . \tag{1.142}$$

We find that $A = PDP^{-1}$ where

$$P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \tag{1.143}$$

and

$$D = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}\ . \tag{1.144}$$

Hence,

$$P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \tag{1.145}$$

and

$$\begin{aligned} A^k &= (PDP^{-1})^k = PD^kP^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}^k \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 5^k & 0 \\ 0 & 3^k \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2\cdot 5^k - 3^k & 5^k - 3^k \\ 2\cdot 3^k - 2\cdot 5^k & 2\cdot 3^k - 5^k \end{pmatrix}\ . \end{aligned} \tag{1.146}$$

■

Chapter 2

Systems of Linear Algebraic Equations

We can find applications of matrices in a wide range of fields including computer science, economics, engineering, physics, and statistics. For example, matrices are used for message encryption and projection of three dimensional image into a two dimensional screen in computer science. These applications often involve solving systems of linear algebraic equations. In general, a system of linear algebraic equations are best represented in terms of matrices and the set of all solutions to the system can then be easily determined. Here, we discuss how to apply matrices to solve system of linear algebraic equations. Let's begin our discussion by introducing some notation and terminology.

2.1 Notation and Terminology

Definition 2.1. *A* **linear algebraic equation** *in n variables $x_1, x_2, \ldots, x_n$ is an equation of the form*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{2.1}$$

where the coefficients $a_1, a_2, \ldots, a_n$ and the constant term b are all constants.

Example 2.2. The following equations are linear

$$\begin{aligned} 5x - 2y &= -1 && (2.2) \\ \frac{\pi}{6}x + \sqrt{3}y - \left(\cos\frac{\pi}{4}\right)z &= 1\ . && (2.3) \end{aligned}$$

Beware that although the coefficients and constant terms are real numbers in these examples, they will be complex numbers in some applications.

The following equations are not linear

$$\begin{aligned} xz - 3y &= 2 && (2.4) \\ 3x_1^2 + 2x_2^3 &= 4 && (2.5) \\ \sin x_1^3 - 4x_2 + 2^{x_3} &= 0\ . && (2.6) \end{aligned}$$

Observe that the variables occur only to the first power and are multiplied only by constants in linear algebraic equations. ■

Definition 2.3. *A* **system of linear algebraic equations** *(or a* **linear system***) is a finite collection of linear algebraic equations with the same variables. For instance, a system of m linear algebraic equations in the n variables $x_1, x_2, \ldots, x_n$ can be written as*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \ . \end{aligned} \tag{2.7}$$

If the constant terms $b_1, b_2, \ldots, b_m$ are all zero, then the system is called **homogeneous***; otherwise it is called* **inhomogeneous***.*

Definition 2.4. *A* **solution** *to the linear system (2.7) is an ordered list of scalars $(s_1, s_2, \ldots, s_n)$ so that all the equations are satisfied simultaneously when we substitute $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$. The set of all solutions of a linear system is called the* **solution set** *of the system. Two linear systems are called* **equivalent** *if they have the same solution set.*

Example 2.5. Solve the following system of linear algebraic equations

$$\begin{aligned} x - y &= 1 \\ x + y &= 3 \ . \end{aligned} \tag{2.8}$$

Easy, $(2, 1)$ is the *unique solution* of this system. ■

Example 2.6. Solve the following system of linear algebraic equations

$$\begin{aligned} x - y &= 2 \\ 2x - 2y &= 4 \ . \end{aligned} \tag{2.9}$$

The second equation in this system is just twice the first. So the solutions are the solutions of the first equation alone. We can write the solution parametrically as $(t+2, t)$. Thus this system has *infinitely many solutions.* ■

Example 2.7.

$$\begin{aligned} x - y &= 1 \\ x - y &= 3 \ . \end{aligned} \tag{2.10}$$

Two numbers x and y cannot simultaneously have a difference of 1 and 3. Thus, this system has *no solution.* ■

A linear system is said to be **consistent** if it has at least one solution while it is said to be **inconsistent** if it has no solution. We have the following theorem.

Theorem 2.8. A system of linear algebraic equations has either

(a) no solution, or

(b) exactly one solution, or

(c) infinitely many solutions.

■

The essential information of a system of linear algebraic equations can be recorded compactly in a matrix.

Definition 2.9. *For the linear system (2.7), the* **coefficient matrix** *is the matrix*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{2.11}$$

and the **augmented matrix** *is the matrix*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}. \tag{2.12}$$

The augmented matrix of a linear system consists of the coefficient matrix with an extra column containing the constant terms of the equations. It completely characterizes the system since it contains all the coefficients and constant terms of the equations. In addition, if A is the coefficient matrix of the linear system (2.7), then the system can be rewritten as the matrix equation

$$A\mathbf{x} = \mathbf{b} \tag{2.13}$$

where $\mathbf{x}$ and $\mathbf{b}$ are the column matrices

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{2.14}$$

2.2 Row Operations and Echelon Forms

In the next section, we will describe two direct methods for solving a system of linear algebraic equations. The basic ideas is to transform the given linear system into an equivalent system that is easier to solve.

The first step to transform a linear system into an equivalent system is to determine the operations that can be performed on the system without altering its solution set. The following example illustrates such operations for a small linear system.

Example 2.10. We are going to show that the solution set of the following system of linear algebraic equations would not be altered by performing appropriate operations on this system

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 2 \\ 2x_1 - 5x_2 + 3x_3 &= 6 \\ 4x_1 + 6x_2 - 7x_3 &= 8 \ . \end{aligned} \tag{2.15}$$

If we interchange, say, the first and third equations, the resulting system is

$$\begin{aligned} 4x_1 + 6x_2 - 7x_3 &= 8 \\ 2x_1 - 5x_2 + 3x_3 &= 6 \\ x_1 + 2x_2 + 4x_3 &= 2 \ , \end{aligned} \tag{2.16}$$

which certainly has the same solution set as the original system. Returning to the original system, if we multiply, say, the second equation by 5, we obtain the system

$$\begin{aligned} 4x_1 + 6x_2 - 7x_3 &= 8 \\ 10x_1 - 25x_2 + 15x_3 &= 30 \\ x_1 + 2x_2 + 4x_3 &= 2 \ , \end{aligned} \tag{2.17}$$

which again has the same solution set as the original system. Finally, if we add, say, twice the first equation to the third equation, we obtain the system

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 2 \\ 2x_1 - 5x_2 + 3x_3 &= 6 \\ 6x_1 + 10x_2 + x_3 &= 12 \ . \end{aligned} \tag{2.18}$$

We can verify that if the above equations are satisfied, then so the original equations, and vice versa. It follows that the above system has the same solution set as the original system. ■

We can use similar argument to show that the following operations can be performed on any linear system without altering its solution set

1. Interchange two equations.
2. Multiply an equation by a non-zero constant.
3. Add a multiple of one equation to another equation.

Since these operations only involve changes in the coefficients and constant terms of the equations, they can be represented by the following operations on the augmented matrix of the system.

Definition 2.11. *The following* **elementary row operations** *can be performed on a matrix*

1. *Interchange two rows;*
2. *Multiply a row by a non-zero constant;*
3. *Add a multiple of one row to another row.*

It is important to note that each elementary row operation is reversible. For example, if two rows are interchanged, they can be returned to the original positions by another interchange. Besides, elementary row operations can be applied to any matrix, not merely to the augmented matrix of a linear system.

We will use the following shorthand notations for the three elementary row operations on a matrix

1. $R_i \leftrightarrow R_j$ means interchanging the ith and jth rows;
2. kR_i means multiplying the ith row by a nonzero constant k;
3. $R_i + kR_j$ means adding the constant k times the jth row to the ith row.

Example 2.12. Solve the following system of linear algebraic equations by performing appropriate elementary row operations on the augmented matrix of the system

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9\ . \end{aligned} \tag{2.19}$$

The augmented matrix of the given linear system is

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{pmatrix} . \tag{2.20}$$

To find the solution of this system, we eliminate the variables by performing elementary row operations on the augmented matrix of the system

$$\begin{aligned} &\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{pmatrix} \\ \xrightarrow{R_3+4R_1 \to R_3} &\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{pmatrix} \xrightarrow{(1/2)R_2 \to R_2} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{pmatrix} \\ \xrightarrow{R_3+3R_2 \to R_3} &\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix} \xrightarrow{R_2+4R_3 \to R_2} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{pmatrix} \\ \xrightarrow{R_1-R_3 \to R_1} &\begin{pmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{pmatrix} \xrightarrow{R_1+2R_2 \to R_1} \begin{pmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{pmatrix} . \end{aligned} \tag{2.21}$$

It shows that the only solution of the original system is $(29, 16, 3)$. Indeed, we can verify that $(29, 16, 3)$ is a solution by substituting these values into the left side of the original system. ■

We introduce a special term for matrices related via elementary row operations.

Definition 2.13. *Two matrices are called* **row equivalent** *if there is a sequence of elementary row operations that transform one matrix into the other.*

Example 2.14. Show that the matrices A and B are row equivalent for

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix} . \tag{2.22}$$

We perform elementary row operations on A as follows

$$\begin{aligned} &\begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \xrightarrow{R_1+R_2\to R_1} \begin{pmatrix} 3 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix} \\ \xrightarrow{3R_2\to R_2} &\begin{pmatrix} 3 & 0 & 1 \\ 6 & 3 & 3 \end{pmatrix} \xrightarrow{R_2-2R_1\to R_2} \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix} . \end{aligned} \tag{2.23}$$

Therefore, we can see that A is transformed into B through a series of elementary row operations. ■

Suppose a linear system is transformed to a new one via elementary row operations. Any solution of the original system will be also a solution of the new system. Conversely, each solution of the new system is also a solution of the original system since original system can be produced via elementary row operations on the new system. Such observation leads us to the following theorem.

Theorem 2.15. If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set. ■

Our algorithm for solving a system of linear algebraic equations is using elementary row operations to transform the augmented matrix of the system into a simple form. To search for the desired simple form, consider the linear system

$$\begin{aligned} x_1 - x_2 &= -5 \\ x_2 - 2x_3 &= -7 \\ x_3 - 3x_4 &= -2 \\ 2x_4 &= 4 . \end{aligned} \tag{2.24}$$

Starting from the last equation and working backward, we find that $x_4 = 2$, $x_3 = 3(2) - 2 = 4$, $x_2 = 2(4) - 7 = 1$, and $x_1 = 1 - 5 = -4$. Thus the solution to the given system is $(-4, 1, 4, 2)$. Such technique is called *back substitution* and could be used because the system has a simple augmented matrix

$$\begin{pmatrix} 1 & -1 & 0 & 0 & -5 \\ 0 & 1 & -2 & 0 & -7 \\ 0 & 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 2 & 4 \end{pmatrix}, \tag{2.25}$$

which corresponds to an upper-triangular coefficient matrix. Indeed, the back substitution method works on any linear system with an augmented matrix of this form. Unfortunately, it is not always possible to reduce the augmented matrix of a linear system to such form. However, we can always transform the augmented matrix to the row echelon form that can be solved (if it has solution) by back substitution.

Definition 2.16. *A matrix is in* **row echelon form** *if it satisfies the following properties*

1. *Any rows consisting entirely of zeros are at the bottom.*

2. *In each nonzero row, the first nonzero entry, called the* **leading entry***, is in a column to the left of any leading entries below it (some texts add the condition that each leading entry must be 1).*

The followings are examples of matrices in row echelon form

$$\begin{pmatrix} 2 & 4 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & -2 & 3 & 7 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 1 & -1 & 3 \\ 0 & -1 & 1 & 2 & 2 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}. \tag{2.26}$$

The process of applying elementary row operations to reduce a matrix into the row echelon form is called **row reduction**. Note that any non-zero matrix may be row-reduced into more than one row echelon forms using different sequences of elementary row operations. However, the leading entries are always in the same positions in any row echelon form obtained from a given matrix.

Definition 2.17. *A* **pivot position** *of a matrix A is a location of entries of A that corresponds to a leading entry in a row echelon form of A. A* **pivot column** *is a column of a matrix A that contains a pivot position.*

Example 2.18. Use elementary row operations to reduce the following matrix

$$\begin{pmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{pmatrix} \tag{2.27}$$

to row echelon form and locate the pivot columns of this matrix.

We work column by column, from left to right and from top to bottom to row reduce the given matrix to the row echelon form. The strategy is to create a leading entry in a column and then use it to create zeros below it. It is implemented by the following procedures.

1. Locate the leftmost nonzero column and create a leading entry at the top of this column. We don't need to create the leading entry for the given matrix as the top entry of the first column is nonzero.

2. Use the leading entry to put zeros to all entries below it. We do this by performing the following elementary row operations on the given matrix

$$\begin{pmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{pmatrix} \xrightarrow[R_4+R_1\to R_4]{\substack{R_2-2R_1\to R_2 \\ R_3-2R_1\to R_3}} \begin{pmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 3 & -1 & 2 & 10 \end{pmatrix}. \tag{2.28}$$

3. Cover up the row containing the leading entry, and go back to step 1 to repeat the procedure on the remaining matrix. Stop when the entire matrix is in row echelon form. So we perform the elementary row operations on the matrix obtained in step 2 as follows

$$\begin{aligned} & \begin{pmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 3 & -1 & 2 & 10 \end{pmatrix} \\ \xrightarrow{R_2\leftrightarrow R_3} & \begin{pmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & 3 & -1 & 2 & 10 \end{pmatrix} \\ \xrightarrow{\substack{(1/8)R_3\to R_3 \\ R_4+3R_2\to R_4}} & \begin{pmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 29 & 29 & -5 \end{pmatrix} \\ \xrightarrow{R_4-29R_3\to R_4} & \begin{pmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{pmatrix}. \end{aligned} \tag{2.29}$$

This is the row echelon form of the original matrix. From the row echelon form, we can see that columns 1, 2, 3, and 5 are pivot columns. ■

As illustrated in Example 2.18, the general procedure for row reduction to the row echelon form using elementary row operations is to start from the top of the leftmost nonzero column of the matrix and proceed downward and to the right of the matrix.

If two matrices A and B are row equivalent, then further elementary row operations could reduce B (and therefore A) to the same row echelon form. Conversely, if A and B have the same row echelon form R, then we can convert A into R and B into R via elementary row operations. Reversing the latter sequence of operations, we can convert R into B and thus the sequence $A \to R \to B$ achieves the desired effect. Such results lead us to the following theorem.

Theorem 2.19. Two matrices are called row equivalent if and only if they can be reduced to the same row echelon form. ∎

Although the row echelon form of a matrix is not unique, the number of non-zero rows is the same in all row echelon forms of a given matrix. Moreover, this number is fundamental in determining the properties of the solution of linear systems. Thus we give this number a special name.

Definition 2.20. *The number of nonzero rows in any row echelon form of a matrix A is called the* **rank** *of A which is denoted as rank(A).*

Example 2.21. Determine the rank of the matrix A for

$$A = \begin{pmatrix} 3 & 1 & 4 \\ 4 & 3 & 5 \\ 2 & -1 & 3 \end{pmatrix} . \tag{2.30}$$

In order to determine rank(A), we must first reduce A to row echelon form as follows

$$\begin{aligned} &\begin{pmatrix} 3 & 1 & 4 \\ 4 & 3 & 5 \\ 2 & -1 & 3 \end{pmatrix} \xrightarrow{R_1-R_3\to R_1} \begin{pmatrix} 1 & 2 & 1 \\ 4 & 3 & 5 \\ 2 & -1 & 3 \end{pmatrix} \\ \xrightarrow[R_3-2R_1\to R_3]{R_2-4R_1\to R_2} &\begin{pmatrix} 1 & 2 & 1 \\ 0 & -5 & 1 \\ 0 & -5 & 1 \end{pmatrix} \xrightarrow{R_3-R_2\to R_3} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -5 & 1 \\ 0 & 0 & 0 \end{pmatrix} . \end{aligned} \tag{2.31}$$

Since there are two nonzero rows in the row echelon form of A, it follows from definition that rank$(A) = 2$. ∎

As we will see later, it's sometimes preferable to solve a system of linear algebraic equations by further reducing the augmented matrix of the system to a special row echelon form which is defined below.

Definition 2.22. *A matrix is in* **reduced row echelon form** *if it satisfies the following properties*

1. *It is in row echelon form.*
2. *The leading entry in each nonzero row is a 1, called a* **leading 1**.

3. Each leading 1 is the only nonzero entry in its column.

The following is an example

$$\begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.32}$$

Although a matrix does not have a unique row echelon form, we are making a particular choice of row-echelon form in reducing it to reduced row echelon form since all the entries above each leading 1 are arranged to be zeros. Thus the following theorem should not be too surprising.

Theorem 2.23. Every matrix is row equivalent to a unique matrix in reduced row echelon form. ■

Example 2.24. Determine the reduced row echelon form of the following matrix

$$\begin{pmatrix} 3 & -1 & 22 \\ -1 & 5 & 2 \\ 2 & 4 & 24 \end{pmatrix}. \tag{2.33}$$

To obtain reduced row echelon form of the given matrix, we first apply row reduction to reduce it to row echelon form as follows

$$\begin{pmatrix} 3 & -1 & 22 \\ -1 & 5 & 2 \\ 2 & 4 & 24 \end{pmatrix}$$

$$\xrightarrow[3R_3-2R_1\to R_3]{3R_2+R_1\to R_2} \begin{pmatrix} 3 & -1 & 22 \\ 0 & 14 & 28 \\ 0 & 14 & 28 \end{pmatrix} \xrightarrow{R_3-R_2\to R_3} \begin{pmatrix} 3 & -1 & 22 \\ 0 & 14 & 28 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.34}$$

Next, beginning from the leading entry in the lowest nonzero row and working upward and to the left, we create zeros above each leading entry. If the leading entry is not 1, make it 1 by a scaling operation. It is achieved by performing the following elementary row operations on the matrix

$$\begin{pmatrix} 3 & -1 & 22 \\ 0 & 14 & 28 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{14R_1+R_2\to R_1} \begin{pmatrix} 42 & 0 & 336 \\ 0 & 14 & 28 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[(1/14)R_2\to R_2]{(1/42)R_1\to R_1} \begin{pmatrix} 1 & 0 & 8 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.35}$$

This is the reduced row echelon form of the original matrix. ■

2.3 Solving Systems of Equations

In this section, we will first discuss two direct methods for solving a system of linear algebraic equations. It is based on the idea of reducing the augmented matrix of the given system to the row echelon or reduced row echelon form using elementary row operations.

The method of reducing the augmented matrix to row echelon form and then using back substitution to solve the equivalent system is called **Gaussian elimination**. The following procedure outlines how to solve a linear system using this method.

Definition 2.25. Gaussian elimination

1. *Write the augmented matrix of the system of linear algebraic equations.*
2. *Use elementary row operations to reduce the augmented matrix to row echelon form.*
3. *If the resulting system is consistent, solve the equivalent system that corresponds to the row-reduced matrix using back substitution.*

Example 2.26. Determine the solution set to the linear system

$$\begin{aligned} 3x_1 - 2x_2 + 2x_3 &= 9 \\ x_1 - 2x_2 + x_3 &= 5 \\ 2x_1 - x_2 - 2x_3 &= -1 \ . \end{aligned} \tag{2.36}$$

The augmented matrix of the linear system is

$$\begin{pmatrix} 3 & -2 & 2 & 9 \\ 1 & -2 & 1 & 5 \\ 2 & -1 & -2 & -1 \end{pmatrix} . \tag{2.37}$$

We proceed to reduce this matrix to row echelon form using elementary row operations

$$\begin{aligned} &\begin{pmatrix} 3 & -2 & 2 & 9 \\ 1 & -2 & 1 & 5 \\ 2 & -1 & -2 & -1 \end{pmatrix} \\ \xrightarrow{R_1 \leftrightarrow R_2} &\begin{pmatrix} 1 & -2 & 1 & 5 \\ 3 & -2 & 2 & 9 \\ 2 & -1 & -2 & -1 \end{pmatrix} \xrightarrow[R_3 - 2R_1 \to R_3]{R_2 - 3R_1 \to R_2} \begin{pmatrix} 1 & -2 & 1 & 5 \\ 0 & 4 & -1 & -6 \\ 0 & 3 & -4 & -11 \end{pmatrix} \\ \xrightarrow{R_2 - R_3 \to R_2} &\begin{pmatrix} 1 & -2 & 1 & 5 \\ 0 & 1 & 3 & 5 \\ 0 & 3 & -4 & -11 \end{pmatrix} \xrightarrow{R_3 - 3R_2 \to R_3} \begin{pmatrix} 1 & -2 & 1 & 5 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & -13 & -26 \end{pmatrix} \\ \xrightarrow{\frac{-1}{13} R_3 \to R_3} &\begin{pmatrix} 1 & -2 & 1 & 5 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{pmatrix} . \end{aligned} \tag{2.38}$$

Note that here we made each leading entry to be a leading 1 for convenient. The linear system corresponding to this row echelon form of the augmented matrix is

$$
\begin{aligned}
x_1 - 2x_2 + x_3 &= 5 \\
x_2 + 3x_3 &= 5 \\
x_3 &= 2 \, ,
\end{aligned} \tag{2.39}
$$

and back substitution gives a unique solution $(1, -1, 2)$. ∎

Example 2.27. Determine the solution set to the linear system

$$
\begin{aligned}
x_1 - x_2 - x_3 + 2x_4 &= 1 \\
2x_1 - 2x_2 - x_3 + 3x_4 &= 3 \\
-x_1 + x_2 - x_3 &= -3 \, .
\end{aligned} \tag{2.40}
$$

The augmented matrix of the linear system is

$$
\begin{pmatrix} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{pmatrix} \, . \tag{2.41}
$$

We proceed to reduce this matrix to row echelon form using elementary row operations

$$
\begin{aligned}
& \begin{pmatrix} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{pmatrix} \\
\xrightarrow[R_3+R_1 \to R_3]{R_2-2R_1 \to R_2} & \begin{pmatrix} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & -2 \end{pmatrix} \\
\xrightarrow{R_3+2R_2 \to R_3} & \begin{pmatrix} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \, .
\end{aligned} \tag{2.42}
$$

The linear system corresponding to this row echelon form of the augmented matrix is

$$
\begin{aligned}
x_1 - x_2 - x_3 + 2x_4 &= 1 \\
x_3 - x_4 &= 1 \, ,
\end{aligned} \tag{2.43}
$$

which has infinitely many solutions. The variables x_1 and x_3 corresponding to the leading entries in the matrix are called the **leading variables** and the other variables x_2 and x_4 are called **free variables**. If we set $x_2 = s$ and $x_4 = t$ where s and t can assume any real value, then back substitution gives

$$
\begin{aligned}
x_3 &= t + 1 \, , \\
x_1 &= s - t + 2 \, .
\end{aligned} \tag{2.44}
$$

Thus the solution set to our original system is the following subset of $\mathbb{R}^3$

$$
S = \{(s - t + 2, s, t + 1, t) : s, t \in \mathbb{R})\} \, . \tag{2.45}
$$

∎

Example 2.28. Determine the solution set to the linear system

$$\begin{aligned} x_1 + x_2 - x_3 + x_4 &= 1 \\ 2x_1 + 3x_2 + x_3 &= 4 \\ 3x_1 + 5x_2 + 3x_3 - x_4 &= 5 \; . \end{aligned} \tag{2.46}$$

The augmented matrix of the linear system is

$$\begin{pmatrix} 1 & 1 & -1 & 1 & 1 \\ 2 & 3 & 1 & 0 & 4 \\ 3 & 5 & 3 & -1 & 5 \end{pmatrix} . \tag{2.47}$$

We proceed to reduce this matrix to row echelon form using elementary row operations

$$\begin{aligned} \begin{pmatrix} 1 & 1 & -1 & 1 & 1 \\ 2 & 3 & 1 & 0 & 4 \\ 3 & 5 & 3 & -1 & 5 \end{pmatrix} &\xrightarrow[R_3-3R_1\to R_3]{R_2-2R_1\to R_2} \begin{pmatrix} 1 & 1 & -1 & 1 & 1 \\ 0 & 1 & 3 & -2 & 2 \\ 0 & 2 & 6 & -4 & 2 \end{pmatrix} \\ &\xrightarrow{R_3-2R_2\to R_3} \begin{pmatrix} 1 & 1 & -1 & 1 & 1 \\ 0 & 1 & 3 & -2 & 2 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix} \end{aligned} \tag{2.48}$$

The last row requires

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -2 \tag{2.49}$$

which is impossible. It tells us that the linear system has no solution, It is inconsistent. Thus the solution set to the system is the empty set. ■

A modification of Gaussian elimination does not require back substitution and it is particularly helpful for solving a linear system with infinitely many solution. This variant is known as **Gauss-Jordan elimination** which involves reducing the augmented matrix to reduced row echelon form. The following procedure outlines how to solve a linear system using this method.

Definition 2.29. Gauss-Jordan elimination

1. *Write the augmented matrix of the system of linear algebraic equations.*
2. *Use elementary row operations to reduce the augmented matrix to reduced row echelon form.*
3. *If the resulting system is consistent, solve for the leading variable in terms of any free variables.*

Example 2.30. Solve the linear system in Example 2.27 by Gauss-Jordan elimination.

The row reduction proceeds as above until we reach the row echelon form

$$\begin{pmatrix} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} . \tag{2.50}$$

We must now put a zero above the leading 1 in the second row, third column. We do this by adding the first row to the second row to obtain

$$\begin{pmatrix} 1 & -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} . \tag{2.51}$$

The linear system corresponding to this reduced row echelon form of the augmented matrix is

$$\begin{aligned} x_1 - x_2 + x_4 &= 2 \\ x_3 - x_4 &= 1 . \end{aligned} \tag{2.52}$$

It is now much easier to solve for the leading variables $x_1 = x_2 - x_4 + 2$ and $x_3 = x_4 + 1$. If we set $x_2 = s$ and $x_4 = t$ as before, then the solution set to the system is the following subset of $\mathbb{R}^3$

$$S = \{(s - t + 2, s, t + 1, t) : s, t \in \mathbb{R})\} . \tag{2.53}$$

■

The Gauss-Jordan Elimination can be also used to compute the inverse of an invertible $n \times n$ matrix A. For such purpose, we construct an augmented matrix for A as

$$\begin{pmatrix} A & I_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{pmatrix} . \tag{2.54}$$

We define the **elementary matrix** E as one that is obtained by performing a single elementary row operation on an identity matrix. Since elementary row operations are reversible, elementary matrices are invertible. If an elementary row operation is performed on an $n \times m$ matrix B, then the resulting matrix can be written as EB where the $n \times n$ elementary matrix E is created by performing the same elementary row operation on the $n \times n$ identity matrix I_n. Thus, any matrix can be reduced to row echelon form by multiplication with a sequence of elementary matrices. In addition, since the unique reduced row echelon form of an invertible $n \times n$ matrix is the identity matrix I_n, it follows that the matrix A can be reduced to I_n by a sequence of elementary row operations. In other words, there exist elementary matrices

$E_1, E_2, \ldots, E_{k-1}, E_k$ such that

$$(E_k E_{k-1} \ldots E_2 E_1)A = I_n \ . \tag{2.55}$$

The sequence of elementary row operations that reduces A to I_n also transforms the augmented matrix $\begin{pmatrix} A & I_n \end{pmatrix}$ to another matrix

$$\begin{pmatrix} I_n & B \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1n} \\ 0 & 1 & \ldots & 0 & b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & b_{n1} & b_{n2} & \ldots & b_{mn} \end{pmatrix} \tag{2.56}$$

where the matrix B can be written as

$$B = (E_k E_{k-1} \ldots E_2 E_1) I_n = E_k E_{k-1} \cdots E_2 E_1 \ . \tag{2.57}$$

Besides, since the matrix product $E_k E_{k-1} \cdots E_2 E_1$ is invertible, Eq. (2.55) leads to

$$\begin{aligned} (E_k E_{k-1} \cdots E_2 E_1)^{-1} (E_k E_{k-1} \cdots E_2 E_1) A &= (E_k E_{k-1} \cdots E_2 E_1)^{-1} I_n \\ A = (E_k E_{k-1} \cdots E_2 E_1)^{-1} &= E_1^{-1} E_2^{-1} \cdots E_{k-1}^{-1} E_k^{-1} \end{aligned} \tag{2.58}$$

which is a product of elementary matrix. Hence, from Eq. (2.57) and (2.58), we obtain

$$AB = (E_1^{-1} E_2^{-1} \cdots E_{k-1}^{-1} E_k^{-1})(E_k E_{k-1} \cdots E_2 E_1) = I_n \ . \tag{2.59}$$

It follows that the matrix B is the inverse of A. As a result, we can compute the inverse of an invertible $n \times n$ matrix by reducing the augmented matrix $\begin{pmatrix} A & I_n \end{pmatrix}$ to the matrix $\begin{pmatrix} I_n & A^{-1} \end{pmatrix}$ using elementary row operations.

Example 2.31. Find the inverse of the invertible matrix

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{pmatrix} \ . \tag{2.60}$$

We perform row reduction on the augmented matrix $\begin{pmatrix} A & I_n \end{pmatrix}$ to the matrix $\begin{pmatrix} I_n & A^{-1} \end{pmatrix}$ using elementary row operations as follows.

$$\begin{aligned}
&\begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{pmatrix} \\
\xrightarrow{R_1 \leftrightarrow R_2} &\begin{pmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{pmatrix} \\
\xrightarrow{R_3 - 4R_1 \to R_3} &\begin{pmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{pmatrix} \\
\xrightarrow{R_3 + 3R_2 \to R_3} &\begin{pmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{pmatrix} \\
\xrightarrow{R_3/2 \to R_3} &\begin{pmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{pmatrix} \\
\xrightarrow[R_2 - 2R_3 \to R_2]{R_1 - 3R_3 \to R_1} &\begin{pmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{pmatrix} .
\end{aligned} \tag{2.61}$$

So the inverse of the matrix A is

$$A^{-1} = \begin{pmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{pmatrix} . \tag{2.62}$$

We can check the final answer by verifying that $AA^{-1} = I$. ■

To determine the nature of the solution set of a system of linear algebraic equations, we need to consider two fundamental questions. First, is the system consistent, does at least one solution exists? Moreover, if the system is consistent, does it has one or infinitely many solutions? We can seek the answers of these questions by studying the (non-reduced) row echelon form of the augmented matrix.

If the row echelon form of the augmented matrix contains no equations of the form $0 = b$ where b is a non-zero constant, every nonzero equation contains a leading variable with a non-zero coefficient. Then either the leading variables are completely determined (with no free variables) or at least one of the leading variables can be expressed in terms of one or more free variables. For the former case, there is a unique solution; for the latter case, there are infinitely many solutions. This gives justification to the next theorem.

Theorem 2.32. Existence and Uniqueness Theorem
A system of linear algebraic equations is consistent if and only if the rightmost column of the augmented matrix is *not* a pivot column, i. e. if and only if

any row echelon form of the augmented matrix has *no* row of the form

$$\begin{pmatrix} 0 & \dots & 0 & b \end{pmatrix} \tag{2.63}$$

with $b \neq 0$. For a consistent linear system, the solution set contains either a unique solution if there are no free variables, or infinitely many solutions if there is at least one free variable. ■

Recall that a homogeneous system of linear algebraic equations can be written in the matrix form $A\mathbf{x} = \mathbf{0}$ where A is an $m \times n$ coefficient matrix and $\mathbf{0}$ is the $n \times 1$ zero matrix. Such system *always* has at least one solution $\mathbf{x} = \mathbf{0}$ which is known as the **trivial solution**. However, how can we determine whether the system has a **nontrivial solution**, a nonzero vector $\mathbf{x}$ that satisfies $A\mathbf{x} = \mathbf{0}$? Theorem 2.32 tells us that *the homogeneous linear system $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if the equation has at least one free variable.*

On the other hand, we can also solve a system of linear algebraic equations $A\mathbf{x} = \mathbf{b}$ via $\mathbf{x} = A^{-1}\mathbf{b}$ provided that the coefficient matrix A is square and invertible. However, such method is seldom used since it's almost always faster to use Gaussian or Gauss-Jordan elimination to find the solution directly. One possible exception is systems with 2×2 coefficient matrices which is illustrated in the following example.

Example 2.33. Use the inverse of the coefficient matrix to solve the linear system

$$\begin{aligned} 3x_1 + 4x_2 &= 3 \\ 5x_1 + 6x_2 &= 7\,. \end{aligned} \tag{2.64}$$

The given system is equivalent to $A\mathbf{x} = \mathbf{b}$ where the coefficient matrix A and the column vector $\mathbf{b}$ are

$$A = \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} \tag{2.65}$$

and

$$\mathbf{b} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}, \tag{2.66}$$

respectively. We can show that the inverse of A is

$$A^{-1} = \begin{pmatrix} -3 & 2 \\ 5/2 & -3/2 \end{pmatrix}. \tag{2.67}$$

Thus the solution of the given linear system is

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{pmatrix} -3 & 2 \\ 5/2 & -3/2 \end{pmatrix} \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}. \tag{2.68}$$

■

A linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x} = A^{-1}\mathbf{b}$ for any $n \times 1$ column matrix $\mathbf{b}$ if the coefficient matrix A is an invertible $n \times n$ matrix. For such case, we can use **Cramer's Rule** to express the unique solution directly in terms of determinants.

Theorem 2.34. Cramer's Rule
Let A be an invertible $n \times n$ matrix. For any $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \qquad i = 1, 2, \ldots, n \tag{2.69}$$

where $A_i(\mathbf{b})$ is the matrix obtained from $A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n \end{pmatrix}$ by replacing column i by the vector $\mathbf{b}$

$$A_i(\mathbf{b}) = \begin{pmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \ldots & \mathbf{a}_n \end{pmatrix} . \tag{2.70}$$

Proof. We denote the columns of the A by $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and the columns of the $n \times n$ identity matrix I by $\mathbf{e}_1, \ldots, \mathbf{e}_n$. If $A\mathbf{x} = \mathbf{b}$, then

$$\begin{aligned} A\,I_i(\mathbf{x}) &= A\begin{pmatrix} \mathbf{e}_1 & \ldots & \mathbf{e}_{i-1} & \mathbf{x} & \mathbf{e}_{i+1} & \ldots & \mathbf{e}_n \end{pmatrix} \\ &= \begin{pmatrix} A\mathbf{e}_1 & \ldots & A\mathbf{e}_{i-1} & A\mathbf{x} & A\mathbf{e}_{i+1} & \ldots & A\mathbf{e}_n \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \ldots & \mathbf{a}_n \end{pmatrix} \\ &= A_i(\mathbf{b}) . \end{aligned} \tag{2.71}$$

Therefore, by Theorem 1.28(h), we have

$$(\det A)(\det I_i(\mathbf{x})) = \det(AI_i(\mathbf{x})) = \det A_i(\mathbf{b}) . \tag{2.72}$$

Making a cofactor expansion along the ith row yields $\det I_i(\mathbf{x}) = x_i$ which implies that

$$\begin{aligned} (\det A)x_i &= \det A_i(\mathbf{b}) \\ x_i &= \frac{\det A_i(\mathbf{b})}{\det A} . \end{aligned} \tag{2.73}$$

∎

So far we have only considered those linear systems with real coefficients and their corresponding real solutions. Nevertheless, the techniques that we have discussed for solving linear systems are also applicable when the linear system has complex coefficients so that the corresponding solutions are also complex.

There are numerous applications of system of linear algebraic equations. Here we will discuss one to illustrate.

Current flow in a simple electrical network can be also described by a system of linear algebraic equations. Electrical networks are a specialized type of network which provides information about power sources and devices powered by these sources. They are governed by a pair of laws called Kirchhoff's laws

- Kirchhoff's current law: The sum of the currents flowing into any point equals to the sum of the currents flowing out from this point.
- Kirchhoff's voltage law: The algebraic sum of the voltage drops around any closed loop is zero.

In an electrical network with multiple loops and power sources, we usually cannot tell in advance the direction of current flow. So we choose arbitrary direction for the current flow in each branch. In addition, Kirchhoff's voltage law requires us to choose arbitrary direction for the current flow in each closed loop known as loop current. We also adopt the following sign conventions for the calculation of the currents

- If the current passing through a resistor is in the same direction as the loop current, a voltage drop occurs at the resistor; otherwise, a voltage rise occurs at the resistor.
- If the loop current flows from the positive terminal to the negative terminal of a power source, a voltage drop occurs at the power source; otherwise, a voltage rise occurs at the power source.

Using Kirchoff's law, we can then set up a system of linear algebraic equations that allows us to determine the currents in an electrical network. If a current turns out to be negative, the actual direction of current flow is opposite to the chosen direction. The next example illustrates how to determine the currents by solving the linear system.

Example 2.35. Fig. 2.1 shows an electric network composed of three power sources and seven resistors. Determine the loop currents I_1, I_2 and I_3 in the network.

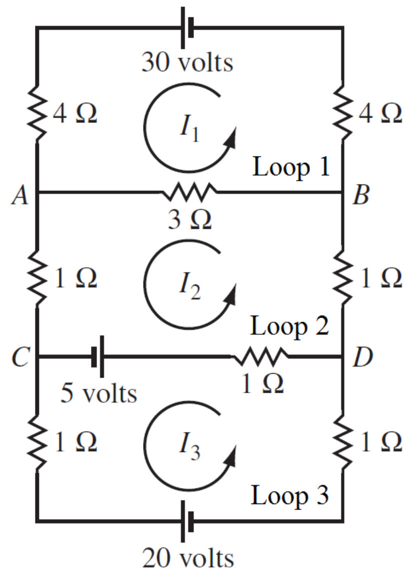


Figure 2.1: A simple electric network.

Applying Kirchoff's voltage law to the three current loops, we obtain for Loop 1, 2 and 3

$$\begin{aligned} 11I_1 - 3I_2 &= 30 \\ -3I_1 + 6I_2 - I_3 &= 5 \\ -I_2 + 3I_3 &= -25 \end{aligned}$$

We use Gauss-Jordan elimination to row reduce the corresponding augmented matrix

$$\begin{aligned}
&\begin{pmatrix} 11 & -3 & 0 & 30 \\ -3 & 6 & -1 & 5 \\ 0 & -1 & 3 & -25 \end{pmatrix} \\
\xrightarrow{11R_2+3R_1\to R_2} &\begin{pmatrix} 11 & -3 & 0 & 30 \\ 0 & 57 & -11 & 145 \\ 0 & -1 & 3 & -25 \end{pmatrix} \\
\xrightarrow[57R_3+R_2\to R_3]{57R_1+3R_2\to R_1} &\begin{pmatrix} 627 & 0 & -33 & 2145 \\ 0 & 57 & -11 & 145 \\ 0 & 0 & 160 & -1280 \end{pmatrix} \\
\xrightarrow[R_3/160\to R_3]{R_1/33\to R_1} &\begin{pmatrix} 19 & 0 & -1 & 65 \\ 0 & 57 & -11 & 145 \\ 0 & 0 & 1 & -8 \end{pmatrix} \\
\xrightarrow[R_2+11R_3\to R_2]{R_1+R_3\to R_1} &\begin{pmatrix} 19 & 0 & 0 & 57 \\ 0 & 57 & 0 & 57 \\ 0 & 0 & 1 & -8 \end{pmatrix} \\
\xrightarrow[R_2/57\to R_2]{R_1/19\to R_1} &\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -8 \end{pmatrix} .
\end{aligned} \tag{2.74}$$

Hence, $I_1 = 3\,\mathrm{A}$, $I_2 = 1\,\mathrm{A}$, and $I_3 = -8\,\mathrm{A}$. The negative value of I_3 indicates that the actual current in Loop 3 flows in the direction opposite to that shown in the figure. ∎

Chapter 3

Vector Spaces

Vector spaces have many applications in physics. The abstract definition is necessary if we want to understand the nature without using coordinates. After all, nature does not need coordinate systems.

3.1 Vector Spaces

Definition 3.1. *A* **vector space** V *(over real numbers* $\mathbb{R}$*) is a non-empty set, in which elements are called* **vectors***, together with two operations, called* **scalar multiplication** $\cdot : \mathbb{R} \times V \to V$ *and* **vector addition** $+ : V \times V \to V$ *satisfying*

1. *For all* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$*;*
2. *For all* $\mathbf{u}, \mathbf{v} \in V, \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$*;*
3. *There exists an element, zero vector,* $\mathbf{0} \in V$ *such that for all* $\mathbf{v} \in V, \mathbf{0} + \mathbf{v} = \mathbf{v}$*;*
4. *For all* $\mathbf{v} \in V$*, there exists an* $\mathbf{u} \in V$ *such that* $\mathbf{v} + \mathbf{u} = \mathbf{0}$*. We denote this* $\mathbf{u}$ *by* $-\mathbf{v}$*, called its negative;*
5. *For all* $a \in \mathbb{R}, \mathbf{u}, \mathbf{v} \in V, a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$*;*
6. *For all* $a, b \in \mathbb{R}, \mathbf{v} \in V, (a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$*;*
7. *For all* $a, b \in \mathbb{R}, \mathbf{v} \in V, a(b\mathbf{u}) = (ab)\mathbf{u}$*;*
8. *For* $1 \in \mathbb{R}$*, for all* $\mathbf{v} \in V, 1\mathbf{v} = \mathbf{v}$*.*

Example 3.2. The first non-trivial example is the plane $\mathbb{R}^2 \equiv \mathbb{R} \times \mathbb{R}$. This is the set of all ordered pairs of two real numbers, almost always interpreted as the x- and y-coordinates of a point on the plane relative to some pre-chosen coordinate system.

This could be generalized to any positive integer n, with the case $n = 3$ being the three dimensional space. Similarly, elements of $\mathbb{R}^n$ is the n-tuple of real numbers, $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$. Scalar multiplication is defined by

$$a\mathbf{v} = a(v_1, v_2, \ldots, v_n) \equiv (av_1, av_2, \ldots, av_n) \tag{3.1}$$

where, for example, av_1 in the right hand side is just multiplication of two real numbers. Vector addition is

$$(u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \ . \tag{3.2}$$

The zero vector is $(0, 0, \ldots, 0)$ and its negative is $(-v_1, -v_2, \ldots, -v_n)$. Students should check that $\mathbb{R}^n$ satisfies all the requirement in Definition 3.1. ■

Example 3.3. Consider the differential equation

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + y = 0 \ . \tag{3.3}$$

The solutions form a vector space. Explicitly, if $y_a = a_1 \cos x + a_2 \sin x$ and $y_b = b_1 \cos x + b_2 \sin x$ are two solutions of the differential equation, then their sum

$$y_a + y_b = (a_1 + b_1) \cos x + (a_2 + b_2) \sin x \tag{3.4}$$

is also a solution.

The solutions of the inhomogeneous differential equation

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + y = 1 \ . \tag{3.5}$$

do *not* form a vector space. The sum of two non-zero solutions is not a solution anymore. ■

Example 3.4. For any fixed n and m, the set of all $n \times m$ matrices, with the obvious operations, forms a vector space. However, the set of all matrices does *not* form a vector space because the addition of two matrices with different sizes is not defined. Note that matrix multiplication is not considered here. We could say that matrices have more structure than a vector space. ■

3.2 Subspaces

Definition 3.5. *A* **vector subspace** *or just* **subspace**, *W, of a vector space V is a subset such that*

1. *The zero vector, $\mathbf{0}$, is in W.*
2. *W is closed under addition: If $\mathbf{u}$ and $\mathbf{v}$ are in W, $\mathbf{u} + \mathbf{v}$ is in W.*
3. *W is closed under scalar multiplication: If $a \in \mathbb{R}$ and $\mathbf{v} \in W$, then $a\mathbf{v} \in W$.*

Example 3.6. The set that contains only one element, the zero vector $\mathbf{0}$, is a subspace. The plane, $\mathbb{R}^2$, is often interpreted as the subspace $z = 0$ of $\mathbb{R}^3$. ■

Example 3.7. For $n \geq 0$, the set, P_n, of all polynomials of degree at most n is a vector space,

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \; . \tag{3.6}$$

This could be understood as the solution space of the differential equation

$$\frac{\mathrm{d}^{n+1} p}{\mathrm{d}x^{n+1}} = 0 \; . \tag{3.7}$$

If $n \leq m$, P_n is a subspace of P_m. The set of all polynomials of degree exactly n ($a_n \neq 0$) is not a vector space because the zero polynomial, $\mathbf{0}$, is not one of them. (There are many other reasons.) ■

Example 3.8. For a vector $\mathbf{v}$ in a vector space, the set of all vectors of the form $a\mathbf{v}$ with real numbers a forms a subspace. For any two vectors $\mathbf{v}$ and $\mathbf{w}$, the set of all vectors of the form $a\mathbf{v} + b\mathbf{w}$ with real numbers a and b also forms a subspace. ■

Definition 3.9. *For vectors* $\mathbf{v}_1$, *...,* $\mathbf{v}_k$ *in a vector space, the* **subspace spanned** *by them, Span*$(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *is the set of all of their linear combinations*

$$a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k \; . \tag{3.8}$$

Example 3.10. Let $\mathbf{v}_1 \equiv (0, 0, 0)$, $\mathbf{v}_2 \equiv (1, 0, 0)$, $\mathbf{v}_3 \equiv (-1, 0, 0)$, $\mathbf{v}_4 \equiv (0, 1, 0)$ and $\mathbf{v}_5 \equiv (1, 1, 0)$ in $\mathbb{R}^3$. Then, $\mathrm{Span}(\mathbf{v}_1) = \{(0, 0, 0)\}$, $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2) = \{(a, 0, 0) : a \in \mathbb{R}\}$, $\mathrm{Span}(\mathbf{v}_2, \mathbf{v}_3) = \{(a, 0, 0) : a \in \mathbb{R}\}$, $\mathrm{Span}(\mathbf{v}_2, \mathbf{v}_4) = \{(a, b, 0) : a, b \in \mathbb{R}\}$ and $\mathrm{Span}(\mathbf{v}_2, \mathbf{v}_5) = \{(a, b, 0) : a, b \in \mathbb{R}\}$. ■

Definition 3.11. *A* **linear transformation**, *T, from a vector space V to a vector space W is a map that assigns to each vector $\mathbf{v}$ of V to a vector $T(\mathbf{v})$ of W such that*

1. *for all $\mathbf{u}$ and $\mathbf{v}$ in V, $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$. Note that the sum in the right hand side is the sum in W.*
2. *for all a in $\mathbb{R}$ and $\mathbf{v}$ in V, $T(a\mathbf{v}) = a\,T(\mathbf{v})$.*

Note that W could be equal to V.

Example 3.12. If the two coordinate systems are related by a rotation, Fig. 3.1, what will be relation between the coordinates? Let the coordinates

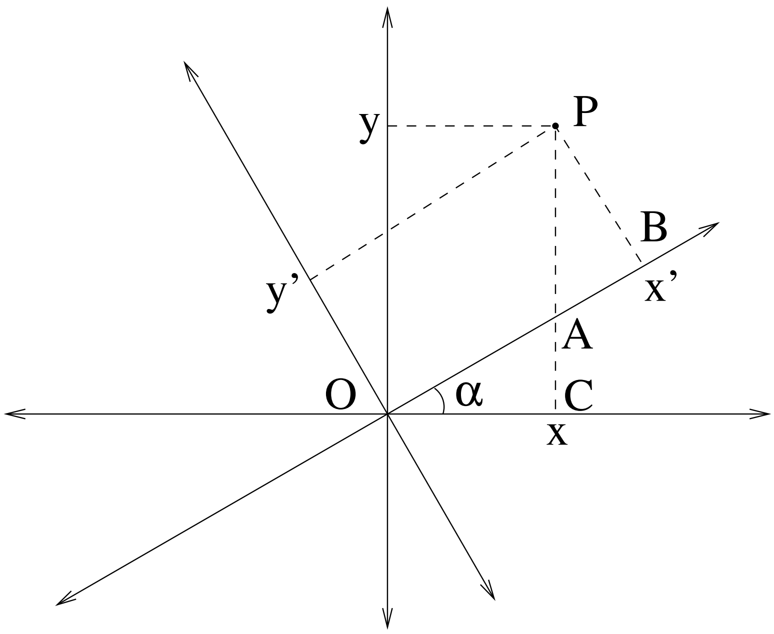


Figure 3.1: Two coordinate systems related by a rotation.

of the same point be (x, y) and (x', y') in the two systems. x' is the sum of the length of OA and AB, we have

$$\begin{aligned} x' &= OA + AB \\ &= \frac{x}{\cos\alpha} + (PA)\sin\alpha \\ &= \frac{x}{\cos\alpha} + (y - AC)\sin\alpha \\ &= \frac{x}{\cos\alpha} + (y - \frac{x}{\cos\alpha}\sin\alpha)\sin\alpha \\ &= \frac{x}{\cos\alpha}(1 - \sin^2\alpha) + y\sin\alpha \\ &= x\cos\alpha + y\sin\alpha\ . \end{aligned} \tag{3.9}$$

Similarly,

$$\begin{aligned} y' &= PB \\ &= (PA)\cos\alpha \\ &= (y - \frac{x}{\cos\alpha}\sin\alpha)\cos\alpha \\ &= -x\sin\alpha + y\cos\alpha\ . \end{aligned} \tag{3.10}$$

In matrix notation, these are

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\ . \tag{3.11}$$

We see that rotation is a linear transformation. ■

Example 3.13. In Example 3.7, we saw that P_n is the solution space of the differential equation

$$\frac{\mathrm{d}^{n+1}p}{\mathrm{d}x^{n+1}} = 0 \ . \tag{3.12}$$

We claim that the differentiation $T \equiv \mathrm{d}/\mathrm{d}x$ is a linear transformation from P_n to P_n.

The differentiation of a polynomial of degree at most n is another polynomial of degree also at most n. Hence, T maps from P_n to P_n. The two conditions in the definition are easy to check. (Yes, after differentiation, the degree is at most $n-1$. We could also say that T maps from P_n to P_{n-1}.) ■

Definition 3.14. *The* **kernel** *of a linear transformation T from V to W is the set of all vectors $\mathbf{v}$ in V such that $T(\mathbf{v}) = \mathbf{0}$, the zero vector in W. We can easily see that the kernel is a subspace of V.*

Example 3.15. In Example 3.12, the kernel of rotation is zero, which means only the zero vector is mapped to zero vector.

In Example 3.13, the kernel of differentiation is all those constant polynomials (polynomials with only constant term). ■

3.3 Linearly Independence and Bases

Definition 3.16. *Given m vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. They are* **linearly independent** *if the solution of the equation*

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m = 0 \tag{3.13}$$

is only $a_1 = a_2 = \cdots = a_m = 0$.

Example 3.17. In $\mathbb{R}^2$, $(1,0)$ and $(0,1)$ are linearly independent. A single vector $(1,0)$ is linearly independent. $(1,0)$, $(0,1)$ and $(1,1)$ are *not* linearly independent. $(1,0)$ and $(2,0)$ are not linearly independent. A single zero vector $(0,0)$ is not linearly independent. ■

Definition 3.18. *A maximum set of linearly independent vectors is a* **basis** *of the vector space. We mean that $B = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis if*

1. *$\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ are linearly independent;*
2. *if $\mathbf{v} \in V$ is any vector, $\mathbf{v}, \mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ are linearly dependent.*

Theorem 3.19. All bases have the same number of the vectors. This is the **dimension** of the vector space. ■

Example 3.20. The dimension of $\mathbb{R}^n$ is n. For $\mathbb{R}^2$, $\{\mathbf{u}_1 \equiv (1,0), \mathbf{u}_2 \equiv (0,1)\}$ is a basis, $\{\mathbf{v}_1 \equiv (1,0), \mathbf{v}_2 \equiv (1,-1)\}$ is also a basis and $\{(2,0),(0,2)\}$ is yet another basis.

The dimension of the vector space of solutions in Example 3.3 is 2.

The space of all continuous functions from real numbers to real numbers $f : \mathbb{R} \to \mathbb{R}$ is also a vector space. (Check this.) Its dimension is infinite.

The dimension of the set of all polynomials of degree at most n is $n+1$ because there are $n+1$ coefficients. ■

Theorem 3.21. If V is a vector space of dimension n and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form a basis, then any vector $\mathbf{v} \in V$ can be expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \ . \tag{3.14}$$

Proof. The vectors $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly dependent, we have for some real numbers $b, b_1, b_2, \ldots, b_n$

$$b\mathbf{v} + b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n = 0 \ . \tag{3.15}$$

$b \neq 0$, otherwise the above equation shows that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linear dependent. Hence, take $a_\mu = -b_\mu/b$, $\mu = 1, \ldots, n$, we have the required result. ■

Definition 3.22. *The numbers $a_1, \ldots, a_n$ in Eq. (3.14) are the* **coordinates** *or* **components** *of the vector $\mathbf{v}$ relative to the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$.*

Example 3.23. For $\mathbb{R}^n$, a standard choice of basis is

$$\begin{array}{ll} \mathbf{e}_1 & = (1, 0, \ldots, 0) \\ \mathbf{e}_2 & = (0, 1, \ldots, 0) \\ \vdots & \\ \mathbf{e}_n & = (0, 0, \ldots, 1) \end{array} \ . \tag{3.16}$$

Then, a vector $\mathbf{v} = (a_1, \ldots, a_n)$ has coordinates relative to the standard basis a_1, ..., a_n. ■

In terms of coordinates, we can represent a linear transformation by a matrix. Suppose $\mathbf{v}_i$ and $\mathbf{w}_j$ are bases of vector spaces V and W respectively, and T is a linear transformation from V to W. Then, $T(\mathbf{v}_i)$ is a vector of W and can be expressed as a linear combination

$$T(\mathbf{v}_i) = \sum_j T_{ji}\mathbf{w}_j \ . \tag{3.17}$$

For a general vector $\mathbf{v} \equiv \sum a_i\mathbf{v}_i$ of V, we have

$$T(\mathbf{v}) = T(\sum_i a_i\mathbf{v}_i) = \sum_i a_iT(\mathbf{v}_i) = \sum_{ij} a_iT_{ji}\mathbf{w}_j \ . = \sum_j \left(\sum_i T_{ji}a_i\right)\mathbf{w}_j \ . \tag{3.18}$$

This means that if a vector $\mathbf{v}$ in V has coordinates (a_i) relative to some basis, then $T(\mathbf{v})$ has coordinates $(\sum_i T_{ji}a_i)$ in W, where T_{ji} are numbers defined in Eq. (3.17).

What will be the coordinates of a vector in the kernel of the linear transformation? From the above paragraph, it is in the kernel if and only if $\sum_i T_{ji}a_i = 0$ for all j, which means that a_i satisfy the system of linear equations

$$\begin{aligned} T_{11}a_1 + T_{12}a_2 + \cdots + T_{1n}a_n &= 0 \\ T_{21}a_1 + T_{22}a_2 + \cdots + T_{2n}a_n &= 0 \\ &\vdots \\ T_{n1}a_1 + T_{n2}a_2 + \cdots + T_{nn}a_n &= 0 \; . \end{aligned} \tag{3.19}$$

We have talked about this in the last chapter.

Example 3.24. Let us work out the matrix representation of differentiation in Example 3.13. For simplicity, we consider only polynomials of degree at most 3. They have the form

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \; . \tag{3.20}$$

We could write it as a column matrix

$$p(x) \sim \begin{pmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} \; . \tag{3.21}$$

These are the coordinates of $p(x)$ in terms of the basis $(x^3, x^2, x, 1)$. We know $p'(x) = 3a_3x^2 + 2a_2x + a_1$. Hence, the matrix representation of differentiation in terms of this basis is

$$\frac{\mathrm{d}}{\mathrm{d}x} \sim \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \; . \tag{3.22}$$

To find out the kernel, we have to solve the matrix equation

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} = 0 \; . \tag{3.23}$$

This is equivalent to $3a_3 = 0$, $2a_2 = 0$, $a_1 = 0$ and no restriction on a_0. We are back to the well known fact that polynomials in the kernel are constants. ■

What happens if we choose another basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$? $\mathbf{v}_i$ can be expressed as linear combinations of $\mathbf{u}_j$,

$$\mathbf{v}_i = \sum_{j=1}^{n} L_{ji}\mathbf{u}_j \tag{3.24}$$

for some real numbers L_{ji}. For a vector $\mathbf{v} \in V$, it has different coordinates relative to different basis,

$$\mathbf{v} = \sum_{i=1}^{n} a_i \mathbf{v}_i = \sum_{j=1}^{n} b_j \mathbf{u}_j \ . \tag{3.25}$$

Applying Eq. (3.24),

$$\sum_{j=1}^{n} b_j \mathbf{u}_j = \sum_{i=1}^{n} a_i \mathbf{v}_i = \sum_{i,j=1}^{n} a_i L_{ji} \mathbf{u}_j \ . \tag{3.26}$$

Comparing coefficients, we have

$$b_i = \sum_{j=1}^{n} L_{ij} a_j \ . \tag{3.27}$$

This is how the coordinates of the same vector transform relative to two bases. On the other hand, $\mathbf{v}_j$ also form a basis

$$\mathbf{u}_j = \sum_{l=1}^{n} K_{lj} \mathbf{v}_l \tag{3.28}$$

for some other real numbers K_{lj}. We have

$$\begin{aligned} \mathbf{v}_i &= \sum_{j=1}^{n} L_{ji} \mathbf{u}_j = \sum_{j,l=1}^{n} L_{ji} K_{lj} \mathbf{v}_l \\ \mathbf{0} &= \sum_{l=1}^{n} \left(\delta_{li} - \sum_{j=1}^{n} K_{lj} L_{ji} \right) \mathbf{v}_l \ , \end{aligned} \tag{3.29}$$

where δ_{li} is the Kronecker delta. It is just the entries of the identity matrix. Back to Eq. (3.29), the $\mathbf{v}_i$ are linear independent. Their coefficients must be zero

$$\delta_{li} = \sum_{j=1}^{n} K_{lj} L_{ji} \ . \tag{3.30}$$

Similarly, we have $\delta_{li} = \sum_{j=1}^{n} L_{lj} K_{ji}$, which means, as matrices, (L_{ij}) and $(K_{ij}) = (L^{-1}{}_{ij})$ are inverse to each other, and their determinants are non-zero.

Conversely, if $L = (L_{ij})$ is an invertible matrix and $\mathbf{u}_i$ form a basis of a vector space, then $\mathbf{v}_i \equiv \sum_j L_{ji} \mathbf{u}_j$ also form a basis of the same vector space.

Example 3.25. Recalled in Example 3.20, for $\mathbb{R}^2$, we have two bases, $\{\mathbf{u}_1 \equiv (1, 0), \mathbf{u}_2 \equiv (0, 1)\}$ and $\{\mathbf{v}_1 \equiv (1, 0), \mathbf{v}_2 \equiv (1, -1)\}$. What is the L

matrix relating them? By Eq. (3.24), we have to find numbers such that

$$\begin{aligned} \mathbf{v}_1 &= L_{11}\mathbf{u}_1 + L_{21}\mathbf{u}_2 \\ \mathbf{v}_2 &= L_{12}\mathbf{u}_1 + L_{22}\mathbf{u}_2 \ , \end{aligned} \tag{3.31}$$

which are

$$\begin{aligned} (1,0) &= (L_{11}, L_{21}) \\ (1,-1) &= (L_{12}, L_{22}) \ . \end{aligned} \tag{3.32}$$

The inverse of L is

$$L^{-1} = \begin{pmatrix} L^{-1}{}_{11} & L^{-1}{}_{12} \\ L^{-1}{}_{21} & L^{-1}{}_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \ . \tag{3.33}$$

Students should check that $\mathbf{u}_j = \sum_i L^{-1}{}_{ij}\mathbf{v}_i$. ■

How will the matrix representation of a linear transformation change under a change of basis? We will consider only the linear transformations of the same vector space. Linear transformations between different vector spaces is only a bit more tedious.

Let T be a linear transformation from V to V. Relative to a basis $\mathbf{v}_i$, define T_{ij} by

$$T(\mathbf{v}_i) = \sum_j T_{ji}\mathbf{v}_j \ . \tag{3.34}$$

For another basis $\mathbf{u}_j$ related to $\mathbf{v}_i$ by Eq. (3.24), we have

$$\begin{aligned} T(\sum_j L_{ji}\mathbf{u}_j) &= \sum_j T_{ji}(\sum_k L_{kj}\mathbf{u}_k) \\ \sum_j L_{ji}\, T(\mathbf{u}_j) &= \sum_{jk} L_{kj}T_{ji}\,\mathbf{u}_k \\ \sum_{ji} L_{ji}L^{-1}{}_{il}\, T(\mathbf{u}_j) &= \sum_{jkl} L_{kj}T_{ji}L^{-1}{}_{il}\,\mathbf{u}_k \\ \sum_j \delta_{jl}\, T(\mathbf{u}_j) &= \sum_{jkl} L_{kj}T_{ji}L^{-1}{}_{il}\,\mathbf{u}_k \\ T(\mathbf{u}_l) &= \sum_{jkl} L_{kj}T_{ji}L^{-1}{}_{il}\,\mathbf{u}_k \ . \end{aligned} \tag{3.35}$$

In the third line, we multiply by $L^{-1}{}_{il}$ and sum over i. This means the matrix representation of T relative to the new basis is given by a similar transformation

$$T'_{kl} = \sum_{jkl} L_{kj}T_{ji}L^{-1}{}_{il} \ . \tag{3.36}$$

In suggestive matrix notations, it is just $T' = LTL^{-1}$. However, students should be careful. We are talking about the same linear transformation relative to two bases.

Example 3.26. Let us denote the rotation in Example 3.12 by R. As in Example 3.25, put

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.37}$$

and

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} . \tag{3.38}$$

We have

$$\begin{aligned} R(\mathbf{u}_1) &= \cos\alpha\,\mathbf{u}_1 - \sin\alpha\,\mathbf{u}_2 && (3.39)\\ R(\mathbf{u}_2) &= \sin\alpha\,\mathbf{u}_1 + \cos\alpha\,\mathbf{u}_2 \ , && (3.40) \end{aligned}$$

and

$$\begin{aligned} R(\mathbf{v}_1) &= \cos\alpha\,\mathbf{u}_1 - \sin\alpha\,\mathbf{u}_2 \\ &= \cos\alpha\,\mathbf{v}_1 - \sin\alpha\,(\mathbf{v}_1 - \mathbf{v}_2) \\ &= (\cos\alpha - \sin\alpha)\mathbf{v}_1 + \sin\alpha\,\mathbf{v}_2 && (3.41)\\ R(\mathbf{v}_2) &= R(\mathbf{u}_1 - \mathbf{u}_2) \\ &= \cos\alpha\,\mathbf{u}_1 - \sin\alpha\,\mathbf{u}_2 - \sin\alpha\,\mathbf{u}_1 - \cos\alpha\,\mathbf{u}_2 \\ &= (\cos\alpha - \sin\alpha)\mathbf{u}_1 - (\sin\alpha + \cos\alpha)\mathbf{u}_2 \\ &= -2\sin\alpha\,\mathbf{v}_1 + (\sin\alpha + \cos\alpha)\mathbf{v}_2 \ . && (3.42) \end{aligned}$$

We could verify the above equations using Eq. (3.36). Relative to $\mathbf{u}_1, \mathbf{u}_2$, the matrix of R is

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} . \tag{3.43}$$

The matrices relating the two bases are given by Eq. (3.32) and Eq. (3.33). Hence,

$$\begin{aligned} & L \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} L^{-1} \\ = & \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \\ = & \begin{pmatrix} \cos\alpha - \sin\alpha & -2\sin\alpha \\ \sin\alpha & \sin\alpha + \cos\alpha \end{pmatrix} . && (3.44) \end{aligned}$$

■

3.4 Inner Product

We now talk about the "length" of a vector. It turns out that the "length square" is a more convenient object.

Definition 3.27. *The* **inner product** *or* **quadratic form** *of a vector space is a real valued function of two vectors,* $(\cdot\,,\cdot) : V \times V \to \mathbb{R}$, *satisfying*

1. *symmetry:* $(\mathbf{v}, \mathbf{u}) = (\mathbf{u}, \mathbf{v})$;

2. *linearity:* $(a\mathbf{v} + b\mathbf{u}, \mathbf{w}) = a(\mathbf{v}, \mathbf{w}) + b(\mathbf{u}, \mathbf{w})$.

Note that in the above equation, the addition in the left hand side is the vector addition, while in the right hand side, it is the addition of real numbers.

Because of the symmetry, we also have

$$(a\mathbf{v} + b\mathbf{u}, c\mathbf{t} + d\mathbf{w}) = ac(\mathbf{v}, \mathbf{t}) + ad(\mathbf{v}, \mathbf{w}) + bc(\mathbf{u}, \mathbf{t}) + bd(\mathbf{u}, \mathbf{w}) \ , \tag{3.45}$$

where $a, b, c, d \in \mathbb{R}$ *and* $\mathbf{u}, \mathbf{v}, \mathbf{t}, \mathbf{w} \in V$. *Note that inner products are linear in each of their arguments. We also note the* **norm** *of a vector* $\mathbf{v}$ *as*

$$||\mathbf{v}|| \equiv \sqrt{(\mathbf{v}, \mathbf{v})} \ . \tag{3.46}$$

Definition 3.28. *For* $\mathbf{u}$ *and* $\mathbf{v}$ *in a vector space* V *with inner product, if* $(\mathbf{u}, \mathbf{v}) = 0$, *we say that the two vectors are* **orthogonal** *to each other.*

If $\mathbf{u}$ *is orthogonal to all vectors in a subspace* W, *we say that* $\mathbf{u}$ *is orthogonal to* W.

If all vectors in a subspace W_1 *are orthogonal to all vectors in subspace* W_2, *we say that* W_1 *is orthogonal to* W_2.

Theorem 3.29. Pythagorean Theorem If two vectors $\mathbf{v}$ and $\mathbf{u}$ are orthogonal to each other, then

$$||\mathbf{v} + \mathbf{u}||^2 = ||\mathbf{v}||^2 + ||\mathbf{u}||^2 \ . \tag{3.47}$$

Proof.

$$\begin{aligned} & ||\mathbf{v} + \mathbf{u}||^2 \\ = \ & (\mathbf{v} + \mathbf{u}, \mathbf{v} + \mathbf{u}) \\ = \ & (\mathbf{v}, \mathbf{v}) + (\mathbf{v}, \mathbf{u}) + (\mathbf{u}, \mathbf{v}) + (\mathbf{u}, \mathbf{u}) \\ = \ & ||\mathbf{v}||^2 + ||\mathbf{u}||^2 \end{aligned} \tag{3.48}$$

because $(\mathbf{v}, \mathbf{u}) = (\mathbf{u}, \mathbf{v}) = 0$. ■

Example 3.30. The standard inner product in $\mathbb{R}^n$ is defined as: for two vectors $\mathbf{v} = (a_1, \ldots, a_n)$ and $\mathbf{u} = (b_1, \ldots, b_n)$

$$(\mathbf{v}, \mathbf{u}) = a_1 b_1 + \cdots + a_n b_n = \sum_{i,j} a_i \, \eta_{ij} b_j \tag{3.49}$$

where the numbers η_{ij} are

$$\eta_{ij} = \left\{ \begin{array}{ll} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{array} \right. . \tag{3.50}$$

$\mathbb{R}^n$ together with this inner product is a **Euclidean space**. In $\mathbb{R}^3$, the inner product is usually called the dot product. In Euclidean space, we could talk about the "length" or "magnitude" of a vector, or the angle between two vectors, etc. Two vectors are orthogonal to each other if they are perpendicular to each other. (Defined like this, η_{ij} is just the Kronecker delta.)

Not all vector spaces with inner product are Euclidean spaces. ■

Suppose $\mathbf{v}_i$ is a basis in a vector space with inner product. Define

$$\eta_{ij} \equiv (\mathbf{v}_i, \mathbf{v}_j) \ . \tag{3.51}$$

Note that $\eta_{ij} = \eta_{ji}$, the matrix $H \equiv (\eta_{ij})$ is symmetric. Then, for two general vectors $\mathbf{v} = \sum a_i \mathbf{v}_i$ and $\mathbf{w} = \sum b_i \mathbf{v}_i$ in V, their inner product in terms of the coordinates is

$$(\mathbf{v}, \mathbf{w}) = \sum_{ij} a_i (\mathbf{v}_i, \mathbf{v}_j) b_j = \sum_{ij} a_i \eta_{ij} b_j \ . \tag{3.52}$$

How does η_{ij} transform in terms of another basis? By Eq. (3.28),

$$\begin{aligned} \eta'_{ij} &\equiv (\mathbf{u}_i, \mathbf{u}_j) \\ &= \left(\sum_l K_{li} \mathbf{v}_l, \sum_m K_{mj} \mathbf{v}_m \right) \\ &= \sum_{lm} K_{li} K_{mj} (\mathbf{v}_l, \mathbf{v}_m) \\ &= \sum_{lm} K_{li}\, \eta_{lm}\, K_{mj} \ . \end{aligned} \tag{3.53}$$

Pay attention to the indices. We sum up the first indices of K. In suggestive matrix notations, it is $H' = K^t H K$, while in Eq. (3.36), it was LTL^{-1}. (Recall that K^t is the transpose of K.)

Theorem 3.31. For a n dimensional vector space with inner product, a basis $\mathbf{v}_i$ can be chosen such that

$$(\mathbf{v}_i, \mathbf{v}_j) = \begin{cases} -1 & \text{if } i = j = 1, \ldots, l \\ 1 & \text{if } i = j = l+1, \ldots, m \\ 0 & \text{otherwise} \end{cases} \tag{3.54}$$

for some l and m. Explicitly, for vectors $\mathbf{v} = \sum a_i \mathbf{v}_i$ and $\mathbf{u} = \sum b_j \mathbf{v}_j$,

$$(\mathbf{v}, \mathbf{u}) = -a_1 b_1 - \cdots - a_l b_l + a_{l+1} b_{l+1} + \cdots + a_m b_m \ . \tag{3.55}$$

Note that the coefficients of $a_{m+1} b_{m+1}$, etc, terms are zero, and there could be no "positive terms" and/or "negative terms" and/or "zero terms."

Proof. We sketch the proof, by mathematical induction. First consider $n = 1$. If $\mathbf{v} \in V$, then all other vectors are multiple of $\mathbf{v}$. If $(\mathbf{v}, \mathbf{v}) \neq 0$, we could normalize it by

$$\mathbf{v}_1 \equiv \frac{\mathbf{v}}{||\mathbf{v}||} \ . \tag{3.56}$$

We have $(\mathbf{v}_1, \mathbf{v}_1) = \pm 1$ and done. If $(\mathbf{v}, \mathbf{v}) = 0$, then the inner product is identically zero. We finish the case $n = 1$.

For $n > 1$, if for all $\mathbf{v} \in V$, $(\mathbf{v}, \mathbf{v}) = 0$, then the inner product is identically zero. We are done. If there exists a vector that $(\mathbf{v}, \mathbf{v}) \neq 0$. We could assume that it is normalized. Consider all the vectors $\mathbf{w}$ that are orthogonal or perpendicular to $\mathbf{v}$,

$$W \equiv \{\mathbf{w} \in V | (\mathbf{v}, \mathbf{w}) = 0\} \ . \tag{3.57}$$

Then, W is a vector space of dimension $n - 1$, the original inner product defines an inner product on W. Hence, by the assumption of induction, we can choose a basis of W such that the inner product of W has the form Eq. (3.55). Put back $\mathbf{v}$, we have an equation of inner product of V with one more term. (Check all the details. How about the case $n = 0$?) ■

We could understand this theorem as follow. For any symmetric matrix H, we could find an invertible matrix K such that

$$K^t H K = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.58}$$

Example 3.32. Not all vectors have non-zero "length." In $\mathbb{R}^4$, consider

$$\eta_{ij} = \begin{cases} -1 & \text{if } i = j = 1 \\ 1 & \text{if } i = j = 2, 3, 4 \\ 0 & \text{otherwise} \end{cases} \ . \tag{3.59}$$

Then, for $\mathbf{v} \equiv (1, 1, 0, 0)$, $(\mathbf{v}, \mathbf{v}) = 0$. This is the inner product we need in special relativity. ■

Definition 3.33. *If there is no "zero terms" in Eq. (3.55), ($m = n$), the inner product is* **non-degenerate***. As a matrix, the determinant of (η_{ij}) is non-zero.*

Theorem 3.34. An inner product is non-degenerate if and only if for any vector $\mathbf{0} \neq \mathbf{v} \in V$, there is another vector $\mathbf{u}$ such that $(\mathbf{v}, \mathbf{u}) \neq 0$.

Proof. If the inner product is degenerate, by Eq. (3.55), we see that $(\mathbf{v}_n, \mathbf{u}) = 0$ for all vector $\mathbf{u} \in V$.

If the inner product is non-degenerate and $\mathbf{v} = \sum a_i \mathbf{v}_i \neq \mathbf{0}$, then one of the a_i is non-zero, and $(\mathbf{v}, \mathbf{v}_i) = a_i \neq 0$. ■

Definition 3.35. *For a Euclidean space V, any basis $\mathbf{v}_i$ such that $(\mathbf{v}_i, \mathbf{v}_j) = \delta_{ij}$ is called an* **orthonormal basis**. *"Ortho" means that they are all orthogonal to each other, and "normal" means that they are all normalized, their length is one.*

Theorem 3.36. Suppose both $\mathbf{v}_i$ and $\mathbf{u}_i$ are orthonormal bases and they are related by

$$\mathbf{v}_i = \sum_{j=1}^{n} L_{ji}\mathbf{u}_j \ . \tag{3.60}$$

Then, the matrix $L \equiv (L_{ij})$ is an orthogonal matrix, $L^t L = 1$.

Proof. We have

$$\begin{aligned} \delta_{ij} &= (\mathbf{v}_i, \mathbf{v}_j) \\ &= \sum_{lk} (L_{li}\mathbf{u}_l, L_{kj}\mathbf{u}_k) \\ &= \sum_{lk} L_{li}L_{kj}(\mathbf{u}_l, \mathbf{u}_k) \\ &= \sum_{lk} L_{li}L_{kj}\delta_{lk} \\ &= \sum_{k} L_{ki}L_{kj} \ , \end{aligned} \tag{3.61}$$

which is just $L^t L = 1$. ∎

Example 3.37. We can easily check that rotation in Example 3.12 is orthogonal. ∎

Example 3.38. Suppose that two particles in outer space are interacting via an attractive force with magnitude proportional to their separation. The potential energy will then be

$$V(\mathbf{r}_1, \mathbf{r}_2) = k||\mathbf{r}_1 - \mathbf{r}_2||^2 \tag{3.62}$$

where $k > 0$. This is a quadratic form of their positions. The kinetic energy is

$$T = \frac{1}{2}m_1\mathbf{v}_1^2 + \frac{1}{2}m_2\mathbf{v}_2^2 \ . \tag{3.63}$$

This is a quadratic form of their velocities. Define the center of mass velocity by

$$\mathbf{V} = \frac{m_1\mathbf{v}_1 + m_2\mathbf{v}_2}{m_1 + m_2} \ , \tag{3.64}$$

their relative velocity

$$\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2 \tag{3.65}$$

and their relative position

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \ . \tag{3.66}$$

We could rewrite

$$
\begin{aligned}
V &= k||\mathbf{r}||^2 && (3.67)\\
T &= \frac{1}{2}(m_1+m_2)\mathbf{V}^2 + \frac{m_1 m_2}{2(m_1+m_2)}\mathbf{v}^2 \ . && (3.68)
\end{aligned}
$$

The potential is a quadratic form of a six-dimensional vector space, three coordinates for each particle. After a change of coordinates, we discover that it depends only on the three coordinates of the relative position. This quadratic form is degenerate, corresponding to $n = 6$, $l = 0$ and $m = 3$ in Theorem 3.31. The physical meaning is that the center of mass can move freely. The kinetic energy, as a quadratic form, is non-degenerate. ■

Example 3.39. When we consider the motion of a rigid body, its kinetic energy can be expressed as a quadratic form of its angular velocity. The object which plays the role of the mass is something called **moment of inertia**. These are necessary to understand why wing nuts move like this: `https://www.youtube.com/watch?v=1VPfZ_XzisU` . ■

3.5 Gram-Schmidt Process

We consider only n-dimensional Euclidean space V in this section. We would like to ask, given a general basis $\mathbf{v}_i$ of V, is it possible to construct an orthonormal basis? The answer is positive and it is called the **Gram-Schmidt process**. Note that the orthonormal basis constructed is not unique, for example, it depends on the order of the vectors in the basis $\mathbf{v}_i$.

The first step is easy. Define

$$\mathbf{w}_1 \equiv \frac{\mathbf{v}_1}{||\mathbf{v}_1||} \ . \tag{3.69}$$

Then, the norm of $\mathbf{w}_1$ is one. For the next step, we need to construct a vector that is orthogonal to $\mathbf{w}_1$. In general, $\mathbf{v}_2$ is not. We need to subtract out the "projection" of $\mathbf{v}_2$ to $\mathbf{w}_1$. Let consider

$$\mathbf{x}_2 \equiv \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{w}_1)\, \mathbf{w}_1 \ . \tag{3.70}$$

$\mathbf{v}_1$ and $\mathbf{v}_2$ are vectors of a basis. They are not proportional to each other. (They are linearly independent.) Hence, $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$ and $\mathbf{x}_2$ is not zero. $\mathbf{x}_2$ is orthogonal to $\mathbf{w}_1$ because

$$(\mathbf{x}_2, \mathbf{w}_1) = (\mathbf{v}_2, \mathbf{w}_1) - (\mathbf{v}_2, \mathbf{w}_1)(\mathbf{w}_1, \mathbf{w}_1) = 0 \tag{3.71}$$

because the norm of $\mathbf{w}_1$ is one. Define

$$\mathbf{w}_2 \equiv \frac{\mathbf{x}_2}{||\mathbf{x}_2||} \ . \tag{3.72}$$

$\mathbf{w}_1$ and $\mathbf{w}_2$ are orthonormal and $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2) = \mathrm{Span}(\mathbf{w}_1, \mathbf{w}_2)$.

For the general step, assume that $\mathbf{w}_1, \ldots, \mathbf{w}_k$ have been constructed such that they are orthonormal and

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \mathrm{Span}(\mathbf{w}_1, \ldots, \mathbf{w}_k) \ . \tag{3.73}$$

We need to subtract out the projection of $\mathbf{v}_{k+1}$ to $\mathbf{w}_1, \ldots, \mathbf{w}_k$. Define

$$\mathbf{x}_{k+1} \equiv \mathbf{v}_{k+1} - (\mathbf{v}_{k+1}, \mathbf{w}_k)\, \mathbf{w}_k - \cdots - (\mathbf{v}_{k+1}, \mathbf{w}_1)\, \mathbf{w}_1 \ . \tag{3.74}$$

$\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ are linearly independent, so are $\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{v}_{k+1}$. $\mathbf{x}_{k+1}$ is not zero. Similar to Eq. (3.71), we could prove that $\mathbf{x}_{k+1}$ is orthogonal to all $\mathbf{w}_1, \ldots, \mathbf{w}_k$. Finally, normalize

$$\mathbf{w}_{k+1} \equiv \frac{\mathbf{x}_{k+1}}{||\mathbf{x}_{k+1}||} \ , \tag{3.75}$$

then $\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}$ are orthonormal and

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}) = \mathrm{Span}(\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}) \ . \tag{3.76}$$

Theorem 3.40. Gram-Schmidt Process Given a basis $\mathbf{v}_i$ of an Euclidean space V, define

$$\begin{aligned}
\mathbf{w}_1 &= \frac{\mathbf{v}_1}{||\mathbf{v}_1||} \\
\mathbf{x}_2 &= \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{w}_1)\, \mathbf{w}_1 \\
\mathbf{w}_2 &= \frac{\mathbf{x}_1}{||\mathbf{x}_1||} \\
&\ \vdots \\
\mathbf{x}_k &= \mathbf{v}_k - (\mathbf{v}_k, \mathbf{w}_{k-1})\, \mathbf{w}_{k-1} - \cdots - (\mathbf{v}_k, \mathbf{w}_1)\, \mathbf{w}_1 \\
\mathbf{w}_k &= \frac{\mathbf{x}_k}{||\mathbf{x}_k||} \ .
\end{aligned} \tag{3.77}$$

Then, $\mathbf{w}_i$ is an orthonormal basis of V and for all k,

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \mathrm{Span}(\mathbf{w}_1, \ldots, \mathbf{w}_k) \ . \tag{3.78}$$

∎

Example 3.41. Find an orthonormal basis from

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ . \tag{3.79}$$

by the Gram-Schmidt process for $\mathbb{R}^2$.

By inspection, $\mathbf{v}_1$ is already normalized. Let $\mathbf{w}_1 = \mathbf{v}_1$. Then,

$$\begin{aligned} \mathbf{x}_2 &= \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{w}_1)\,\mathbf{w}_1 \\ &= \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ -1 \end{pmatrix} . \end{aligned} \tag{3.80}$$

This is also normalized. The answer is

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix} . \tag{3.81}$$

■

Example 3.42. Find an orthonormal basis of $\mathbb{R}^3$ from

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} -5 \\ 2 \\ 0 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 7 \end{pmatrix} . \tag{3.82}$$

We have

$$\mathbf{w}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} . \tag{3.83}$$

Note that $(\mathbf{v}_2, \mathbf{w}_1) = -5/\sqrt{10}$. Hence,

$$\begin{aligned} \mathbf{x}_2 &= \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{w}_1)\,\mathbf{w}_1 \\ &= \begin{pmatrix} -5 \\ 2 \\ 0 \end{pmatrix} + \frac{5}{10} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} -9/2 \\ 2 \\ 3/2 \end{pmatrix} . \end{aligned} \tag{3.84}$$

$$\mathbf{w}_2 = \frac{1}{\sqrt{106}} \begin{pmatrix} -9 \\ 4 \\ 3 \end{pmatrix} . \tag{3.85}$$

$(\mathbf{v}_3, \mathbf{w}_1) = 22/\sqrt{10}$, $(\mathbf{v}_3, \mathbf{w}_2) = 8/\sqrt{106}$ and

$$\mathbf{w}_3 = \frac{1}{\sqrt{265}} \begin{pmatrix} -6 \\ -15 \\ 2 \end{pmatrix} . \tag{3.86}$$

Students could see that the calculation is quite tedious. ■

Chapter 4

Line, Surface and Volume Integrals

We have assumed that students have studied the action of various differential operators on continuously varying scalar and vector fields. Indeed, these differential operations are often involved in the integration of field quantities along lines, over surfaces, and throughout volume. Here, we study the line integrals, surface integrals, as well as volume integrals. We also discuss three fundamental theorems of vector calculus: Green's Theorem, Stoke's theorem, and the Divergence theorem, which tells us the relationships of these new types of integrals to the single, double, and triple integrals. The main points of this chapter are

- Line integral with respect to arc length, p.70, p.77
$$\int_C f(x, y, z)\,\mathrm{d}s \tag{4.1}$$
(total mass of a wire)

- Line integral with respect to x, p.74
$$\int_C f(x, y, z)\,\mathrm{d}x \tag{4.2}$$
(work done by the x-component of a force)

- Line integral of $\mathbf{F}$ along C, p.81
$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} \tag{4.3}$$
(work done by a force)

- Fundamental Theorem for Line Integrals, p.84
$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a)) \tag{4.4}$$

- Green's Theorem in the Plane, p.94

$$\oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A \tag{4.5}$$

- Parametric Surfaces, p.104
- Surface Area of a Smooth Parametric Surface, p.114

$$A(S) = \iint_D ||\mathbf{r}_u \times \mathbf{r}_v||\,\mathrm{d}A \tag{4.6}$$

- Surface Integral of a Scalar Field, p.119

$$\iint_S f(x, y, z)\,\mathrm{d}S \tag{4.7}$$

(total mass of a metal sheet)

- Surface Integral of **F** over S, p.127

$$\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} \tag{4.8}$$

(the flux of **F** across S)

- Stokes' Theorem, p.132

$$\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \iint_S \text{curl } \mathbf{F} \cdot \mathrm{d}\mathbf{S} \tag{4.9}$$

- Volume Integrals, p.139

$$\iiint_E f(x, y, z)\,\mathrm{d}V \qquad \iiint_E \mathbf{F}(x, y, z)\,\mathrm{d}V \tag{4.10}$$

- Divergence Theorem, p.141

$$\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iiint_E \operatorname{div} \mathbf{F}\,\mathrm{d}V \tag{4.11}$$

4.1 Line Integrals

Suppose a thin wire lying on the x-axis from $x = a$ to $x = b$ has linear and varying mass density $\lambda(x)$. Obviously, the mass of the wire is given by the single integral $\int_a^b \lambda(x)\,\mathrm{d}x$. However, if the wire lies along a curve in the plane or space, then its mass is given by an integral over the curve rather than an integral over a line segment on the x-axis. Such integrals of functions over curves are called **line integrals**.

To begin, consider a thin wire takes the shape of an **oriented curve** C in the plane, which is a curve where a consistent direction is defined along the curve. Assume that C is a smooth curve defined by the vector equation

$$\mathbf{r}(t) = x(t)\,\hat{\imath} + y(t)\,\hat{\jmath}\ , \qquad a \le t \le b \tag{4.12}$$

so that $\mathbf{r}'(t)$ is continuous and never equal to $\mathbf{0}$. Moreover, suppose the linear mass density of the wire is given by a continuous function $\lambda(x, y)$ defined along the curve C.

We first divide the parameter interval $[a, b]$ into n subintervals $[t_{i-1}, t_i]$ of equal widths. If we denote $x_i = x(t_i)$ and $y_i = y(t_i)$, then the points $P_i(x_i, y_i)$ divide the curve C into n segments with lengths $\Delta s_1, \Delta s_2, \ldots, \Delta s_n$, respectively. In each segment, we choose an arbitrary point $P_i^*(x_i^*, y_i^*)$ which corresponds to a point t_i^* in the subinterval $[t_{i-1}, t_i]$. Then the product $\lambda(x_i^*, y_i^*)\,\Delta s_i$ gives the approximate mass of the ith segment.

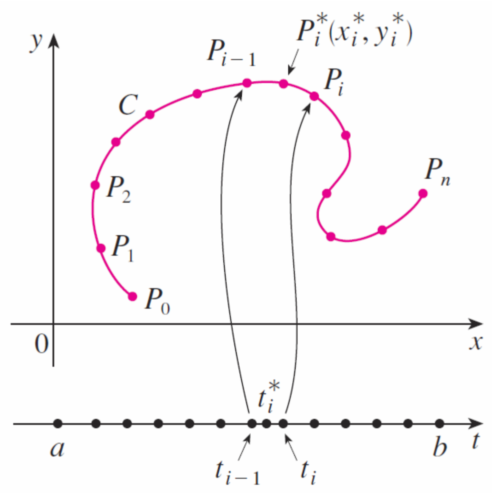


Figure 4.1: A curve on the plane.

So the total mass M of the wire is approximately given by the sum

$$\sum_{i=1}^{n} \lambda(x_i^*, y_i^*)\,\Delta s_i\ , \tag{4.13}$$

which is similar to a Riemann sum. If we let n approach infinity so that each segment becomes infinitesimally short, then this sum would be expected to approach the total mass of the wire

$$M = \lim_{n\to\infty} \sum_{i=1}^{n} \lambda(x_i^*, y_i^*)\,\Delta s_i\ , \tag{4.14}$$

provided that this limit exists. This observation leads us to the following definitions.

Definition 4.1. *If f is a scalar field defined on a smooth oriented curve C in the plane, then the* **line integral of f along C with respect to arc length** *is*

$$\int_C f(x,y)\,\mathrm{d}s = \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i^*, y_i^*)\,\Delta s_i \ , \tag{4.15}$$

provided that this limit exists.

As shown in Fig. 4.1, we can approximate the length Δs_i of the ith segment of the curve C by the straight-line distance

$$\Delta s_i \approx \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \tag{4.16}$$

Since C is assumed to be a smooth curve, both $x(t)$ and $y(t)$ have continuous first-order derivatives. By the Mean Value Theorem, there exists some numbers t_i^{**} and t_i^{***} in (t_{i-1}, t_i) such that

$$\Delta s_i \approx \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \approx \sqrt{[x'(t_i^{**})]^2 + [y'(t_i^{***})]^2}\Delta t_i \ . \tag{4.17}$$

Therefore, the length of curve C is given by

$$\begin{aligned} L &= \lim_{n\to\infty}\sum_{i=1}^{n}\Delta s_i = \lim_{n\to\infty}\sum_{i=1}^{n}\sqrt{[x'(t_i^{**})]^2 + [y'(t_i^{***})]^2}\,\Delta t_i \\ &= \int_a^b \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t \ . \end{aligned} \tag{4.18}$$

Similar arguments can be used to show that if f is a continuous function, then the limit in Eq. (4.15) always exists and the following formula can be used to evaluate the line integral of f along C with respect to arc length

$$\int_C f(x,y)\,\mathrm{d}s = \int_a^b f(x(t), y(t))\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t \ . \tag{4.19}$$

Example 4.2. Evaluate $\int_C (x^2y + 2)\,\mathrm{d}s$, where C is the upper half of a unit circle from $(1,0)$ to $(-1,0)$.

Note that the upper half of the unit circle can be described by the parametric equations

$$x = \cos t, \qquad y = \sin t, \qquad 0 \le t \le \pi \ . \tag{4.20}$$

Therefore, using Eq. (4.19), we obtain

$$\begin{aligned}
\int_C (x^2 y + 2)\, \mathrm{d}s &= \int_0^\pi (\cos^2 t \sin t + 2)\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\, \mathrm{d}t \\
&= \int_0^\pi (\cos^2 t \sin t + 2)\sqrt{(-\sin t)^2 + (\cos t)^2}\, \mathrm{d}t \\
&= \int_0^\pi (\cos^2 t \sin t + 2)\, \mathrm{d}t \\
&= \left[-\frac{1}{3}\cos^3 t + 2t\right]_0^\pi \\
&= 2\pi + \frac{2}{3}\ . \qquad (4.21)
\end{aligned}$$

■

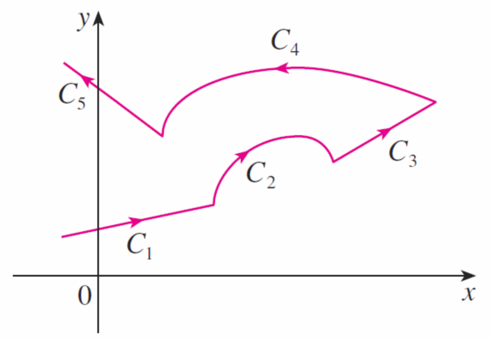


Figure 4.2: A piecewise-smooth curve.

A curve C is a **piecewise-smooth** curve if it is a union of a finite number of smooth curves $C_1, C_2, \ldots, C_n$ where the terminal point of C_i is the same as the initial point of C_{i+1} for $i = 1, 2, \ldots, n-1$ as shown in Fig. 2. Then the line integral of f along the curve C is the sum of the integrals of f along each of the smooth pieces of C

$$\int_C f(x,y)\, \mathrm{d}s = \int_{C_1} f(x,y)\, \mathrm{d}s + \int_{C_2} f(x,y)\, \mathrm{d}s + \cdots + \int_{C_n} f(x,y)\, \mathrm{d}s\ . \quad (4.22)$$

Example 4.3. Evaluate $\int_C 2x\, \mathrm{d}s$, where C consists of the arc C_1 of the parabola $y = x^2$ from $(0,0)$ to $(1,1)$ followed by the vertical segment C_2 from $(1,1)$ to $(1,2)$.

The curve C is shown in Fig. 4.3. Since the arc C_1 is the graph of a function of x, we can set $x = t$ and the parametric equations for C_1 become

$$x = t, \qquad y = t^2, \qquad 0 \le t \le 1\ . \quad (4.23)$$

For the vertical segment C_2, we set $y = t$ and so the parametric equations

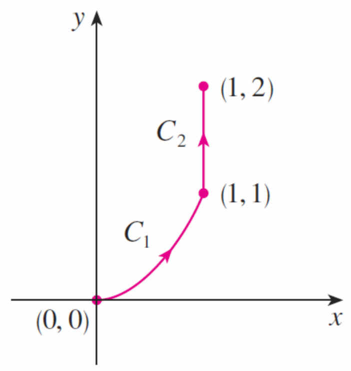


Figure 4.3: Part of a parabola and a segment of straight line.

for C_2 are

$$x = 1, \qquad y = t, \qquad 1 \leq t \leq 2 \ . \tag{4.24}$$

Therefore, Eq. (4.19) gives

$$\begin{aligned}
\int_{C_1} 2x \, \mathrm{d}s &= \int_0^1 2t \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} \, \mathrm{d}t \\
&= \int_0^1 2t \sqrt{(1)^2 + (2t)^2} \, \mathrm{d}t \\
&= \int_0^1 2t \sqrt{4t^2 + 1} \, \mathrm{d}t \\
&= \left[\frac{1}{6}(4t^2 + 1)^{3/2}\right]_0^1 \\
&= \frac{5\sqrt{5} - 1}{6} \ .
\end{aligned} \tag{4.25}$$

$$\begin{aligned}
\int_{C_2} 2x \, \mathrm{d}s &= \int_1^2 2 \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} \, \mathrm{d}t \\
&= \int_1^2 2 \sqrt{(0)^2 + (1)^2} \, \mathrm{d}t \\
&= 2 \ .
\end{aligned} \tag{4.26}$$

Hence, applying Eq. (4.22) yields

$$\int_C 2x \, \mathrm{d}s = \frac{5\sqrt{5} + 11}{6} \ . \tag{4.27}$$

■

As we have discussed earlier, the mass of a thin wire represented by a smooth curve C in the plane that has linear mass density $\lambda(x, y)$ is given by

$$M = \int_C \lambda(x, y)\, \mathrm{d}s\ . \tag{4.28}$$

The **center of mass** of the wire is located at the point $(\bar{x}, \bar{y})$ where

$$\bar{x} = \frac{1}{M}\int_C x\,\lambda(x, y)\, \mathrm{d}s\ , \qquad \bar{y} = \frac{1}{M}\int_C y\,\lambda(x, y)\, \mathrm{d}s\ . \tag{4.29}$$

Example 4.4. A thin wire has the shape of a semicircle of radius a. The linear mass density of the wire is proportional to the distance from the diameter that joins the two endpoints of the wire. Find the mass of the wire and the location of its center of mass.

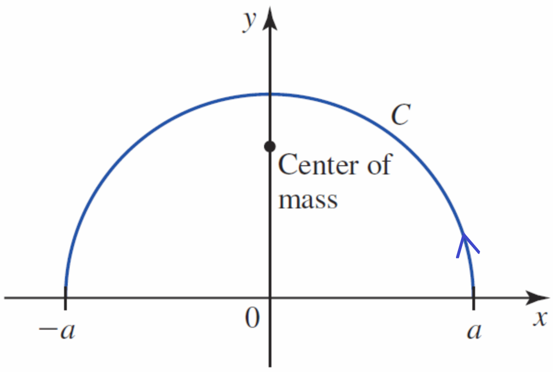


Figure 4.4: A semicircle.

If the wire is placed on a coordinate system as shown in Fig.4.4, then it coincides with the curve C described by the parametric equations

$$x = a\cos t\ , \qquad y = a\sin t\ , \qquad 0 \le t \le \pi\ . \tag{4.30}$$

The linear mass density of the wire is given by $\lambda(x, y) = ky$, where k is a positive constant. Then using Eq. (4.28), we obtain the mass of the wire

$$\begin{aligned} M &= \int_C \lambda(x, y)\, \mathrm{d}s \\ &= \int_C ky\, \mathrm{d}s \\ &= \int_0^\pi ka\sin t\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\, \mathrm{d}t \\ &= \int_0^\pi ka\sin t\sqrt{(-a\sin t)^2 + (a\cos t)^2}\, \mathrm{d}t \\ &= ka^2\int_0^\pi \sin t\, \mathrm{d}t \\ &= 2ka^2\ . \end{aligned} \tag{4.31}$$

By symmetry, we see that the center of mass of the wire is located on the

y-axis, $\bar{x} = 0$. From Eq. (4.29), we have

$$\begin{aligned} \bar{y} &= \frac{1}{M}\int_C y\,\lambda(x,y)\,\mathrm{d}s \\ &= \frac{1}{2ka^2}\int_C ky^2\,\mathrm{d}s \\ &= \frac{1}{2ka^2}\int_0^{\pi}(ka^2\sin^2 t)(a)\,\mathrm{d}t \\ &= \frac{a}{4}\int_0^{\pi}(1-\cos 2t)\,\mathrm{d}t \\ &= \frac{a}{4}\left[t-\frac{1}{2}\sin 2t\right]_0^{\pi} \\ &= \frac{\pi a}{4}\,. \end{aligned} \tag{4.32}$$

Therefore, the center of mass of the wire is located at $(0, \pi a/4)$. ■

If we replace Δs_i by either $\Delta x_i = x_i - x_{i-1}$ or $\Delta y_i = y_i - y_{i-1}$ in Eq. (4.15), two other line integrals are obtained. They are known as the **line integrals of f along C with respect to x and y**

$$\int_C f(x,y)\,\mathrm{d}x = \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i^*, y_i^*)\,\Delta x_i \tag{4.33}$$

$$\int_C f(x,y)\,\mathrm{d}y = \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i^*, y_i^*)\,\Delta y_i\,. \tag{4.34}$$

Line integrals with respect to x and y can be also evaluated by expressing everything in terms of t: $x = x(t)$, $y = y(t)$, $\mathrm{d}x = x'(t)\,\mathrm{d}t$, and $\mathrm{d}y = y'(t)\,\mathrm{d}t$. This leads to the following formulas

$$\int_C f(x,y)\,\mathrm{d}x = \int_a^b f(x(t),y(t))\,x'(t)\,\mathrm{d}t \tag{4.35}$$

$$\int_C f(x,y)\,\mathrm{d}y = \int_a^b f(x(t),y(t))\,y'(t)\,\mathrm{d}t\,. \tag{4.36}$$

It is quite often that the line integrals with respect to x and y appear together. When this happens, it's customary to abbreviate by writing

$$\int_C P(x,y)\,\mathrm{d}x + \int_C Q(x,y)\,\mathrm{d}y = \int_C P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y\,. \tag{4.37}$$

Example 4.5. Evaluate $\int_C y\,\mathrm{d}x + x^2\,\mathrm{d}y$, where

(a) C is the line segment C_1 from $(1,-1)$ to $(4,2)$,

(b) C is the arc C_2 of the parabola $x = y^2$ from $(1,-1)$ to $(4,2)$, and

(c) C is the arc C_3 of the parabola $x = y^2$ from $(4,2)$ to $(1,-1)$.

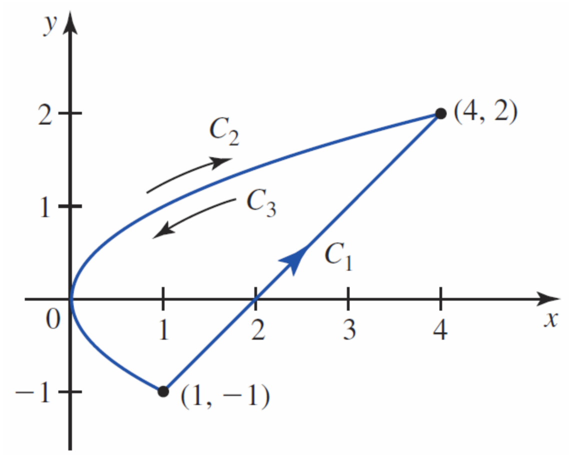


Figure 4.5: Curves for Example 4.5.

The curves C_1, C_2, and C_3 are depicted in Fig. 4.5.

(a) C_1 can be described by the parametric equations

$$x = 3t + 1 \;, \qquad y = 3t - 1 \;, \qquad 0 \leq t \leq 1 \;, \tag{4.38}$$

which implies $\mathrm{d}x = 3\,\mathrm{d}t$ and $\mathrm{d}y = 3\,\mathrm{d}t$. Then, Eq. (4.36) gives

$$\begin{aligned} \int_C y\,\mathrm{d}x + x^2\,\mathrm{d}y &= \int_0^1 (3t-1)(3\,\mathrm{d}t) + (3t+1)^2(3\,\mathrm{d}t) \\ &= \int_0^1 27(t^2+t)\,\mathrm{d}t \\ &= \frac{45}{2} \;. \end{aligned} \tag{4.39}$$

(b) By letting $y = t$, we obtain the parametric equations of C_2

$$x = t^2 \;, \qquad y = t \;, \qquad -1 \leq t \leq 2 \;, \tag{4.40}$$

which implies $\mathrm{d}x = 2t\,\mathrm{d}t$ and $\mathrm{d}y = \mathrm{d}t$. Then, Eq. (4.36) gives

$$\begin{aligned} \int_C y\,\mathrm{d}x + x^2\,\mathrm{d}y &= \int_{-1}^2 (t)(2t\,\mathrm{d}t) + (t^2)^2(\mathrm{d}t) \\ &= \int_{-1}^2 (t^4 + 2t^2)\,\mathrm{d}t \\ &= \frac{63}{5} \;. \end{aligned} \tag{4.41}$$

(c) By letting $y = -t$, we obtain the parametric equations of C_3

$$x = t^2 \;, \qquad y = -t \;, \qquad -2 \leq t \leq 1 \;. \tag{4.42}$$

which implies $\mathrm{d}x = 2t\,\mathrm{d}t$ and $\mathrm{d}y = -\mathrm{d}t$. Then, Eq. (4.36) gives

$$\begin{aligned}\int_C y\,\mathrm{d}x + x^2\,\mathrm{d}y &= \int_{-2}^{1} (-t)(2t\,\mathrm{d}t) + (t^2)^2(-\mathrm{d}t) \\ &= -\int_{-2}^{1} (t^4 + 2t^2)\,\mathrm{d}t \\ &= -\frac{63}{5}\ . \end{aligned} \tag{4.43}$$

■

Notice that we got different answers in parts (a) and (b) even though the two curves had the same endpoints. It suggests that the value of a line integral depends not only on the endpoints, but also on the path joining these points. In addition, the difference in the answers to parts (b) and (c) suggests that the value of a line integral also depends on the **orientation** of the curve — the direction in which the curve is traced as the parameter t increases. Let $-C$ denote the curve with the same points as C but with the orientation reversed as shown in Fig. 4.6. Then, we have

$$\int_{-C} f(x,y)\,\mathrm{d}x = -\int_C f(x,y)\,\mathrm{d}x \tag{4.44}$$

$$\int_{-C} f(x,y)\,\mathrm{d}y = -\int_C f(x,y)\,\mathrm{d}y\ . \tag{4.45}$$

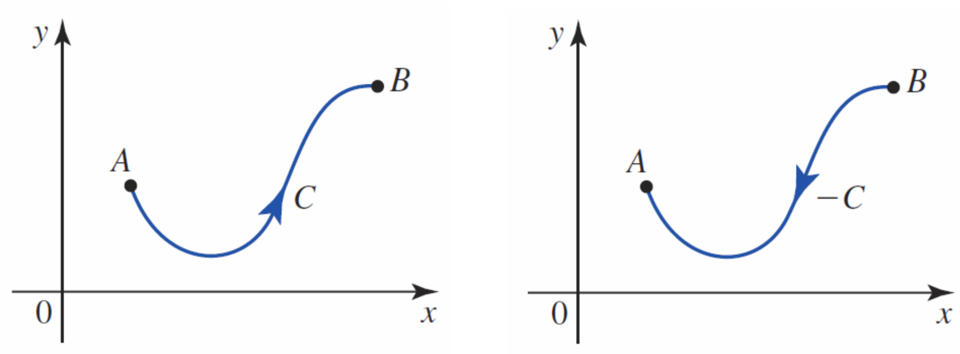


Figure 4.6: Two curves with opposite orientations.

However, if we integrate with respect to arc length, the value of the line integral doesn't change when we reverse the orientation of the curve

$$\int_{-C} f(x,y)\,\mathrm{d}s = \int_C f(x,y)\,\mathrm{d}s\ . \tag{4.46}$$

It is because Δs_i is always positive while Δx_i and Δy_i change sign when we reverse the orientation of C.

The line integrals in two-dimensional space can be extended to line integrals in three-dimensional space. Suppose C is a smooth curve described by the vector equation

$$\mathbf{r}(t) = x(t)\,\hat{\imath} + y(t)\,\hat{\jmath} + z(t)\,\hat{k}\ , \qquad a \le t \le b\ . \tag{4.47}$$

If f is a function of three variables that is defined and continuous on some region containing C, then the line integral of f along C with respect to arc length is defined as follows.

Definition 4.6. *If f is a scalar field defined on a smooth oriented curve C in space, then the line integral of f along C with respect to arc length is*

$$\int_C f(x,y,z)\,\mathrm{d}s = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*, y_i^*, z_i^*)\,\Delta s_i \tag{4.48}$$

provided that this limit exists.

Just like the two-dimensional case, since f is a continuous function, the following formula can be used to evaluate the line integral of f along C with respect to arc length

$$\int_C f(x,y,z)\,\mathrm{d}s = \int_a^b f(x(t),y(t),z(t))\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}\,\mathrm{d}t\ . \tag{4.49}$$

Example 4.7. Evaluate $\int_C y\sin z\,\mathrm{d}s$, where C is the circular helix given by the parametric equations $x = \cos t$, $y = \sin t$ and $z = t$ where $0 \le t \le 2\pi$.

Eq. (4.49) gives

$$\begin{aligned}
\int_C y\sin z\,\mathrm{d}s &= \int_0^{2\pi} \sin^2 t\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}\,\mathrm{d}t \\
&= \int_0^{2\pi} \sin^2 t\sqrt{(-\sin t)^2 + (\cos t)^2 + (1)^2}\,\mathrm{d}t \\
&= \int_0^{2\pi} \frac{1}{\sqrt{2}}(1-\cos 2t)\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2}}\left[t - \frac{1}{2}\sin 2t\right]_0^{2\pi} \\
&= \sqrt{2}\pi\ .
\end{aligned} \tag{4.50}$$

■

Line integrals of f along a curve C in space with respect to x, y, and z are defined in a similar manner as line integrals along a plane curve. For example,

$$\int_C f(x,y,z)\,\mathrm{d}x = \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i^*, y_i^*, z_i^*)\,\Delta x_i = \int_a^b f(x(t),y(t),z(t))\,x'(t)\,\mathrm{d}t\ . \tag{4.51}$$

Thus we can evaluate line integral of the form

$$\int_C P(x,y,z)\,\mathrm{d}x + Q(x,y,z)\,\mathrm{d}y + R(x,y,z)\,\mathrm{d}z \tag{4.52}$$

by expressing everything $(x, y, z, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$ in terms of the parameter t.

Example 4.8. Evaluate $\int_C y\,\mathrm{d}x + z\,\mathrm{d}y + x\,\mathrm{d}z$, where the curve C consists of part of the twisted cubic C_1 given by the parametric equations $x = t$, $y = t^2$, $z = t^3$ with $0 \le t \le 1$, followed by the line segment C_2 from $(1,1,1)$ to $(0,1,0)$.

The curve C is shown in Fig. 4.7. For the twisted cubic C_1, we have $\mathrm{d}x = \mathrm{d}t$, $\mathrm{d}y = 2t\,\mathrm{d}t$ and $\mathrm{d}z = 3t^2\,\mathrm{d}t$. Therefore,

$$\begin{aligned}
\int_{C_1} y\,\mathrm{d}x + z\,\mathrm{d}y + x\,\mathrm{d}z &= \int_0^1 (t^2)(\mathrm{d}t) + (t^3)(2t\,\mathrm{d}t) + (t)(3t^2\,\mathrm{d}t) \\
&= \int_0^1 (2t^4 + 3t^3 + t^2)\,\mathrm{d}t \\
&= \frac{89}{60}\ .
\end{aligned} \tag{4.53}$$

Next, we write the parametric equations of the line segment C_2

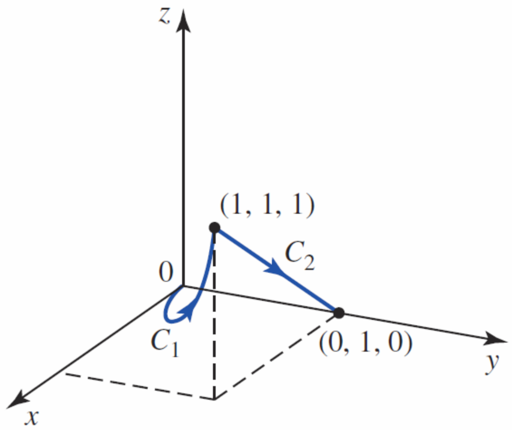


Figure 4.7: A twisted curve.

$$x = -t + 1\ , \qquad y = 1\ , \qquad z = -t + 1\ , \qquad 0 \le t \le 1\ . \tag{4.54}$$

Then, $\mathrm{d}x = -\mathrm{d}t$, $\mathrm{d}y = 0$, and $\mathrm{d}z = -\mathrm{d}t$. Therefore,

$$\begin{aligned}\int_{C_2} y\,\mathrm{d}x + z\,\mathrm{d}y + x\,\mathrm{d}z &= \int_0^1 (1)(-\mathrm{d}t) + (1-t)(0) + (1-t)(-\mathrm{d}t) \\ &= \int_0^1 (t-2)\,\mathrm{d}t \\ &= -\frac{3}{2}\ . \end{aligned} \tag{4.55}$$

Finally, putting these results together, we have

$$\int_C y\,\mathrm{d}x + z\,\mathrm{d}y + x\,\mathrm{d}z = -\frac{1}{60}\ . \tag{4.56}$$

■

In analogy to a thin wire in the plane, the mass of a thin wire represented by a smooth curve C in space that has linear mass density $\lambda(x, y, z)$ is given by

$$M = \int_C \lambda(x, y, z)\,\mathrm{d}s\ . \tag{4.57}$$

The center of mass of the wire is located at the point $(\bar{x}, \bar{y}, \bar{z})$ where, for example,

$$\bar{x} = \frac{1}{M}\int_C x\,\lambda(x, y, z)\,\mathrm{d}s\ . \tag{4.58}$$

Example 4.9. A coil spring lies along the helix $\mathbf{r}(t) = \cos 4t\,\hat{\imath} + \sin 4t\,\hat{\jmath} + t\,\hat{k}$ where $0 \le t \le 2\pi$. It has a constant linear mass density $\lambda(x, y, z) = 1$. Find the mass of the spring.

Figure 4.8 is the sketch of the spring. Because of the symmetries involved, the center of mass of the spring lies at the point $(0, 0, \pi)$ on the z-axis. For the remaining calculations, we first find the arc length element

$$\begin{aligned}\mathrm{d}s &= \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}\,\mathrm{d}t \\ &= \sqrt{(-4\sin 4t)^2 + (4\cos 4t)^2 + 1^2}\,\mathrm{d}t \\ &= \sqrt{17}\,\mathrm{d}t\ . \end{aligned} \tag{4.59}$$

We then use Eq. (4.57) to find the mass of the spring

$$\begin{aligned} M &= \int_{\text{Helix}} \lambda(x, y, z)\,\mathrm{d}s \\ &= \int_0^{2\pi} (1)(\sqrt{17}\,\mathrm{d}t) \\ &= 2\pi\sqrt{17}\ . \end{aligned} \tag{4.60}$$

■

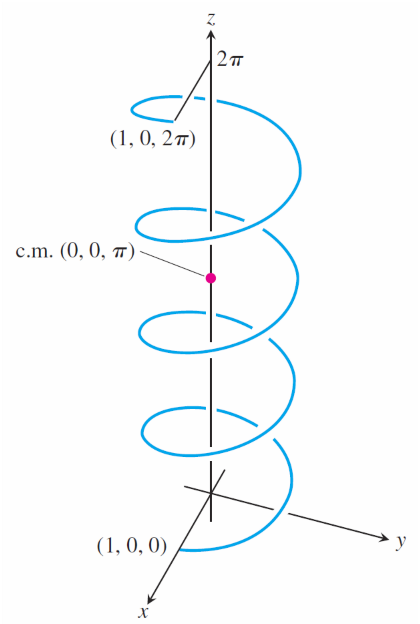


Figure 4.8: A coil spring.

Up to now, we have only considered line integrals involving a scalar field f. We now turn our attention to the line integrals of vector fields. Suppose we want to find the work done by a continuous force field $\mathbf{F}(x, y, z) = P(x, y, z)\,\hat{\imath} + Q(x, y, z)\,\hat{\jmath} + R(x, y, z)\,\hat{k}$ in moving a particle along a smooth curve C in space. Let C be described by the vector equation

$$\mathbf{r}(t) = x(t)\,\hat{\imath} + y(t)\,\hat{\jmath} + z(t)\,\hat{k}\ , \qquad a \leq t \leq b\ . \tag{4.61}$$

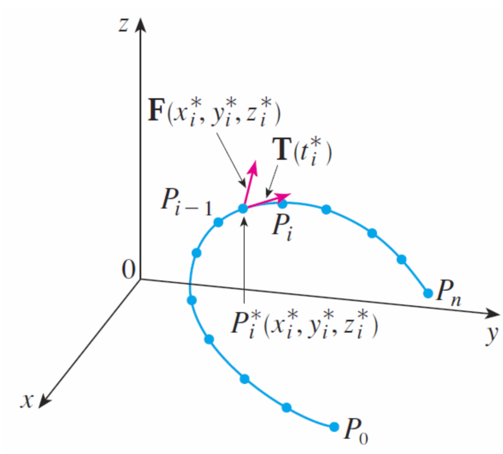


Figure 4.9: Line integral of vector field.

We divide the parameter interval $[a, b]$ into n subintervals $[t_{i-1}, t_i]$ of equal widths. If we denote $x_i = x(t_i)$, $y_i = y(t_i)$, and $z_i = z(t_i)$, then the points $P_i(x_i, y_i, z_i)$ divide the curve C into n segments with lengths $\Delta s_1, \Delta s_2, \ldots, \Delta s_n$, respectively (see Fig. 4.9). Pick a point $P_i^*(x_i^*, y_i^*, z_i^*)$

on the ith segment corresponding to the parameter value $t_i^* \in [t_{i-1}, t_i]$. If Δs_i is small, the particle moves approximately in the direction of the unit tangent vector $\mathbf{T}(t_i^*)$ when it moves from P_{i-1} to P_i along the smooth curve C. Moreover, since at t_i^* the force $\mathbf{F}$ is continuous, it is approximated by $\mathbf{F}(x_i^*, y_i^*, z_i^*)$ for $t_{i-1} \leq t \leq t_i$. Thus we can approximate the work done by $\mathbf{F}$ in moving the particle along the curve C from P_{i-1} to P_i by

$$\Delta W_i = \mathbf{F}(x_i^*, y_i^*, z_i^*) \cdot \mathbf{T}(t_i^*)\, \Delta s_i \ . \tag{4.62}$$

So the total work done in moving the particle along the curve C is

$$W \approx \sum_{i=1}^{n} \mathbf{F}(x_i^*, y_i^*, z_i^*) \cdot \mathbf{T}(t_i^*)\, \Delta s_i \ . \tag{4.63}$$

We can see that this approximation should become better as n becomes larger. Therefore, we define the work done W by the force field $\mathbf{F}$ as

$$W = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbf{F}(x_i^*, y_i^*, z_i^*) \cdot \mathbf{T}(t_i^*)\, \Delta s_i = \int_C \mathbf{F} \cdot \mathbf{T}\, \mathrm{d}s \ . \tag{4.64}$$

Since the unit tangent vector $\mathbf{T} = \mathbf{r}'(t)/||\mathbf{r}'(t)||$, the above equation can be rewritten as

$$W = \int_a^b \left[\mathbf{F}(\mathbf{r}(t)) \cdot \frac{\mathbf{r}'(t)}{||\mathbf{r}'(t)||} \right] ||\mathbf{r}'(t)||\, \mathrm{d}t = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\, \mathrm{d}t \ . \tag{4.65}$$

It is often abbreviated as $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$, which occurs in other areas of physics as well. Thus we make the following definition for the line integral of any continuous vector field.

Definition 4.10. *Let* $\mathbf{F}$ *be a continuous vector field defined on a smooth curve* C *described by a vector function* $\mathbf{r}(t)$*,* $a \leq t \leq b$*. Then the* **line integral of F along** C *is*

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{T}\, \mathrm{d}s = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\, \mathrm{d}t \ . \tag{4.66}$$

Example 4.11. Find the work done by the force field $\mathbf{F}(x, y) = x^2\, \hat{\imath} - xy\, \hat{\jmath}$ in moving a particle along the quarter-circle $\mathbf{r}(t) = \cos t\, \hat{\imath} + \sin t\, \hat{\jmath}$ where $0 \leq t \leq \pi/2$.

Since $x = \cos t$ and $y = \sin t$, we have

$$\begin{aligned} \mathbf{F}(\mathbf{r}(t)) &= \cos^2 t\, \hat{\imath} - \cos t \sin t\, \hat{\jmath} \ , && (4.67) \\ \mathbf{r}'(t) &= -\sin t\, \hat{\imath} + \cos t\, \hat{\jmath} \ . && (4.68) \end{aligned}$$

Thus, the work done by the force is

$$\begin{aligned}
\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_0^{\pi/2} \mathbf{F}(\mathbf{r}(t))\cdot \mathbf{r}'(t)\,\mathrm{d}t \\
&= \int_0^{\pi/2} (\cos^2 t\,\hat{\imath} - \cos t \sin t\,\hat{\jmath})\cdot(-\sin t\,\hat{\imath} + \cos t\,\hat{\jmath})\,\mathrm{d}t \\
&= \int_0^{\pi/2} (-2\cos^2 t \sin t)\,\mathrm{d}t \\
&= -2\left[-\frac{1}{3}\cos^3 t\right]_0^{\pi/2} \\
&= -\frac{2}{3}\,. \qquad (4.69)
\end{aligned}$$

Observe that the work done by the force field $\mathbf{F}$ on the particle is negative because $\mathbf{F}$ opposes the motion of the particle. ∎

Example 4.12. Find the work done by the force field $\mathbf{F}(x,y,z) = (y - x^2)\,\hat{\imath} + (z - y^2)\,\hat{\jmath} + (x - z^2)\,\hat{k}$ in moving a particle along the curve C described by the vector equation $\mathbf{r}(t) = t\,\hat{\imath} + t^2\,\hat{\jmath} + t^3\,\hat{k}$ where $0 \le t \le 1$, Fig. 4.10.

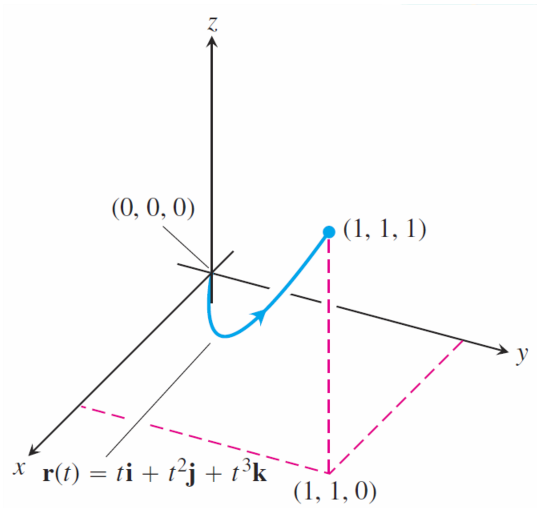


Figure 4.10: The curve of Example 4.12.

Since $x = t$, $y = t^2$, and $z = t^3$, we have

$$\begin{aligned}
\mathbf{F}(\mathbf{r}(t)) &= (t^3 - t^4)\,\hat{\jmath} + (t - t^6)\,\hat{k}\,, \\
\mathbf{r}'(t) &= \hat{\imath} + 2t\,\hat{\jmath} + 3t^2\,\hat{k}\,. \qquad (4.70)
\end{aligned}$$

Thus the work done by the force is

$$\begin{aligned}
\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_0^1 \mathbf{F}(\mathbf{r}(t))\cdot \mathbf{r}\,'(t)\,\mathrm{d}t \\
&= \int_0^1 [(t^3-t^4)\,\hat{\jmath} + (t-t^6)\,\hat{k}]\cdot(\hat{\imath}+2t\,\hat{\jmath}+3t^2\,\hat{k})\,\mathrm{d}t \\
&= \int_0^1 (-3t^8-2t^5+2t^4+3t^3)\,\mathrm{d}t \\
&= \left[-\frac{1}{3}t^9-\frac{1}{3}t^6+\frac{2}{5}t^5+\frac{3}{4}t^4\right]_0^1 \\
&= \frac{29}{60}\,. \qquad (4.71)
\end{aligned}$$

■

Finally, we explore the connection between line integrals of vector fields and line integrals of scalar fields. Suppose a vector field $\mathbf{F}$ in space is defined by $\mathbf{F}(x,y,z) = P(x,y,z)\,\hat{\imath} + Q(x,y,z)\,\hat{\jmath} + R(x,y,z)\,\hat{k}$. We use Eq. (4.66) to compute its line integral along the curve C

$$\begin{aligned}
\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_a^b \mathbf{F}(\mathbf{r}(t))\cdot \mathbf{r}\,'(t)\,\mathrm{d}t \\
&= \int_a^b (P\,\hat{\imath}+Q\,\hat{\jmath}+R\,\hat{k})\cdot[x'(t)\,\hat{\imath}+y'(t)\,\hat{\jmath}+z'(t)\,\hat{k}]\mathrm{d}t \\
&= \int_a^b [P\,x'(t)+Q\,y'(t)+R\,z'(t)]\,\mathrm{d}t\,. \qquad (4.72)
\end{aligned}$$

But the last integral is just the line integral $\int_C P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z$. So we have

$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \int_C P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z\,. \qquad (4.73)$$

The above equation implies that $\int_{-C} \mathbf{F}\cdot \mathrm{d}\mathbf{r} = -\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r}$ because the unit tangent vector $\mathbf{T}$ is replaced by its negative when C is replaced by $-C$. It is in contrast with the fact that the line integral of a scalar field with respect to arc length doesn't change sign when the orientation of the curve is reversed.

4.2 Independence of Path and Conservative Fields

Suppose C_1 and C_2 are two piecewise-smooth curves (which are called **paths**) with the same initial point A and terminal point B in an open region D in space. In general, the line integrals of a vector field $\mathbf{F}$ defined on D along these two curves are not equal, $\int_{C_1} \mathbf{F}\cdot \mathrm{d}\mathbf{r} \neq \int_{C_2} \mathbf{F}\cdot \mathrm{d}\mathbf{r}$. However, for some special vector fields, the value of the integral $\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r}$ is the same for all paths from A to B.

Definition 4.13. *Let* $\mathbf{F}$ *be a continuous vector field on an open region* D. *A line integral* $\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r}$ *is said to be* **independent of path** *in* D *if*

$$\int_{C_1} \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \int_{C_2} \mathbf{F}\cdot \mathrm{d}\mathbf{r} \tag{4.74}$$

for any two paths in D *that have the same initial and terminal points.*

Recall that a vector field $\mathbf{F}$ is a conservative vector field on a region D if $\mathbf{F} = \nabla\phi$ for some scalar function ϕ on D known as a potential function for $\mathbf{F}$. The following theorem tells us that the line integral of any conservative vector field is independent of path.

Theorem 4.14. Fundamental Theorem for Line Integrals
Let $\mathbf{F} = \nabla\phi$ be a continuous vector field on an open region D where ϕ is a differentiable potential function for $\mathbf{F}$. If C is a piecewise-smooth curve lying in D given by the vector function $\mathbf{r}(t) = x(t)\,\hat{\imath} + y(t)\,\hat{\jmath} + z(t)\,\hat{k}$, $a \le t \le b$, then

$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \int_C \nabla\phi\cdot \mathrm{d}\mathbf{r} = \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a))\ . \tag{4.75}$$

Note that this theorem is valid in any number of dimensions.

Proof. For a smooth curve C in space, $\phi(x,y,z)$ is a differentiable function of t. Since ϕ_x, ϕ_y, and ϕ_z were assumed to be continuous, applying the chain rule yields

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\partial\phi}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial\phi}{\partial y}\frac{\mathrm{d}\phi}{\mathrm{d}t} + \frac{\partial\phi}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t} = \nabla\phi\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\ . \tag{4.76}$$

Therefore, we have

$$\begin{aligned}
\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_C \nabla\phi\cdot \mathrm{d}\mathbf{r} \\
&= \int_a^b \nabla\phi\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t \\
&= \int_a^b \frac{\mathrm{d}}{\mathrm{d}t}[\phi(\mathbf{r}(t))]\,\mathrm{d}t \\
&= [\phi(\mathbf{r}(t))]_a^b \\
&= \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a))
\end{aligned} \tag{4.77}$$

by the Fundamental Theorem of Calculus. The theorem is also true for piecewise-smooth curves. This can be seen by subdividing into a finite number of smooth curves and adding the resulting integrals. ∎

Example 4.15.

(a) Prove that $\mathbf{F}(x,y) = 2xy\,\hat{\imath} + x^2\,\hat{\jmath}$ is conservative by showing that it is the gradient of the potential function $\phi(x,y) = x^2 y$.

(b) Use the result of part (a) to evaluate the line integral $\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r}$, where C is any piecewise-smooth curve from $(1,-2)$ to $(3,1)$.

(a) The gradient of ϕ is equal to

$$\nabla\phi(x,y) = \frac{\partial}{\partial x}(x^2y)\,\hat\imath + \frac{\partial}{\partial y}(x^2y)\,\hat\jmath = 2xy\,\hat\imath + x^2\,\hat\jmath = \mathbf{F}(x,y)\ . \tag{4.78}$$

We conclude that $\mathbf{F}$ is conservative.

(b) By the Fundamental Theorem for Line Integrals, the line integral $\int_C \mathbf{F}\cdot d\mathbf{r}$ depends only on the endpoints of the curve C since $\mathbf{F}$ is conservative. Using Eq. (4.75), we have

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \phi(3,1) - \phi(-1,2) = (3)^2(1) - (-1)^2(2) = 7\ . \tag{4.79}$$

■

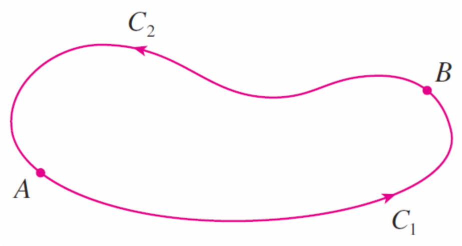


Figure 4.11: A closed curve.

A curve is called **closed** if its initial and terminal points are the same. Suppose $\int_C \mathbf{F}\cdot d\mathbf{r}$ is independent of path in an open region D and C is any closed path in D. We can pick any two points A and B on C and regard C as being composed of the path C_1 from A to B followed by the path C_2 from B to A, Fig. 4.11). Then

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} + \int_{C_2} \mathbf{F}\cdot d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} - \int_{-C_2} \mathbf{F}\cdot d\mathbf{r} = 0 \tag{4.80}$$

since C_1 and $-C_2$ have the same initial and terminal points.

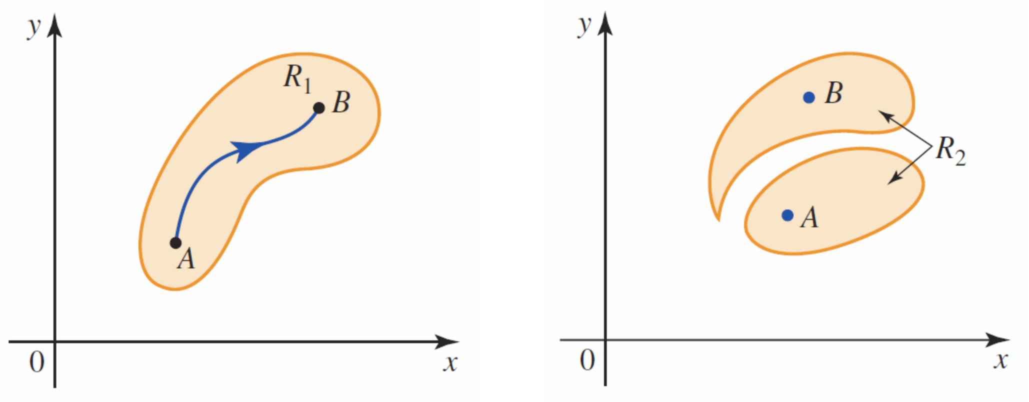
Conversely, suppose $\int_C \mathbf{F}\cdot d\mathbf{r} = 0$ for every closed path C in D. Take any two paths C_1 and C_2 from A to B in D and let C be the curve consisting of C_1 followed by $-C_2$. Then

$$0 = \int_C \mathbf{F}\cdot d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} + \int_{-C_2} \mathbf{F}\cdot d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} - \int_{C_2} \mathbf{F}\cdot d\mathbf{r}\ , \tag{4.81}$$

which implies $\int_{C_1} \mathbf{F}\cdot d\mathbf{r} = \int_{C_2} \mathbf{F}\cdot d\mathbf{r}$. Thus we have proved the following theorem

Theorem 4.16. Let $\mathbf{F}$ be a continuous vector field on an open region D. The line integral $\int_C \mathbf{F}\cdot d\mathbf{r}$ is independent of path if and only if $\int_C \mathbf{F}\cdot d\mathbf{r} = 0$ for every closed path C in D. ■

The following theorem says that a vector field whose line integral is independent of path must be a conservative vector field. In this theorem, we assume that the region D is **open**, which means that every point in D is the center of a disk that lies entirely in D. In addition, we assume that D is **connected**, which means that any two points in D can be connected by a path that lies in D, Fig. 4.12).



(**a**) The plane region R_1 is connected. (**b**) The region R_2 is not connected.

Figure 4.12: Connected region and disconnected region.

Theorem 4.17. Independence of Path and Conservative Vector Fields
Let $\mathbf{F}$ be a continuous vector field on an open connected region D. The line integral $\int_C \mathbf{F}\cdot d\mathbf{r}$ is independent of path if and only if $\mathbf{F}$ is a conservative vector field on D, $\mathbf{F} = \nabla\phi$ for some scalar function ϕ.

Proof. If $\mathbf{F}$ is conservative, then the Fundamental Theorem for Line Integrals implies that the line integral is independent of path. We will prove the converse for the case that D is a plane region; the proof for the three-dimensional case is similar. Let (a, b) be a fixed point in D and (x, y) be any point in D. If C is any path from (a, b) to (x, y), we define the potential function by

$$\phi(x,y) = \int_{(a,b)}^{(x,y)} \mathbf{F}\cdot d\mathbf{r}\ . \tag{4.82}$$

Since D is open, there exists a disk contained in D with center (x, y). Pick any point (x_1, y) in the disk with $x_1 < x$. Since the line integral is assumed to be independent of path, we can choose C to be the path consisting of any path C_1 from (a, b) to (x_1, y) followed by the horizontal line segment C_2 from (x_1, y) to (x, y) as shown in Fig. 4.13. Then we have

$$\phi(x,y) = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} + \int_{C_2} \mathbf{F}\cdot d\mathbf{r} = \int_{(a,b)}^{(x_1,y)} \mathbf{F}\cdot d\mathbf{r} + \int_{C_2} \mathbf{F}\cdot d\mathbf{r}\ . \tag{4.83}$$

Since the first of the two integrals of the right is independent of x, we obtain

$$\frac{\partial}{\partial x}\phi(x,y) = 0 + \frac{\partial}{\partial x}\int_{C_2} \mathbf{F}\cdot d\mathbf{r}\ . \tag{4.84}$$

If we let $\mathbf{F}(x,y) = P(x,y)\,\hat{\imath} + Q(x,y)\,\hat{\jmath}$, then

$$\int_{C_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{C_2} P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y \ . \tag{4.85}$$

We can represent C_2 by the parametric equations $x = t$, $y = y$ where $x_1 \leq t \leq x$ and y is a constant. This gives $\mathrm{d}x = \mathrm{d}t$ and $\mathrm{d}y = 0$. Thus,

$$\begin{aligned} \frac{\partial}{\partial x}\phi(x,y) &= \frac{\partial}{\partial x}\left[\int_{C_2} P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y\right] \\ &= \frac{\partial}{\partial x}\int_{x_1}^{x} P(t,y)\,\mathrm{d}t \\ &= P(x,y) \end{aligned} \tag{4.86}$$

by the Fundamental Theorem of Calculus.

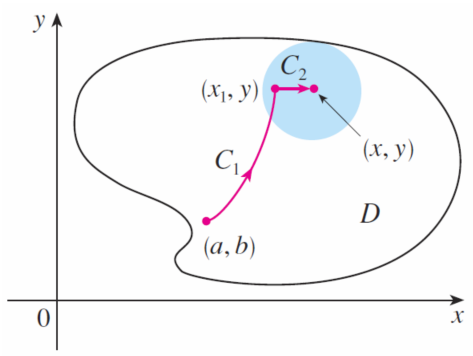


Figure 4.13: A curve with a horizontal segment.

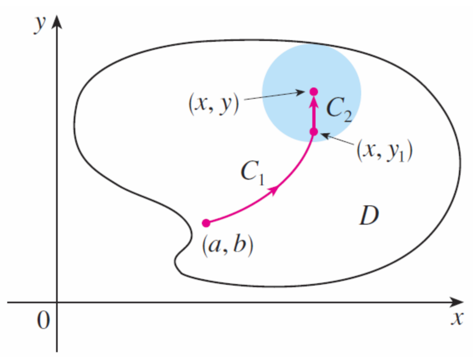


Figure 4.14: A curve with a vertical segment.

Similarly, if C is the path with a vertical line segment as shown in Fig. 4.14, then we can show that

$$\frac{\partial}{\partial y}\phi(x,y) = Q(x,y) \ . \tag{4.87}$$

Hence,

$$\mathbf{F}(x,y) = P(x,y)\,\hat{\imath} + Q(x,y)\,\hat{\jmath} = \frac{\partial \phi}{\partial x}\hat{\imath} + \frac{\partial \phi}{\partial y}\hat{\jmath} \ , \tag{4.88}$$

which implies that $\mathbf{F}$ is conservative. ∎

Now we see that it is convenient to evaluate line integrals of conservative vector fields. However, how can we determine whether a vector field is conservative? Suppose $\mathbf{F}(x, y) = P(x, y)\,\hat{\imath} + Q(x, y)\,\hat{\jmath}$ is a conservative vector field on an open region D, where P and Q have continuous first-order partial derivatives. Then there exists a scalar function ϕ such that $\mathbf{F} = \nabla\phi$

$$P = \frac{\partial \phi}{\partial x} \quad \text{and} \quad Q = \frac{\partial \phi}{\partial y}\ . \tag{4.89}$$

Since $\partial P/\partial y$ and $\partial Q/\partial x$ were assumed to be continuous, it follows from Clairaut's Theorem that

$$\frac{\partial P}{\partial y} = \frac{\partial^2 \phi}{\partial y \partial x} = \frac{\partial^2 \phi}{\partial x \partial y} = \frac{\partial Q}{\partial x}\ . \tag{4.90}$$

Thus we have proved the following theorem.

Theorem 4.18. Let $\mathbf{F}(x, y) = P(x, y)\,\hat{\imath} + Q(x, y)\,\hat{\jmath}$ be a conservative vector field on an open region D in the plane, where P and Q have continuous first-order partial derivatives. Then

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{4.91}$$

for all (x, y) in D. ■

The converse of this theorem is true only for a special type of region. To describe such region, we need the concept of a **simple curve**, which is a curve that does not intersect itself anywhere except at the endpoints (see Fig. 4.15).

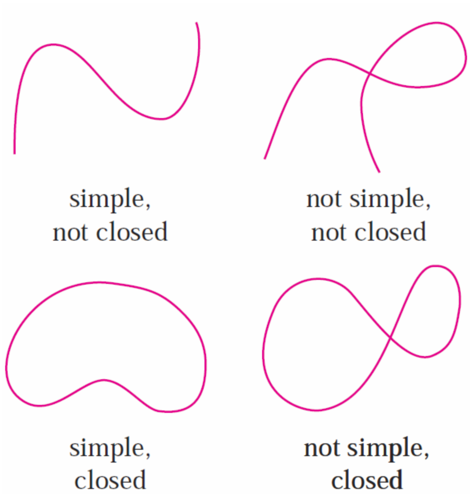


Figure 4.15: Simple curves.

The following theorem, which is a partial converse, requires a stronger condition than an open connected region. A connected region D in the plane is a **simply-connected region** if every simple closed curve C in D encloses

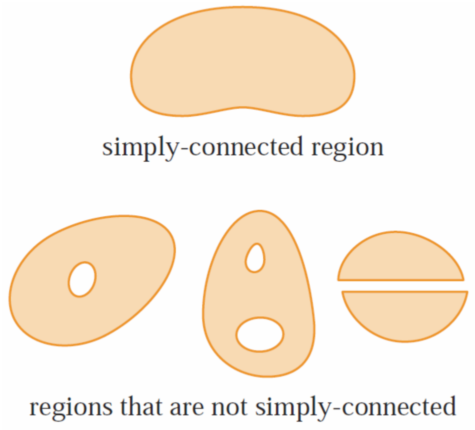


Figure 4.16: Simply-connected region.

only points that are in D. It means that a simply-connected region is a connected region contains no hole, as illustrated in Fig. 4.16.

Theorem 4.19. Let $\mathbf{F}(x, y) = P(x, y)\,\hat{\imath} + Q(x, y)\,\hat{\jmath}$ be a vector field on an open simply-connected region D in the plane. If P and Q have continuous first-order partial derivatives on D and

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{4.92}$$

for all (x, y) in D, then $\mathbf{F}$ is conservative. ■

This theorem provides us a test to determine whether a vector field in the plane is conservative. The proof will be discussed in the next section as a consequence of Green's Theorem.

Example 4.20. Determine whether the following vector fields are conservative.

(a) $\mathbf{F}(x, y) = (3 + 2xy)\,\hat{\imath} + (x^2 - 3y^2)\,\hat{\jmath}$

(b) $\mathbf{F}(x, y) = 2xy^2\,\hat{\imath} + x^2 y\,\hat{\jmath}$

(a) Let $P(x, y) = 3 + 2xy$ and $Q(x, y) = x^2 - 3y^2$. Then,

$$\frac{\partial P}{\partial y} = 2x = \frac{\partial Q}{\partial x} \ . \tag{4.93}$$

So we see that Eq. (4.92) is satisfied for all points (x, y) in the plane, which is open and simply-connected. Therefore, we conclude that $\mathbf{F}$ is conservative.

(b) Let $P(x, y) = 2xy^2$ and $Q(x, y) = x^2 y$. Then

$$\frac{\partial P}{\partial y} = 4xy \quad \text{and} \quad \frac{\partial Q}{\partial x} = 2xy \ . \tag{4.94}$$

Since $\partial P/\partial y \neq \partial Q/\partial x$ in general, we see that Eq. (4.92) is not satisfied for all points (x, y) in any open simply connected region in the plane. Therefore, we conclude that $\mathbf{F}$ is not conservative.

■

Theorem 4.19 told us the criterion for determining whether a vector field $\mathbf{F}$ is conservative, but it did not tell us how to find the potential function f such that $\mathbf{F} = \nabla\phi$. The proof of Theorem 4.17 gives us a hint for finding ϕ. We can find ϕ using partial integration as illustrated in the next example.

Example 4.21.

(a) For $\mathbf{F}(x, y) = (3 + 2xy)\,\hat{\imath} + (x^2 - 3y^2)\,\hat{\jmath}$, find a potential function ϕ such that $\mathbf{F} = \nabla\phi$.

(b) Use the result of part (a) to evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is the curve given by $\mathbf{r}(t) = e^t \sin t\,\hat{\imath} + e^t \cos t\,\hat{\jmath}$ with $0 \le t \le \pi$.

(a) From Example 4.20, we know that $\mathbf{F}$ is conservative. So there exists a function ϕ such that $\mathbf{F} = \nabla\phi$

$$(3 + 2xy)\,\hat{\imath} + (x^2 - 3y^2)\,\hat{\jmath} = \frac{\partial\phi}{\partial x}\,\hat{\imath} + \frac{\partial\phi}{\partial y}\,\hat{\jmath}\ . \tag{4.95}$$

This will occur if and only if

$$\frac{\partial\phi}{\partial x} = 3 + 2xy \tag{4.96}$$

$$\frac{\partial\phi}{\partial y} = x^2 - 3y^2\ . \tag{4.97}$$

Integrating Eq. (4.96) with respect to x, we have

$$\phi(x, y) = 3x + x^2 y + \gamma(y) \tag{4.98}$$

where $\gamma(y)$ is the constant of integration. Next we differentiate both sides of Eq. (4.98) with respect to y and obtain

$$\frac{\partial\phi}{\partial y} = x^2 + \gamma'(y)\ . \tag{4.99}$$

Comparing it with (4.97) yields

$$\gamma'(y) = -3y^2\ . \tag{4.100}$$

Integrating $g'(y)$ with respect to y gives

$$\gamma(y) = -y^3 + K \tag{4.101}$$

where K is a constant. Putting this in Eq. (4.98), we have

$$\phi(x, y) = 3x + x^2 y - y^3 + K\ . \tag{4.102}$$

(b) To use the Fundamental Theorem for Line Integrals, all we need to know are the initial and terminal points of C, namely, $\mathbf{r}(0) = (0, 1)$ and $\mathbf{r}(\pi) = (0, -e^{\pi})$. Using Eq. (4.75), we have

$$\begin{aligned} \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \phi(0, -e^{\pi}) - \phi(0, 1) \\ &= e^{3\pi} + 1 \ . \end{aligned} \tag{4.103}$$

■

The following theorem gives us a test to determine whether a vector field in space is conservative.

Theorem 4.22. Test for a Conservative Vector Field in Space
Let $\mathbf{F}(x, y, z) = P(x, y, z)\,\hat{\imath} + Q(x, y, z)\,\hat{\jmath} + R(x, y, z)\,\hat{k}$ be a vector field on an open simply-connected region D in space. If P, Q, and R have continuous first-order partial derivatives on D and

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \ , \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y} \ , \qquad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z} \tag{4.104}$$

for all (x, y, z) in D, then $\mathbf{F}$ is conservative. In other words, $\mathbf{F}$ is conservative if $\operatorname{curl}\mathbf{F} = \mathbf{0}$ for all points in D. ■

The next example shows that the technique of finding the potential function for a conservative vector field in space is similar to that for a conservative vector field in the plane.

Example 4.23. Show that $\mathbf{F}(x, y, z) = (e^x \cos y + yz)\,\hat{\imath} + (xz - e^x \sin y)\,\hat{\jmath} + (xy + z)\,\hat{k}$ is conservative, and find a potential function ϕ such that $\mathbf{F} = \nabla\phi$.

Let $P(x, y, z) = e^x \cos y + yz$, $Q(x, y, z) = xz - e^x \sin y$ and $R(x, y, z) = xy + z$. Then,

$$\frac{\partial P}{\partial y} = -e^x \sin y + z = \frac{\partial Q}{\partial x} \ , \qquad \frac{\partial Q}{\partial z} = x = \frac{\partial R}{\partial y} \ , \qquad \frac{\partial R}{\partial x} = y = \frac{\partial P}{\partial z} \ . \tag{4.105}$$

So we see that Eq. (4.104) is satisfied for all points (x, y, z) in space, which is open and simply-connected. Therefore, we conclude that $\mathbf{F}$ is conservative.

If there exists a function ϕ such that $\mathbf{F} = \nabla\phi$, then

$$\frac{\partial \phi}{\partial x} = e^x \cos y + yz \tag{4.106}$$

$$\frac{\partial \phi}{\partial y} = xz - e^x \sin y \tag{4.107}$$

$$\frac{\partial \phi}{\partial z} = xy + z \ . \tag{4.108}$$

Integrating Eq. (4.106) with respect to x, we obtain

$$\phi(x, y, z) = e^x \cos y + xyz + \gamma(y, z) \tag{4.109}$$

where $\gamma(y, z)$ is the constant of integration. Then differentiating Eq. (4.109) with respect to y, we have

$$\frac{\partial \phi}{\partial y} = -e^x \sin y + xz + \frac{\partial \gamma}{\partial y} \ . \tag{4.110}$$

and comparing with Eq. (4.107) yields

$$\frac{\partial \gamma}{\partial y} = 0 \ . \tag{4.111}$$

Thus g is a function of z alone, which implies

$$\phi(x, y, z) = e^x \cos y + xyz + \eta(z) \ . \tag{4.112}$$

Finally, differentiating Eq. (4.112) with respect to z and comparing with Eq. (4.108), we obtain

$$\frac{\partial \eta}{\partial z} = z \ , \tag{4.113}$$

and therefore

$$\eta(z) = \frac{1}{2}z^2 + K \ , \tag{4.114}$$

where K is a constant. Putting this into Eq. (4.112), we have

$$\phi(x, y, z) = e^x \cos y + xyz + \frac{1}{2}z^2 + K \ . \tag{4.115}$$

We have infinitely many potential functions of $\mathbf{F}$, one for each value of K. ■

We can use Fundamental Theorem for Line Integrals to derive one of the most important laws of physics: the Law of Conservation of Energy. Suppose an object of mass m is subjected to the action of a continuous force field $\mathbf{F}$ so that it moves from A to B along the piecewise smooth curve C, given by $\mathbf{r}(t)$ where $a \leq t \leq b$. According to Newton's Second Law of Motion, $\mathbf{F} = m\mathbf{r}''(t)$ where $\mathbf{a}(t) = \mathbf{r}''(t)$ is the acceleration of the particle. So the work done by the force $\mathbf{F}$ on the object as it moves from A to B along C is

$$\begin{aligned} W &= \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} \\ &= \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,\mathrm{d}t \\ &= \int_a^b m\mathbf{r}''(t) \cdot \mathbf{r}'(t)\,\mathrm{d}t \\ &= \frac{m}{2}\int_a^b \frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{r}'(t) \cdot \mathbf{r}'(t)]\,\mathrm{d}t \\ &= \frac{m}{2}\int_a^b \frac{\mathrm{d}}{\mathrm{d}t}||\mathbf{r}'(t)||^2\,\mathrm{d}t \\ &= \frac{m}{2}(||\mathbf{r}'(b)||^2 - ||\mathbf{r}'(a)||^2) \\ &= \frac{1}{2}m||\mathbf{v}(b)||^2 - \frac{1}{2}m||\mathbf{v}(a)||^2 \end{aligned} \tag{4.116}$$

where $\mathbf{v}(t) = \mathbf{r}'(t)$ is the velocity of the particle.

Recall that the kinetic energy of an object of mass m and speed v is $\frac{1}{2}mv^2$. So we can rewrite the above equation as

$$W = K(B) - K(A) \tag{4.117}$$

which says that the work done by the force $\mathbf{F}$ along the curve C is equal to the change in kinetic energy of the object at the endpoints of C.

Now let's further assume that $\mathbf{F}$ is a conservative force field, $\mathbf{F} = \nabla\phi$ for some scalar function ϕ. In physics, the potential energy of an object at the point (x, y, z) is defined as $U(x, y, z) = -\phi(x, y, z)$ so that $\mathbf{F} = -\nabla U$. Then, by the Fundamental Theorem for Line Integrals, we have

$$\begin{aligned} W &= \int_C \mathbf{F} \cdot d\mathbf{r} \\ &= -\int_C \nabla U \cdot d\mathbf{r} \\ &= -[U(\mathbf{r}(b)) - U(\mathbf{r}(a))] \\ &= U(A) - U(B)\ . \end{aligned} \tag{4.118}$$

Comparing the above results, we obtain $K(A) + U(A) = K(B) + U(B)$. It states that as an object moves from one point to another in a conservative force field, then the sum of its kinetic energy and potential energy keeps constant. This is known as the Law of Conservation of Energy and it is the reason why the vector field is called conservative.

Before moving on to the next section, we summarize the results we have developed for three-dimensional vector fields $\mathbf{F}(x, y, z) = P(x, y, z)\,\hat{\imath} + Q(x, y, z)\,\hat{\jmath} + R(x, y, z)\,\hat{k}$, where $P(x, y, z)$, $Q(x, y, z)$, and $R(x, y, z)$ are assumed to have continuous first-order partial derivatives on an open simply-connected region D in space. In such case, the following five statements are equivalent.

1. $\mathbf{F}(x, y, z)$ is conservative in D.
2. There exists a scalar function $\phi(x, y, z)$ on D such that $\mathbf{F} = \nabla\phi$.
3. $\int_C \mathbf{F} \cdot d\mathbf{r}$ is independent of path in D.
4. $\int_C \mathbf{F} \cdot d\mathbf{r} = 0$ for every piecewise-smooth closed curve C in D.
5. $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}\ ,\ \dfrac{\partial Q}{\partial z} = \dfrac{\partial R}{\partial y}\ ,\ \dfrac{\partial R}{\partial x} = \dfrac{\partial P}{\partial z}$ for all (x, y, z) in D.

4.3 Green's Theorem in the Plane

Green's theorem gives the connection between a line integral around a simple closed plane curve C and a double integral over the plane region D bounded by C as shown in Fig. 4.17. (So D is a simply-connected region since it must have no holes as it is bounded by a simple closed curve.) It

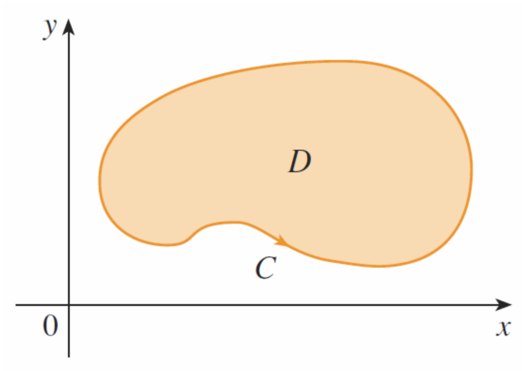


Figure 4.17: A region D.

is used extensively in many fields of physics such as in the analysis of fluid flows and in the theories of electricity and magnetism.

In stating Green's theorem, we adopt the convention that a simple closed curve C has a **positive orientation** if it is traversed in a counterclockwise direction. That is to say, if C is defined by the vector function $\mathbf{r}(t)$ where $a \leq t \leq b$, then the region D enclosed by C is always on the left as the curve C is traced out by the tip of the vector $\mathbf{r}(t)$. For illustration, Fig. 4.18 shows a simple closed curve with positive orientation and one with negative orientation, respectively.

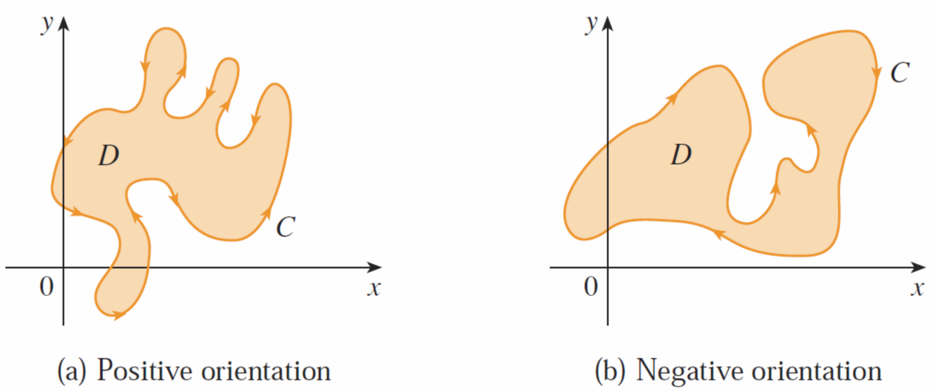


Figure 4.18: Orientations of boundaries of regions.

Note that here we use the notation

$$\oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y \tag{4.119}$$

to denote a line integral over a simple closed curve C with positive orientation.

Theorem 4.24. Green's Theorem Let C be a piecewise-smooth, simple closed curve in the plane with positive orientation and D be the region enclosed by C. If $P(x,y)$ and $Q(x,y)$ have continuous first-order partial derivatives on an open region that contains D, then

$$\oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A \,. \tag{4.120}$$

Proof. It is not easy to prove Green's theorem for general regions. So here we will prove it for the special case that the region is both an x-simple and a y-simple region. Such regions are called **simple regions**.

Let D be a simple region in the plane enclosed by a piecewise-smooth, simple closed curve C. By viewing D as a y-simple region, it can be described as

$$D = \{(x, y) \mid a \le x \le b,\ f_1(x) \le y \le f_2(x)\} \tag{4.121}$$

where f_1 and f_2 are continuous functions on $[a, b]$, as illustrated in Fig. 4.19. So we can write

$$\begin{aligned} \iint_D \frac{\partial P}{\partial y}\,\mathrm{d}A &= \int_a^b \int_{f_1(x)}^{f_2(x)} \frac{\partial}{\partial y}[P(x,y)]\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_a^b [P(x, f_2(x)) - P(x, f_1(x))]\,\mathrm{d}x \end{aligned} \tag{4.122}$$

where the last step follows from the Fundamental Theorem of Calculus.

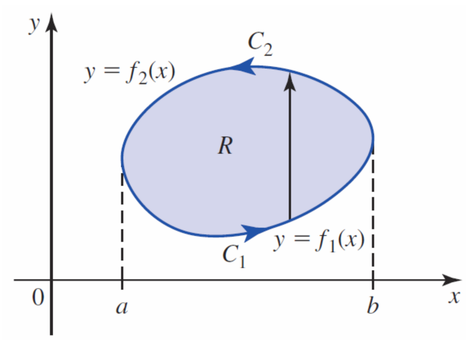


Figure 4.19: A simple region.

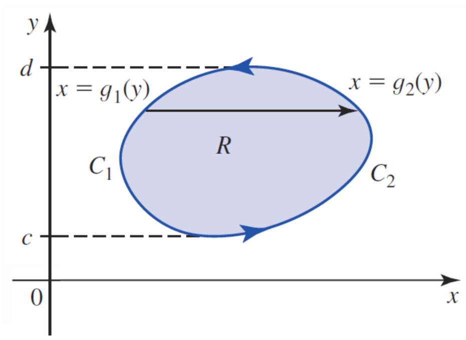


Figure 4.20: A simple region.

We can split the curve C into two parts C_1 and C_2 which are the graphs of the functions $f_1(x)$ and $f_2(x)$ with the orientations as shown in Fig. 4.19. Notice that C_1 goes from left to right whereas C_2 goes from right to left. Therefore, we have

$$\begin{aligned}
\oint_C P(x,y)\,\mathrm{d}x &= \int_{C_1} P(x,y)\,\mathrm{d}x + \int_{C_2} P(x,y)\,\mathrm{d}x \\
&= \int_a^b P(x,f_1(x))\,\mathrm{d}x + \int_b^a P(x,f_2(x))\,\mathrm{d}x \\
&= \int_a^b P(x,f_1(x))\,\mathrm{d}x - \int_a^b P(x,f_2(x))\,\mathrm{d}x \\
&= \int_a^b [P(x,f_1(x)) - P(x,f_2(x))]\,\mathrm{d}x \ . \qquad (4.123)
\end{aligned}$$

Comparing this with Eq. (4.122), we obtain

$$\oint_C P(x,y)\,\mathrm{d}x = -\iint_D \frac{\partial P}{\partial y}\,\mathrm{d}A \ . \qquad (4.124)$$

On the other hand, by viewing D as an x-simple region, it can be described as

$$D = \{(x,y) \mid g_1(y) \le x \le g_2(y),\ c \le y \le d\} \qquad (4.125)$$

where g_1 and g_2 are continuous functions on $[c,d]$, as illustrated in Fig. 4.20. In a similar manner, we can show that

$$\oint_C Q(x,y)\,\mathrm{d}y = \iint_D \frac{\partial Q}{\partial x}\,\mathrm{d}A \ . \qquad (4.126)$$

Adding together Eq. (4.124) and (4.126), we obtain

$$\begin{aligned}
\oint_C P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y &= \oint_C P(x,y)\,\mathrm{d}x + \oint_C Q(x,y)\,\mathrm{d}y \\
&= -\iint_D \frac{\partial P}{\partial y}\,\mathrm{d}A + \iint_D \frac{\partial Q}{\partial x}\,\mathrm{d}A \\
&= \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}A \ , \qquad (4.127)
\end{aligned}$$

which is Green's Theorem for the simple region D with boundary C. ■

Example 4.25. Evaluate $\int_C x^4\,\mathrm{d}x + xy\,\mathrm{d}y$, where C is the triangular curve made up of the line segments from $(0,0)$ to $(1,0)$, from $(1,0)$ to $(0,1)$, and from $(0,1)$ to $(0,0)$.

As shown in Fig. 4.21, the region D enclosed by the curve C is simple and C is a piecewise-smooth, simple closed curve with positive orientation. So we can use Green's Theorem to evaluate the line integral. Using Green's Theorem with $P(x,y) = x^4$ and $Q(x,y) = xy$, we have

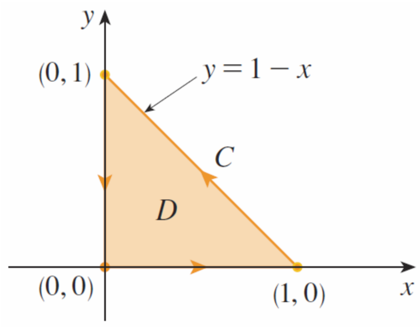


Figure 4.21: A triangular curve.

$$\begin{aligned}
\oint_C x^4\,\mathrm{d}x + xy\,\mathrm{d}y &= \iint_D \left[\frac{\partial}{\partial x}(xy) - \frac{\partial}{\partial y}(x^4)\right]\mathrm{d}A \\
&= \int_0^1 \int_0^{1-x} y\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_0^1 \left[\frac{1}{2}y^2\right]_0^{1-x} \mathrm{d}x \\
&= \frac{1}{2}\int_0^1 (1-x)^2\,\mathrm{d}x \\
&= \frac{1}{6}\;. \qquad (4.128)
\end{aligned}$$

■

Example 4.26. Evaluate $\int_C (y^2 + \tan x)\,\mathrm{d}x + (x^3 + 2xy + \sqrt{y})\,\mathrm{d}y$, where C is the circle $x^2 + y^2 = 4$ oriented counterclockwise.

Obviously, the circle C is a piecewise-smooth, simple closed curve with positive orientation. The simple region D bounded by C is the disk $D = \{(x, y) \mid x^2 + y^2 \leq 4\}$. So we can use Green's Theorem to evaluate the line integral. Using Green's Theorem with $P(x, y) = y^2 + \tan x$ and $Q(x, y) = x^3 + 2xy + \sqrt{y}$, we have

$$\begin{aligned}
\oint_C x^4\,\mathrm{d}x + xy\,\mathrm{d}y &= \iint_D \left[\frac{\partial}{\partial x}(x^3 + 2xy + \sqrt{y}) - \frac{\partial}{\partial y}(y^2 + \tan x)\right]\mathrm{d}A \\
&= \iint_D 3x^2\,\mathrm{d}A \\
&= 3\int_0^{2\pi}\int_0^2 (r\cos\theta)^2\, r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= 3\left(\int_0^2 r^3\,\mathrm{d}r\right)\left(\int_0^{2\pi}\cos^2\theta\,\mathrm{d}\theta\right) \\
&= 3\left[\frac{1}{4}r^4\right]_0^2 \left[\frac{1}{2}\theta + \frac{1}{4}\sin 2\theta\right]_0^{2\pi} \\
&= 12\pi\;. \qquad (4.129)
\end{aligned}$$

■

In last two examples, we see that evaluating the line integral directly was more complicated than evaluating the corresponding double integral. However, sometimes it's easier to evaluate the line integral, and Green's Theorem is used in the reverse direction. For example, we can apply Green's Theorem in the reverse direction to compute the area of a plane region D. Since the area of D is equal to $\iint_D \mathrm{d}A$, we want to choose P and Q such that

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1 \ . \tag{4.130}$$

There are several choices of P and Q satisfying the above requirement

$$P(x,y) = 0 \ , \qquad Q(x,y) = x \ ; \tag{4.131}$$
$$P(x,y) = -y \ , \qquad Q(x,y) = 0 \ ; \tag{4.132}$$
$$P(x,y) = -\frac{1}{2}y \ , \qquad Q(x,y) = \frac{1}{2}x \ . \tag{4.133}$$

Applying Green's Theorem, we find that the area of the region D is given by

$$A = \oint_C x \,\mathrm{d}y = -\oint_C y \,\mathrm{d}x = \frac{1}{2}\oint_C x \,\mathrm{d}y - y \,\mathrm{d}x \ . \tag{4.134}$$

Example 4.27. Find the area enclosed by the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$.

The ellipse C can be described by the parametric equations

$$x = a\cos t, \qquad y = b\sin t, \qquad 0 \le t \le 2\pi \ . \tag{4.135}$$

Observe that the ellipse C is a piecewise-smooth, simple closed curve traversed in the counterclockwise direction as t increases from 0 to 2π. Using Eq. (4.134), the area A of the ellipse is given by

$$\begin{aligned} A &= \frac{1}{2}\oint_C x \,\mathrm{d}y - y \,\mathrm{d}x \\ &= \frac{1}{2}\int_0^{2\pi} [(a\cos t)(b\cos t) - (b\sin t)(-a\sin t)] \,\mathrm{d}t \\ &= \frac{ab}{2}\int_0^{2\pi} \mathrm{d}t \\ &= \pi ab \ . \end{aligned} \tag{4.136}$$

■

We have proved Green's Theorem only for the case in which D is a simple region. However, the theorem can be extended to the case in which D is a finite union of simple regions. For example, the region D shown in Fig. 4.22 is the union of two simple regions D_1 and D_2 so that we can write $D = D_1 \cup D_2$.

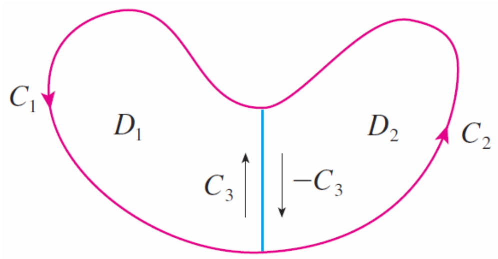


Figure 4.22: A region divided into two simple regions.

The boundary of D_1 is $C_1 \cup C_3$ and the boundary of D_2 is $C_2 \cup (-C_3)$. Applying Green's Theorem to D_1 and D_2, we have

$$\begin{aligned} \oint_{C_1 \cup C_3} P\,\mathrm{d}x + Q\,\mathrm{d}y &= \iint_{D_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A \\ \oint_{C_2 \cup (-C_3)} P\,\mathrm{d}x + Q\,\mathrm{d}y &= \iint_{D_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A\,. \end{aligned} \tag{4.137}$$

Adding these two equations and noting that the line integrals along C_3 and $-C_3$ cancel each other, we obtain

$$\begin{aligned} &\oint_{C_1 \cup C_3} P\,\mathrm{d}x + Q\,\mathrm{d}y + \oint_{C_2 \cup (-C_3)} P\,\mathrm{d}x + Q\,\mathrm{d}y \\ &= \oint_{C_1 \cup C_2} P\,\mathrm{d}x + Q\,\mathrm{d}y \\ &= \iint_{D} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A\,, \end{aligned} \tag{4.138}$$

which is Green's Theorem for the region $D = D_1 \cup D_2$ with boundary $C = C_1 \cup C_2$.

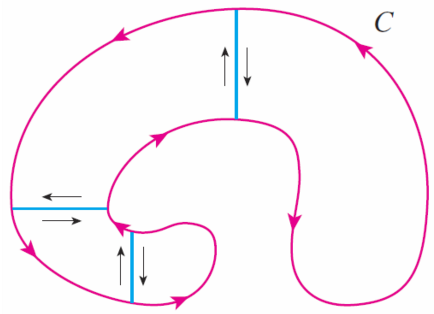


Figure 4.23: A region divided into three simple regions.

We can use similar arguments to prove Green's Theorem for the case in which D is the union of any finite number of nonoverlapping simple regions as shown in Fig. 4.23.

Example 4.28. Evaluate $\int_C y^2\,\mathrm{d}x + 3xy\,\mathrm{d}y$, where C is the positively-oriented boundary of the semiannular region D bounded by the semicircles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$ and the x-axis as shown in Fig. 4.24.

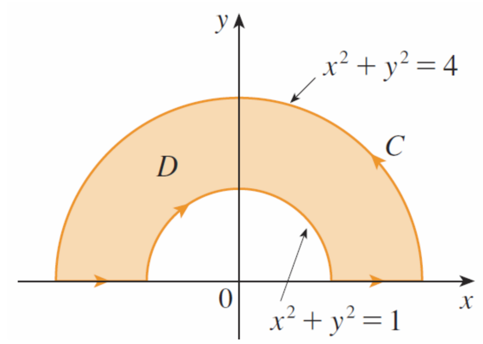


Figure 4.24: Region for example 4.28.

The region D is not simple, but the y-axis divides it into two simple regions. It can be expressed in polar coordinates as

$$D = \{(r, \theta) \mid 1 \leq r \leq 2,\ 0 \leq \theta \leq \pi\}\ . \tag{4.139}$$

Using Green's Theorem with $P(x, y) = y^2$ and $Q(x, y) = 3xy$, we have

$$\begin{aligned}
\oint_C y^2\,\mathrm{d}x + 3xy\,\mathrm{d}y &= \iint_D \left[\frac{\partial}{\partial x}(3xy) - \frac{\partial}{\partial y}(y^2)\right] \mathrm{d}A \\
&= \iint_D y\,\mathrm{d}A \\
&= \int_0^\pi \int_1^2 (r\sin\theta)\, r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \left(\int_1^2 r^2\,\mathrm{d}r\right)\left(\int_0^\pi \sin\theta\,\mathrm{d}\theta\right) \\
&= \left[\frac{1}{3}r^3\right]_1^2 [-\cos\theta]_0^\pi \\
&= \frac{14}{3}\ .
\end{aligned} \tag{4.140}$$

■

Green's Theorem can be also extended to regions with holes, which are regions that are not simply-connected. For example, the annular region D shown in Fig. 4.25 has a boundary C consists of two separate simple closed curve C_1 and C_2. By convention, the region D is always on the left as the boundary curve C is traversed in the positive direction. Thus the positive direction is counterclockwise for the outer curve C_1 but clockwise for the inner curve C_2.

The region D can be divided into two simple regions D' and D'' by means of two boundary lines as shown in Fig. 27. Applying Green's Theorem to D' and D'' separately, we obtain

$$\begin{aligned}
\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A &= \iint_{D'} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A + \iint_{D''} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A \\
&= \oint_{\partial D'} P\,\mathrm{d}x + Q\,\mathrm{d}y + \oint_{\partial D''} P\,\mathrm{d}x + Q\,\mathrm{d}y\ ,
\end{aligned} \tag{4.141}$$

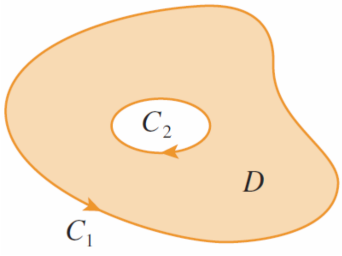


Figure 4.25: A non-simply connected region.

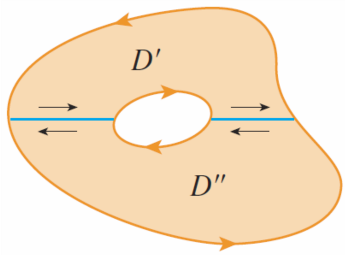


Figure 4.26: Dividing the region of Fig. 4.25.

where $\partial D'$ and $\partial D''$ denote the boundaries of D' and D'', respectively. Since the line integrals along the common boundary lines are traversed in opposite direction, they cancel out and we obtain

$$\begin{aligned}\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A &= \oint_{C_1} P\,\mathrm{d}x + Q\,\mathrm{d}y + \oint_{C_2} P\,\mathrm{d}x + Q\,\mathrm{d}y \\ &= \oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y\ , \end{aligned} \tag{4.142}$$

which is the Green's Theorem for the region D.

Example 4.29. Let C be an arbitrary piecewise-smooth, simple closed curve in the plane with positive orientation that does not pass through the origin. Show that

$$\oint_C -\frac{y}{(x^2+y^2)}\,\mathrm{d}x + \frac{x}{(x^2+y^2)}\,\mathrm{d}y \tag{4.143}$$

is equal to zero if C does not enclose the origin but it is equal to 2π if C encloses the origin.

Because C is an arbitrary curve, we cannot evaluate the given line integral directly. Thus we will evaluate it by using Green's Theorem.

We first consider the case that C does not enclose the origin. Let D be the region bounded by C. The functions $P(x,y) = -y/(x^2+y^2)$ and $Q(x,y) = x/(x^2+y^2)$ have continuous first-order partial derivatives in D. Moreover,

$$\begin{aligned}\frac{\partial Q}{\partial x} &= \frac{(x^2+y^2)(1)-(x)(2x)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2} \\ \frac{\partial P}{\partial y} &= \frac{(x^2+y^2)(-1)-(-y)(2y)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{\partial Q}{\partial x}\ . \end{aligned} \tag{4.144}$$

Applying Green's Theorem, we have

$$\oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}A = \iint_D 0\,\mathrm{d}A = 0\ . \tag{4.145}$$

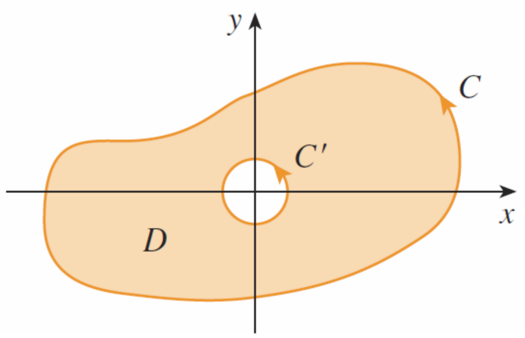


Figure 4.27: A region with a circular hole.

Next, we consider the case that C encloses the origin. In such case, we cannot apply Green's Theorem directly to the region enclosed by C since P and Q are not continuous in this region. So let's consider a counterclockwise-oriented circle C' with center at the origin and radius a, where a is chosen small enough so that C' lies inside C as shown in Fig. 4.27. Let D be the annular region bounded by C and C' whose positively-oriented boundary is $C\cup(-C')$. Then both P and Q have continuous first-order partial derivatives in D. Applying Green's Theorem, we have

$$\oint_{C\cup -C'} P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}A = \iint_D 0\,\mathrm{d}A = 0\ . \tag{4.146}$$

Hence,

$$\oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y = \oint_{C'} P\,\mathrm{d}x + Q\,\mathrm{d}y\ . \tag{4.147}$$

To evaluate the line integral around C', we represent the circle C' by the parametric equations $x = a\cos t$ and $y = a\sin t$ where $0 \le t \le 2\pi$. Then, we obtain

$$\begin{aligned}
\oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y &= \oint_{C'} P\,\mathrm{d}x + Q\,\mathrm{d}y \\
&= \int_0^{2\pi} \left[\frac{(a\cos t)(a\cos t) - (a\sin t)(-a\sin t)}{(a\cos t)^2 + (a\sin t)^2}\right]\mathrm{d}t \\
&= \int_0^{2\pi} \mathrm{d}t \\
&= 2\pi\ .
\end{aligned} \tag{4.148}$$

■

We are now ready to give the proof for Theorem 4.19 which was discussed in the previous section. Let $\mathbf{F}(x, y) = P(x, y)\,\hat{\imath} + Q(x, y)\,\hat{\mathbf{j}}$ be a vector field on an open simply-connected region D in the plane. Suppose P and Q have continuous first-order partial derivatives on D and

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{4.149}$$

for all (x, y) in D. If C is any piecewise-smooth, simple closed curve in D with positive orientation and R is the region enclosed by C, then Green's Theorem gives

$$\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A = \iint_R 0\,\mathrm{d}A = 0 \ . \tag{4.150}$$

Recall that a curve that is not simple intersects itself at one or more points. So it can be split into a number of simple curves. We have shown that the line integrals of $\mathbf{F}$ around these simple curves are all 0. By adding these integrals, we see that $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ for any closed curve C in D. Thus $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ is independent of path in D by Theorem 4.16. Hence, it follows from Theorem 4.17 that $\mathbf{F}$ is a conservative vector field on D.

Using the curl and divergence operator, we can rewrite Green's Theorem in vector form that will be useful for later work. Suppose a plane region D has a boundary curve C and the functions P and Q satisfy the assumptions of Green's Theorem. Let $\mathbf{F}(x, y) = P(x, y)\,\hat{\imath} + Q(x, y)\,\hat{\jmath}$ be a vector field. The its line integral is

$$\oint_C \mathbf{F} \cdot \mathbf{T}\,\mathrm{d}s = \oint_C P\,\mathrm{d}x + Q\,\mathrm{d}y \ , \tag{4.151}$$

which measures the counterclockwise circulation of $\mathbf{F}$ around C. If we regard $\mathbf{F}$ as a vector field in $\mathbb{R}^3$ whose z-component equal to zero, then we have

$$\operatorname{curl} \mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & 0 \end{vmatrix} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \hat{k} \ . \tag{4.152}$$

Then we can rewrite Green's Theorem in the vector form

$$\oint_C \mathbf{F} \cdot \mathbf{T}\,\mathrm{d}s = \iint_D (\operatorname{curl}\ \mathbf{F}) \cdot \hat{k}\,\mathrm{d}A \ . \tag{4.153}$$

It states that the line integral of the tangential component of $\mathbf{F}$ along a closed curve C is equal to the double integral of the normal component of $\operatorname{curl} \mathbf{F}$ over the region D enclosed by the curve C.

Next we derive a similar formula involving the normal component of $\mathbf{F}$. Let the curve C be represented by the vector equation $\mathbf{r}(t) = x(t)\,\hat{\imath} + y(t)\,\hat{\jmath}$ where $a \le t \le b$. Then the unit tangent vector and unit normal vector to C

are

$$\mathbf{T}(t) = \frac{x'(t)}{||\mathbf{r}'(t)||}\,\hat{\imath} + \frac{y'(t)}{||\mathbf{r}'(t)||}\,\hat{\jmath}\,, \tag{4.154}$$

$$\mathbf{n}(t) = \frac{y'(t)}{||\mathbf{r}'(t)||}\,\hat{\imath} - \frac{x'(t)}{||\mathbf{r}'(t)||}\,\hat{\jmath}\,. \tag{4.155}$$

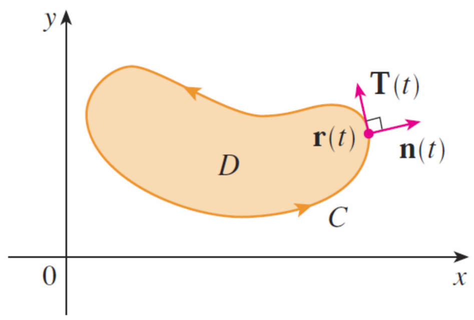


Figure 4.28: Tangent and normal.

We can easily check that $\mathbf{T}(t)\cdot\mathbf{n}(t) = 0$ (see Fig. 4.28). Thus, using Eq. (4.19) and Green's Theorem, we obtain

$$\begin{aligned}
\oint_C \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}s &= \int_a^b (\mathbf{F}\cdot\mathbf{n})(t)||\mathbf{r}'(t)||\,\mathrm{d}t \\
&= \int_a^b \left[\frac{P(x(t),y(t))y'(t)}{||\mathbf{r}'(t)||} - \frac{Q(x(t),y(t))x'(t)}{||\mathbf{r}'(t)||}\right]||\mathbf{r}'(t)||\,\mathrm{d}t \\
&= \int_a^b P(x(t),y(t))y'(t)\,\mathrm{d}t - \int_a^b Q(x(t),y(t))x'(t)\,\mathrm{d}t \\
&= \oint_C P\,\mathrm{d}y - Q\,\mathrm{d}x \\
&= \iint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right)\mathrm{d}A\,,
\end{aligned} \tag{4.156}$$

which measures the rate of outward flow of $\mathbf{F}$ across C. Noting that the integrand of the last integral is just the divergence of $\mathbf{F}$, we have the second vector form of Green's Theorem

$$\oint_C \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}s = \iint_D \mathrm{div}\,\mathbf{F}\,\mathrm{d}A\,. \tag{4.157}$$

It states that the line integral of the normal component of $\mathbf{F}$ along a closed curve C is equal to the double integral of the divergence of $\mathbf{F}$ over the region D enclosed by C.

4.4 Parametric Surfaces

Curves in planes can be defined in explicit form $y = f(x)$, implicit form $F(x,y) = 0$, or parametric form $\mathbf{r}(t) = f(t)\,\hat{\imath} + g(t)\,\hat{\jmath}$ where $a \le t \le b$.

In analogy, surfaces in space can be defined in explicit form $z = f(x, y)$ or implicit form $F(x, y, z) = 0$. There is also a parametric form giving the position of a point on the surface as a vector function of two variables.

Definition 4.30. *Let*

$$\mathbf{r}(u, v) = x(u, v)\,\hat{\imath} + y(u, v)\,\hat{\jmath} + z(u, v)\,\hat{k} \tag{4.158}$$

be a continuous vector function that is defined on a region D in the uv-plane and one-to-one on the interior of D. The set of all points (x, y, z) in $\mathbb{R}^3$ such that

$$x = x(u, v)\ , \qquad y = y(u, v)\ , \qquad z = z(u, v) \tag{4.159}$$

as (u, v) ranges over D is called a **parametric surface** *S defined by* $\mathbf{r}$*. Eq. (4.158) together with the domain D constitute a* **parametrization** *(also called parametric representations) of S and Eq. (4.159) are called* **parametric equations** *of S. We call the variables u and v the* **parameters***, and the region D the* **parameter domain***.*

Note that the requirement that $\mathbf{r}$ be one-to-one on the interior of D ensures that S does not cross itself. Moreover, as (u, v) varies throughout D, the tip of the vector $\mathbf{r}(u, v)$ traces out the surface S (see Fig. 4.29). That is to say, we can regard $\mathbf{r}(u, v)$ as mapping each point (u, v) in D onto a point $(x(u, v), y(u, v), z(u, v))$ on S in such a way that the surface S is the image of the plane region D.

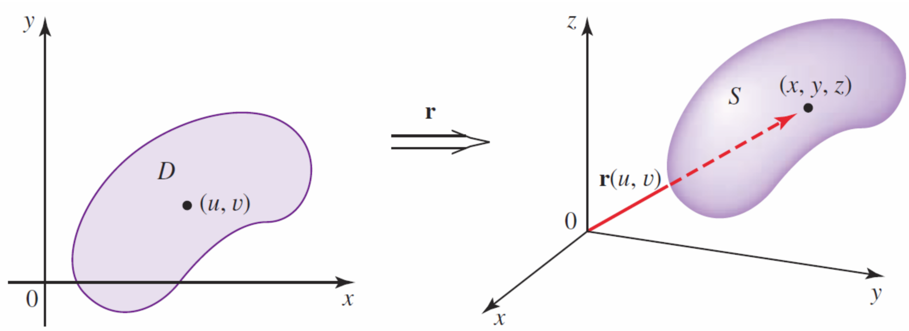


Figure 4.29: Parametric surface.

Example 4.31. Identify and sketch the surface represented by

$$\mathbf{r}(u, v) = 2\cos u\,\hat{\imath} + 2\sin u\,\hat{\jmath} + v\,\hat{k} \tag{4.160}$$

with the parameter domain $D = \{(u, v) \mid 0 \le u \le 2\pi,\ 0 \le v \le 3\}$.

The parametric equations for the surface are

$$x = 2\cos u\ , \qquad y = 2\sin u\ , \qquad z = v\ . \tag{4.161}$$

Eliminating the parameters u and v in the first two equations gives

$$x^2 + y^2 = 4\cos^2 u + 4\sin^2 u = 4\ . \tag{4.162}$$

Observe that the variable z is missing in this equation and it indicates that the trace in the xy-plane is a circle of radius 2. So it represents a circular cylinder with the z-axis as its axis of symmetry. Moreover, the third equation $z = v$ tells us that $0 \leq z \leq 3$ because $0 \leq v \leq 3$. Thus the required surface is a cylinder of finite height as shown in Fig. 4.30. ∎

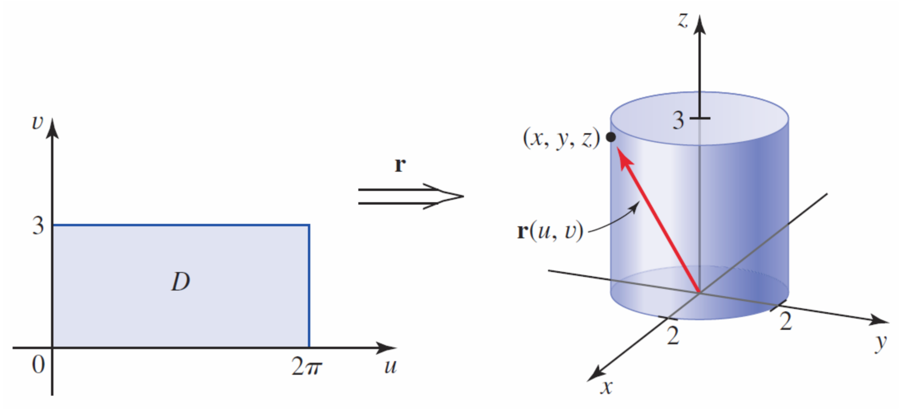


Figure 4.30: A cylinder.

There is another way of visualizing how $\mathbf{r}$ maps the domain D onto the surface S. If we keep u constant by setting $u = u_0$, where u_0 is a constant, then the tip of the resulting vector $\mathbf{r}(u_0, v)$ traces the curve C_1 lying on S as v is allowed to vary along the vertical line $u = u_0$ in D (see Fig. 4.31). Similarly, if we keep v constant by setting $v = v_0$, where v_0 is a constant, then the tip of the resulting vector $\mathbf{r}(u, v_0)$ traces the curve C_2 lying on S as u is allowed to vary along the horizontal line $v = v_0$ in D. The curves C_1 and C_2 are called **grid curves**. We can determine the shape of a parametric surface by considering the grid curves of the surface.

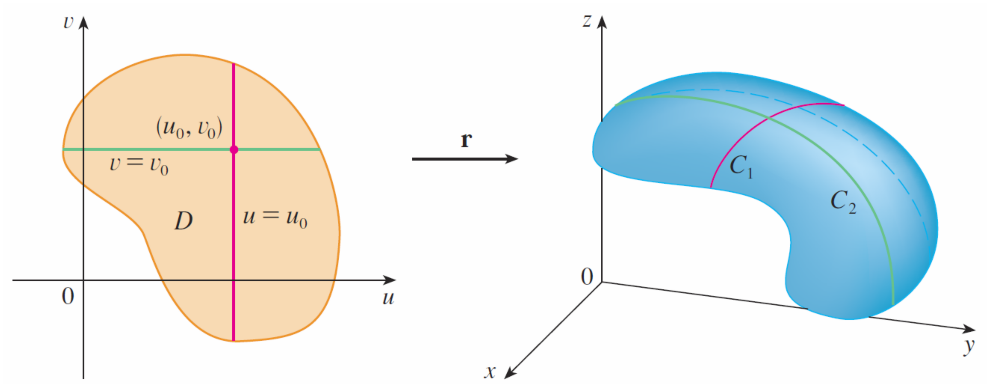


Figure 4.31: Surface as varying curves.

Example 4.32. Sketch a graph of the surface defined by

$$\mathbf{r}(u, v) = \sin u \cos v \,\hat{\imath} + \sin u \sin v \,\hat{\jmath} + \cos u \,\hat{k} \tag{4.163}$$

with the parameter domain $D = \{(u, v) \mid 0 \leq u \leq \pi,\ 0 \leq v \leq 2\pi\}$. Identify the grid curves on the surface that correspond to the curves with u held constant and those with v held constant.

The parametric equations for the surface are

$$x = \sin u \cos v \ , \qquad y = \sin u \sin v \ , \qquad z = \cos u \ . \tag{4.164}$$

Eliminating the parameters u and v in these equations, we obtain

$$x^2+y^2+z^2 = \sin^2 u \cos^2 v + \sin^2 u \sin^2 v + \cos^2 u = \sin^2 u + \cos^2 u = 1 \tag{4.165}$$

which represents a unit sphere centered at the origin (see Fig. 4.32(a)). Fixing

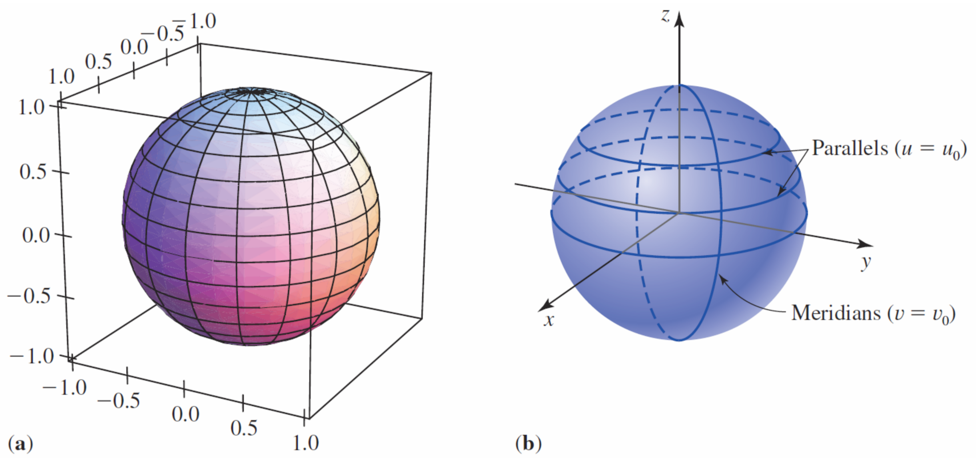


Figure 4.32: The surface of a sphere.

$u = u_0$, where u_0 is a constant, the parametric equations become

$$x = \sin u_0 \cos v \ , \qquad y = \sin u_0 \sin v \ , \qquad z = \cos u_0 \ . \tag{4.166}$$

Then we have

$$x^2 + y^2 = \sin^2 u_0 \cos^2 v + \sin^2 u_0 \sin^2 v = \sin^2 u_0 \ . \tag{4.167}$$

So the system of equations

$$x^2 + y^2 = \sin^2 u_0, \qquad z = \cos u_0 \tag{4.168}$$

for a fixed $u_0 \in [0, \pi]$ is equivalent to the vector function

$$\mathbf{r}(u_0, v) = \sin u_0 \cos v \, \hat{\imath} + \sin u_0 \sin v \, \hat{\jmath} + \cos u_0 \, \hat{k} \ . \tag{4.169}$$

It represents a circle of radius $\sin u_0$ on the sphere that is parallel to the xy-plane. Therefore, if we think of the sphere as a globe, then the horizontal lines in the domain of $\mathbf{r}$ are mapped onto the parallels as shown in Fig. 4.32(b). Similarly, we can show that the vertical lines in the domain of $\mathbf{r}$ with $v = v_0$, where v_0 is a constant, are mapped by

$$\mathbf{r}(u, v_0) = \sin u \cos v_0 \, \hat{\imath} + \sin u \sin v_0 \, \hat{\jmath} + \cos u \, \hat{k} \tag{4.170}$$

onto the meridians — great circles on the surface of the globe passing through the poles. ■

In last two examples, we were asked to graph the corresponding parametric surfaces for given vector functions. Indeed, we often need to do the opposite task, finding the vector function representing a given parametric surface. The technique for finding the parametrization of a surface is illustrated in the following examples.

Example 4.33.

(a) Find a parametrization for the graph of a function $z = f(x, y)$.

(b) Use part (a) to find a parametrization for the elliptic paraboloid $z = 4x^2 + y^2$.

(a) Suppose S is the graph of a function $z = f(x, y)$ defined on a domain D in the xy-plane. We simply pick x and y to be the parameters. That is to say, we write the parametric equations of S as

$$x = x(u, v) = u \ , \qquad y = y(u, v) = v \ , \qquad z = z(u, v) = f(u, v) \ . \tag{4.171}$$

and take the domain of f to be the parameter domain. Alternatively, we write the vector function representing the surface S as

$$\mathbf{r}(u, v) = u\,\hat{\imath} + v\,\hat{\jmath} + f(u, v)\,\hat{k} \ . \tag{4.172}$$

(b) We consider the surface S which is the graph of the function $z = f(x, y) = 4x^2 + y^2$. Using the result of part (a), the required parametric equations are

$$x = x(u, v) = u \ , \qquad y = y(u, v) = v \ , \qquad z = z(u, v) = 4u^2 + v^2 \ . \tag{4.173}$$

with the parameter domain $D = \{(u, v) \mid -\infty < u < \infty, \ -\infty < v < \infty\}$. So the corresponding vector function is

$$\mathbf{r}(u, v) = u\,\hat{\imath} + v\,\hat{\jmath} + (4u^2 + v^2)\,\hat{k} \ . \tag{4.174}$$

■

Example 4.34.

(a) Find a parametrization of the plane that passes through the point P_0 with position vector $\mathbf{r}_0$ and contains two non-parallel vectors $\mathbf{a}$ and $\mathbf{b}$.

(b) Using the result of part (a), find a parametrization of the plane passing through point $P_0(3, -1, 1)$ and contains the vectors $\mathbf{a} = -2\,\hat{\imath} + 5\,\hat{\jmath} + \hat{k}$ and $\mathbf{b} = -3\,\hat{\imath} + 2\,\hat{\jmath} + 3\,\hat{k}$.

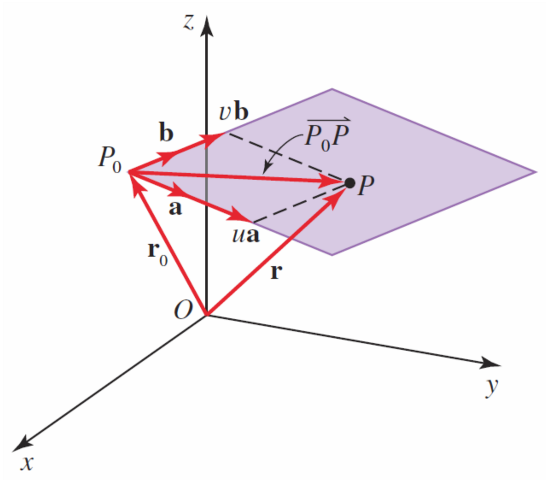


Figure 4.33: For Example 4.34(a).

(a) Let P be a point lying on the plane with position vector $\mathbf{r} = \overrightarrow{OP}$. Since $\overrightarrow{P_0P}$ lies on the plane determined by $\mathbf{a}$ and $\mathbf{b}$, there exist real numbers u and v such that

$$\overrightarrow{P_0P} = u\mathbf{a} + v\mathbf{b} \ , \tag{4.175}$$

as shown in Fig. 4.33. It implies that

$$\mathbf{r} = \mathbf{r}_0 + \overrightarrow{P_0P} = \mathbf{r}_0 + u\mathbf{a} + v\mathbf{b} \ . \tag{4.176}$$

In addition, any point on the plane is located at the tip of $\mathbf{r}$ for an appropriate choice of u and v. Thus the required parametrization is

$$\mathbf{r}(u, v) = \mathbf{r}_0 + u\mathbf{a} + v\mathbf{b} \tag{4.177}$$

with the parameter domain $D = \{(u, v) \mid -\infty < u < \infty,\ -\infty < v < \infty\}$.

(b) Using the result of part (a), we find that the required parametrization is

$$\begin{aligned} \mathbf{r}(u, v) &= (3\,\hat{\imath} - \hat{\jmath} + \hat{k}) + u(-2\,\hat{\imath} + 5\,\hat{\jmath} + \hat{k}) + v(-3\,\hat{\imath} + 2\,\hat{\mathbf{j}} + 3\,\hat{k}) \\ &= (-2u - 3v + 3)\,\hat{\imath} + (5u + 2v - 1)\,\hat{\jmath} + (u + 3v + 1)\,\hat{k} \end{aligned} \tag{4.178}$$

with the parameter domain $D = \{(u, v) \mid -\infty < u < \infty,\ -\infty < v < \infty\}$.

■

Example 4.35. Find a parametrization for the sphere $x^2 + y^2 + z^2 = a^2$.

The sphere has a simple representation $r = a$ in spherical coordinates (r, θ, ϕ). This suggests us to choose the angles θ and ϕ as the parameters. Denoting θ by u and ϕ by v, we obtain the required parametric equations

$$x = a \sin u \cos v \ , \qquad y = a \sin u \sin v \ , \qquad z = a \cos u \ . \tag{4.179}$$

So the corresponding vector function is

$$\mathbf{r}(u, v) = a \sin u \cos v\,\hat{\imath} + a \sin u \sin v\,\hat{\jmath} + a \cos u\,\hat{k} \tag{4.180}$$

The parameter domain is $D = \{(u, v) \mid 0 \le u \le \pi,\ 0 \le v \le 2\pi\}$. ■

Example 4.36. Find a parametrization for the cylinder

$$x^2 + (y-3)^2 = 9 \ , \qquad 0 \le z \le 5 \ . \tag{4.181}$$

Rewriting the equation $x^2+(y-3)^2 = 9$ in cylindrical coordinates (r, θ, z), we have

$$\begin{aligned} (r\cos\theta)^2 + (r\sin\theta - 3)^2 &= 9 \\ r^2 - 6r\sin\theta &= 0 \\ r &= 6\sin\theta \end{aligned} \tag{4.182}$$

where $0 \le \theta \le \pi$. Thus the cylinder has a simple representation $r = 6\sin\theta$ in cylindrical coordinates. This suggests us to choose θ and z as the parameters. Denoting θ by u and z by v, we obtain the required parametric equations

$$x = 6\sin u\cos u = 3\sin 2u \ , \qquad y = 6\sin^2 u \ , \qquad z = v \ . \tag{4.183}$$

So the corresponding vector function is

$$\mathbf{r}(u, v) = 3\sin 2u\,\hat{\imath} + 6\sin^2 u\,\hat{\jmath} + z\,\hat{k} \ . \tag{4.184}$$

The parameter domain is $D = \{(u, v) \mid 0 \le u \le \pi, \ 0 \le v \le 5\}$. ■

Parametric representations of surfaces are not unique. For example, an equation of the plane in Example 4.34(b) is $13x + 3y + 11z = 47$. Viewing the plane as the graph of the function $z = f(x, y) = (47 - 13x - 3y)/11$, we know that the plane can be represented by the parametrization

$$\mathbf{r}(u, v) = u\,\hat{\imath} + v\,\hat{\jmath} + \left(\frac{47 - 13u - 3v}{11}\right)\hat{k} \tag{4.185}$$

with parameter domain $D = \{(u, v) \mid -\infty < u < \infty, \ -\infty < v < \infty\}$. The next example shows two ways to parametrize a cone.

Example 4.37. Find a parametrization for the surface

$$z = 2\sqrt{x^2 + y^2} \ , \tag{4.186}$$

which is the top half of the cone $z^2 = 4x^2 + 4y^2$.

One possible parametrization is obtained by choosing x and y as the parameters. Then, the required parametric equations are

$$x = u \ , \qquad y = v \ , \qquad z = 2\sqrt{u^2 + v^2} \tag{4.187}$$

and the corresponding vector function is

$$\mathbf{r}(u, v) = u\,\hat{\imath} + v\,\hat{\jmath} + 2\sqrt{u^2 + v^2}\,\hat{k} \ . \tag{4.188}$$

The parameter domain is $D = \{(u, v) \mid -\infty < u < \infty, \ -\infty < v < \infty\}$. Another parametrization results from choosing the polar coordinates r and

θ as the parameters. Denoting r by u and θ by v, we obtain the required parametric equations

$$x = u\cos v\ , \qquad y = u\sin v\ , \qquad z = 2\sqrt{(u\cos v)^2 + (u\sin v)^2} = 2u\ . \tag{4.189}$$

So the corresponding vector function is

$$\mathbf{r}(u,v) = u\cos v\,\hat{\imath} + u\sin v\,\hat{\jmath} + 2u\,\hat{\mathbf{k}}\ . \tag{4.190}$$

The parameter domain is $D = \{(u,v) \mid -\infty < u < \infty,\ 0 \le v \le 2\pi\}$. ∎

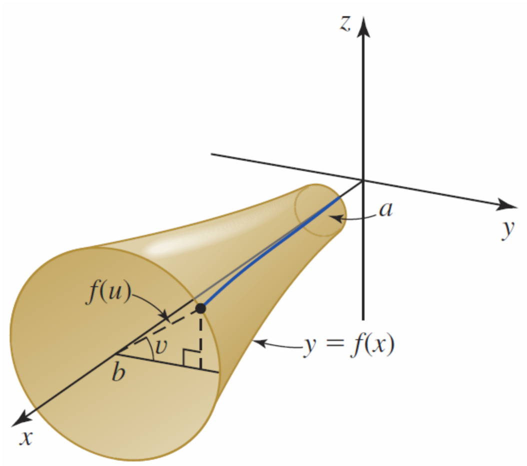


Figure 4.34: A surface of revolution.

If we rotate a curve on a plane about an axis on the same plane, then a three-dimensional surface known as surface of revolution is formed. Such a surface can be represented parametrically and thus graphed using a computer. For instance, suppose a surface S is obtained by revolving the graph of the function $y = f(x)$ for $a \le x \le b$ about the x-axis, where $f(x) \ge 0$. We denote x by u and the angle of rotation by v as shown in Fig. 4.34. If (x,y,z) is any point on S, then

$$x = u\ , \qquad y = f(u)\cos v\ , \qquad z = f(u)\sin v\ , \tag{4.191}$$

which can be regarded as the parametric equations of S. The parameter domain is $D = \{(u,v) \mid a \le u \le b,\ 0 \le v \le 2\pi\}$. We can modify Eq. (4.191) to represent a surface of revolution obtained by rotating about the y- or z-axis.

Example 4.38. Find parametric equations for the surface generated by rotating the curve $y = \sin x$, $0 \le x \le 2\pi$ about the x-axis. Use these equations to graph the surface of revolution.

According to Eq. (4.191), the required parametric equations are

$$x = u\ , \qquad y = \sin u\cos v\ , \qquad z = \sin u\sin v\ , \tag{4.192}$$

and the parameter domain is $D = \{(u,v) \mid 0 \le u \le 2\pi,\ 0 \le v \le 2\pi\}$. Using a computer to plot these equations, we obtain the graph in Fig. 4.35. ∎

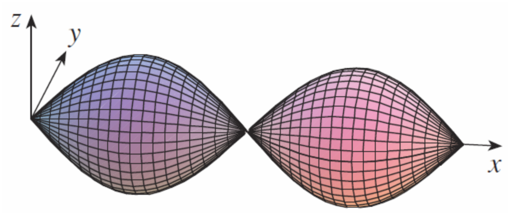


Figure 4.35: For Example 4.38.

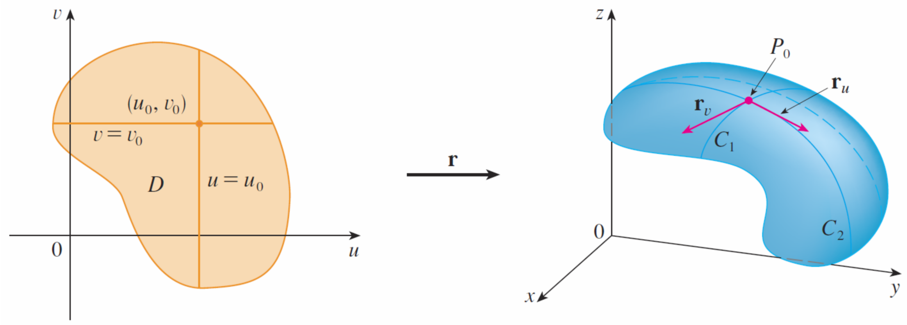


Figure 4.36: The tangents on a surface.

Suppose S is a parametric surface represented by the vector function

$$\mathbf{r}(u, v) = x(u, v)\,\hat{\imath} + y(u, v)\,\hat{\jmath} + z(u, v)\,\hat{k} \tag{4.193}$$

with parameter domain D. Let P_0 be a point on the surface S with position vector $\mathbf{r}(u_0, v_0)$, where u_0 and v_0 are constants. As shown in Fig. 4.36, if we keep u constant by setting $u = u_0$, then $\mathbf{r}(u_0, v)$ becomes a vector function of the single parameter v and defines a grid curve C_1 that lies on S. Taking partial derivatives of $\mathbf{r}$ with respect to v, we obtain the tangent vector to C_1 at P_0

$$\mathbf{r}_v(u_0, v_0) = \frac{\partial x}{\partial v}(u_0, v_0)\,\hat{\imath} + \frac{\partial y}{\partial v}(u_0, v_0)\,\hat{\jmath} + \frac{\partial z}{\partial v}(u_0, v_0)\,\hat{k} \tag{4.194}$$

Similarly, if we keep v constant by setting $v = v_0$, then we get a grid curve C_2 lying on S defined by $\mathbf{r}(u, v_0)$ and its tangent vector at P_0 is given by

$$\mathbf{r}_u(u_0, v_0) = \frac{\partial x}{\partial u}(u_0, v_0)\,\hat{\imath} + \frac{\partial y}{\partial u}(u_0, v_0)\,\hat{\jmath} + \frac{\partial z}{\partial u}(u_0, v_0)\,\hat{k}\ . \tag{4.195}$$

If $\mathbf{r}_u(u, v)$ and $\mathbf{r}_v(u, v)$ are continuous and $\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v) \neq \mathbf{0}$ for any (u, v) in D, then the surface S is said to be **smooth** (no cusps or sharp points). If S is a smooth surface, the plane that contains the tangent vectors $\mathbf{r}_u(u_0, v_0)$ and $\mathbf{r}_v(u_0, v_0)$ is the **tangent plane** to S at P_0 which has a normal vector

$$\mathbf{n} = \mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)\ . \tag{4.196}$$

Example 4.39. Find an equation of the tangent plane to the helicoid defined by the vector function

$$\mathbf{r}(u, v) = u\cos v\,\hat{\imath} + u\sin v\,\hat{\jmath} + v\,\hat{\mathbf{k}} \tag{4.197}$$

at the point where $u = 1/2$ and $v = \pi/4$.

Note that the point $(u, v) = (1/2, \pi/4)$ in the parameter domain corresponds to the point on the helicoid with coordinates

$$x = \frac{1}{2}\cos\frac{\pi}{4} = \frac{1}{2\sqrt{2}}\,, \qquad y = \frac{1}{2}\sin\frac{\pi}{4} = \frac{1}{2\sqrt{2}}\,, \qquad z = \frac{\pi}{4}\,. \tag{4.198}$$

Moreover, the partial derivatives of $\mathbf{r}$ are

$$\mathbf{r}_u(u, v) = \cos v\,\hat{\imath} + \sin v\,\hat{\jmath}\,, \qquad \mathbf{r}_v(u, v) = -u\sin v\,\hat{\imath} + u\cos v\,\hat{\jmath} + \hat{k}\,. \tag{4.199}$$

Then, we have

$$\mathbf{r}_u\left(\frac{1}{2}, \frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}\,\hat{\imath} + \frac{1}{\sqrt{2}}\,\hat{\jmath}\,, \qquad \mathbf{r}_v\left(\frac{1}{2}, \frac{\pi}{4}\right) = -\frac{1}{2\sqrt{2}}\,\hat{\imath} + \frac{1}{2\sqrt{2}}\,\hat{\jmath} + \hat{k}\,. \tag{4.200}$$

So the tangent plane to the helicoid at $(1/2\sqrt{2}, 1/2\sqrt{2}, \pi/4)$ has a normal vector

$$\mathbf{n} = \mathbf{r}_u\left(\frac{1}{2}, \frac{\pi}{4}\right) \times \mathbf{r}_v\left(\frac{1}{2}, \frac{\pi}{4}\right) = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} & 0 \\ -\dfrac{1}{2\sqrt{2}} & \dfrac{1}{2\sqrt{2}} & 1 \end{vmatrix} = \frac{\hat{\imath}}{\sqrt{2}} - \frac{\hat{\jmath}}{\sqrt{2}} + \frac{\hat{k}}{2}\,. \tag{4.201}$$

Since any normal vector will do, we take $\mathbf{n} = \sqrt{2}\,\hat{\imath} - \sqrt{2}\,\hat{\jmath} + \hat{k}$. Therefore, an equation of the tangent plane at the point $(1/(2\sqrt{2}), 1/(2\sqrt{2}), \pi/4)$ to the helicoid is

$$\begin{aligned} \sqrt{2}\left(x - \frac{1}{2\sqrt{2}}\right) - \sqrt{2}\left(y - \frac{1}{2\sqrt{2}}\right) + \left(z - \frac{\pi}{4}\right) &= 0 \\ \sqrt{2}x - \sqrt{2}y + z - \frac{\pi}{4} &= 0\,. \end{aligned} \tag{4.202}$$

■

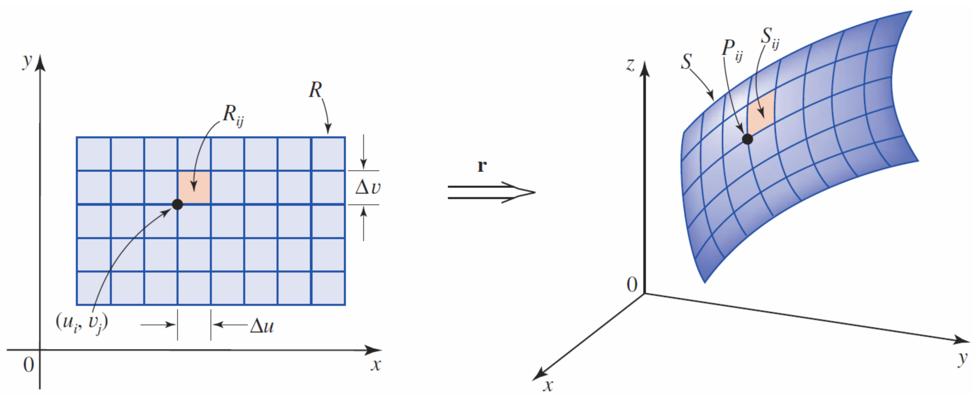


Figure 4.37: To calculate the area of a surface.

How can we find the surface area of a surface defined parametrically by Eq. (4.158)? To answer this question, let's consider a parametric surface S

defined by $\mathbf{r}(u, v)$ with parameter domain D that is a rectangle as shown in Fig. 38. We divide D into subrectangles $R_{11}, R_{12}, \ldots, R_{mn}$. Let (u_i, v_j) be the lower left corner of R_{ij}.

The part S_{ij} of the surface S that corresponds to R_{ij} is called a patch and the point P_{ij} with position vector $\mathbf{r}(u_i, v_j)$ is one of its corners. Then the tangent vectors to S at P_{ij} are $\mathbf{r}_u(u_i, v_j)$ and $\mathbf{r}_v(u_i, v_j)$. As shown in

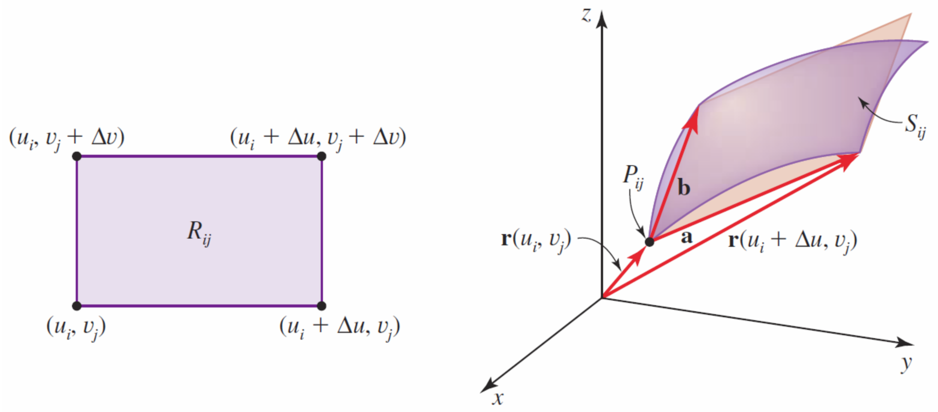


Figure 4.38: A very small part of a surface.

Fig. 4.38, the two edges of the patch S_{ij} meeting at P_{ij} can be approximated by the vectors

$$\begin{aligned} \mathbf{a} &= \mathbf{r}(u_i + \Delta u, v_j) - \mathbf{r}(u_i, v_j) \approx \mathbf{r}_u(u_i, v_j)\,\Delta u\ , \\ \mathbf{b} &= \mathbf{r}(u_i, v_j + \Delta v) - \mathbf{r}(u_i, v_j) \approx \mathbf{r}_v(u_i, v_j)\,\Delta v\ . \end{aligned} \tag{4.203}$$

So we can approximate S_{ij} by the parallelogram with the sides $\mathbf{r}_u(u_i, v_j)\,\Delta u$ and $\mathbf{r}_v(u_i, v_j)\,\Delta v$ as shown in Fig. 4.38. This parallelogram lies in the tangent plane to S at P_{ij} and its area is equal to

$$\Delta S_{ij} \approx ||[\mathbf{r}_u(u_i, v_j)\,\Delta u] \times [\mathbf{r}_v(u_i, v_j)\,\Delta v]|| = ||\mathbf{r}_u(u_i, v_j) \times \mathbf{r}_v(u_i, v_j)||\,\Delta u\,\Delta v\ . \tag{4.204}$$

Therefore, the area of the surface S can be approximated by

$$\sum_{i=1}^{m}\sum_{j=1}^{n} ||\mathbf{r}_u(u_i, v_j) \times \mathbf{r}_v(u_i, v_j)||\,\Delta u\,\Delta v \tag{4.205}$$

which is the Riemann sum for the double integral $\iint_D ||\mathbf{r}_u \times \mathbf{r}_v||\,du\,dv$. Intuitively, this approximation should gets better as we increase the number of subrectangles. This motivates the following definition.

Definition 4.40. Surface Area of a Smooth Parametric Surface
Let S be a smooth parametric surface defined by

$$\mathbf{r}(u, v) = x(u, v)\,\hat{\imath} + y(u, v)\,\hat{\jmath} + z(u, v)\,\hat{k} \tag{4.206}$$

with parameter domain D. If S is covered just once as (u, v) varies throughout

D, then the **surface area** *of S is*

$$A(S) = \iint_D ||\mathbf{r}_u \times \mathbf{r}_v||\,\mathrm{d}A \tag{4.207}$$

where

$$\mathbf{r}_u = \frac{\partial x}{\partial u}\,\hat{\imath} + \frac{\partial y}{\partial u}\,\hat{\jmath} + \frac{\partial z}{\partial u}\,\hat{k}\ , \qquad \mathbf{r}_v = \frac{\partial x}{\partial v}\,\hat{\imath} + \frac{\partial y}{\partial v}\,\hat{\jmath} + \frac{\partial z}{\partial v}\,\hat{k}\ . \tag{4.208}$$

Example 4.41. Find the surface area of a sphere of radius a centered at the origin.

In Example 4.35, we found that a parametrization of the sphere is

$$\mathbf{r}(u,v) = a\sin u\cos v\,\hat{\imath} + a\sin u\sin v\,\hat{\jmath} + a\cos u\,\hat{k} \tag{4.209}$$

with the parameter domain $D = \{(u,v)\mid 0 \le u \le \pi,\ 0 \le v \le 2\pi\}$.

So the tangent vectors to the sphere are

$$\begin{aligned}
\mathbf{r}_u(u,v) &= a\cos u\cos v\,\hat{\imath} + a\cos u\sin v\,\hat{\mathbf{j}} - a\sin u\,\hat{k}\ ,\\
\mathbf{r}_v(u,v) &= -a\sin u\sin v\,\hat{\imath} + a\sin u\cos v\,\hat{\mathbf{j}}\ ,
\end{aligned} \tag{4.210}$$

and their cross product is equal to

$$\begin{aligned}
\mathbf{r}_u \times \mathbf{r}_v &= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ a\cos u\cos v & a\cos u\sin v & -a\sin u \\ -a\sin u\sin v & a\sin u\cos v & 0 \end{vmatrix}\\
&= a^2\sin^2 u\cos v\,\hat{\imath} + a^2\sin^2 u\sin v\,\hat{\jmath} + a^2\sin u\cos u\,\hat{k}
\end{aligned} \tag{4.211}$$

which points outward from the sphere. Therefore,

$$\begin{aligned}
||\mathbf{r}_u \times \mathbf{r}_v|| &= \sqrt{(a^2\sin^2 u\cos v)^2 + (a^2\sin^2 u\sin v)^2 + (a^2\sin u\cos u)^2}\\
&= \sqrt{a^4\sin^4 u(\cos^2 v + \sin^2 v) + a^4\sin^2 u\cos^2 u}\\
&= a^2\sin u\ ,
\end{aligned} \tag{4.212}$$

since $\sin u \ge 0$ for $0 \le u \le \pi$. Using Eq. (4.207), the area of the sphere is then given by

$$\begin{aligned}
A &= \iint_D ||\mathbf{r}_u \times \mathbf{r}_v||\,\mathrm{d}A\\
&= \int_0^{2\pi}\int_0^{\pi} a^2\sin u\,du\,\mathrm{d}v\\
&= a^2\left(\int_0^{\pi}\sin u\,du\right)\left(\int_0^{2\pi}\mathrm{d}v\right)\\
&= a^2[-\cos u]_0^{\pi}[v]_0^{2\pi}\\
&= 4\pi a^2\ .
\end{aligned} \tag{4.213}$$

■

Example 4.42. Find the area of one complete turn of the helicoid represented by the vector function

$$\mathbf{r}(u,v) = u\cos v\,\hat{\imath} + u\sin v\,\hat{\jmath} + v\,\hat{k} \tag{4.214}$$

with the parameter domain $D = \{(u,v) \mid 0 \le u \le 1,\ 0 \le v \le 2\pi\}$.

The tangent vectors to the helicoid are

$$\mathbf{r}_u(u,v) = \cos v\,\hat{\imath} + \sin v\,\hat{\jmath}\ , \qquad \mathbf{r}_v(u,v) = -u\sin v\,\hat{\imath} + u\cos v\,\hat{\jmath} + \hat{k}\ . \tag{4.215}$$

So their cross product is equal to

$$\mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \cos v & \sin v & 0 \\ -u\sin v & u\cos v & 1 \end{vmatrix} = \sin v\,\hat{\imath} - \cos v\,\hat{\jmath} + u\,\hat{k} \tag{4.216}$$

and

$$||\mathbf{r}_u \times \mathbf{r}_v|| = \sqrt{(\sin v)^2 + (-\cos v)^2 + (u)^2} = \sqrt{1+u^2}\ . \tag{4.217}$$

Using Eq. (4.207), the required area is then given by

$$\begin{aligned} A &= \iint_D ||\mathbf{r}_u \times \mathbf{r}_v||\,\mathrm{d}A \\ &= \int_0^{2\pi}\int_0^1 \sqrt{1+u^2}\,du\,\mathrm{d}v \\ &= \left(\int_0^1 \sqrt{1+u^2}\,du\right)\left(\int_0^{2\pi}\mathrm{d}v\right) \\ &= \left[\frac{u}{2}\sqrt{1+u^2} + \frac{1}{2}\log(u+\sqrt{1+u^2})\right]_0^1 [v]_0^{2\pi} \\ &= \pi[\sqrt{2} + \log(1+\sqrt{2})]\ . \end{aligned} \tag{4.218}$$

■

Suppose a surface S is the graph of a function $z = g(x,y)$, where (x,y) lies in the domain D and g has continuous first-order partial derivative. We cannot apply Eq. (4.207) to compute the surface area of S as we don't know the parametric equations of S. Taking x and y as the parameters, the surface S can be represented by the vector function

$$\mathbf{r}(x,y) = x\,\hat{\imath} + y\,\hat{\jmath} + g(x,y)\,\hat{k}\ . \tag{4.219}$$

So the tangent vectors to S at the point (x,y,z) are

$$\mathbf{r}_x(x,y) = \hat{\imath} + \frac{\partial g}{\partial x}\,\hat{k}\ , \qquad \mathbf{r}_y(x,y) = \hat{\jmath} + \frac{\partial g}{\partial y}\,\hat{k}\ . \tag{4.220}$$

Therefore,

$$\mathbf{r}_x \times \mathbf{r}_y = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 1 & 0 & \dfrac{\partial g}{\partial x} \\ 0 & 1 & \dfrac{\partial g}{\partial y} \end{vmatrix} = -\frac{\partial g}{\partial x}\,\hat{\imath} - \frac{\partial g}{\partial y}\,\hat{\jmath} + \hat{k}\ . \tag{4.221}$$

Thus, we have

$$||\mathbf{r}_x \times \mathbf{r}_y|| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}\ . \tag{4.222}$$

Then, the surface area formula in Eq. (4.207) becomes

$$A(S) = \iint_D \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}\ \mathrm{d}A\ . \tag{4.223}$$

Similar formula applies if S is the graph of the function $y = g(x, z)$ or $x = g(y, z)$.

Example 4.43. Find the area of the part of the paraboloid $z = x^2 + y^2$ that lies under the plane $z = 9$.

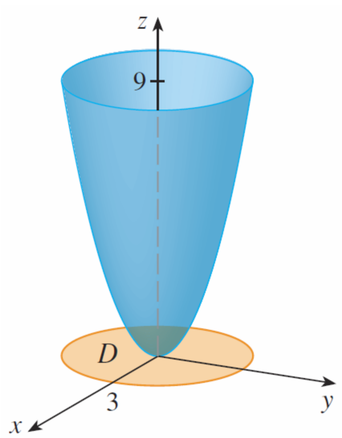


Figure 4.39: A paraboloid.

The intersection of the plane and the paraboloid is the circle $x^2 + y^2 = 9$. So the given surface lies above the disk D with center the origin and radius 3 as shown in Fig. 40. Using Eq. (4.223) with $z = g(x, y) = x^2 + y^2$, we find the area of the given surface to be

$$\begin{aligned} A &= \iint_D \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}\,\mathrm{d}A \\ &= \iint_D \sqrt{(2x)^2 + (2y)^2 + 1}\,\mathrm{d}A \\ &= \iint_D \sqrt{4x^2 + 4y^2 + 1}\,\mathrm{d}A\ . \end{aligned} \tag{4.224}$$

Rewriting the integral in polar coordinates, we obtain

$$
\begin{aligned}
A &= \int_0^{2\pi}\int_0^3 \sqrt{4r^2+1}\, r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \left(\int_0^3 r\sqrt{4r^2+1}\,\mathrm{d}r\right)\left(\int_0^{2\pi}\mathrm{d}\theta\right) \\
&= \left[\frac{1}{12}(4r^2+1)^{3/2}\right]_0^3 [\theta]_0^{2\pi} \\
&= \frac{\pi}{6}(37\sqrt{37}-1)\ . \qquad (4.225)
\end{aligned}
$$

■

4.5 Surface Integrals

The mass of a thin plate lying in a plane region R is given by the double integral $\iint_R \sigma(x,y)\,\mathrm{d}A$ where $\sigma(x,y)$ is the area mass density of the plate at any point (x,y) in R. However, if the thin plate is in the form of a curved surface, how can we find its mass?

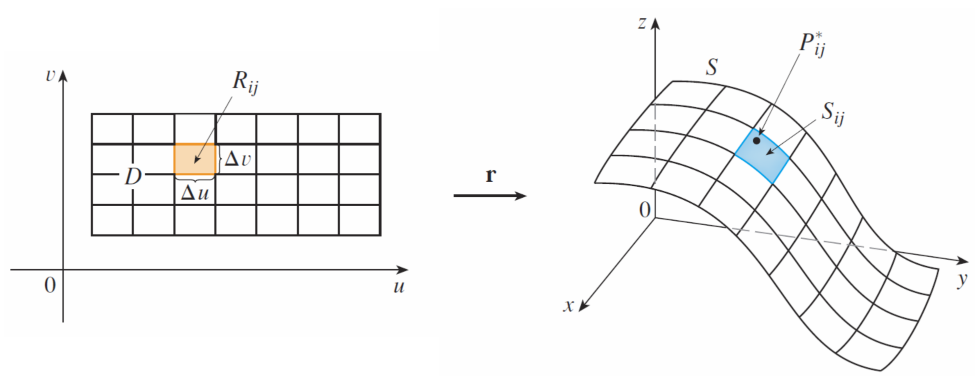


Figure 4.40: Surface integral.

For simplicity, suppose a thin plate has the shape of the parametric surface S defined by the vector function

$$
\mathbf{r}(u,v) = x(u,v)\,\hat{\imath} + y(u,v)\,\hat{\jmath} + z(u,v)\,\hat{k} \qquad (4.226)
$$

with parameter domain D that is a rectangle. Let $\sigma(x,y,z)$ be the area mass density of the plate at any point (x,y,z) on S, where σ is a non-negative continuous function defined on an open region containing S. We divide D into subrectangles $R_{11}, R_{12}, \ldots, R_{mn}$ with dimensions Δu and Δv. As shown in Fig. 41, the surface S is thus divided into corresponding patches. Choose an arbitrary point P_{ij}^* with position vector $\mathbf{r}(u_i^*, v_j^*)$ in each patch. Then, the mass of the part of the plate that lies on the patch S_{ij} is

$$\Delta m \approx \sigma(x(u_i^*, v_j^*), y(u_i^*, v_j^*), z(u_i^*, v_j^*))\, \Delta S_{ij} \ . \tag{4.227}$$

Therefore, the total mass M of the plate can be approximated by the sum

$$\sum_{i=1}^{m} \sum_{j=1}^{m} \sigma(x(u_i^*, v_j^*), y(u_i^*, v_j^*), z(u_i^*, v_j^*))\, \Delta S_{ij} \ , \tag{4.228}$$

which is similar to a Riemann sum. Obviously, this approximation should improve as the number of subrectangles increases. This observation leads us to the following definition.

Definition 4.44. Surface Integral of a Scalar Field
Let S be a parametric surface defined by

$$\mathbf{r}(u,v) = x(u,v)\,\hat{\imath} + y(u,v)\,\hat{\jmath} + z(u,v)\,\hat{k} \tag{4.229}$$

with parameter domain D. If f is a scalar function of three variables defined on a region in space containing S, then the **surface integral of f over S** *is*

$$\iint_S f(x,y,z)\,\mathrm{d}S = \lim_{m,n\to\infty} \sum_{i=1}^{m} \sum_{j=1}^{n} f(\mathbf{r}(u_i^*, v_j^*))\, \Delta S_{ij} \tag{4.230}$$

provided that this limit exists.

In the previous section, we have shown that the area of the patch S_{ij} is

$$\Delta S_{ij} \approx ||\mathbf{r}_u(u_i, v_j) \times \mathbf{r}_v(u_i, v_j)||\, \Delta u\, \Delta v \tag{4.231}$$

where $\mathbf{r}_u(u_i, v_j)$ and $\mathbf{r}_v(u_i, v_j)$ are the tangent vectors to the surface S at a corner of S_{ij}. If $\mathbf{r}(u,v)$ has continuous components and $\mathbf{r}_u$ and $\mathbf{r}_v$ are nonzero and nonparallel in the interior of D, then it can be shown from Eq. (4.230) that

$$\iint_S f(x,y,z)\,\mathrm{d}S = \iint_D f(\mathbf{r}(u,v))\, ||\mathbf{r}_u \times \mathbf{r}_v||\,\mathrm{d}A \tag{4.232}$$

even when D is not a rectangle. Using Eq. (4.232), we can compute a surface integral by converting it into a double integral over the parameter domain D. When using this formula, remember that $f(\mathbf{r}(u,v))$ is evaluated by writing $x = x(u,v)$, $y = y(u,v)$, and $z = z(u,v)$ in the formula for $f(x,y,z)$. Notice that if we put $f(x,y,z) = 1$ into Eq. (4.232), we have

$$\iint_S 1\,\mathrm{d}S = \iint_D ||\mathbf{r}_u \times \mathbf{r}_v||\,\mathrm{d}A = A(S) \ , \tag{4.233}$$

which gives the area of the surface S.

Example 4.45. Evaluate $\iint_S x^2\,\mathrm{d}S$, where S is a unit sphere centered at the origin.

As shown in Example 4.35, a unit sphere centered at the origin can be represented by the vector function

$$\mathbf{r}(u,v) = \sin u \cos v\, \hat{\imath} + \sin u \sin v\, \hat{\jmath} + \cos u\, \hat{k} \tag{4.234}$$

with the parameter domain $D = \{(u,v) \mid 0 \le u \le \pi,\ 0 \le v \le 2\pi\}$. Just like in Example 4.41, we can show that

$$||\mathbf{r}_u \times \mathbf{r}_v|| = \sin u \tag{4.235}$$

Using Eq. (4.232) with $f(x,y,z) = x^2$, we have

$$\begin{aligned}
\iint_S x^2\, \mathrm{d}S &= \iint_D (\sin u \cos v)^2 ||\mathbf{r}_u \times \mathbf{r}_v||\, \mathrm{d}A \\
&= \int_0^{2\pi} \int_0^{\pi} \sin^2 u \cos^2 v\ \sin u\, du\, \mathrm{d}v \\
&= \left(\int_0^{\pi} \sin^3 u\, du \right) \left(\int_0^{2\pi} \cos^2 v\, \mathrm{d}v \right) \\
&= \left[\int_0^{\pi} (\cos^2 u - 1)\, \mathrm{d}\cos u \right] \left[\int_0^{2\pi} \frac{1}{2}(1 + \cos 2v)\, \mathrm{d}v \right] \\
&= \left[\frac{1}{3}\cos^3 u - \cos u \right]_0^{\pi} \left[\frac{1}{2}v + \frac{1}{4}\sin 2v \right]_0^{2\pi} \\
&= \frac{4\pi}{3}\ .
\end{aligned} \tag{4.236}$$

■

Recall that any surface S that is the graph of a function $z = g(x,y)$ whose projection onto the xy-plane is D can be considered as a parametric surface represented by the vector function

$$\mathbf{r}(x,y) = x\, \hat{\imath} + y\, \hat{\jmath} + g(x,y)\, \hat{k}\ , \tag{4.237}$$

So the tangent vectors to S at the point (x,y,z) are

$$\mathbf{r}_x(x,y) = \hat{\imath} + \frac{\partial g}{\partial x}\, \hat{k}\ , \qquad \mathbf{r}_y(x,y) = \hat{\jmath} + \frac{\partial g}{\partial y}\, \hat{k}\ . \tag{4.238}$$

Then, we can show that

$$\mathbf{r}_x \times \mathbf{r}_y = -\frac{\partial g}{\partial x}\, \hat{\imath} - \frac{\partial g}{\partial y}\, \hat{\jmath} + \hat{k}\ , \tag{4.239}$$

and so

$$||\mathbf{r}_x \times \mathbf{r}_y|| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}\ . \tag{4.240}$$

Therefore, the surface integral formula in Eq. (4.232) becomes

$$\iint_S f(x,y,z)\, \mathrm{d}S = \iint_D f(x,y,g(x,y)) \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}\ \mathrm{d}A\ . \tag{4.241}$$

Similar formula applies when S is projected onto the xz-plane or yz-plane. For example, if S is defined by $y = g(x,z)$ and the projection of S onto the

xz-plane is D, then

$$\iint_S f(x,y,z)\,\mathrm{d}S = \iint_D f(x,g(x,z),z)\sqrt{\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial z}\right)^2+1}\;\mathrm{d}A\,. \tag{4.242}$$

Example 4.46. Evaluate $\iint_S x\,\mathrm{d}S$, where S is the part of the plane $2x+3y+z=6$ in the first octant.

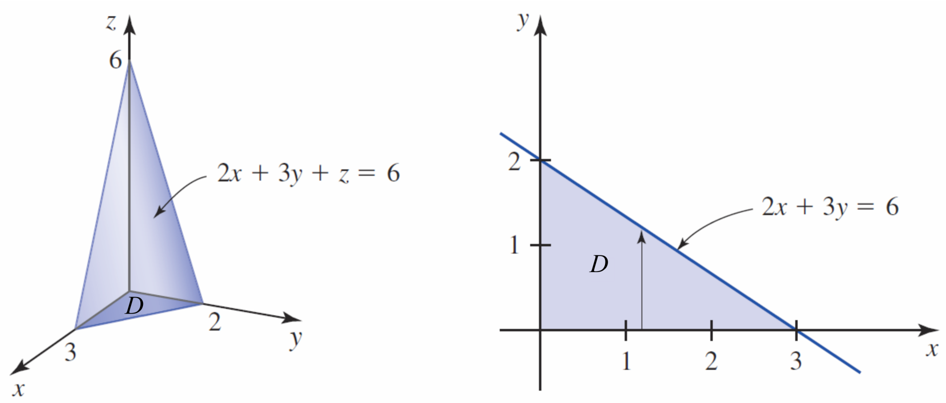


Figure 4.41: For Example 4.46.

Figure 4.41 shows the plane $2x+3y+z=6$ and its projection D onto the xy-plane. Using Eq. (4.241) with $f(x,y,z)=x$ and $z=g(x,y)=6-2x-3y$, we have

$$\begin{aligned}
\iint_S x\,\mathrm{d}S &= \iint_D x\sqrt{\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2+1}\,\mathrm{d}A\\
&= \iint_D x\sqrt{(-2)^2+(-3)^2+1}\,\mathrm{d}A\\
&= \sqrt{14}\iint_D x\,\mathrm{d}A\\
&= \sqrt{14}\int_0^3\int_0^{2-2x/3} x\,\mathrm{d}y\,\mathrm{d}x\\
&= \sqrt{14}\int_0^3\left(2x-\frac{2}{3}x^2\right)\mathrm{d}x\\
&= 3\sqrt{14}\,.
\end{aligned} \tag{4.243}$$

■

Surface integrals have applications similar to those for line integrals. For example, if a thin metallic sheet has the shape of a surface S and the area mass density at the point (x,y,z) is $\sigma(x,y,z)$, then the total mass of the sheet is equal to

$$M = \iint_S \sigma(x,y,z)\,\mathrm{d}S \tag{4.244}$$

and the center of mass is located at the point $(\bar{x}, \bar{y}, \bar{z})$, where, for example,

$$\bar{x} = \frac{1}{M} \iint_S x\, \sigma(x, y, z)\, \mathrm{d}S \;. \tag{4.245}$$

Example 4.47. Find the mass of the surface S that is made up of the part of the paraboloid $y = x^2 + z^2$ between the planes $y = 1$ and $y = 4$ if the area mass density at any point P on S in inversely proportional to the distance from P to the axis of symmetry of S.

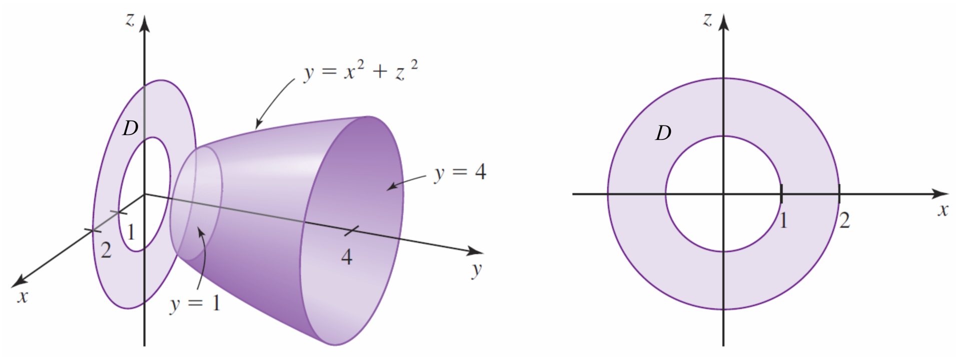


Figure 4.42: A curve on the plane.

Figure 4.42 shows the surface S and its projection D onto the xz-plane. According to the given information, the area mass density of S is $\sigma(x, y, z) = k/\sqrt{x^2 + z^2}$ where k is a proportionality constant. Using Eq. (4.242) with $f(x, y, z) = \sigma(x, y, z) = k/\sqrt{x^2 + z^2}$ and $y = g(x, z) = x^2 + z^2$, we find the mass of S to be

$$\begin{aligned} M &= \iint_S \sigma(x, y, z)\, \mathrm{d}S \\ &= k \iint_S \frac{1}{\sqrt{x^2 + z^2}}\, \mathrm{d}S \\ &= k \iint_D \frac{1}{\sqrt{x^2 + z^2}} \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial z}\right)^2 + 1}\, \mathrm{d}A \\ &= k \iint_D \frac{1}{\sqrt{x^2 + z^2}} \sqrt{(2x)^2 + (2z)^2 + 1}\, \mathrm{d}A \\ &= k \iint_D \frac{1}{\sqrt{x^2 + z^2}} \sqrt{4x^2 + 4z^2 + 1}\, \mathrm{d}A \;. \end{aligned} \tag{4.246}$$

Rewriting the integral in polar coordinates $x = r\cos\theta$ and $z = r\sin\theta$, we obtain

$$
\begin{aligned}
M &= k\int_0^{2\pi}\int_1^2\left(\frac{1}{r}\right)\sqrt{4r^2+1}\,r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= 2k\left(\int_1^2\sqrt{r^2+\frac{1}{4}}\,\mathrm{d}r\right)\left(\int_0^{2\pi}\mathrm{d}\theta\right) \\
&= 2k\left[\frac{r}{2}\sqrt{r^2+\frac{1}{4}}+\frac{1}{8}\log\left(r+\sqrt{r^2+\frac{1}{4}}\right)\right]_1^2[\theta]_0^{2\pi} \\
&= k\pi\left[2\sqrt{17}-\sqrt{5}+\frac{1}{2}\log\left(\frac{4+\sqrt{17}}{2+\sqrt{5}}\right)\right]. \qquad (4.247)
\end{aligned}
$$

■

If S is a piecewise-smooth surface, which is a finite union of smooth surfaces $S_1, S_2, \ldots, S_n$ that intersect only along their boundaries, then the surface integral of f over S is

$$
\begin{aligned}
&\iint_S f(x,y,z)\,\mathrm{d}S \\
= &\iint_{S_1} f(x,y,z)\,\mathrm{d}S+\iint_{S_2} f(x,y,z)\,\mathrm{d}S+\cdots+\iint_{S_n} f(x,y,z)\,\mathrm{d}S\,.
\end{aligned}
$$

The following example illustrates how to compute the surface integral for a piecewise-smooth surface.

Example 4.48. Evaluate $\iint_S z\,\mathrm{d}S$, where S is the surface whose side S_1 is given by the cylinder $x^2+y^2=1$, whose bottom S_2 is the disk $x^2+y^2\leq 1$ in the plane $z=0$, and whose top S_3 is the part of the plane $z=1+x$ that lies above S_2.

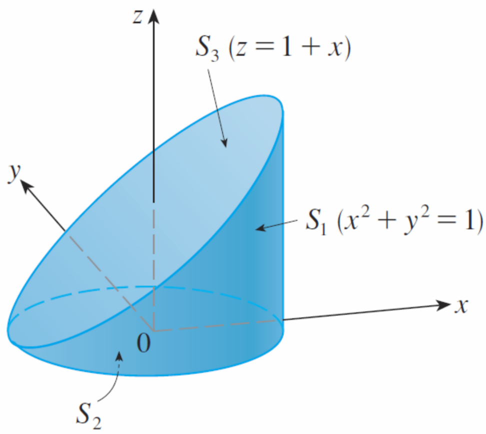


Figure 4.43: For Example 4.48.

The surface S is shown in Fig. 4.43. For S_1, we take $u=\theta$ and $v=z$ as the parameters and then obtain a parametrization of S

$$\mathbf{r}(u,v) = \cos u\,\hat{\imath} + \sin u\,\hat{\jmath} + v\,\hat{k} \tag{4.248}$$

with the parameter domain $D = \{(u,v) \mid 0 \le u \le 2\pi,\ 0 \le v \le 1+\cos u\}$. So the tangent vectors to S are

$$\mathbf{r}_u = -\sin u\hat{\imath} + \cos u\,\hat{\jmath}\ , \qquad \mathbf{r}_v = \hat{k}\ . \tag{4.249}$$

Therefore,

$$\begin{aligned}
\mathbf{r}_u \times \mathbf{r}_v &= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ -\sin u & \cos u & 0 \\ 0 & 0 & 1 \end{vmatrix} = \cos u\,\hat{\imath} + \sin u\,\hat{\jmath}\ , && (4.250) \\
||\mathbf{r}_u \times \mathbf{r}_v|| &= \sqrt{(\cos u)^2 + (\sin u)^2} = 1\ . && (4.251)
\end{aligned}$$

Using Eq. (4.232) with $f(x,y,z) = z$, we obtain

$$\begin{aligned}
\iint_{S_1} z\,\mathrm{d}S &= \iint_D v||\mathbf{r}_u \times \mathbf{r}_v||\,\mathrm{d}A \\
&= \int_0^{2\pi} \int_0^{1+\cos u} v\,\mathrm{d}v\,du \\
&= \int_0^{2\pi} \left[\frac{1}{2}v^2\right]_0^{1+\cos u} du \\
&= \frac{1}{2}\int_0^{2\pi} (1+\cos u)^2 du \\
&= \frac{1}{2}\left[\frac{3}{2}u + 2\sin u + \frac{1}{4}\sin 2u\right]_0^{2\pi} \\
&= \frac{3\pi}{2}\ . && (4.252)
\end{aligned}$$

Since S_2 lies in the plane $z = 0$, we have

$$\iint_{S_2} z\,\mathrm{d}S = \iint_{S_2} 0\,\mathrm{d}S = 0\ . \tag{4.253}$$

The top surface S_3 lies above the unit disk D' centered at the origin and is part of the plane $z = 1 + x$. So it can be regarded as the graph of this plane with projection D' onto the xy-plane. Therefore, using Eq. (4.241) with $f(x,y,z) = z$ and $z = g(x,y) = 1 + x$ and converting to polar coordinates, we have

$$\begin{aligned}
\iint_{S_3} z\,\mathrm{d}S &= \iint_{D'} (1+x)\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}\,\mathrm{d}A \\
&= \iint_{D'} (1+x)\sqrt{(1)^2+(0)^2+1}\,\mathrm{d}A \\
&= \sqrt{2}\iint_{D'} (1+x)\,\mathrm{d}A \\
&= \sqrt{2}\int_0^{2\pi}\int_0^1 (1+r\cos\theta)\,r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \sqrt{2}\int_0^{2\pi}\left[\frac{1}{2}r^2 + \frac{1}{3}r^3\cos\theta\right]_0^1 \mathrm{d}\theta \\
&= \sqrt{2}\int_0^{2\pi}\left(\frac{1}{2}+\frac{1}{3}\cos\theta\right)\mathrm{d}\theta \\
&= \sqrt{2}\left[\frac{1}{2}\theta + \frac{1}{3}\sin\theta\right]_0^{2\pi} \\
&= \sqrt{2}\pi\ . \qquad (4.254)
\end{aligned}$$

Hence,

$$\begin{aligned}
\iint_S z\,\mathrm{d}S &= \iint_{S_1} z\,\mathrm{d}S + \iint_{S_2} z\,\mathrm{d}S + \iint_{S_3} z\,\mathrm{d}S \\
&= \frac{3\pi}{2} + 0 + \sqrt{2}\pi \\
&= \left(\frac{3}{2}+\sqrt{2}\right)\pi\ . \qquad (4.255)
\end{aligned}$$

■

To define surface integrals of vector fields, we first need to introduce some terminology to describe the orientation of a surface. We call a smooth surface S **orientable** or **two-sided** if it is possible to define a unit normal vector $\hat{\mathbf{n}}$ at each point (x, y, z) on S except at any boundary point so that $\hat{\mathbf{n}}$ varies continuously over S. Smooth closed surfaces (smooth surfaces that encloses solids) such as spheres are examples of orientable surfaces. Once $\hat{\mathbf{n}}$ has been chosen, S is called an **oriented surface**. We say that the given choice of $\hat{\mathbf{n}}$ provides S with an orientation. For any orientable surface, there are always two possible orientations, one pointing outward and one pointing inward (see Fig. 4.44). By convention, the **positive orientation** for a closed surface S is the one for which the unit normal vector $\hat{\mathbf{n}}$ points outward from S.

An example of a nonorientable surface is the Möbius strip, which can be constructed by taking a long rectangular strip of paper, giving it a half-twist, and then taping the short edges together to produce the surface as shown in Fig. 4.45. No matter at which point P you start to construct a unit normal vector $\hat{\mathbf{n}}$, moving $\hat{\mathbf{n}}$ around the surface in the manner shown will return it to

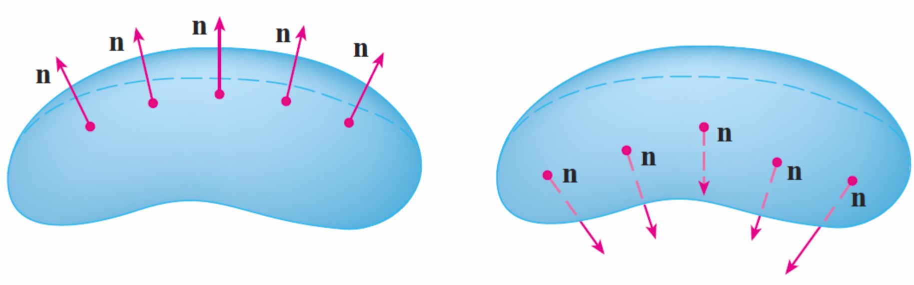


Figure 4.44: Orientations of surfaces.

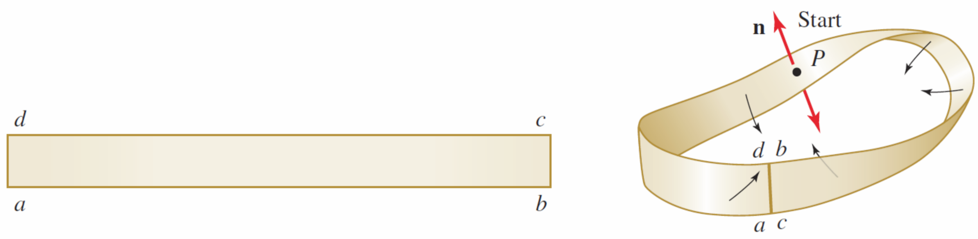


Figure 4.45: The Möbius strip.

the point P with a direction precisely opposite to its initial direction. This shows that $\hat{\mathbf{n}}$ does not vary continuously on a Möbius strip. So we conclude that the strip is not orientable.

Suppose a fluid with mass density $\rho(x, y, z)$ and velocity field $\mathbf{v}(x, y, z)$ flowing in space through an oriented surface S with unit normal vector $\hat{\mathbf{n}}$. Then, the rate of flow (mass per unit time) per unit area is $\rho\mathbf{v}$ and its component in the direction of $\hat{\mathbf{n}}$ is $\rho\mathbf{v}\cdot\hat{\mathbf{n}}$. If S is divided into a number of small patches S_{ij} as in Fig. 4.46, then S_{ij} is almost flat and so we can approximate the mass of fluid per unit time crossing S_{ij} in the direction of the normal $\hat{\mathbf{n}}$ by the quantity

$$(\rho\mathbf{v}\cdot\hat{\mathbf{n}})\,A(S_{ij}) \tag{4.256}$$

where ρ, $\mathbf{v}$, and $\hat{\mathbf{n}}$ are evaluated at some point on S_{ij}. By summing these quantities and taking the limit we obtain, according to Eq. (4.230), the surface integral of the function $\rho\mathbf{v}\cdot\hat{\mathbf{n}}$ over S is

$$\iint_S \rho\mathbf{v}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \iint_D \rho(x,y,z)\mathbf{v}(x,y,z)\cdot\hat{\mathbf{n}}(x,y,z)\,\mathrm{d}S\ . \tag{4.257}$$

This is interpreted physically as the volume of fluid passing through S per unit time.

Writing $\mathbf{F} = \rho\mathbf{v}$, the surface integral in the above equation becomes

$$\iint_S \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}S\ . \tag{4.258}$$

Such form of surface integrals occurs frequently in physics, even when $\mathbf{F}$ is not equal to $\rho\mathbf{v}$. In general, we have the following definition for this form of surface integrals.

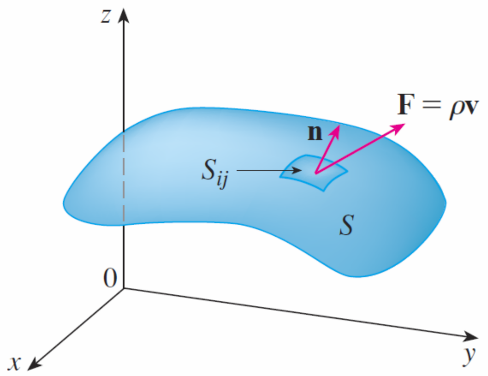


Figure 4.46: Fluid flowing through a surface.

Definition 4.49. *Let* **F** *be a continuous vector field defined on an oriented surface* S *with unit normal vector* **n**. *The* **surface integral of F over** S *is*

$$\iint_S \mathbf{F}\cdot \mathrm{d}\mathbf{S} = \iint_S \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}S \tag{4.259}$$

This integral is also called the **flux** *of* **F** *across* S.

If an oriented surface S is a smooth surface defined parametrically by the vector function $\mathbf{r}(u,v)$ with parameter domain D, then it is automatically supplied with the orientation of the unit normal vector

$$\hat{\mathbf{n}} = \frac{\mathbf{r}_u\times\mathbf{r}_v}{||\mathbf{r}_u\times\mathbf{r}_v||} \tag{4.260}$$

and the opposite orientation is given by $-\hat{\mathbf{n}}$. Putting $\hat{\mathbf{n}}$ into Eq. (4.259), we obtain

$$\begin{aligned}
&\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S}\\
\equiv\ &\iint_S \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}S\\
=\ &\iint_S \mathbf{F}\cdot\frac{(\mathbf{r}_u\times\mathbf{r}_v)}{||\mathbf{r}_u\times\mathbf{r}_v||}\,\mathrm{d}S\\
=\ &\iint_D \left[\mathbf{F}(\mathbf{r}(u,v))\cdot\frac{(\mathbf{r}_u\times\mathbf{r}_v)}{||\mathbf{r}_u\times\mathbf{r}_v||}\right]||\mathbf{r}_u\times\mathbf{r}_v||\,\mathrm{d}A\ .
\end{aligned} \tag{4.261}$$

Thus, we have

$$\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iint_D \mathbf{F}(\mathbf{r}(u,v))\cdot(\mathbf{r}_u\times\mathbf{r}_v)\,\mathrm{d}A\ . \tag{4.262}$$

Example 4.50. Find the flux of the vector field $\mathbf{F}(x,y,z) = y\,\hat{\imath} + x\,\hat{\jmath} + 2z\,\hat{k}$ across the outward-oriented unit sphere S given by $x^2+y^2+z^2=1$.

The unit sphere S can be described by the parametrization

$$\mathbf{r}(u,v) = \sin u \cos v\,\hat{\imath} + \sin u \sin v\,\hat{\jmath} + \cos u\,\hat{k} \tag{4.263}$$

with the parameter domain $D = \{(u,v) \mid 0 \le u \le \pi,\ 0 \le v \le 2\pi\}$. Using the result in Example 4.41 and putting $a = 1$, we obtain

$$\mathbf{r}_u \times \mathbf{r}_v = \sin^2 u \cos v\hat{\imath} + \sin^2 u \sin v\,\hat{\jmath} + \sin u \cos u\,\hat{k}\ , \tag{4.264}$$

which points outward from S. Therefore,

$$\begin{aligned}
\mathbf{F}(\mathbf{r}(u,v)) \cdot (\mathbf{r}_u \times \mathbf{r}_v) &= (\sin u \sin v\,\hat{\imath} + \sin u \cos v\,\hat{\jmath} + 2\cos u\,\hat{k}) \\
&\quad \cdot (\sin^2 u \cos v\hat{\imath} + \sin^2 u \sin v\,\hat{\jmath} + \sin u \cos u\,\hat{k}) \\
&= \sin^3 u \sin v \cos v + \sin^3 u \cos v \sin v + 2\cos^2 u \sin u \\
&= 2(\sin^3 u \sin v \cos v + \sin u \cos^2 u)\ .
\end{aligned} \tag{4.265}$$

Using Eq. (4.262), we find the flux across the unit sphere

$$\begin{aligned}
&\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} \\
=\ & \iint_D \mathbf{F}(\mathbf{r}(u,v)) \cdot (\mathbf{r}_u \times \mathbf{r}_v)\,\mathrm{d}A \\
=\ & 2\int_0^{2\pi}\int_0^{\pi} (\sin^3 u \sin v \cos v + \sin u \cos^2 u)\,du\,\mathrm{d}v \\
=\ & 2\left(\int_0^{\pi} \sin^3 u\,du\right)\left(\int_0^{2\pi} \sin v \cos v\,\mathrm{d}v\right) \\
& +2\left(\int_0^{\pi} \sin u \cos^2 u\,du\right)\left(\int_0^{2\pi} \mathrm{d}v\right) \\
=\ & 2\left[\frac{1}{3}\cos^3 u - \cos u\right]_0^{\pi}\left[\frac{1}{2}\sin^2 v\right]_0^{2\pi} + 2\left[-\frac{1}{3}\cos^3 u\right]_0^{\pi}[v]_0^{2\pi} \\
=\ & \frac{8\pi}{3}\ .
\end{aligned} \tag{4.266}$$

■

If an oriented surface S is the graph of the function $z = g(x,y)$ whose projection onto the xy-plane is D, we can choose x and y as the parameters and use Eq. (4.239) to obtain

$$\mathbf{F} \cdot (\mathbf{r}_x \times \mathbf{r}_y) = (P\,\hat{\imath} + Q\,\hat{\jmath} + R\,\hat{k}) \cdot \left(-\frac{\partial g}{\partial x}\,\hat{\imath} - \frac{\partial g}{\partial y}\,\hat{\jmath} + \hat{k}\right)\ . \tag{4.267}$$

where $\mathbf{F} = P\,\hat{\imath} + Q\,\hat{\jmath} + R\,\hat{k}$. Therefore, Eq. (4.262) becomes

$$\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iint_D \left(-P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R\right)\mathrm{d}A \tag{4.268}$$

Note that S is assumed to have an upward orientation in deriving Eq. (4.268); and we simply multiply it by -1 for a downward orientation. Similar formula can be obtained if S is defined by $y = g(x,z)$ or $x = g(y,z)$.

Just like surface integrals of scalar fields, if S is a finite union of smooth surfaces $S_1, S_2, \ldots, S_n$ that intersect only along their boundaries, then the surface integral of $\mathbf{F}$ over S is defined by

$$\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iint_{S_1} \mathbf{F} \cdot \mathrm{d}\mathbf{S} + \iint_{S_2} \mathbf{F} \cdot \mathrm{d}\mathbf{S} + \cdots + \iint_{S_n} \mathbf{F} \cdot \mathrm{d}\mathbf{S} . \tag{4.269}$$

Example 4.51. Evaluate $\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$ where $\mathbf{F}(x, y, z) = x\,\hat{\imath} + y\,\hat{\jmath} + z\,\hat{k}$ and S is the boundary of the solid region E enclosed by the paraboloid $z = 1 - x^2 - y^2$ and the plane $z = 0$.

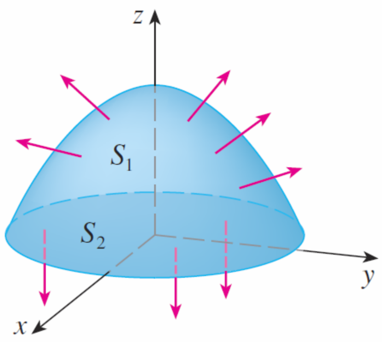


Figure 4.47: A paraboloid.

The surface S is composed of a parabolic top surface S_1 and a circular bottom surface S_2 as shown in Fig. 4.47. Since S is a closed surface, we use the convention that the normal points outward for a positive orientation. It implies that S_1 is oriented upward while S_2 is oriented downward.

Notice that S_1 is the graph of the function $z = g(x, y) = 1 - x^2 - y^2$ and its projection D onto the xy-plane is the disk $x^2 + y^2 \leq 1$. Then we have

$$\frac{\partial g}{\partial x} = \frac{\partial}{\partial x}(1 - x^2 - y^2) = -2x \ , \qquad \frac{\partial g}{\partial y} = \frac{\partial}{\partial y}(1 - x^2 - y^2) = -2y \ . \tag{4.270}$$

Moreover, we can rewrite $\mathbf{F}$ as $\mathbf{F}(x, y, z) = x\,\hat{\mathbf{i}} + y\,\hat{\jmath} + R(x, y, z)\,\hat{k}$ where

$$R(x, y, z) = z = 1 - x^2 - y^2 \tag{4.271}$$

on S_1. Therefore, using Eq. (4.268), we obtain

$$\begin{aligned}
\iint_{S_1} \mathbf{F}\cdot \mathrm{d}\mathbf{S} &= \iint_D \left[-(x)(-2x)-(y)(-2y)+(1-x^2-y^2)\right] \mathrm{d}A \\
&= \iint_D (1+x^2+y^2)\,\mathrm{d}A \\
&= \int_0^{2\pi}\int_0^1 (1+r^2)\,r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \left[\int_0^1 (r+r^3)\,\mathrm{d}r\right]\left(\int_0^{2\pi}\mathrm{d}\theta\right) \\
&= \left[\frac{1}{2}r^2+\frac{1}{4}r^4\right]_0^1 [\theta]_0^{2\pi} \\
&= \frac{3\pi}{2}\,. \qquad (4.272)
\end{aligned}$$

Next, observe that the unit normal vector for S_2 is $\hat{\mathbf{n}} = -\hat{k}$. So we have

$$\iint_{S_2} \mathbf{F}\cdot \mathrm{d}\mathbf{S} = \iint_{S_2} \mathbf{F}\cdot(-\hat{k})\,\mathrm{d}S = \iint_D (-z)\,\mathrm{d}A = \iint_D 0\,\mathrm{d}A = 0 \qquad (4.273)$$

since $z = 0$ on S_2. Hence,

$$\iint_{S} \mathbf{F}\cdot \mathrm{d}\mathbf{S} = \iint_{S_1} \mathbf{F}\cdot \mathrm{d}\mathbf{S} + \iint_{S_2} \mathbf{F}\cdot \mathrm{d}\mathbf{S} = 0 + \frac{3\pi}{2} = \frac{3\pi}{2}\,. \qquad (4.274)$$

■

We have introduced the concept of the surface integral of a vector field using the example of a fluid flow. These integrals also have applications in other fields of physics. For example, if $\mathbf{E}$ is an electric field and S is an oriented surface, then the surface integral

$$\iint_S \mathbf{E}\cdot \mathrm{d}\mathbf{S} \qquad (4.275)$$

is called the electric flux of $\mathbf{E}$ through the surface S. Gauss's Law, which is one of the fundamental laws of electromagnetism, tells us that the net charge enclosed by a closed oriented surface S is given by

$$Q = \epsilon_0 \iint_S \mathbf{E}\cdot \mathrm{d}\mathbf{S} \qquad (4.276)$$

where ϵ_0 is a physical constant known as the permittivity of free space.

Another application of surface integrals arises in the study of heat flow. Suppose the temperature at a point (x, y, z) in a homogeneous body is $T(x, y, z)$. It can be shown that the flow of heat can be described by the vector field

$$\mathbf{q} = -k\nabla T \qquad (4.277)$$

where k is a proportionality constant called the thermal conductivity of the body. The rate of heat flow across a surface S in the body is then given by the surface integral

$$\iint_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = -k\iint_S \nabla T\cdot\hat{\mathbf{n}}\,\mathrm{d}S\ . \tag{4.278}$$

Example 4.52. The temperature T at a point $P(x, y, z)$ in a medium with thermal conductivity k is inversely proportional to the distance between P and the origin. Find the rate of heat flow outward across a sphere S of radius a centered at the origin.

According to the given information, the temperature in the medium is equal to

$$T(x, y, z) = \frac{c}{\sqrt{x^2+y^2+z^2}} \tag{4.279}$$

where c is the proportionality constant. Then the flow of heat is

$$\mathbf{q} = -k\nabla T = \frac{ck}{(x^2+y^2+z^2)^{3/2}}(x\,\hat{\imath}+y\,\hat{\jmath}+z\,\hat{k})\ . \tag{4.280}$$

The outward unit normal vector to the sphere $x^2+y^2+z^2=a^2$ at the point (x, y, z) is

$$\hat{\mathbf{n}} = \frac{1}{a}(x\,\hat{\imath}+y\,\hat{\jmath}+z\,\hat{k})\ . \tag{4.281}$$

Thus, the rate at which heat flows across S is given by

$$\begin{aligned}
&\iint_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S\\
=\;& \iint_S \left[\frac{ck}{(x^2+y^2+z^2)^{3/2}}(x\,\hat{\imath}+y\,\hat{\jmath}+z\,\hat{k})\right]\cdot\left[\frac{1}{a}(x\,\hat{\imath}+y\,\hat{\jmath}+z\,\hat{k})\right]\mathrm{d}S\\
=\;& \frac{ck}{a}\iint_S \frac{1}{\sqrt{x^2+y^2+z^2}}\,\mathrm{d}S\\
=\;& \frac{ck}{a^2}\iint_S \mathrm{d}S\\
=\;& \frac{ck}{a^2}A(S)\\
=\;& 4\pi ck\ .
\end{aligned} \tag{4.282}$$

■

4.6 Stokes' Theorem

Recall that Green's Theorem can be expressed in the vector form as

$$\oint_C \mathbf{F}\cdot\mathbf{T}\,\mathrm{d}s = \iint_D (\text{curl }\mathbf{F})\cdot\hat{k}\,\mathrm{d}A\ , \tag{4.283}$$

where the plane curve C is a piecewise-smooth, simple closed curve with positive orientation that encloses a region D. The theorem states that the line integral of the tangential component of a vector field in two-dimensional space around a simple closed curve is equal to the double integral of the normal component of the curl of the vector field over the plane region bounded by the curve.

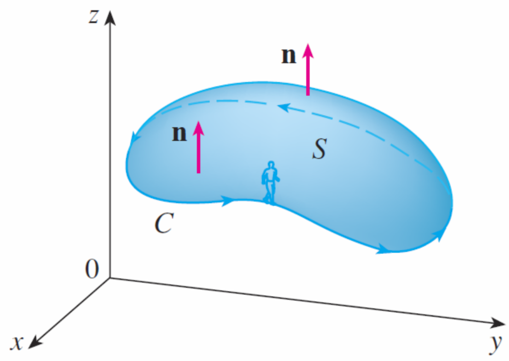


Figure 4.48: Stokes' theorem.

Stokes' Theorem generalizes this form of Green's Theorem to three-dimensional space. This theorem relates the line integral of the tangential component of a vector field in three-dimensional space around a simple closed curve in space to a surface integral of the normal component of the curl of the vector field over any surface bounded by the curve. Figure 4.48 shows an oriented surface S with unit normal vector $\hat{\mathbf{n}}$.

The orientation of S induces the **positive orientation of the boundary curve** C as shown in the figure. That is to say, if you walk in the positive direction around C with your head pointing in the direction of $\hat{\mathbf{n}}$, then S will be always on your left.

Theorem 4.53. Stokes' Theorem
Let S be an oriented piecewise-smooth surface that has a unit normal vector $\hat{\mathbf{n}}$ and is bounded by a piecewise-smooth, simple closed curve C with positive orientation. If $\mathbf{F}$ is a vector field whose components have continuous partial derivatives on a open region in space that contains S, then

$$\oint_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \oint_C \mathbf{F}\cdot\mathbf{T}\,\mathrm{d}s = \iint_S \mathrm{curl}\ \mathbf{F}\cdot \mathrm{d}\mathbf{S} = \iint_S \mathrm{curl}\ \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}S\ . \tag{4.284}$$

Proof. Stokes' Theorem tells us that the work done by a force field $\mathbf{F}$ along a path C is equal to the flux of $\mathrm{curl}\,\mathbf{F}$ across the surface S with C as its boundary.

In the special case that S is a flat surface lying on the xy-plane with upward orientation, its unit normal vector is $\hat{k}$ and the surface integral becomes a double integral. Then, Stokes' Theorem gives

$$\oint_C \mathbf{F}\cdot\mathbf{T}\,\mathrm{d}s = \iint_S \mathrm{curl}\,\mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iint_D (\mathrm{curl}\,\mathbf{F})\cdot\hat{k}\,\mathrm{d}A\ , \tag{4.285}$$

which is the vector form of Green's Theorem stated above. We can see that Green's Theorem is indeed a special case of Stokes' Theorem.

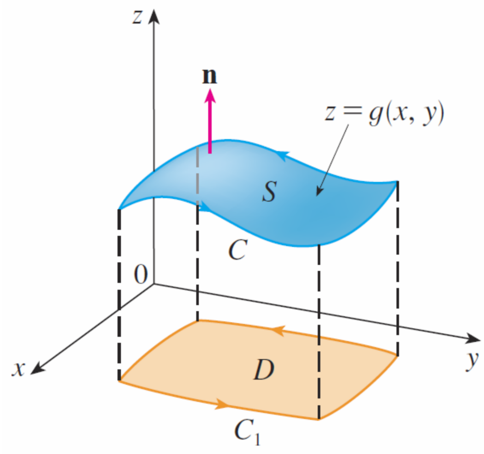


Figure 4.49: Surface as the graph of a function.

It is very difficult to prove Stokes' Theorem in general. So here we give a proof for the special case where S is the graph of a function and $\mathbf{F}$, S, and C are well behaved. Suppose S is the graph of a function $z = g(x, y)$ where g has continuous second-order partial derivatives and the parameter domain D is a simple plane region enclosed by a curve C_1 corresponding to C, Fig. 4.49. If we orient S upward, then C has the same positive orientation as C_1. Assume the vector field $\mathbf{F} = P\,\hat{\imath} + Q\,\hat{\jmath} + R\,\hat{k}$ where P, Q, and R have continuous partial derivatives.

Since S is the graph of a function, we can apply Eq. (4.268) with $\mathbf{F}$ replaced by $\operatorname{curl}\mathbf{F}$. Then we obtain

$$\begin{aligned}&\iint_S \operatorname{curl}\mathbf{F}\cdot \mathrm{d}\mathbf{S}\\ &= \iint_D \left[-\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\frac{\partial z}{\partial x}-\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\frac{\partial z}{\partial y}+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\right]\mathrm{d}A\ ,\end{aligned}$$

where the partial derivatives of P, Q, and R are evaluated at the point $(x, y, g(x, y))$. If a parametrization of C_1 is

$$x = x(t)\ , \qquad y = y(t)\ , \qquad a \le t \le b\ , \tag{4.286}$$

then C can be parametrized by

$$x = x(t)\ , \qquad y = y(t)\ , \qquad z = g(x(t), y(t))\ , \qquad a \le t \le b\ . \tag{4.287}$$

So we can evaluate the above line integral with the aid of the Chain Rule as follows

$$\begin{aligned}
\oint_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_a^b \left(P\frac{\mathrm{d}x}{\mathrm{d}t} + Q\frac{\mathrm{d}y}{\mathrm{d}t} + R\frac{\mathrm{d}z}{\mathrm{d}t}\right)\mathrm{d}t \\
&= \int_a^b \left[P\frac{\mathrm{d}x}{\mathrm{d}t} + Q\frac{\mathrm{d}y}{\mathrm{d}t} + R\left(\frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}\right)\right]\mathrm{d}t \\
&= \int_a^b \left[\left(P + R\frac{\partial z}{\partial x}\right)\frac{\mathrm{d}x}{\mathrm{d}t} + \left(Q + R\frac{\partial z}{\partial y}\right)\frac{\mathrm{d}y}{\mathrm{d}t}\right]\mathrm{d}t \\
&= \int_{C_1} \left(P + R\frac{\partial z}{\partial x}\right)\mathrm{d}x + \left(Q + R\frac{\partial z}{\partial y}\right)\mathrm{d}y \\
&= \iint_D \left[\frac{\partial}{\partial x}\left(Q + R\frac{\partial z}{\partial y}\right) - \frac{\partial}{\partial y}\left(P + R\frac{\partial z}{\partial x}\right)\right]\mathrm{d}A \ , \qquad (4.288)
\end{aligned}$$

where the last step follows from Green's Theorem. Using the Chain Rule again and recall that P, Q, and R are functions of x, y, and z with z itself being a function of x and y, we obtain

$$\begin{aligned}
&\oint_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} \\
= &\iint_D \left[\left(\frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z}\frac{\partial z}{\partial x} + \frac{\partial R}{\partial x}\frac{\partial z}{\partial y} + \frac{\partial R}{\partial z}\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} + R\frac{\partial^2 z}{\partial x\partial y}\right)\right. \\
&\qquad \left. - \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial y} + \frac{\partial R}{\partial y}\frac{\partial z}{\partial x} + \frac{\partial R}{\partial z}\frac{\partial z}{\partial y}\frac{\partial z}{\partial x} + R\frac{\partial^2 z}{\partial y\partial x}\right)\right]\mathrm{d}A \\
= &\iint_D \left[-\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\frac{\partial z}{\partial x} - \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\frac{\partial z}{\partial y} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\right]\mathrm{d}A \\
= &\iint_S \operatorname{curl}\mathbf{F}\cdot \mathrm{d}\mathbf{S} \ . \qquad (4.289)
\end{aligned}$$

■

Example 4.54. Verify Stokes' Theorem for the case in which $\mathbf{F}(x,y,z) = 3z\,\hat{\imath} + 2x\,\hat{\jmath} + y^2\,\hat{k}$, S is the part of the paraboloid $z = 4 - x^2 - y^2$ for $z \geq 0$ with upward orientation, and C is the trace of S on the xy-plane.

Figure 4.50 shows the sketches of the surface S and the curve C. Since S is oriented upward, C is traversed counterclockwise when viewed from the positive z-axis.

First, we calculate the curl of $\mathbf{F}$

$$\operatorname{curl}\mathbf{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 3z & 2x & y^2 \end{vmatrix} = 2y\,\hat{\imath} + 3\,\hat{\jmath} + 2\,\hat{k} \ . \qquad (4.290)$$

Notice that S can be regarded as the graph of the function $z = g(x,y) = 4 - x^2 - y^2$ and its projection onto the xy-plane is $D = \{(x,y) \mid x^2 + y^2 \leq 4\}$.

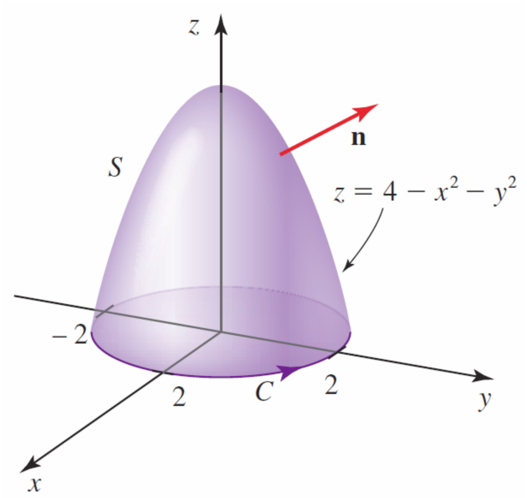


Figure 4.50: For Example 4.54

Then we find that

$$\frac{\partial g}{\partial x} = -2x \ , \qquad \frac{\partial g}{\partial y} = -2y \ . \tag{4.291}$$

Therefore, by Eq. (4.268) with $P(x, y, z) = 2y$, $Q(x, y, z) = 3$ and $R(x, y, z) = 3$, we obtain

$$\begin{aligned}
\iint_S \operatorname{curl} \mathbf{F} \cdot \mathrm{d}\mathbf{S} &= \iint_D [-(2y)(-2x) - (3)(-2y) + 2] \, \mathrm{d}A \\
&= \iint_D (4xy + 6y + 2) \, \mathrm{d}A \\
&= \int_0^{2\pi} \int_0^2 (4r^2 \cos\theta \sin\theta + 6r \sin\theta + 2) \, r \, \mathrm{d}r \, \mathrm{d}\theta \\
&= \int_0^{2\pi} \left[r^4 \cos\theta \sin\theta + 2r^3 \sin\theta + r^2\right]_0^2 \mathrm{d}\theta \\
&= \int_0^{2\pi} (16 \cos\theta \sin\theta + 16 \sin\theta + 4) \mathrm{d}\theta \\
&= \left[8 \sin^2\theta - 16 \cos\theta + 4\theta\right]_0^{2\pi} \\
&= 8\pi \ .
\end{aligned} \tag{4.292}$$

On the other hand, we note that C can be parametrized by

$$\mathbf{r}(t) = 2 \cos t \, \hat{\imath} + 2 \sin t \, \hat{\jmath} \ , \qquad 0 \le t \le 2\pi \ . \tag{4.293}$$

So we have

$$\begin{aligned}
\mathbf{r}'(t) &= -2 \sin t \, \hat{\imath} + 2 \cos t \, \hat{\jmath} \ , \\
\mathbf{F}(\mathbf{r}(t)) &= 4 \cos t \, \hat{\jmath} + 4 \sin^2 t \, \hat{k} \ .
\end{aligned} \tag{4.294}$$

Thus, using Eq. (4.66), we obtain

$$\begin{aligned}
\oint_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_0^{2\pi} \mathbf{F}(\mathbf{r}(t))\cdot \mathbf{r}'(t)\,\mathrm{d}t \\
&= \int_0^{2\pi} (4\cos t\,\hat{\jmath} + 4\sin^2 t\,\hat{k})\cdot(-2\sin t\,\hat{\imath} + 2\cos t\,\hat{\jmath})\,\mathrm{d}t \\
&= \int_0^{2\pi} 8\cos^2 t\,\mathrm{d}t \\
&= 4\left[t + \frac{1}{2}\sin 2t\right]_0^{2\pi} \\
&= 8\pi\ , \qquad (4.295)
\end{aligned}$$

which is the same as the surface integral. ■

Example 4.55. Evaluate $\oint_c \mathbf{F}\cdot \mathrm{d}\mathbf{r}$, where $\mathbf{F}(x,y,z) = -y^2\,\hat{\imath} + x\,\hat{\jmath} + z^2\,\hat{k}$ and C is the curve of intersection of the plane $y+z=2$ and the cylinder $x^2+y^2=1$. Assume C is oriented so that it is traversed counterclockwise when viewed from the positive z-axis.

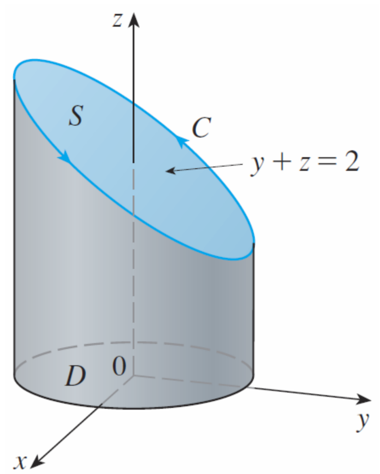


Figure 4.51: For Example 4.55.

The curve C is an ellipse as shown in Fig. 4.51. We can evaluate $\oint_c \mathbf{F}\cdot \mathrm{d}\mathbf{r}$ directly. But we instead use Stokes' Theorem to evaluate the integral as it's much easier. We first compute

$$\operatorname{curl}\mathbf{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y^2 & x & z^2 \end{vmatrix} = (1+2y)\,\hat{k}\ . \qquad (4.296)$$

There are many surfaces with C as their boundary. But the most convenient choice is the elliptical region S in the plane $y+z=2$ that is bounded by C. Note that the orientation of S is upward for the positive orientation of C to be the same as the given orientation. The projection of S onto the xy-plane is the disk D defined by $x^2 + y^2 \leq 1$. Therefore, using Stokes' Theorem and Eq. (4.268) with $P(x, y, z) = Q(x, y, z) = 0$, $R(x, y, z) = 1+2y$ and $z = g(x, y) = 2 - y$, we have

$$\begin{aligned}
\oint_c \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \iint_S \mathrm{curl}\ \mathbf{F} \cdot \mathrm{d}\mathbf{S} \\
&= \iint_D [-(0)(0) - (0)(-1) + (1 + 2y)]\, \mathrm{d}A \\
&= \iint_D (1 + 2y)\, \mathrm{d}A \\
&= \int_0^{2\pi} \int_0^1 (1 + 2r \sin\theta)\, r\, \mathrm{d}r\, \mathrm{d}\theta \\
&= \int_0^{2\pi} \left[\frac{r^2}{2} + \frac{2r^3}{3} \sin\theta \right]_0^1 \mathrm{d}\theta \\
&= \int_0^{2\pi} \left(\frac{1}{2} + \frac{2}{3} \sin\theta \right) \mathrm{d}\theta \\
&= \left[\frac{1}{2}\theta - \frac{2}{3} \cos\theta \right]_0^{2\pi} \\
&= \pi\ . \qquad (4.297)
\end{aligned}$$

■

Example 4.56. Use Stokes' Theorem to compute the integral $\iint_S \mathrm{curl}\, \mathbf{F} \cdot \mathrm{d}\mathbf{S}$, where $\mathbf{F}(x, y, z) = yz\, \hat{\imath} - xz\, \hat{\jmath} + z^3\, \hat{k}$, and S is the part of the sphere $x^2 + y^2 + z^2 = 8$ that lies inside the cone $z = \sqrt{x^2 + y^2}$ with upward orientation, Fig. 4.52.

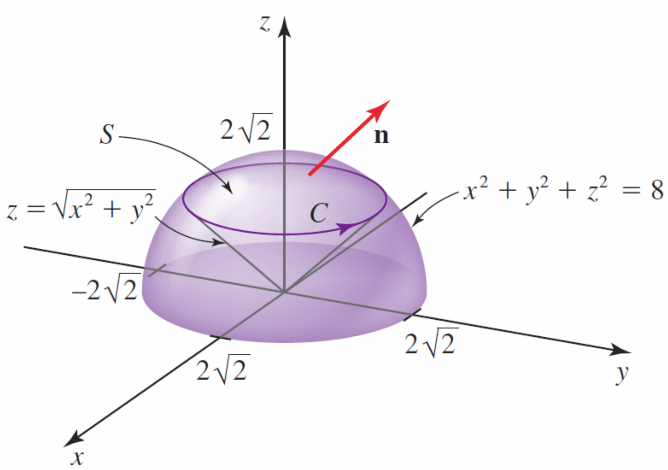


Figure 4.52: For Example 4.56.

The boundary of the surface S is the curve C whose equation is obtained by solving the equations $x^2 + y^2 + z^2 = 8$ and $z = \sqrt{x^2 + y^2}$. Squaring the second equation and substituting this result into the first equation give $2z^2 = 8$ and so $z = 2$ since $z \geq 0$. Thus C is the circle defined by the equations $x^2 + y^2 = 4$ and $z = 2$. So it can be represented by the vector function

$$\mathbf{r}(t) = 2\cos t\,\hat{\imath} + 2\sin t\,\hat{\jmath} + 2\,\hat{k}, \qquad 0 \leq t \leq 2\pi\ . \tag{4.298}$$

So we have

$$\mathbf{r}'(t) = -2\sin t\,\hat{\imath} + 2\cos t\,\hat{\jmath}\ , \qquad \mathbf{F}(\mathbf{r}(t)) = 4\sin t\,\hat{\imath} - 4\cos t\,\hat{\jmath} + 8\,\hat{k}\ . \tag{4.299}$$

Using Stokes' Theorem (Eq. (4.284)), we have

$$\begin{aligned}
\iint_S \operatorname{curl}\mathbf{F}\cdot \mathrm{d}\mathbf{S} &= \oint_C \mathbf{F}\cdot\mathrm{d}\mathbf{r} \\
&= \int_0^{2\pi} \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,\mathrm{d}t \\
&= \int_0^{2\pi} (4\sin t\,\hat{\imath} - 4\cos t\,\hat{\jmath} + 8\,\hat{k})\cdot(-2\sin t\,\hat{\imath} + 2\cos t\,\hat{\jmath})\,\mathrm{d}t \\
&= -8\int_0^{2\pi} \mathrm{d}t \\
&= -16\pi\ .
\end{aligned} \tag{4.300}$$

■

We can use Stokes' Theorem to give a physical interpretation of the curl vector. Suppose C is an oriented closed curve and $\mathbf{v}(x, y, z)$ represents the velocity field in fluid flow. If $\mathbf{T}$ is the unit tangent vector, then $\mathbf{v}\cdot\mathbf{T}$ is the component of $\mathbf{v}$ in the direction of $\mathbf{T}$. That is to say, the value of $\mathbf{v}\cdot\mathbf{T}$ increases if the direction of $\mathbf{v}$ gets closer to that of $\mathbf{T}$. Thus, the line integral

$$\oint_C \mathbf{v}\cdot\mathrm{d}\mathbf{r} = \oint_C \mathbf{v}\cdot\hat{\mathbf{T}}\,\mathrm{d}s \tag{4.301}$$

is a measure of the tendency of the fluid to move around C and is called the **circulation of v around** C, Fig. 4.53.

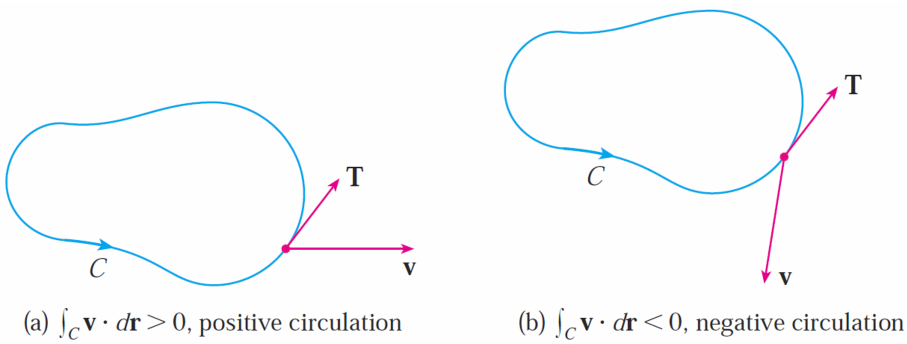


Figure 4.53: Circulation of a vector field.

Next, let $P_0(x_0, y_0, z_0)$ be a point in the fluid and S_a be a disk with center P_0 and very small radius a. Then, $\operatorname{curl}\mathbf{v}(P) \approx \operatorname{curl}\mathbf{v}(P_0)$ for all points P in S_a since $\operatorname{curl}\mathbf{F}$ is continuous. Therefore, by Stokes' Theorem, we get the following approximation to the circulation of $\mathbf{v}$ around the boundary curve C_a

$$\begin{aligned}
\oint_{C_a} \mathbf{v}\cdot \mathrm{d}\mathbf{r} &= \iint_{S_a} \operatorname{curl}\,\mathbf{v}\cdot \mathrm{d}\mathbf{S} \\
&\approx \iint_{S_a} \operatorname{curl}\mathbf{v}(P_0)\cdot \hat{\mathbf{n}}(P_0)\,\mathrm{d}S \\
&= \operatorname{curl}\mathbf{v}(P_0)\cdot \hat{\mathbf{n}}(P_0)\pi a^2 \ . \qquad (4.302)
\end{aligned}$$

This approximation improves as $a \to 0$ which suggests that

$$\operatorname{curl}\mathbf{v}(P_0)\cdot \hat{\mathbf{n}}(P_0) = \lim_{a\to 0} \frac{1}{\pi a^2} \oint_{C_a} \mathbf{v}\cdot \mathrm{d}\mathbf{r} \ . \qquad (4.303)$$

The above equation gives the relationship between the curl and the circulation. It tells us that $\operatorname{curl}\mathbf{v}\cdot\hat{\mathbf{n}}$ is a measure of the rotating effect of the fluid about the axis $\hat{\mathbf{n}}$ which is greatest about the axis parallel to $\operatorname{curl}\mathbf{v}$.

4.7 Volume Integrals

In general, volume integrals are simpler than line or surface integrals since the volume element $\mathrm{d}V$ is a scalar quantity. There are two types of volume integrals

$$\iiint_E f(x,y,z)\,\mathrm{d}V \qquad \iiint_E \mathbf{F}(x,y,z)\,\mathrm{d}V \ , \qquad (4.304)$$

where E is a region in space. Obviously, the first type yields a scalar while the second type yields a vector. For example, the mass of a fluid in a solid region E is given by $\iiint_E \rho(x,y,z)\,\mathrm{d}V$ where $\rho(x,y,z)$ is the density of the fluid. Moreover, the total linear momentum of that same fluid is given by $\iiint_E \rho(x,y,z)\mathbf{v}(x,y,z)\,\mathrm{d}V$ where $\mathbf{v}(x,y,z)$ is the velocity field in the fluid. We have already encountered the first type of volume integrals. The second type of volume integrals is defined as follows.

Definition 4.57. *Let $\mathbf{F}(x,y,z) = P(x,y,z)\,\hat{\imath} + Q(x,y,z)\,\hat{\jmath} + R(x,y,z)\,\hat{k}$ be a vector field defined on an open region E in space where P, Q, and R are integrable. Then,* **the volume integral of F over** *E is*

$$\iiint_E \mathbf{F}\,\mathrm{d}V = \hat{\imath}\iiint_E P\,\mathrm{d}V + \hat{\jmath}\iiint_E Q\,\mathrm{d}V + \hat{k}\iiint_E R\,\mathrm{d}V \ . \qquad (4.305)$$

Of course, we can express the vector field $\mathbf{F}$ in other coordinate systems. However, since the unit vectors in such systems are not in general constant, they cannot be taken out of the integral as in Eq. (4.305).

Example 4.58. Evaluate $\iiint_E \mathbf{F}\,\mathrm{d}V$, where $\mathbf{F}(x,y,z) = 2xz\,\hat{\imath} - x\,\hat{\jmath} + y^2\,\hat{k}$ and E is the region bounded by the surfaces $x = 0$, $x = 2$, $y = 0$, $y = 6$, $z = x^2$, and $z = 4$.

Using Eq. (4.305), we have

$$\begin{aligned}
&\iiint_E \mathbf{F}\,\mathrm{d}V \\
= \;& \hat{\imath}\iiint_E P\,\mathrm{d}V + \hat{\jmath}\iiint_E Q\,\mathrm{d}V + \hat{k}\iiint_E R\,\mathrm{d}V \\
= \;& \hat{\imath}\int_0^2\int_0^6\int_{x^2}^4 2xz\,\mathrm{d}z\mathrm{d}y\mathrm{d}x - \hat{\jmath}\int_0^2\int_0^6\int_{x^2}^4 x\,\mathrm{d}z\mathrm{d}y\mathrm{d}x \\
&+\hat{k}\int_0^2\int_0^6\int_{x^2}^4 y^2\,\mathrm{d}z\mathrm{d}y\mathrm{d}x \\
= \;& \hat{\imath}\int_0^2\int_0^6 [xz^2]_{x^2}^4\,\mathrm{d}y\,\mathrm{d}x - \hat{\jmath}\int_0^2\int_0^6 [xz]_{x^2}^4\,\mathrm{d}y\,\mathrm{d}x + \hat{k}\int_0^2\int_0^6 [y^2 z]_{x^2}^4\,\mathrm{d}y\,\mathrm{d}x \\
= \;& \hat{\imath}\int_0^2\int_0^6 (16x - x^5)\,\mathrm{d}y\,\mathrm{d}x - \hat{\jmath}\int_0^2\int_0^6 (4x - x^3)\,\mathrm{d}y\,\mathrm{d}x \\
&+\hat{k}\int_0^2\int_0^6 (4y^2 - x^2y^2)\,\mathrm{d}y\,\mathrm{d}x \\
= \;& \hat{\imath}\int_0^2 [16xy - x^5y]_0^6\,\mathrm{d}x - \hat{\jmath}\int_0^2 [4xy - x^3y]_0^6\,\mathrm{d}x \\
&+\hat{k}\int_0^2 \left[\frac{4}{3}y^3 - \frac{1}{3}x^2y^3\right]_0^6 \mathrm{d}x \\
= \;& \hat{\imath}\int_0^2 (96x - 6x^5)\,\mathrm{d}x - \hat{\jmath}\int_0^2 (24x - 6x^3)\,\mathrm{d}x + \hat{k}\int_0^2 (288 - 72x^2)\,\mathrm{d}x \\
= \;& [48x^2 - x^6]_0^2\,\hat{\imath} - \left[12x^2 - \frac{3}{2}x^4\right]_0^2 \hat{\jmath} + [288x - 24x^3]_0^2\,\hat{k} \\
= \;& 128\,\hat{\imath} - 24\,\hat{\jmath} + 384\,\hat{k}\ .
\end{aligned} \tag{4.306}$$

■

4.8 The Divergence Theorem

In this section, we consider another generalization of Green's Theorem to higher dimensional space. We start from the following form of Green's Theorem

$$\oint_C \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}s = \iint_D \operatorname{div}\mathbf{F}\,\mathrm{d}A \tag{4.307}$$

where C is an oriented, piecewise-smooth, simple closed curve that bounds a plane region D. The theorem states that the line integral of the normal component of a vector field in two-dimensional space around a simple closed curve is equal to the double integral of the divergence of the vector field over the plane region bounded by the curve.

The Divergence Theorem generalizes this form of Green's Theorem to three-dimensional space. This theorem, also called Gauss's Theorem, relates the surface integral of the normal component of a vector field in three-dimensional over a closed surface to a volume integral of the divergence of the vector field over the solid region bounded by the surface.

The Divergence Theorem is valid for general surfaces. However, we will focus our attention to **simple solid regions**, which is the solid regions that are simultaneously x-, y-, and z-simple. For examples, regions bounded by ellipsoids, cubes and tetrahedrons are simple solid regions.

Theorem 4.59. The Divergence Theorem
Let E be a simple solid region bounded by a closed piecewise-smooth surface S that has a outward unit normal vector $\hat{\mathbf{n}}$. If $\mathbf{F}$ is a vector field whose components have continuous partial derivatives on a open region in space that contains E, then

$$\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \iiint_E \operatorname{div} \mathbf{F} \, \mathrm{d}V \ . \tag{4.308}$$

Proof. We prove the Divergence Theorem for simple solid regions as follows. Let $\mathbf{F} = P\,\hat{\imath} + Q\,\hat{\jmath} + R\,\hat{k}$. Then,

$$\operatorname{div} \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \ . \tag{4.309}$$

Thus, the right-hand side of Eq. (4.308) has the form

$$\iiint_E \operatorname{div} \mathbf{F} \, \mathrm{d}V = \iiint_E \frac{\partial P}{\partial x} \, \mathrm{d}V + \iiint_E \frac{\partial Q}{\partial y} \, \mathrm{d}V + \iiint_E \frac{\partial R}{\partial z} \, \mathrm{d}V \ . \tag{4.310}$$

If $\hat{\mathbf{n}}$ is outward unit normal vector to S, then the left-hand side of Eq. (4.308) is

$$\begin{aligned} & \iint_S \mathbf{F} \cdot \hat{\mathbf{n}} \, \mathrm{d}S \\ = & \iint_S (P\,\hat{\imath} + Q\,\hat{\jmath} + R\,\hat{k}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S \\ = & \iint_S P\,\hat{\imath} \cdot \hat{\mathbf{n}} \, \mathrm{d}S + \iint_S Q\,\hat{\jmath} \cdot \hat{\mathbf{n}} \, \mathrm{d}S + \iint_S R\,\hat{k} \cdot \hat{\mathbf{n}} \, \mathrm{d}S \ . \end{aligned} \tag{4.311}$$

Therefore, the Divergence Theorem will be proved if we can show that

$$\iint_S P\,\hat{\imath} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \iiint_E \frac{\partial P}{\partial x} \, \mathrm{d}V \tag{4.312}$$

$$\iint_S Q\,\hat{\jmath} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \iiint_E \frac{\partial Q}{\partial y} \, \mathrm{d}V \tag{4.313}$$

$$\iint_S R\,\hat{k} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \iiint_E \frac{\partial R}{\partial z} \, \mathrm{d}V \ . \tag{4.314}$$

To prove Eq. (4.314), recall that E is a z-simple region and so it can be described by

$$E = \{(x, y, z) \mid (x, y) \in D, k_1(x, y) \leq z \leq k_2(x, y)\} \tag{4.315}$$

where D is the projection of E onto the xy-plane, and k_1 and k_2 are continuous function of x and y. Then, we have

$$\begin{aligned}\iiint_E \frac{\partial R}{\partial z}\,\mathrm{d}V &= \iint_D \left[\int_{k_1(x,y)}^{k_2(x,y)} \frac{\partial R}{\partial z}\,\mathrm{d}z\right] \mathrm{d}A \\ &= \iint_D [R(x, y, k_2(x, y)) - R(x, y, k_1(x, y))]\,\mathrm{d}A \end{aligned} \tag{4.316}$$

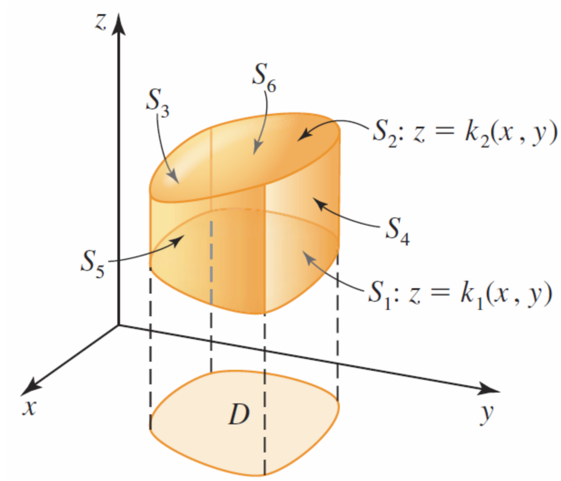


Figure 4.54: A z-simple region.

As shown in Fig. 4.54, the boundary surface S of E may consists of up to six surfaces. On each of the vertical surfaces, $\hat{k} \cdot \hat{\mathbf{n}} = 0$ and so

$$\iint_{S_3} R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S = \iint_{S_4} R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S = \iint_{S_5} R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S = \iint_{S_6} R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S = 0\,. \tag{4.317}$$

Therefore, no matter how many vertical surfaces S has, we can write

$$\iint_S R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S = \iint_{S_1} R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S + \iint_{S_2} R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S\,. \tag{4.318}$$

Note that S_2 is the graph of the function $z = k_2(x, y)$ and its outward normal $\hat{\mathbf{n}}$ points upward. Therefore, using Eq. (4.268) with $\mathbf{F}$ replaced by $R\,\hat{k}$, we obtain

$$\iint_{S_2} R\,\hat{k} \cdot \hat{\mathbf{n}}\,\mathrm{d}S = \iint_D R(x, y, k_2(x, y))\,\mathrm{d}A\,. \tag{4.319}$$

On the other hand, S_1 is the graph of the function $z = k_1(x, y)$ and its outward normal $\hat{\mathbf{n}}$ points downward. Therefore, using Eq. (4.268) with $\mathbf{F}$

replaced by $R\,\hat{k}$ and multiplying it by -1, we obtain

$$\iint_{S_1} R\,\hat{k}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = -\iint_D R(x,y,k_1(x,y))\,\mathrm{d}A\ . \tag{4.320}$$

As a result, Eq. (4.318) yields

$$\iint_S R\,\hat{k}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \iint_D [R(x,y,k_2(x,y)) - R(x,y,k_1(x,y))]\,\mathrm{d}A\ . \tag{4.321}$$

Comparing with Eq. (4.316), we have

$$\iint_S R\,\hat{k}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \iiint_E \frac{\partial R}{\partial z}\,\mathrm{d}V\ . \tag{4.322}$$

We can prove Eq. (4.312) and (4.313) in a similar manner by viewing E as a x-simple or y-simple region. ■

Example 4.60. Find the flux of the vector field $\mathbf{F}(x,y,z) = z\,\hat{\imath} + y\,\hat{\jmath} + x\,\hat{k}$ over the outward-oriented unit sphere S given by $x^2+y^2+z^2=1$.

First, we compute the divergence of $\mathbf{F}$

$$\operatorname{div}\mathbf{F} = \frac{\partial}{\partial x}(z) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}(x) = 0+1+0 = 1\ . \tag{4.323}$$

The unit sphere S is the boundary of the unit ball B given by $x^2+y^2+z^2 \le 1$. Thus, the flux is given by the Divergence Theorem, Eq. (4.308), as

$$\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iiint_B \operatorname{div}\mathbf{F}\,\mathrm{d}V = \iiint_B 1\,\mathrm{d}V = V(B) = \frac{4\pi}{3}\ . \tag{4.324}$$

■

Example 4.61. Let S be the outward-oriented surface of a solid E bounded by the cylinder $x^2+y^2=4$ and the planes $z=0$ and $z=3$. Calculate the outward flux of the vector field $\mathbf{F}(x,y,z) = xy^2\,\hat{\imath} + yz^2\,\hat{\jmath} + zx^2\,\hat{k}$ over S.

The surface S is shown in Fig. 4.55. The flux of $\mathbf{F}$ over S is given by $\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S}$. By the Divergence Theorem, the flux can also be found by evaluating $\iiint_E \operatorname{div}\mathbf{F}\,\mathrm{d}V$. The divergence of $\mathbf{F}$ is

$$\operatorname{div}\mathbf{F} = \frac{\partial}{\partial x}(xy^2) + \frac{\partial}{\partial y}(yz^2) + \frac{\partial}{\partial z}(zx^2) = x^2+y^2+z^2\ . \tag{4.325}$$

Therefore,

$$\iint_S \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iiint_E \operatorname{div}\mathbf{F}\,\mathrm{d}V = \iiint_E (x^2+y^2+z^2)\,\mathrm{d}V\ . \tag{4.326}$$

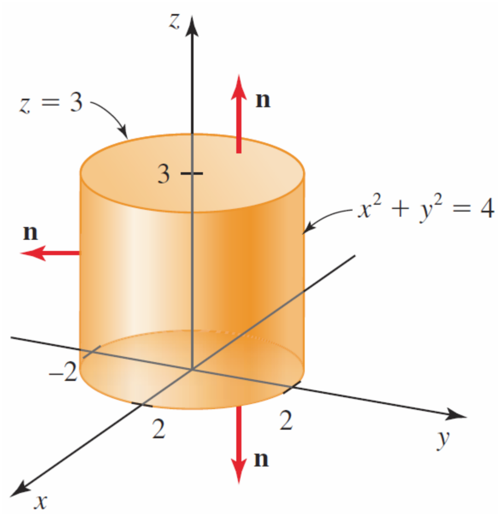


Figure 4.55: For Example 4.61.

Using cylindrical coordinates to evaluate the triple integral, we obtain

$$\begin{aligned}
\iint_S \mathbf{F}\cdot \mathrm{d}\mathbf{S} &= \int_0^{2\pi}\int_0^2\int_0^3 (r^2+z^2)\, r\, \mathrm{d}z\, \mathrm{d}r \mathrm{d}\theta \\
&= \int_0^{2\pi}\int_0^2 \left[r^3 z + \frac{1}{3} r z^3\right]_0^3 \mathrm{d}r\, \mathrm{d}\theta \\
&= \int_0^{2\pi}\int_0^2 (3r^3+9r)\mathrm{d}r\mathrm{d}\theta \\
&= \int_0^{2\pi} \left[\frac{3}{4} r^4 + \frac{9}{2} r^2\right]_0^2 \mathrm{d}\theta \\
&= 30\int_0^{2\pi} \mathrm{d}\theta \\
&= 60\pi\ .
\end{aligned} \tag{4.327}$$

■

Example 4.62. Evaluate $\iint_S \mathbf{F}\cdot \mathrm{d}\mathbf{S}$, where $\mathbf{F}(x,y,z) = xy\,\hat{\imath} + (y^2 + e^{xz^2})\,\hat{\jmath} + \sin(xy)\,\hat{k}$ and S is the outward-oriented surface of the region E bounded by the parabolic cylinder $z = 1 - x^2$ and the planes $y = 0$, $z = 0$, and $y + z = 2$, Fig. 4.56.

It would be extremely difficult to evaluate the given surface integral directly since we would have to evaluate four surface integrals corresponding to the four pieces of S. Furthermore, the divergence of $\mathbf{F}$ is much less complicated than itself

$$\operatorname{div}\mathbf{F} = \frac{\partial}{\partial x}(xy) + \frac{\partial}{\partial y}(y^2 + e^{xz^2}) + \frac{\partial}{\partial z}[\sin(xy)] = y + 2y + 0 = 3y\ . \tag{4.328}$$

Thus, we use the Divergence Theorem to transform the given surface integral into a triple integral. The easiest way to evaluate the triple integral is to

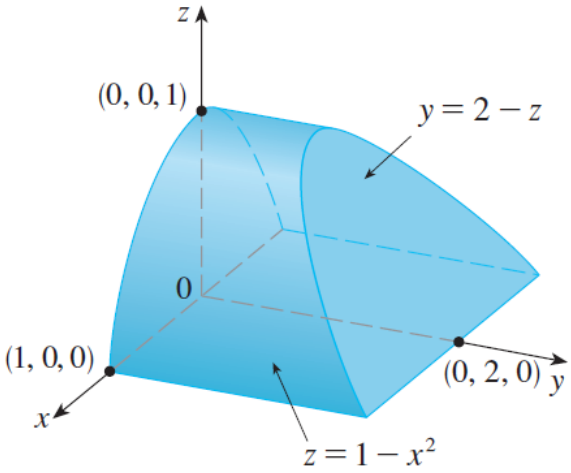


Figure 4.56: For Example 4.62.

express the region E as

$$E = \{(x, y, z) \mid -1 \leq x \leq 1, 0 \leq z \leq 1 - x^2, 0 \leq y \leq 2 - z\} \,. \qquad (4.329)$$

Then, we have

$$\begin{aligned}
\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} &= \iiint_E \operatorname{div} \mathbf{F} \, \mathrm{d}V \\
&= \iiint_E 3y \, \mathrm{d}V \\
&= 3 \int_{-1}^{1} \int_0^{1-x^2} \int_0^{2-z} y \, \mathrm{d}y \, \mathrm{d}z \, \mathrm{d}x \\
&= \frac{3}{2} \int_{-1}^{1} \int_0^{1-x^2} (2-z)^2 \mathrm{d}z \, \mathrm{d}x \\
&= \frac{3}{2} \int_{-1}^{1} \left[-\frac{1}{3}(2-z)^3 \right]_0^{1-x^2} \mathrm{d}x \\
&= -\frac{1}{2} \int_{-1}^{1} [(x^2+1)^3 - 8] \, \mathrm{d}x \\
&= -\int_0^1 (x^6 + 3x^4 + 3x^2 - 7) \, \mathrm{d}x \\
&= \frac{184}{35} \,. \qquad (4.330)
\end{aligned}$$

■

We have proved the Divergence Theorem only for simple solid regions. However, it can be proved for regions that are finite unions of simple solid regions. For example, let E be the region that lies between the closed surfaces S_1 and S_2 where S_1 lies inside S_2 as shown in Fig. 4.57. Then, the boundary surface of E is $S = S_1 \cup S_2$ and its normal is given by $\hat{\mathbf{n}} = -\hat{\mathbf{n}}_1$ on S_1 and

$\hat{\mathbf{n}} = \hat{\mathbf{n}}_2$ on S_2. Applying the Divergence Theorem to E, we have

$$\begin{aligned}\iiint_E \operatorname{div}\mathbf{F}\,\mathrm{d}V &= \iint_S \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}S \\ &= \iint_{S_1} \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}S + \iint_{S_2} \mathbf{F}\cdot\hat{\mathbf{n}}\,\mathrm{d}S \\ &= \iint_{S_1} \mathbf{F}\cdot(-\hat{\mathbf{n}}_1)\,\mathrm{d}S + \iint_{S_2} \mathbf{F}\cdot\hat{\mathbf{n}}_2\,\mathrm{d}S \\ &= -\iint_{S_1} \mathbf{F}\cdot\mathrm{d}\mathbf{S} + \iint_{S_2} \mathbf{F}\cdot\mathrm{d}\mathbf{S}\ . \end{aligned} \tag{4.331}$$

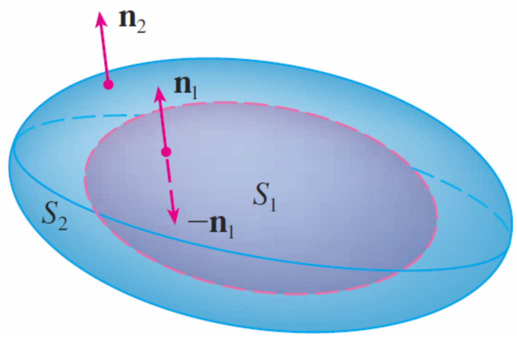


Figure 4.57: One surface inside another.

Example 4.63. Consider the electric field

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0}\frac{\mathbf{r}}{||\mathbf{r}||^3} \tag{4.332}$$

induced by a point charge q placed at the origin of a three-dimensional coordinate system, where $\mathbf{r} = x\,\hat{\imath} + y\,\hat{\jmath} + z\,\hat{k}$. Find the outward flux of $\mathbf{E}$ across a smooth surface S that encloses the origin.

We cannot apply the Divergence Theorem directly to find the outward flux of $\mathbf{E}$ across S because $\mathbf{E}$ is not continuous at the origin. To avoid such difficulty, let's construct a sphere with center at the origin and a radius a that is small enough to ensure that the sphere lies completely inside S. If we denote this sphere by S_1, then $\mathbf{E}$ satisfies the conditions of the Divergence Theorem for the solid E that lies between S_1 and S. Using Eq. (4.331), we obtain

$$\iiint_E \operatorname{div}\mathbf{E}\,\mathrm{d}V = -\iint_{S_1} \mathbf{E}\cdot\mathrm{d}\mathbf{S} + \iint_{S} \mathbf{E}\cdot\mathrm{d}\mathbf{S}\ . \tag{4.333}$$

Besides, the divergence of $\mathbf{E}$ is

$$\begin{aligned}
&\operatorname{div}\mathbf{E} \\
=\;& \frac{\partial}{\partial x}\left[\frac{x}{(x^2+y^2+z^2)^{3/2}}\right]+\frac{\partial}{\partial y}\left[\frac{y}{(x^2+y^2+z^2)^{3/2}}\right] \\
&+\frac{\partial}{\partial z}\left[\frac{z}{(x^2+y^2+z^2)^{3/2}}\right] \\
=\;& \left[\frac{-2x^2+y^2+z^2}{(x^2+y^2+z^2)^{5/2}}\right]+\left[\frac{x^2-2y^2+z^2}{(x^2+y^2+z^2)^{5/2}}\right]+\left[\frac{x^2+y^2-2z^2}{(x^2+y^2+z^2)^{5/2}}\right] \\
=\;& 0\;.
\end{aligned} \tag{4.334}$$

So we have

$$\iint_S \mathbf{E}\cdot \mathrm{d}\mathbf{S} = \iint_{S_1} \mathbf{E}\cdot \mathrm{d}\mathbf{S} = \iint_{S_1} \mathbf{E}\cdot \hat{\mathbf{n}}\,\mathrm{d}S\;. \tag{4.335}$$

To evaluate the integral on the right, note that the outward unit normal to the sphere S_1 is $\hat{\mathbf{n}} = \mathbf{r}/||\mathbf{r}||$. Therefore, on the sphere S_1,

$$\mathbf{E}\cdot\hat{\mathbf{n}} = \left(\frac{q}{4\pi\epsilon_0}\frac{\mathbf{r}}{||\mathbf{r}||^3}\right)\cdot\left(\frac{\mathbf{r}}{||\mathbf{r}||}\right) = \frac{q}{4\pi\epsilon_0}\frac{\mathbf{r}\cdot\mathbf{r}}{||\mathbf{r}||^4} = \frac{q}{4\pi\epsilon_0||\mathbf{r}||^2} = \frac{q}{4\pi\epsilon_0 a^2} \tag{4.336}$$

because $||\mathbf{r}|| = a$ on the sphere S_1. Thus, we have

$$\iint_S \mathbf{E}\cdot \mathrm{d}\mathbf{S} = \frac{q}{4\pi\epsilon_0 a^2}\iint_{S_1}\mathrm{d}S = \frac{q}{4\pi\epsilon_0 a^2}(4\pi a^2) = \frac{q}{\epsilon_0}\;. \tag{4.337}$$

■

Just as Stokes' Theorem can be used to gain some insight into the meaning of the curl vector, the Divergence Theorem can be used to gain some insight into the meaning of the divergence vector. Suppose $\mathbf{v}(x, y, z)$ is the velocity field of a fluid with constant mass density ρ. Then, $\mathbf{F} = \rho\mathbf{v}$ is the rate of flow of the fluid per unit area. Let $P_0(x_0, y_0, z_0)$ be a point in the fluid and B_a be a ball with center P_0 and very small radius a. Then, $\operatorname{div}\mathbf{F}(P) \approx \operatorname{div}\mathbf{F}(P_0)$ for all points P in B_a since $\operatorname{div}\mathbf{F}$ is continuous. Therefore, by the Divergence Theorem, we obtain the following approximation to the outward flux of $\mathbf{F}$ over the boundary sphere S_a

$$\iint_{S_a}\mathbf{F}\cdot\mathrm{d}\mathbf{S} = \iiint_{B_a}\operatorname{div}\mathbf{F}\,\mathrm{d}V \approx \iiint_{B_a}\operatorname{div}\mathbf{F}(P_0)\,\mathrm{d}V = \operatorname{div}\mathbf{F}(P_0)\,V(B_a)\;. \tag{4.338}$$

This approximation improves as $a \to 0$ which suggests that

$$\operatorname{div}\mathbf{F}(P_0) = \lim_{a\to 0}\frac{1}{V(B_a)}\iint_{S_a}\mathbf{F}\cdot\mathrm{d}\mathbf{S}\;. \tag{4.339}$$

The above equation tells us that $\operatorname{div}\mathbf{F}(P_0)$ can be regarded as the net rate of outward flow of the fluid per unit volume at P_0 — hence the name divergence. If $\operatorname{div}\mathbf{F}(P) > 0$, the net flow is outward near P and P is called a **source**. If $\operatorname{div}\mathbf{F}(P) < 0$, the net flow is inward near P and P is called a **sink**. In addition, if the fluid is incompressible and there are no sources or sinks present, then no fluid exits or enters B_a and thus $\operatorname{div}\mathbf{F}(P) = 0$ at every point P.

Chapter 5

Fourier Series and Delta Function

When we study oscillations in physics, we usually focus on the sine and cosine functions. One reason is that they are relatively simple. However, the main reason is that a "general" periodic function is a linear combination of them, maybe infinitely many. This fact may come as a surprise, and the topic is called the Fourier analysis. In this chapter, we will discuss the basics and mainly follow Jeffrey, *Advanced Engineering Mathematics*.

5.1 Fourier Series

We will consider periodic functions, $f(x+L) = f(x)$, with period L on the interval $0 \le x \le L$. The most important ones are $\sin\frac{2\pi nx}{L}$ and $\cos\frac{2\pi nx}{L}$, where n are non-negative integers. First, note that for $n, m \ge 1$, we have

$$\frac{1}{L}\int_0^L 1\,\mathrm{d}x = 1 \tag{5.1}$$

$$\frac{2}{L}\int_0^L \sin\frac{2\pi nx}{L}\,\mathrm{d}x = 0 \tag{5.2}$$

$$\frac{2}{L}\int_0^L \cos\frac{2\pi nx}{L}\,\mathrm{d}x = 0 \tag{5.3}$$

$$\frac{2}{L}\int_0^L \sin\frac{2\pi nx}{L}\sin\frac{2\pi mx}{L}\,\mathrm{d}x = \delta_{nm} \tag{5.4}$$

$$\frac{2}{L}\int_0^L \cos\frac{2\pi nx}{L}\cos\frac{2\pi mx}{L}\,\mathrm{d}x = \delta_{nm} \tag{5.5}$$

$$\frac{2}{L}\int_0^L \sin\frac{2\pi nx}{L}\cos\frac{2\pi mx}{L}\,\mathrm{d}x = 0\,. \tag{5.6}$$

These can be easily verified using, for example, $\sin A\sin B = (\cos(A-B) - \cos(A+B))/2$.

If a function f is defined on the interval $0 \le x \le L$ and it is a combination of the sine and cosine functions

$$f(x) = a_0 + \sum_{n=1}^{N} \left(a_n \cos \frac{2\pi n x}{L} + b_n \sin \frac{2\pi n x}{L} \right) \ , \tag{5.7}$$

we can find out the coefficients as

$$\begin{aligned}
& \frac{1}{L} \int_0^L f \, \mathrm{d}x \\
= \ & \frac{1}{L} \int_0^L a_0 + \sum_{m=1}^{N} \left(a_m \cos \frac{2\pi m x}{L} + b_m \sin \frac{2\pi m x}{L} \right) \mathrm{d}x \\
= \ & a_0 \ , \qquad (5.8) \\
& \frac{2}{L} \int_0^L f \cos \frac{2\pi n x}{L} \, \mathrm{d}x \\
= \ & \frac{2}{L} \int_0^L \left[a_0 + \sum_{m=1}^{N} \left(a_n \cos \frac{2\pi m x}{L} + b_m \sin \frac{2\pi m x}{L} \right) \right] \cos \frac{2\pi n x}{L} \, \mathrm{d}x \\
= \ & a_n \ , \qquad (5.9) \\
& \frac{2}{L} \int_0^L f \sin \frac{2\pi n x}{L} \, \mathrm{d}x \\
= \ & \frac{2}{L} \int_0^L \left[a_0 + \sum_{m=1}^{N} \left(a_n \cos \frac{2\pi m x}{L} + b_m \sin \frac{2\pi m x}{L} \right) \right] \sin \frac{2\pi n x}{L} \, \mathrm{d}x \\
= \ & b_n \ . \qquad (5.10)
\end{aligned}$$

In summary, if

$$f(x) = a_0 + \sum_{n=1}^{N} \left(a_n \cos \frac{2\pi n x}{L} + b_n \sin \frac{2\pi n x}{L} \right) \ , \tag{5.11}$$

then

$$\begin{aligned}
a_0 &= \frac{1}{L} \int_0^L f \, \mathrm{d}x \ , & (5.12) \\
a_n &= \frac{2}{L} \int_0^L f \cos \frac{2\pi n x}{L} \, \mathrm{d}x \ , & (5.13) \\
b_n &= \frac{2}{L} \int_0^L f \sin \frac{2\pi n x}{L} \, \mathrm{d}x \ . & (5.14)
\end{aligned}$$

If we ignore that convergence problem (for the moment), we could take $N \to \infty$ in above formulas. The resulting expression, Eq. (5.11), is the **Fourier series** of the function f and the a_n and b_n are the **Fourier coefficients**.

Example 5.1. The function $f(x) = \sin\left(\frac{2\pi x}{L} + \alpha\right)$ is periodic of period L.

We have

$$f(x) = \sin\left(\frac{2\pi x}{L} + \alpha\right) = \sin\alpha\cos\frac{2\pi x}{L} + \cos\alpha\sin\frac{2\pi x}{L} \tag{5.15}$$

by trigonometrical formulas. Hence, we find that $a_1 = \sin\alpha$ and $b_1 = \cos\alpha$, and all other Fourier coefficients are zero. ■

Example 5.2. In this example, we consider the square waves, Figure 5.1, and we put $L = 2\pi$ to simplify some calculations. The function is

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x < \pi \\ -1 & \text{if } \pi \le x < 2\pi \end{cases} . \tag{5.16}$$

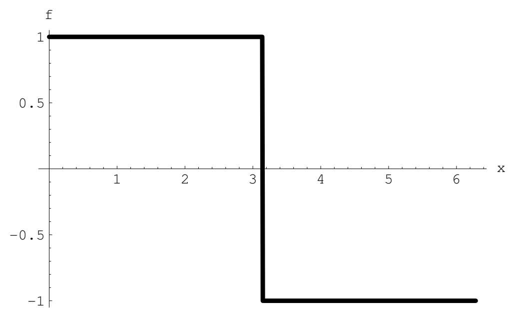


Figure 5.1: The square waves.

We have

$$\begin{aligned} a_0 &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}x = 0 , \\ a_n &= \frac{1}{\pi}\int_0^{\pi} \cos nx\,\mathrm{d}x - \frac{1}{\pi}\int_\pi^{2\pi} \cos nx\,\mathrm{d}x = 0 , \\ b_n &= \frac{1}{\pi}\int_0^{\pi} \sin nx\,\mathrm{d}x - \frac{1}{\pi}\int_\pi^{2\pi} \sin nx\,\mathrm{d}x = \frac{2(1-(-1)^n)}{n\pi} . \end{aligned} \tag{5.17}$$

The Fourier series is

$$\sum_{n=0}^{\infty} \frac{4}{(2n+1)\pi}\sin(2n+1)x . \tag{5.18}$$

Let the partial sum be

$$f_N(x) \equiv \sum_{n=0}^{N} \frac{4}{(2n+1)\pi}\sin(2n+1)x . \tag{5.19}$$

We have plotted the $f_0(x)$, $f_3(x)$ and $f_9(x)$ in Figure 5.2. The convergence looks good. ■

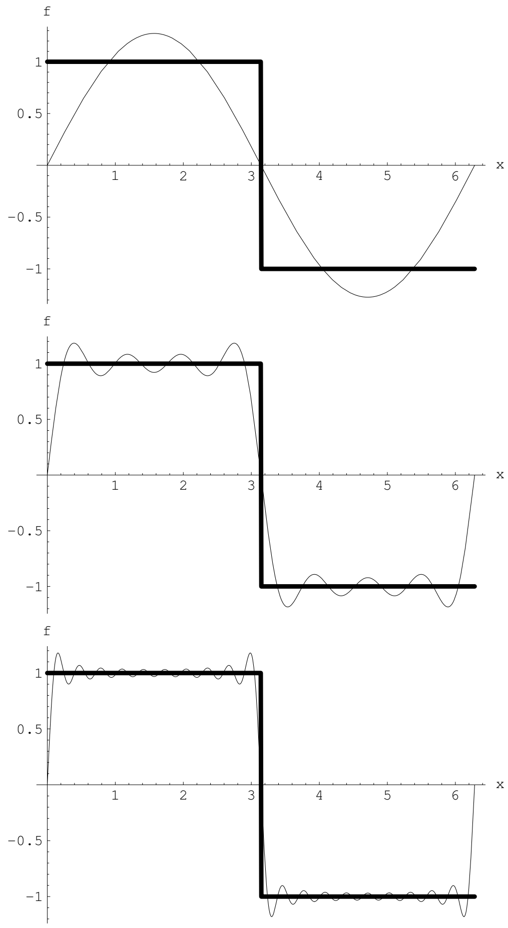


Figure 5.2: Partial sum of Fourier series.

Example 5.3. Consider the sawtooth waves, Figure 5.3, and we still put $L = 2\pi$. The function is

$$f(x) = x \qquad 0 \leq x \leq 2\pi. \tag{5.20}$$

We have

$$\begin{aligned} a_0 &= \frac{1}{2\pi}\int_0^{2\pi} x \,\mathrm{d}x = \pi \ , \\ a_n &= \frac{1}{\pi}\int_0^{2\pi} x \cos nx \,\mathrm{d}x = 0 \ , \\ b_n &= \frac{1}{\pi}\int_0^{2\pi} x \sin nx \,\mathrm{d}x = \frac{-2}{n} \ . \end{aligned} \tag{5.21}$$

The Fourier series is

$$\pi - \sum_{n=1}^{\infty} \frac{2}{n} \sin nx \ . \tag{5.22}$$

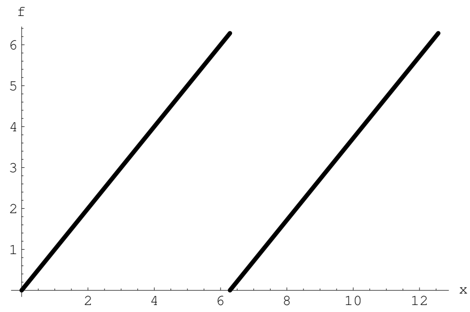


Figure 5.3: The sawtooth waves.

Let the partial sum be

$$f_N(x) \equiv \pi - \sum_{n=1}^{N} \frac{2}{n} \sin nx \ . \tag{5.23}$$

We have plotted the $f_1(x)$, $f_3(x)$ and $f_9(x)$ in Figure 5.4. ■

Theorem 5.4. Assume that

1. a function is periodic with period L, $f(x+L) = f(x)$;
2. both $f(x)$ and its derivative $f'(x)$ are piecewise continuous, except at finite number (could be zero) of points in $0 \le x \le L$;
3. at those exceptional points, it has finite discontinuities. This means that
$$\lim_{\epsilon_1, \epsilon_2 \to 0} \left(f(x+\epsilon_1) - f(x-\epsilon_2) \right) \tag{5.24}$$
is finite;

then, the partial sum of its Fourier series

$$f_N(x) \equiv a_0 + \sum_{n=1}^{N} \left(a_n \cos \frac{2\pi n x}{L} + b_n \sin \frac{2\pi n x}{L} \right) \tag{5.25}$$

converges to

$$\lim_{N\to\infty} f_N(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) + f(x-\epsilon)}{2} \ . \tag{5.26}$$

■

A few remarks are in order.

1. Point 3 of the theorem means that we do not allow discontinuities like $1/x$.
2. If f is continuous at a point x, then the limit in Eq. (5.26) converges to $f(x)$.

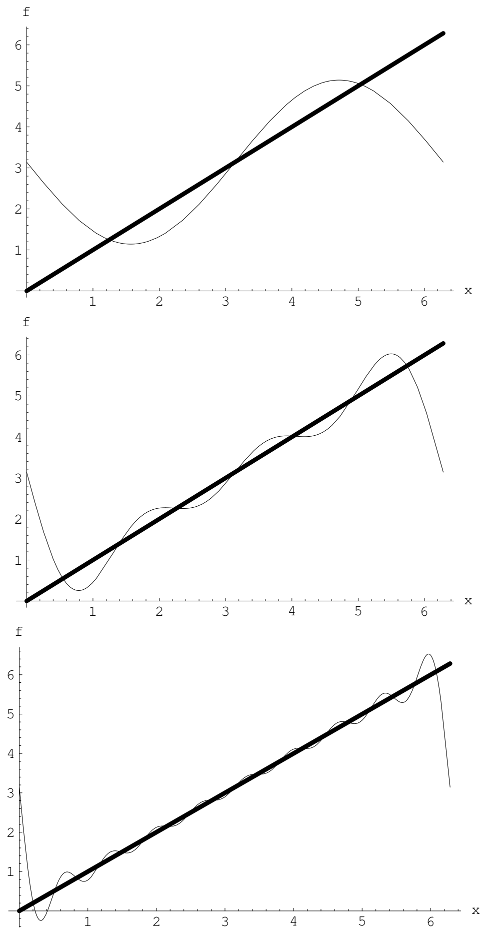


Figure 5.4: Partial sum of Fourier series.

3. If f is not continuous at a point x, then the limit in Eq. (5.26) converges to the "average". In Example 5.2, since the function is periodic, its value at $x = 0$ is $f(x = 0) = f(x = 2\pi) = 1$ according to Eq. (5.16). However, the partial sum, Eq. (5.19), at $x = 0$ is 1/2. While in Example 5.3, the partial sum, Eq. (5.23), at $x = 0$ is π. This also shows up in the corresponding graphs, Figure 5.2 and Figure 5.4.

4. Fourier series expansion can be applied to functions more general than those described in this theorem.

Theorem 5.5. Parseval's Theorem

$$\frac{2}{L}\int_0^L (f(x))^2\,\mathrm{d}x = 2a_0^2 + \sum_{n=1}^{\infty}\left(a_n^2 + b_n^2\right) \ . \tag{5.27}$$

Proof. Ignoring the problem of convergence, by Eq. (5.1) to Eq. (5.6), we have

$$
\begin{aligned}
& \frac{2}{L}\int_0^L (f_N(x))^2\,\mathrm{d}x \\
= & \frac{2}{L}\int_0^L \left(a_0 + \sum_{n=1}^{N}\left(a_n \cos\frac{2\pi n x}{L} + b_n \sin\frac{2\pi n x}{L}\right)\right)^2 \mathrm{d}x \\
= & 2a_0^2 + \sum_{n=1}^{N}\left(a_n^2 + b_n^2\right) \ . \qquad (5.28)
\end{aligned}
$$

Take N to infinity and we have the theorem. ∎

Students with sharp eyes have seen that the partial sum will "overshoot" the original function at discontinuities, for example, in Figure 5.2 near $x = 0$ and $x = \pi$. If we add in more terms in the partial sum, the "overshooting" does not improve. It only moves closer and closer to the discontinuity. This is known as **Gibbs phenomenon**. For square waves, it overshoots by 18 percent. See Arfken and Weber, p. 910 for details.

If the function is well behaved (see references for details), we can integrate or differentiate the Fourier series term by term.

Example 5.6. Consider the triangular waves, Figure 5.5,

$$
f(x) = \begin{cases} x & 0 \le x < \pi \\ 2\pi - x & \pi \le x < 2\pi \end{cases} \ . \qquad (5.29)
$$

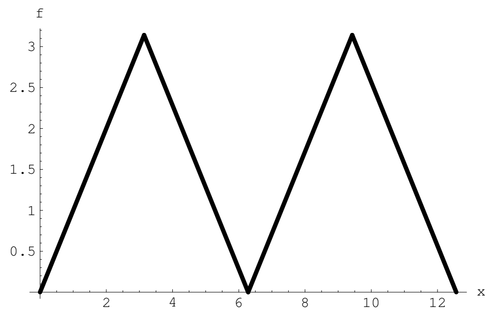


Figure 5.5: The triangular waves.

We have

$$
\begin{aligned}
a_0 &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x = \frac{\pi}{2} \ , \\
a_n &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,\mathrm{d}x = \frac{-2(1-(-1)^n)}{n^2\pi} \ , \\
b_n &= \frac{1}{\pi}\int_0^{2\pi} x \sin nx\,\mathrm{d}x = 0 \ . \qquad (5.30)
\end{aligned}
$$

The Fourier series is

$$\frac{\pi}{2} - \sum_{n=0}^{\infty} \frac{4}{(2n+1)^2\pi} \cos(2n+1)x \ . \tag{5.31}$$

If we differentiate the Fourier series term by term, we have

$$\sum_{n=0}^{\infty} \frac{4}{(2n+1)\pi} \sin(2n+1)x \ , \tag{5.32}$$

which is exactly Eq. (5.18). ■

If a function is defined only in the range $0 \le x < L$, we could extend it to $-L \le x < L$ by defining for $-L \le x < 0$ $f(x) = \pm f(-x)$ and treat it as a periodic function with period $2L$. The first choice, $f(x) = f(-x)$, gives us a symmetric function. If we expand it as Fourier series, the Fourier coefficients for the sine terms will be zero, $b_n = 0$ for all n. The second choice, $f(x) = -f(-x)$, gives us an antisymmetric function. The Fourier coefficients for the cosine terms will be zero, $a_n = 0$ for all n.

Let $f(x)$ be a complex valued function for $-L/2 \le x < L/2$. (We shift the domain for later convenience.) We can expand its real and imaginary parts as Fourier series and then combine them to form a Fourier series with complex coefficients. Using the Euler identity

$$e^{ix} = \cos x + i \sin x \ , \tag{5.33}$$

we could express the resulting series in terms of e^{inx}, and obtain the **complex Fourier series**. It is easier to do it directly. The proof is similar to what we have given in the beginning of this chapter. We just quote the result.

For m and n integers (could be positive, negative or zero), we have

$$\frac{1}{L}\int_{-L/2}^{L/2} \exp\frac{2\pi i(n-m)x}{L}\,\mathrm{d}x = \delta_{nm} \ . \tag{5.34}$$

If $f(x)$ is a periodic complex valued function with period L, define

$$c_n \equiv \frac{1}{L}\int_{-L/2}^{L/2} f(x) e^{-2\pi i n x/L}\,\mathrm{d}x \ , \tag{5.35}$$

then

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/L} \ . \tag{5.36}$$

Example 5.7. Consider the square wave in Example 5.2 again.

$$\begin{aligned} c_n &= \frac{1}{2\pi}\left(\int_0^{\pi} e^{-inx}\,\mathrm{d}x - \int_{-\pi}^{0} e^{-inx}\,\mathrm{d}x\right) \\ &= \frac{1}{2\pi}\left(\frac{e^{-in\pi}-1}{-in} - \frac{1-e^{-in\pi}}{-in}\right) \\ &= \frac{1-(-1)^n}{in\pi} \ . \end{aligned} \tag{5.37}$$

The Fourier series is then

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} \frac{1-(-1)^n}{in\pi} e^{inx} &= \sum_{n \text{ odd}} \frac{2}{in\pi} e^{inx} \\
&= \sum_{n \text{ odd},=1}^{\infty} \frac{2}{in\pi} e^{inx} - \sum_{n \text{ odd},=1}^{\infty} \frac{2}{in\pi} e^{-inx} \\
&= \sum_{n \text{ odd},=1}^{\infty} \frac{2}{in\pi} (e^{inx} - e^{-inx}) \\
&= \sum_{n \text{ odd},=1}^{\infty} \frac{4}{n\pi} \sin nx \\
&= \sum_{m=0}^{\infty} \frac{4}{(2m+1)\pi} \sin(2m+1)x \ , \qquad (5.38)
\end{aligned}$$

which is Eq. (5.18). ■

5.2 Fourier Transform

With the notation $\omega_n \equiv 2n\pi/L$ and $\Delta\omega = 2\pi/L$, Eq. (5.35) and Eq. (5.36) suggest the following.

$$\begin{aligned}
f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{2\pi inx/L} \\
&= \sum_{n=-\infty}^{\infty} \frac{1}{L} \int_{-L/2}^{L/2} f(y) e^{-2\pi iny/L} \, \mathrm{d}y \, e^{2\pi inx/L} \\
&= \sum_{n=-\infty}^{\infty} \frac{\Delta\omega}{2\pi} \left(\int_{-L/2}^{L/2} f(y) e^{-i\omega_n y} \, \mathrm{d}y \right) e^{i\omega_n x} \ . \qquad (5.39)
\end{aligned}$$

This looks like the integral of some function.

Definition 5.8. *For a well-behaved function f defined on the real line, its* **Fourier transform** *is defined to be*

$$\tilde{f}(\omega) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y) e^{-i\omega y} \, \mathrm{d}y \ . \qquad (5.40)$$

The **inverse Fourier transform** *for a function $\tilde{g}$ is defined to be*

$$g(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{g}(\omega) e^{i\omega x} \, \mathrm{d}\omega \ . \qquad (5.41)$$

Theorem 5.9. Fourier Integral Theorem
For a well-behaved function f defined on the real line, we have

$$\begin{aligned}
&\lim_{\epsilon_1,\epsilon_2\to 0}\frac{f(x+\epsilon_1)+f(x-\epsilon_2)}{2} \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(y)e^{-i\omega y}\,\mathrm{d}y\right)e^{i\omega x}\,\mathrm{d}\omega \ . \qquad (5.42)
\end{aligned}$$

(We have to worry about the convergence, discontinuity, etc. Please read Jeffrey, *Advanced Engineering Mathematics*.) ■

If we interpret the left hand side of the above equation as $f(x)$, this theorem tells us that the original function is a linear combinations of waves $e^{i\omega x}$. The Fourier transform gives us the "amplitude" of the waves at some particular angular frequency ω.

Example 5.10. Find out the Fourier transform of $f(x) = e^{-|x|}$.
We calculate

$$\begin{aligned}
\tilde{f}(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-|y|}e^{-i\omega y}\,\mathrm{d}y \\
&= \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} e^{-y-i\omega y}\,\mathrm{d}y \\
&= \sqrt{\frac{2}{\pi}}\,\frac{1}{1+i\omega} \ . \qquad (5.43)
\end{aligned}$$

■

Example 5.11. Find out the Fourier transform of $f(x) = e^{-x^2}$ and verify the Fourier integral theorem.

$$\begin{aligned}
\tilde{f}(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2}e^{-i\omega y}\,\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2-i\omega y}\,\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(y+i\omega/2)^2-\omega^2/4}\,\mathrm{d}y \\
&= \frac{e^{-\omega^2/4}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(y+i\omega/2)^2}\,\mathrm{d}y \\
&= \frac{e^{-\omega^2/4}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2}\,\mathrm{d}z \\
&= \frac{e^{-\omega^2/4}}{\sqrt{2}} \ . \qquad (5.44)
\end{aligned}$$

Apply inverse Fourier transform, we have

$$\begin{aligned}\frac{1}{2\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-\omega^2/4}\, e^{i\omega x}\,\mathrm{d}\omega &= \frac{1}{2\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(\omega/2-ix)^2-x^2}\,\mathrm{d}\omega \\ &= \frac{e^{-x^2}}{2\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(\omega/2-ix)^2}\,\mathrm{d}\omega \\ &= e^{-x^2}\ . \qquad (5.45)\end{aligned}$$

■

5.3 Dirac Delta Function

Dirac δ-function is not a function.

Recall the meaning of **impulse** in mechanics. Imagine a tennis player hits the tennis ball with his racket. We could denote the incoming and outgoing speed of the ball be u and v respectively. The momenta will be $-mu$ and mv. The change in momentum is $\Delta p \equiv m(u+v)$.

How hard must the player hit the ball? This question is not well defined because he could use a weak force and apply for a long time or he could use a strong force and short duration, as long as the change in momentum is the given value

$$\Delta p = \int \mathbf{F}\,\mathrm{d}t\ . \qquad (5.46)$$

It is convenient to consider the limiting case that the force is extremely strong and the duration is extremely short, and yet, the above integration gives some definite value. In the tennis ball example, this means the force is so strong that the ball changes its velocity instantaneously. In this section, we follow Riley, et.al., *Mathematical Methods for Physics and Engineering.*

Definition 5.12. *The Dirac δ-function has the properties that*

$$\delta(x) = 0 \qquad \textit{for } x \neq 0\ , \qquad (5.47)$$

and

$$\int_{-\infty}^{\infty} f(x)\,\delta(x-a)\,\mathrm{d}x = f(a) \qquad (5.48)$$

for well-behaved function $f(x)$. (In rigorous mathematics, Dirac δ-function is defined as a distribution.)

Thinking an integral as area under the curve, we could say that the curve of the δ-function is infinitely narrow and infinitely tall such that the area is still 1. Define

$$\delta_\epsilon(x) \equiv \begin{cases} 1/\epsilon & \text{if } |x| < \epsilon/2 \\ 0 & \text{otherwise} \end{cases}\ , \qquad (5.49)$$

then we could think

$$\delta(x) \sim \lim_{\epsilon\to 0}\delta_\epsilon(x)\ . \qquad (5.50)$$

Take $f(x) \equiv 1$ in Eq. (5.48), we have

$$\int_{-\infty}^{\infty} \delta(x-a)\,\mathrm{d}x = 1 \ . \tag{5.51}$$

In fact, because of Eq. (5.47), for all $a, b > 0$,

$$\int_{-b}^{a} \delta(x)\,\mathrm{d}x = 1 \ . \tag{5.52}$$

If $b > 0$, then, let $y = bx$

$$\int_{-\infty}^{\infty} f(x)\,\delta(bx)\,\mathrm{d}x = \frac{1}{b}\int_{-\infty}^{\infty} f(y/b)\,\delta(y)\,\mathrm{d}y = \frac{f(0)}{b} \ . \tag{5.53}$$

If $b < 0$, then, let $y = -bx$

$$\int_{-\infty}^{\infty} f(x)\,\delta(bx)\,\mathrm{d}x = \frac{-1}{b}\int_{-\infty}^{\infty} f(-y/b)\,\delta(y)\,\mathrm{d}y = \frac{-f(0)}{b} \ . \tag{5.54}$$

Hence, we have for $b \neq 0$

$$\delta(bx) = \frac{1}{|b|}\delta(x) \ . \tag{5.55}$$

Consider $\delta((x-a)(x-b))$ for $a \neq b$. This is non-zero only when $x = a$ or $x = b$. Near a, we have

$$\begin{aligned}
\int_a f(x)\delta((x-a)(x-b))\,\mathrm{d}x &= \int_a f(x)\delta((x-a)(a-b))\,\mathrm{d}x \\
&= \frac{1}{|a-b|}\int_a f(x)\delta(x-a)\,\mathrm{d}x \\
&= \frac{f(a)}{|a-b|} \ .
\end{aligned} \tag{5.56}$$

Similarly, near b, we have $\int_b f(x)\delta((x-a)(x-b))\,\mathrm{d}x = f(b)/|a-b|$. Hence,

$$\int_{-\infty}^{\infty} f(x)\delta((x-a)(x-b))\,\mathrm{d}x = \frac{f(a)}{|a-b|} + \frac{f(b)}{|a-b|} \ . \tag{5.57}$$

If a function f is continuous at x, it is tempting to rewrite Eq. (5.42) as

$$\begin{aligned}
f(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)e^{-i\omega y}\,\mathrm{d}y\, e^{i\omega x}\,\mathrm{d}\omega \\
&= \int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega(y-x)}\,\mathrm{d}\omega\right) f(y)\,\mathrm{d}y \ .
\end{aligned} \tag{5.58}$$

Comparing with Eq. (5.48), we have

$$\delta(x) \sim \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega x}\,\mathrm{d}\omega \ . \tag{5.59}$$

However, note that the interchange of integrations in Eq. (5.58) is mathematically not allowed.

The Fourier transform of the δ-function is

$$\tilde{\delta}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta(y) e^{-i\omega y} \, \mathrm{d}y = \frac{1}{\sqrt{2\pi}} \; . \tag{5.60}$$

We could say that the δ-function is the linear combination of waves with all frequencies of equal amplitude.

References

- C. H. Edwards and D. E. Penney, Elementary Differential Equations, Prentice Hall, 2008.
- A. Jeffrey, Advanced Engineering Mathematics, Harcourt, 2002.
- G. B. Arfken and H. J. Weber, Mathematical Methods for Physicists, Elsevier, 2005.
- K. F. Riley, M. P. Hobson and S. J. Bence, Mathematical Methods for Physics and Engineering, Cambridge, 2006.
- D. Lay, Linear Algebra and Its Applications, Pearson, 2017.
- D. Zill D and M. Cullen, Differential Equations, With Boundary-Value Problems, Brooks/Cole, Cengage Learning, 2009.

Index