Chapter 1

Mathematical Preliminaries

This chapter serves to improve the completeness and systemization of theory introduced in this book. First of all, the basic concepts and basic elements of mathematics relevant to this book are briefly introduced. It includes the Euclidean space, symplectic space, and Legendre transformation, etc. In addition, it also briefly reviews the Hamilton principle and the Hamilton canonical equations in analytical mechanics, as well as the reciprocal theorems which are closely related to the contents of this book. Readers who are familiar with the mathematical preliminaries may omit this chapter and go directly to Chapter 2.

1.1. Linear Space

Linear space is one of the most basic concepts in linear algebra. It has been not only extensively applied in many fields in modern mathematics, but also exists as a common mathematical structure in various models in physics. As the basic concepts and basic elements have been widely described and proved in detail in many teaching materials[1–3], only contents relevant to this book will be particularly discussed without proofs in this section.

Definition 1.1. Let a linear space V in a real number field R has n linearly independent vectors (generalized vectors) $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ and every vector $\boldsymbol{\alpha}$ in V can be expressed as a linear combination of the vectors $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ as

$$\boldsymbol{\alpha} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n \tag{1.1.1}$$

Then $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ is called a **basis** of V, denoted as $\{\boldsymbol{\alpha}_i\}$ in brief, and $\{x_1, x_2, \ldots, x_n\}^{\mathrm{T}}$ are the **coordinates** of $\boldsymbol{\alpha}$ referring to basis $\{\boldsymbol{\alpha}_i\}$. Here, V is regarded as a n-dimensional **linear space**.

Incidentally, the above definition indicates that a problem in an abstract n-dimensional linear space can be completely described by general n-dimensional vectors in a real number field R via a basis. In reality, many properties and operations in linear space are eventually transformed to corresponding properties and operations of general vectors and matrices by way of a basis. It is conducive to a better understanding and application if discussion is done in a general n-dimensional vector space. It will become clear through further discussion introduced as follows.

A basis in a n-dimensional linear space is not unique. Coordinate systems of a vector referring to different bases are different.

Let $\{\boldsymbol{\alpha}_i\}$ and $\{\boldsymbol{\beta}_j\}$ are two bases in V and they are related by

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n\} = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}\boldsymbol{A} \tag{1.1.2}$$

where

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{1.1.3}$$

is the **transformation matrix** from basis $\{\boldsymbol{\alpha}_i\}$ to basis $\{\boldsymbol{\beta}_j\}$. It must be a nonsingular matrix. If $\{x_1, x_2, \ldots, x_n\}^{\mathrm{T}}$ and $\{y_1, y_2, \ldots, y_n\}^{\mathrm{T}}$ are the coordinates of a vector $\boldsymbol{\gamma}$ referring to bases $\{\boldsymbol{\alpha}_i\}$ and $\{\boldsymbol{\beta}_j\}$, respectively, then we have

$$\begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} = \boldsymbol{A} \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{Bmatrix} \quad \text{or} \quad \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{Bmatrix} = \boldsymbol{A}^{-1} \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \tag{1.1.4}$$

The transformation of a basis to another basis in a general vector space is actually the transformation of coordinates.

The self-mapping of a linear space V in a linear space is usually regarded as a transformation of V. In this regard, linear transformation is the most basic and simplest transformation.

Definition 1.2. A transformation $\tilde{\boldsymbol{A}}$ in a linear space V in a real number field R is called a **linear transformation** if, for any two vectors $\boldsymbol{\xi}, \boldsymbol{\eta}$ in V and any constant k in R, we have

$$\tilde{\boldsymbol{A}}(\boldsymbol{\xi}+\boldsymbol{\eta}) = \tilde{\boldsymbol{A}}(\boldsymbol{\xi}) + \tilde{\boldsymbol{A}}(\boldsymbol{\eta}) \tag{1.1.5}$$

$$\tilde{\boldsymbol{A}}(k\boldsymbol{\xi}) = k\tilde{\boldsymbol{A}}(\boldsymbol{\xi}) \tag{1.1.6}$$

Here and subsequently, $\tilde{\boldsymbol{A}}(\boldsymbol{\xi})$ is denoted briefly as $\tilde{\boldsymbol{A}}\boldsymbol{\xi}$.

Let $\{\boldsymbol{\alpha}_i\}$ be a basis in a linear space V, then the image $\{\tilde{\boldsymbol{A}}\boldsymbol{\alpha}_i\}$ of the vector upon linear transformation $\tilde{\boldsymbol{A}}$ can be linearly represented by basis $\{\boldsymbol{\alpha}_i\}$ as

$$\left.\begin{aligned} \tilde{\boldsymbol{A}}\boldsymbol{\alpha}_1 &= a_{11}\boldsymbol{\alpha}_1 + a_{21}\boldsymbol{\alpha}_2 + \cdots + a_{n1}\boldsymbol{\alpha}_n \\ \tilde{\boldsymbol{A}}\boldsymbol{\alpha}_2 &= a_{12}\boldsymbol{\alpha}_1 + a_{22}\boldsymbol{\alpha}_2 + \cdots + a_{n2}\boldsymbol{\alpha}_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \tilde{\boldsymbol{A}}\boldsymbol{\alpha}_n &= a_{1n}\boldsymbol{\alpha}_1 + a_{2n}\boldsymbol{\alpha}_2 + \cdots + a_{nn}\boldsymbol{\alpha}_n \end{aligned}\right\} \tag{1.1.7}$$

or in matrix form, it is

$$\{\tilde{\boldsymbol{A}}\boldsymbol{\alpha}_1, \tilde{\boldsymbol{A}}\boldsymbol{\alpha}_2, \ldots, \tilde{\boldsymbol{A}}\boldsymbol{\alpha}_n\} = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}\boldsymbol{A} \tag{1.1.8}$$

where

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{1.1.9}$$

Matrix $\boldsymbol{A}$ is the matrix resulted from linear transformation $\tilde{\boldsymbol{A}}$ referring to basis $\{\boldsymbol{\alpha}_i\}$. Obviously, $\{\tilde{\boldsymbol{A}}\boldsymbol{\alpha}_i\}$ is another basis of the linear space and matrix $\boldsymbol{A}$ is the transformation matrix if the latter is nonsingular.

Definition 1.3. Let $\tilde{\boldsymbol{A}}$ be a linear transformation in a linear space $\boldsymbol{V}$ in a real number field R. If there exists a nonzero vector $\boldsymbol{x}$ with respect to a number μ (inclusive a complex number) such that

$$\tilde{\boldsymbol{A}}\boldsymbol{x} = \mu\boldsymbol{x} \tag{1.1.10}$$

then μ is the **eigenvalue** of the linear transformation $\tilde{\boldsymbol{A}}$ and $\boldsymbol{x}$ the **eigenvector** of the linear transformation $\tilde{\boldsymbol{A}}$ corresponding to eigenvalue μ.

Let $\boldsymbol{A}$ be the matrix of linear transformation $\tilde{\boldsymbol{A}}$ referring to basis $\{\boldsymbol{\alpha}_i\}$ in linear space V, μ be an eigenvalue of $\tilde{\boldsymbol{A}}$, and $\{x_1, x_2, \ldots, x_n\}^{\mathrm{T}}$ be the coordinates of the corresponding eigenvector $\boldsymbol{x}$ referring to basis $\{\boldsymbol{\alpha}_i\}$.

Substituting into Eq. (1.1.10) yields

$$\tilde{\boldsymbol{A}}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\} \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} = \mu\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\} \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \tag{1.1.11}$$

Substituting Eq. (1.1.8) into the above equation and applying the linear independence characteristics of $\{\boldsymbol{\alpha}_i\}$ yield

$$\boldsymbol{A} \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} = \mu \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \tag{1.1.12}$$

or, equivalently, as

$$(\mu \boldsymbol{I}_n - \boldsymbol{A}) \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} = \mathbf{0} \tag{1.1.13}$$

where $\boldsymbol{I}_n$ is a n-order identity matrix, denoted briefly as $\boldsymbol{I}$. Equation (1.1.13) will have a nonzero (nontrivial) vector solution, if and only if, the determinant of the coefficient matrix is zero. Equivalently, μ is the root of

$$f(\mu) = |\mu \boldsymbol{I} - \boldsymbol{A}| = 0 \tag{1.1.14}$$

The eigenvalue problem of a linear transformation can be transformed to the eigenvalue problem of a matrix by the application of a basis. Equation (1.1.14) is the **characteristic polynomial** of matrix $\boldsymbol{A}$. Therefore, we will focus on the eigenvalue problem of a matrix in the following discussion on eigenvalue problems. The properties can be extended to any general linear transformation. An eigenvalue problem of a matrix has the following properties.

Theorem 1.1. Let $\mu_1, \mu_2, \ldots, \mu_t$ be t distinct eigenvalues of matrix $\mathbf{A}$, and $\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2}, \ldots, \boldsymbol{\alpha}_{im_i}$ $(i = 1, \ldots, t)$ be the linearly independent eigenvectors corresponding to eigenvalue μ_i, then all eigenvectors $\boldsymbol{\alpha}_{11}, \boldsymbol{\alpha}_{12}, \ldots, \boldsymbol{\alpha}_{1m_1}$; $\boldsymbol{\alpha}_{21}, \boldsymbol{\alpha}_{22}, \ldots, \boldsymbol{\alpha}_{2m_2}; \cdots; \boldsymbol{\alpha}_{t1}, \boldsymbol{\alpha}_{t2}, \ldots, \boldsymbol{\alpha}_{tm_t}$ of $\boldsymbol{A}$ are linearly independent.

Theorem 1.2. For every $n \times n$ matrix $\boldsymbol{A}$, there exists a nonsingular $n \times n$ matrix $\boldsymbol{X}$ (inclusive complex elements) such that matrix $\boldsymbol{A}$ can be transformed to the **Jordan canonical form**

$$\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} = \operatorname{diag}(\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_t) \tag{1.1.15}$$

where

$$\boldsymbol{D}_i = \begin{bmatrix} \mu_i & 1 & 0 & \cdots & 0 \\ 0 & \mu_i & 1 & \cdots & 0 \\ 0 & 0 & \mu_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mu_i \end{bmatrix} \tag{1.1.16}$$

is the Jordan part, and $m_1 + \cdots + m_t = n$. Also, we have

$$\boldsymbol{X} = \left\{\boldsymbol{\psi}_1^{(0)}, \ldots, \boldsymbol{\psi}_i^{(0)}, \boldsymbol{\psi}_i^{(1)}, \ldots, \boldsymbol{\psi}_i^{(m_i-1)}, \ldots, \boldsymbol{\psi}_t^{(m_t-1)}\right\} \tag{1.1.17}$$

or, in other words, there are m_i vectors corresponding to the Jordan part $\boldsymbol{D}_i$ in matrix $\boldsymbol{X}$, i.e. the basic eigenvector $\boldsymbol{\psi}_i^{(0)}$ and the Jordan form eigenvectors $\boldsymbol{\psi}_i^{(k)}$ $(k = 1, 2, \ldots, m_i - 1)$ corresponding to eigenvalue μ_i where k denotes the kth-order Jordan form eigenvector.

Throughout this book, the order of Jordan form vectors is indicated as a superscript as above.

Equation (1.1.15) can be written as[1,2]

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X} \cdot \operatorname{diag}(\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_t) \tag{1.1.18}$$

Expanding Eq. (1.1.18) yields

$$\left.\begin{array}{l} \boldsymbol{A}\boldsymbol{\psi}_i^{(0)} = \mu_i\boldsymbol{\psi}_i^{(0)} \\ \boldsymbol{A}\boldsymbol{\psi}_i^{(1)} = \mu_i\boldsymbol{\psi}_i^{(1)} + \boldsymbol{\psi}_i^{(0)} \\ \cdots\cdots\cdots\cdots \\ \boldsymbol{A}\boldsymbol{\psi}_i^{(m_i-1)} = \mu_i\boldsymbol{\psi}_i^{(m_i-1)} + \boldsymbol{\psi}_i^{(m_i-2)} \end{array}\right\} \quad (i = 1, 2, \ldots, t) \tag{1.1.19}$$

Equation (1.1.19) shows the general method for solving the basic eigenvector and the Jordan form eigenvectors.

The discussion above is not restricted to a Cartesian coordinate system.

1.2. Euclidean Space

Vector addition and scalar multiplication are the two basic linear operations of vectors in a linear space. However, such linear operations cannot be used to describe the metric properties of vectors, such as length, orthogonality, etc. The metric concept can be introduced into linear space via the operation of inner product. This concept will be particularly discussed without proofs in this section.

Definition 1.4. Let V be a linear space defined in a real number filed R. For any two arbitrary vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ in V, there exists a real number according to a specified rule, termed the **inner product** and denoted as $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. The inner product operation satisfies the following four conditions:

$$\text{(1)}\quad (\boldsymbol{\alpha}, \boldsymbol{\alpha}) \geq 0, (\boldsymbol{\alpha}, \boldsymbol{\alpha}) = \mathbf{0} \text{ if and only if } \boldsymbol{\alpha} = \mathbf{0} \tag{1.2.1a}$$

$$\text{(2)}\quad (\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\boldsymbol{\beta}, \boldsymbol{\alpha}) \tag{1.2.1b}$$

$$\text{(3)}\quad (\boldsymbol{\alpha} + \boldsymbol{\gamma}, \boldsymbol{\beta}) = (\boldsymbol{\alpha}, \boldsymbol{\beta}) + (\boldsymbol{\gamma}, \boldsymbol{\beta}), \boldsymbol{\gamma} \text{ is an arbitrary vector in } V \tag{1.2.1c}$$

$$\text{(4)}\quad (k\boldsymbol{\alpha}, \boldsymbol{\beta}) = k(\boldsymbol{\alpha}, \boldsymbol{\beta}), k \text{ is an arbitrary real number} \tag{1.2.1d}$$

A linear space satisfying the above conditions of inner product is called a **Euclidean space**.

Having defined inner product, it is possible to measure the length, orthogonality, unit vector, etc. of vectors related to metric concept in Euclidean space.

Let V be a Euclidean space. The **norm** of an arbitrary vector $\boldsymbol{\alpha}$ is defined as

$$\|\boldsymbol{\alpha}\| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})} \tag{1.2.2}$$

$\boldsymbol{\alpha}$ is a **unit vector** if the norm of $\boldsymbol{\alpha}$ is $\|\boldsymbol{\alpha}\| = 1$.

Example 1.1. For arbitrary vectors $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}^{\mathrm{T}}$, $\boldsymbol{y} = \{y_1, y_2, \ldots, y_n\}^{\mathrm{T}}$ in a n-dimensional real vector space R^n, the inner product is defined as

$$(\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{y} (= \boldsymbol{x}^{\mathrm{T}} \boldsymbol{I} \boldsymbol{y}) \tag{1.2.3}$$

It is obvious that the operation satisfy the four conditions of inner product in Eqs. (1.2.1) and therefore they form an n-dimensional Euclidean space.

The magnitude of vector $\boldsymbol{x}$ in R^n is

$$\|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{1.2.4}$$

For a particular linear space, there are various definitions of inner product and therefore there exist different Euclidean spaces. The inner product in Example 1.1 is the normal inner product of a n-dimensional real vector space R^n. It is also the most common definition of inner product. The following discussion on R^n always refers to this normal inner product.

Definition 1.5. Vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ are **orthogonal**, denoted as $\boldsymbol{\alpha} \perp \boldsymbol{\beta}$, if their inner product $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0$.

If any two vectors in a nonzero vector set $\{\boldsymbol{\alpha}_i\}$ are orthogonal, then the vector set $\{\boldsymbol{\alpha}_i\}$ is called an **orthogonal vector set**. If the vectors are all unit vectors, then the vector set $\{\boldsymbol{\alpha}_i\}$ is called a **normal orthogonal vector set**. A basis formed by n (normal) orthogonal vectors in a n-dimensional Euclidean space is called a **(normal) orthogonal basis**.

From Definitions 1.4 and 1.5, it is obvious that

Theorem 1.3. A zero vector is orthogonal to any vector. Conversely, a vector which is orthogonal to any vector in the space must be a zero vector.

Theorem 1.4. An orthogonal vector set is a linearly independent vector set.

Theorem 1.5. Any arbitrary (normal) orthogonal vector set in a n-dimensional Euclidean space can be extended to a set of (normal) orthogonal bases.

Let V be a n-dimensional Euclidean space, $\{\boldsymbol{\alpha}_i\}$ be a set of normal orthogonal bases, then the coordinates $\{x_1, x_2, \ldots, x_n\}^{\mathrm{T}}$ of an arbitrary vector $\boldsymbol{\beta}$ referring to basis $\{\boldsymbol{\alpha}_i\}$ can be expressed (expansion theorem) as:

$$x_i = (\boldsymbol{\beta}, \boldsymbol{\alpha}_i) \quad (i = 1,\ 2, \ldots, n) \tag{1.2.5}$$

Let $\{y_1, y_2, \ldots, y_n\}^{\mathrm{T}}$ be the coordinates of another vector $\boldsymbol{\gamma}$ referring to basis $\{\boldsymbol{\alpha}_i\}$, then the inner product of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ is

$$(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{i=1}^{n} x_i y_i = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{y} (= \boldsymbol{x}^{\mathrm{T}} \boldsymbol{I} \boldsymbol{y}) = (\boldsymbol{\gamma}, \boldsymbol{\beta}) \tag{1.2.6}$$

where

$$\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}^{\mathrm{T}}, \quad \boldsymbol{y} = \{y_1, y_2, \ldots, y_n\}^{\mathrm{T}} \tag{1.2.7}$$

Through the use of a normal orthogonal basis, the inner product operation in a n-dimensional Euclidean space can be transformed to the normal inner product operation in a n-dimensional real vector space R^n.

To discuss on the transformation equation of bases for a normal orthogonal basis, we introduce.

Definition 1.6. If a $n \times n$ matrix $\boldsymbol{Q}$ satisfies

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{I} \tag{1.2.8}$$

then $\boldsymbol{Q}$ is an **orthogonal matrix**.

An orthogonal matrix has the following properties:

(1) The inverse matrix (i.e. transpose matrix) of an orthogonal matrix is an orthogonal matrix.
(2) The determinant of an orthogonal matrix is equal to either 1 or −1.
(3) The product of two orthogonal matrices is an orthogonal matrix.

From Definition 1.6, it is obviously that

Theorem 1.6. The transformation matrix for normal orthogonal bases is an orthogonal matrix.

The following discussion focuses on the most fundamental linear operator (transformation) in Euclidean space, i.e. the symmetrization operator.

Definition 1.7. Let V be a n-dimensional Euclidean space. If the linear transformation $\tilde{\boldsymbol{A}}$ on arbitrary vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ in V satisfies

$$(\boldsymbol{\alpha}, \tilde{\boldsymbol{A}}\boldsymbol{\beta}) = (\boldsymbol{\beta}, \tilde{\boldsymbol{A}}\boldsymbol{\alpha}) \tag{1.2.9}$$

then $\tilde{\boldsymbol{A}}$ is a symmetric operator of the Euclidean space V.

Obviously, the matrix of symmetric operator $\tilde{\boldsymbol{A}}$ for any normal orthogonal basis $\{\boldsymbol{\alpha}_i\}$ is a $n \times n$ real symmetric matrix $\boldsymbol{A}$. A real symmetric matrix (**symmetric operator**) is self-adjoint. The eigenvalue problem of a self-adjoint operator is discussed in detail because of its needs in vibration theory and other problems in mathematical physics. There are some theorems as follows regarding the eigenvalue problem of a real symmetric matrix.

Theorem 1.7. The eigenvalues of a real symmetric matrix are all real numbers.

Theorem 1.8. Let $\boldsymbol{A}$ be a real symmetric matrix, then the eigenvectors corresponding to different eigenvalues of $\boldsymbol{A}$ in R^n are mutually orthogonal.

Theorem 1.9. For an arbitrary n order real symmetric matrix $\boldsymbol{A}$, there exists a n-order orthogonal matrix $\boldsymbol{Q}$ such that $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ is diagonal.

Theorem 1.9 shows that for an arbitrary symmetric matrix (operator), there exists a normal orthogonal basis composed of eigenvectors. The corresponding eigenvalues are all real and there will be no Jordan form even if there are repeated eigenvalues. The eigenvectors are all mutually orthogonal. As a result, the eigen-solutions span to a complete space and every vector in this space can be constructed from a linear combination of the eigenvectors (expansion theorem).

1.3. Symplectic Space

All conservative real physical processes can be described by a suitable Hamiltonian system whose common mathematical fundamentals are the symplectic spaces. A symplectic space is different from a Euclidean space which studies the metric properties such as length, etc. It focuses on the study of area, or the study of work and this is a mathematical structure present throughout this book. Using a finite-dimensional symplectic space as an example in this section, the basic concepts and basic properties of a symplectic space are described and proved in detail[4,5]. The discussion lays a sound mathematical foundation for the study in the following chapters.

Definition 1.8. Let V be a n-dimensional linear space defined in a real number field R, and V' be the corresponding n-dimensional dual linear space. We define

$$W = V \times V' = \left\{ \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} \middle| \boldsymbol{q} \in V, \boldsymbol{p} \in V' \right\} \tag{1.3.1}$$

where the linear space W is called a $2n$-dimensional **phase space** in a real number field R constructed by V and V'.

It is emphasized here that the linear spaces V and V' have absolutely different dimensions in actual problems, and there usually exists no direct relation between the spaces. However, the product of their corresponding components has a specific physical meaning. In this book, for instance, one is usually the displacement and the other the stress, and the product of their corresponding components has the dimension of work.

Definition 1.9. Let W be a $2n$-dimensional phase space in a real number field R. For any two vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ in W, there exists a real number according to a specified rule. This real number is termed the **symplectic inner product**, denoted as $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle$, and it satisfies the following four conditions:

$$\text{(1)}\quad \langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = -\langle \boldsymbol{\beta}, \boldsymbol{\alpha} \rangle \tag{1.3.2a}$$

$$\text{(2)}\quad \langle k\boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = k\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle, k \text{ is an arbitrary real number} \tag{1.3.2b}$$

$$\text{(3)}\quad \langle \boldsymbol{\alpha} + \boldsymbol{\gamma}, \boldsymbol{\beta} \rangle = \langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle + \langle \boldsymbol{\gamma}, \boldsymbol{\beta} \rangle, \boldsymbol{\gamma} \text{ is an arbitrary vector in } W \tag{1.3.2c}$$

$$\text{(4)}\quad \boldsymbol{\alpha} = \mathbf{0} \text{ if } \langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = 0 \text{ for every vector } \boldsymbol{\beta} \text{ in } W \tag{1.3.2d}$$

A phase space satisfying the above conditions of symplectic inner product is called a **symplectic space**.

From Eq. (1.3.2a), the symplectic self inner product of every vector must vanish, i.e. for every vector $\boldsymbol{\alpha}$

$$\langle \boldsymbol{\alpha}, \boldsymbol{\alpha} \rangle = 0 \tag{1.3.3}$$

Example 1.2. For any two vectors $\boldsymbol{x} = \{x_1, x_2\}^{\mathrm{T}}$, $\boldsymbol{y} = \{y_1, y_2\}^{\mathrm{T}}$ in a two-dimensional real vector space R^2, the symplectic inner product is defined as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_2 - x_2 y_1 \tag{1.3.4}$$

It is obvious that Eq. (1.3.4) satisfies the four conditions of symplectic inner product in Eqs. (1.3.2) and therefore it forms a 2-dimensional symplectic space. Here, the symplectic inner product (1.3.4) represents the area of a parallelogram constructed by $\boldsymbol{x}, \boldsymbol{y}$ as its adjacent sides.

Obviously, we can generalize Eq. (1.3.4) to a $2n$-dimensional real vector space R^{2n}. For any two vectors $\boldsymbol{x} = \{x_1, x_2, \ldots, x_{2n}\}^{\mathrm{T}}$, $\boldsymbol{y} = \{y_1, y_2, \ldots, y_{2n}\}^{\mathrm{T}}$, the symplectic inner product is defined as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle \overset{\text{def}}{=} (\boldsymbol{x}, \boldsymbol{J}_{2n}\boldsymbol{y}) = \sum_{i=1}^{n} (x_i y_{n+i} - x_{n+i} y_i) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{J}_{2n} \boldsymbol{y} \tag{1.3.5}$$

where

$$\boldsymbol{J}_{2n} = \begin{bmatrix} \mathbf{0} & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \mathbf{0} \end{bmatrix} \tag{1.3.6}$$

is called the **unit symplectic matrix**, denoted briefly as $\boldsymbol{J}$. It is obvious that Eq. (1.3.5) satisfies the four conditions of symplectic inner product in Eqs. (1.3.2), and therefore it forms a $2n$-dimensional symplectic space.

The determinant of a unit symplectic matrix is equal to 1. A unit symplectic matrix has the following properties:

$$\boldsymbol{J}^2 = -\boldsymbol{I}, \quad \boldsymbol{J}^{\mathrm{T}} = \boldsymbol{J}^{-1} = -\boldsymbol{J} \tag{1.3.7}$$

Similarly, there are various definitions of symplectic inner product for a phase space and therefore there exist different symplectic spaces. The symplectic inner product defined in Eq. (1.3.5) is called the normal symplectic inner product in a $2n$-dimensional real vector space R^{2n}. The following discussion on real vector space R^{2n} always refers to this normal symplectic inner product.

Definition 1.10. Vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ are **symplectic orthogonal** if their symplectic inner product $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = 0$. Otherwise, they are **symplectic adjoint**.

Hence from Eq. (1.3.2d), there exists a symplectic adjoint nonzero vector for every nonzero vector. Virtually, $\boldsymbol{\alpha}$ and $\boldsymbol{J\alpha}$ must be symplectic adjoint if $\boldsymbol{\alpha} \neq \mathbf{0}$.

If the vectors in a vector set $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_r\}$ $(r \leq n)$ satisfy

$$\left.\begin{aligned} &\langle \boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j \rangle = \langle \boldsymbol{\beta}_i, \boldsymbol{\beta}_j \rangle = 0 \\ &\langle \boldsymbol{\alpha}_i, \boldsymbol{\beta}_j \rangle = \begin{cases} k_{ii} \neq 0 & (i = j) \\ 0 & (i \neq j) \end{cases} \end{aligned}\right\} \quad (i, j = 1,\ 2, \dots, r) \tag{1.3.8}$$

then the vector set $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_r\}$ is called an **adjoint symplectic orthonormal vector set**. If $k_{ii} \equiv 1 (i = 1, 2, \dots, r)$ in Eq. (1.3.8), then the vector set $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_r\}$ is called **normal adjoint symplectic orthonormal vector set.** From the definition, it is obvious that:

Theorem 1.10. An adjoint symplectic orthonormal vector set is a linearly independent vector set.

Proof. Using proof by contradiction, assume $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r\}$ is an adjoint symplectic orthonormal vector set and it is linearly dependent. Then there exists a vector, denoted as $\boldsymbol{\alpha}_1$, which can be expressed as a linear combination of the other vectors as

$$\boldsymbol{\alpha}_1 = s_2\boldsymbol{\alpha}_2 + \cdots + s_r\boldsymbol{\alpha}_r + t_1\boldsymbol{\beta}_1 + \cdots + t_r\boldsymbol{\beta}_r$$

Hence, we have

$$\begin{aligned}\langle \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1 \rangle = s_2\langle \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1 \rangle &+ \cdots + s_r\langle \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1 \rangle \\ &+ t_1\langle \boldsymbol{\beta}_1, \boldsymbol{\beta}_1 \rangle + \cdots + t_r\langle \boldsymbol{\beta}_r, \boldsymbol{\beta}_1 \rangle = 0\end{aligned}$$

which contradicts $\langle \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1 \rangle \neq 0$. Therefore, the original proposition is tenable.

The basis formed by $2n$ (normal) adjoint symplectic orthonormal vectors in a $2n$-dimensional symplectic space is called a **(normal) adjoint symplectic orthonormal basis**.

Theorem 1.11. Every adjoint symplectic orthonormal vector set in a $2n$-dimensional symplectic space can be extended to an adjoint symplectic orthonormal basis.

Proof. Let $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r\}$ be an adjoint symplectic orthonormal vector set. A proof for $n - r$ by mathematical induction is established.

(1) For $n - r = 0$, $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r\}$ is an adjoint symplectic orthonormal basis. The theorem is tenable for $n - r = 0$.
(2) Assume that the theorem is tenable for $n - r = k$, then consider the case of $n - r = k + 1$.

As $r < n$, there exists a vector $\boldsymbol{\gamma}$ which cannot be expressed as a linear combination of $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r\}$. Denoting

$$\boldsymbol{\alpha}_{r+1} = \boldsymbol{\gamma} - \sum_{i=1}^{r} s_i\boldsymbol{\alpha}_i - \sum_{i=1}^{r} t_i\boldsymbol{\beta}_i$$

where

$$s_i = \frac{\langle \boldsymbol{\gamma}, \boldsymbol{\beta}_i \rangle}{\langle \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i \rangle} \quad t_i = -\frac{\langle \boldsymbol{\gamma}, \boldsymbol{\alpha}_i \rangle}{\langle \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i \rangle}$$

It is obvious that $\boldsymbol{\alpha}_{r+1}$ is symplectic orthogonal to $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r\}$. Besides, $\boldsymbol{\alpha}_{r+1}$ is a nonzero vector from the definition and, hence, there exists a symplectic adjoint nonzero vector $\tilde{\boldsymbol{\gamma}}$, or $\langle \boldsymbol{\alpha}_{r+1}, \tilde{\boldsymbol{\gamma}} \rangle \neq 0$. Obviously, $\tilde{\boldsymbol{\gamma}}$ cannot be expressed as a linear combination of $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r\}$. Denoting

$$\boldsymbol{\beta}_{r+1} = \tilde{\boldsymbol{\gamma}} - \sum_{i=1}^{r} \tilde{s}_i \boldsymbol{\alpha}_i - \sum_{i=1}^{r} \tilde{t}_i \boldsymbol{\beta}_i$$

where

$$\tilde{s}_i = \frac{\langle \tilde{\boldsymbol{\gamma}}, \boldsymbol{\beta}_i \rangle}{\langle \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i \rangle} \quad \tilde{t}_i = -\frac{\langle \tilde{\boldsymbol{\gamma}}, \boldsymbol{\alpha}_i \rangle}{\langle \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i \rangle}$$

It is obvious that $\boldsymbol{\beta}_{r+1}$ is symplectic orthogonal to $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r\}$, and also symplectic adjoint with $\boldsymbol{\alpha}_{r+1}$, or $\langle \boldsymbol{\alpha}_{r+1}, \boldsymbol{\beta}_{r+1} \rangle = \langle \boldsymbol{\alpha}_{r+1}, \tilde{\boldsymbol{\gamma}} \rangle \neq 0$. Then $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r, \boldsymbol{\beta}_{r+1}\}$ is an adjoint symplectic orthonormal vector set. According to the assumption of mathematical induction, the theorem is tenable for $n-(r+1) = k$. In other words, $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_r, \boldsymbol{\beta}_{r+1}\}$ can be extended to an adjoint symplectic orthonormal basis and the theorem is tenable for $n - r = k + 1$. According to mathematical induction, therefore, the theorem is tenable for any $n - r$. Hence, the proposition is proven.

Deduction. Every normal adjoint symplectic orthonormal vector set in a $2n$-dimensional symplectic space can be extended to a normal adjoint symplectic orthonormal basis.

The theorem and deduction above indicate that there exists a normal adjoint symplectic orthonormal basis in a $2n$-dimensional symplectic space, but it is not unique. Based on a normal adjoint symplectic orthonormal basis and its properties, the expansion theorem of a symplectic space can be obtained directly.

Theorem 1.12. Let W be a $2n$-dimensional symplectic space and $\{\boldsymbol{\alpha}_i\}$ be a normal adjoint symplectic orthonormal basis, then the coordinates $\{x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}\}^{\mathrm{T}}$ of any vector $\boldsymbol{\beta}$ in W referring to the basis $\{\boldsymbol{\alpha}_i\}$ can be expressed as:

$$x_i = \langle \boldsymbol{\beta}, \boldsymbol{\alpha}_{n+i} \rangle, x_{n+i} = -\langle \boldsymbol{\beta}, \boldsymbol{\alpha}_i \rangle \quad (i = 1, 2, \ldots, n) \tag{1.3.9}$$

Let $\{y_1, \ldots, y_n, y_{n+1}, \ldots, y_{2n}\}^{\mathrm{T}}$ be the coordinates of another vector $\boldsymbol{\gamma}$ referring to the basis $\{\boldsymbol{\alpha}_i\}$, then the symplectic inner product of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ is

$$\langle \boldsymbol{\beta}, \boldsymbol{\gamma} \rangle = \sum_{i=1}^{n} (x_i y_{n+i} - x_{n+i} y_i) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{J}_{2n} \boldsymbol{y} \tag{1.3.10}$$

where

$$\boldsymbol{x} = \{x_1, x_2, \ldots, x_{2n}\}^{\mathrm{T}}, \quad \boldsymbol{y} = \{y_1, y_2, \ldots, y_{2n}\}^{\mathrm{T}} \tag{1.3.11}$$

Through the use of a normal adjoint symplectic orthonormal basis, the symplectic inner product operation in a $2n$-dimensional symplectic space can be transformed to the matrix operation of ordinary vectors (matrices).

To discuss the transformation equation of bases for a normal adjoint symplectic orthonormal basis, we introduce

Definition 1.11. If a $2n \times 2n$ matrix $\boldsymbol{S}$ satisfies

$$\boldsymbol{S}^{\mathrm{T}}\boldsymbol{J}\boldsymbol{S} = \boldsymbol{J} \tag{1.3.12}$$

then $\boldsymbol{S}$ is a **symplectic matrix**, where $\boldsymbol{J}$ is the unit symplectic matrix.

A symplectic matrix has the following properties:

(1) The inverse matrix of a symplectic matrix is a symplectic matrix.
(2) The transpose matrix of a symplectic matrix is a symplectic matrix.
(3) The determinant of a symplectic matrix is equal to either 1 or −1.
(4) The product of two symplectic matrices is a symplectic matrix.

For a normal adjoint symplectic orthonormal basis, it is obviously that:

Theorem 1.13. The transformation matrix for normal adjoint symplectic orthonormal bases is a symplectic matrix.

The following discussion focuses on the most fundamental linear operator in a symplectic space, i.e. the Hamiltonian operator.

Definition 1.12. Let W be a $2n$-dimensional symplectic space. If a linear operator $\tilde{\boldsymbol{H}}$ acting on arbitrary vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ satisfies

$$\langle \boldsymbol{\alpha}, \tilde{\boldsymbol{H}}\boldsymbol{\beta} \rangle = \langle \boldsymbol{\beta}, \tilde{\boldsymbol{H}}\boldsymbol{\alpha} \rangle \tag{1.3.13}$$

then linear transformation $\tilde{\boldsymbol{H}}$ is a **Hamiltonian operator** of the symplectic space W.

Definition 1.13. If a $2n \times 2n$ matrix $\boldsymbol{H}$ acting on arbitrary $2n$-dimensional vectors $\boldsymbol{x}, \boldsymbol{y}$ satisfies

$$\langle \boldsymbol{x}, \boldsymbol{H}\boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{H}\boldsymbol{x} \rangle \tag{1.3.14}$$

then matrix $\boldsymbol{H}$ is a **Hamiltonian matrix**.

It is obvious that the definition of Hamiltonian matrix in Eq. (1.3.14) is equivalent to the following definition

$$(\boldsymbol{JH})^{\mathrm{T}} = \boldsymbol{JH}, \quad \text{or} \quad \boldsymbol{JHJ} = \boldsymbol{H}^{\mathrm{T}} \tag{1.3.14$'$}$$

Obviously, the matrix of Hamiltonian operator $\tilde{\boldsymbol{H}}$ referring to a normal adjoint symplectic orthonormal basis $\{\boldsymbol{\alpha}_i\}$ is a Hamiltonian matrix. The eigenvalue problem of a Hamiltonian matrix (Hamiltonian operator) is non-self-adjoint, hence it is possible to have complex eigenvalues or repeated eigenvalues. However, the said eigenvalue problem of a Hamiltonian matrix (Hamiltonian operator) has its specific characteristics. Hamiltonian matrices are used as examples in the following discussion. Certainly, the relevant conclusion can be directly generalized to the Hamiltonian operator of a finite-dimensional symplectic space.

Theorem 1.14. If μ is an eigenvalue of a Hamiltonian matrix with multiplicity m, then $-\mu$ is also an eigenvalue with multiplicity m. If zero is an eigenvalue of a Hamiltonian matrix $\boldsymbol{H}$, then the multiplicity number is even.

Proof. Let the characteristic polynomial of a Hamiltonian matrix $\boldsymbol{H}$ be

$$f(\mu) = |\mu\boldsymbol{I} - \boldsymbol{H}|$$

then according to the definitions of unit symplectic matrix in Eq. (1.3.6) and Hamiltonian matrix in Eq. (1.3.14′), we have

$$\begin{aligned} f(\mu) &= |\boldsymbol{J}(\mu\boldsymbol{I} - \boldsymbol{H})\boldsymbol{J}| = |\mu\boldsymbol{JJ} - \boldsymbol{JHJ}| \\ &= |-\mu\boldsymbol{I} - \boldsymbol{H}^{\mathrm{T}}| = |-\mu\boldsymbol{I} - \boldsymbol{H}| = f(-\mu) \end{aligned}$$

As the expression above is true for every μ, the theorem is therefore proven.

Subsequently, the two eigenvalues $\pm\mu$ are called **mutually symplectic adjoint eigenvalues** of a Hamiltonian matrix. The nonzero eigenvalues of a Hamiltonian matrix are usually divided into two sets:

$$\left.\begin{array}{ll} (\alpha) & \mu_i, \quad \mathrm{Re}(\mu_i) < 0 \quad \text{or} \quad \mathrm{Re}(\mu_i) = 0 \wedge \mathrm{Im}(\mu_i) < 0 \\ (\beta) & \mu_{n+i} = -\mu_i \end{array}\right\} \tag{1.3.15}$$

The eigenvalues in the (α)-set can be further arranged according to the absolute values of μ_i, in an ascending order, for instance. Note that Eq. (1.3.15) does not include the zero eigenvalue which is a special symplectic eigenvalue with itself the mutually symplectic adjoint eigenvalue.

Theorem 1.15. Let $\boldsymbol{H}$ be a Hamiltonian matrix, $\boldsymbol{\psi}_i^{(0)}, \boldsymbol{\psi}_i^{(1)}, \ldots, \boldsymbol{\psi}_i^{(m)}$ and $\boldsymbol{\psi}_j^{(0)}, \boldsymbol{\psi}_j^{(1)}, \ldots, \boldsymbol{\psi}_j^{(n)}$ be the basic eigenvectors and Jordan form eigenvectors corresponding to the eigenvalues μ_i, μ_j, respectively. For $\mu_i + \mu_j \neq 0$, there exists symplectic orthogonality between the eigenvectors as follows:

$$\langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{\psi}_j^{(t)} \rangle = \boldsymbol{\psi}_i^{(s)\mathrm{T}} \boldsymbol{J} \boldsymbol{\psi}_j^{(t)} = 0 \quad (s = 0, 1, \ldots, m;\ t = 0,\, 1, \ldots, n) \tag{1.3.16}$$

Proof. A proof for $r = s + t$ by mathematical induction is established.

(1) For $r = 0$, i.e. $s = 0$ and $t = 0$, we have

$$\boldsymbol{H}\boldsymbol{\psi}_i^{(0)} = \mu_i \boldsymbol{\psi}_i^{(0)}, \quad \boldsymbol{H}\boldsymbol{\psi}_j^{(0)} = \mu_j \boldsymbol{\psi}_j^{(0)}$$

because $\boldsymbol{\psi}_i^{(0)}, \boldsymbol{\psi}_j^{(0)}$ are the basic eigenvectors corresponding to the eigenvalues μ_i, μ_j, respectively. Hence,

$$\langle \boldsymbol{\psi}_i^{(0)}, \boldsymbol{H}\boldsymbol{\psi}_j^{(0)} \rangle = \langle \boldsymbol{\psi}_i^{(0)}, \mu_j \boldsymbol{\psi}_j^{(0)} \rangle = \mu_j \langle \boldsymbol{\psi}_i^{(0)}, \boldsymbol{\psi}_j^{(0)} \rangle$$

Similarly,

$$\langle \boldsymbol{\psi}_j^{(0)}, \boldsymbol{H}\boldsymbol{\psi}_i^{(0)} \rangle = \mu_i \langle \boldsymbol{\psi}_j^{(0)}, \boldsymbol{\psi}_i^{(0)} \rangle = -\mu_i \langle \boldsymbol{\psi}_i^{(0)}, \boldsymbol{\psi}_j^{(0)} \rangle$$

As $\boldsymbol{H}$ is a Hamiltonian matrix, the left-hand-sides of the two expressions above are equal. After substituting and rearranging the expressions, we have

$$(\mu_i + \mu_j) \langle \boldsymbol{\psi}_i^{(0)}, \boldsymbol{\psi}_j^{(0)} \rangle = 0$$

Since $\mu_i + \mu_j \neq 0$, Eq. (1.3.16) is tenable for $r = 0$.

(2) Assume that Eq. (1.3.16) is tenable for $r = k$, consider the case of $r = s + t = k + 1$.

Firstly, from Eq. (1.1.19), the eigenvectors $\boldsymbol{\psi}_i^{(s)}$ and $\boldsymbol{\psi}_j^{(t)}$ satisfy these equations

$$\boldsymbol{H}\boldsymbol{\psi}_i^{(s)} = \mu_i \boldsymbol{\psi}_i^{(s)} + \boldsymbol{\psi}_i^{(s-1)} \quad \text{and} \quad \boldsymbol{H}\boldsymbol{\psi}_j^{(t)} = \mu_j \boldsymbol{\psi}_j^{(t)} + \boldsymbol{\psi}_j^{(t-1)}$$

respectively. Hence

$$\langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{H}\boldsymbol{\psi}_j^{(t)} \rangle = \mu_j \langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{\psi}_j^{(t)} \rangle + \langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{\psi}_j^{(t-1)} \rangle$$

In accordance with the assumption of mathematical induction, Eq. (1.3.16) is tenable for $r = k$. As a result, we have

$$\langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{\psi}_j^{(t-1)} \rangle = 0$$

hence,

$$\langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{H}\boldsymbol{\psi}_j^{(t)} \rangle = \mu_j \langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{\psi}_j^{(t)} \rangle$$

Similarly,

$$\langle \boldsymbol{\psi}_j^{(t)}, \boldsymbol{H}\boldsymbol{\psi}_i^{(s)} \rangle = \mu_i \langle \boldsymbol{\psi}_j^{(t)}, \boldsymbol{\psi}_i^{(s)} \rangle = -\mu_i \langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{\psi}_j^{(t)} \rangle$$

As $\boldsymbol{H}$ is a Hamiltonian matrix, the left-hand-sides of the two expressions above are equal. After substituting and rearranging the expression, we have

$$(\mu_i + \mu_j)\langle \boldsymbol{\psi}_i^{(s)}, \boldsymbol{\psi}_j^{(t)} \rangle = 0$$

Since $\mu_i + \mu_j \neq 0$, Eq. (1.3.16) is tenable for $r = k + 1$ too. In accordance with mathematical induction, Eq. (1.3.16) is tenable for any s and t, and the proposition is proven.

The theorem above indicates that the basic eigenvectors and Jordan form eigenvectors corresponding to the non-symplectic adjoint eigenvalues are symplectic orthogonal. Subsequently, we discuss the relations among the eigenvectors corresponding to the mutually symplectic adjoint eigenvalues. For brevity of proofs in the remaining parts in this section, only one Jordan chain is assumed to exist for each eigenvalue.

Theorem 1.16. Let $\pm\mu \neq 0$ be a pair of mutually symplectic adjoint eigenvalues of a Hamiltonian matrix $\boldsymbol{H}$ with multiplicity m, then there exists an adjoint symplectic orthonormal vector set

$$\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(m-1)}, \boldsymbol{\phi}^{(m-1)}, \boldsymbol{\phi}^{(m-2)}, \ldots, \boldsymbol{\phi}^{(0)}\}$$

such that

$$\langle \boldsymbol{\psi}^{(i)}, \boldsymbol{\phi}^{(j)} \rangle \begin{cases} = (-1)^i a \neq 0; & \text{when} \quad i + j = m - 1 \\ = 0; & \text{when} \quad i + j \neq m - 1 \end{cases} \tag{1.3.17}$$

where $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(m-1)}\}$ and $\{\boldsymbol{\phi}^{(0)}, \boldsymbol{\phi}^{(1)}, \ldots, \boldsymbol{\phi}^{(m-1)}\}$ are, respectively, the basic eigenvectors and Jordan form eigenvectors corresponding to μ and $-\mu$.

Proof. A proof for i by mathematical induction is established.

(1) For $i = 0$, let $\boldsymbol{\psi}^{(0)}$ be the basic eigenvector corresponding to the eigenvalue μ, and $\{\boldsymbol{\phi}^{(0)}, \ldots, \boldsymbol{\phi}^{(m-1)}\}$ be a set of arbitrary eigenvectors corresponding to the eigenvalue $-\mu$.

Firstly, for any $j \leq m - 2$, we have

$$\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{H}\boldsymbol{\phi}^{(j+1)} \rangle = -\mu \langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\phi}^{(j+1)} \rangle + \langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\phi}^{(j)} \rangle$$

and

$$\langle \boldsymbol{\phi}^{(j+1)}, \boldsymbol{H}\boldsymbol{\psi}^{(0)} \rangle = \mu \langle \boldsymbol{\phi}^{(j+1)}, \boldsymbol{\psi}^{(0)} \rangle = -\mu \langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\phi}^{(j+1)} \rangle$$

As $\boldsymbol{H}$ is a Hamiltonian matrix, the left-hand-sides of the two expressions above are equal. Hence, we have

$$\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\phi}^{(j)} \rangle = 0 \quad (j \leq m-2)$$

Next, according to Theorem 1.15, $\boldsymbol{\psi}^{(0)}$ is symplectic orthogonal to all other basic eigenvectors and Jordan form eigenvectors except that corresponding to the eigenvalue $-\mu$. Hence, $\boldsymbol{\psi}^{(0)}$ must be symplectic adjoint with $\boldsymbol{\phi}^{(m-1)}$. Otherwise $\boldsymbol{\psi}^{(0)}$ will be symplectic orthogonal to all eigenvectors, then $\boldsymbol{\psi}^{(0)} \equiv \mathbf{0}$. The conclusion is contradictory, i.e.

$$\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\phi}^{(m-1)} \rangle = a \neq 0$$

Hence, Eq. (1.3.17) is true for $i = 0$.

(2) Assuming there exists a set of basic eigenvectors and Jordan form eigenvectors $\{\boldsymbol{\psi}^{(0)}, \ldots, \boldsymbol{\psi}^{(m-1)}\}$ and $\{\boldsymbol{\phi}^{(0)}, \ldots, \boldsymbol{\phi}^{(m-1)}\}$ corresponding to μ, $-\mu$ such that Eq. (1.3.17) is true for $i \leq k$. Now consider the case of $i \leq k+1$.

Firstly, denote that

$$t = -\frac{1}{a} \langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\phi}^{(m-1)} \rangle$$

and

$$\tilde{\boldsymbol{\psi}}^{(k+1+p)} = \boldsymbol{\psi}^{(k+1+p)} + t\boldsymbol{\psi}^{(p)} \quad (p = 0, 1, \ldots, (m-k-2))$$

Obviously, $\{\boldsymbol{\psi}^{(0)}, \ldots, \boldsymbol{\psi}^{(k)}, \tilde{\boldsymbol{\psi}}^{(k+1)}, \ldots, \tilde{\boldsymbol{\psi}}^{(m-1)}\}$ remains as a set of basic eigenvectors and Jordan form eigenvectors corresponding to the eigenvalue μ, and we have

$$\left\langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \boldsymbol{\phi}^{(m-1)} \right\rangle = \langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\phi}^{(m-1)} \rangle + t \langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\phi}^{(m-1)} \rangle = 0$$

Next, for every $j < m-1$, we have

$$\langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \boldsymbol{H}\boldsymbol{\phi}^{(j+1)} \rangle = -\mu \langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \boldsymbol{\phi}^{(j+1)} \rangle + \langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \boldsymbol{\phi}^{(j)} \rangle$$

and

$$\begin{aligned} \left\langle \boldsymbol{\phi}^{(j+1)}, \boldsymbol{H}\tilde{\boldsymbol{\psi}}^{(k+1)} \right\rangle &= \mu \langle \boldsymbol{\phi}^{(j+1)}, \tilde{\boldsymbol{\psi}}^{(k+1)} \rangle + \langle \boldsymbol{\phi}^{(j+1)}, \boldsymbol{\psi}^{(k)} \rangle \\ &= -\mu \langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \boldsymbol{\phi}^{(j+1)} \rangle - \langle \boldsymbol{\psi}^{(k)}, \boldsymbol{\phi}^{(j+1)} \rangle \end{aligned}$$

As $\boldsymbol{H}$ is a Hamiltonian matrix, the left-hand-sides of the two expressions above are equal. Hence, we have

$$\langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \boldsymbol{\phi}^{(j)} \rangle = -\langle \tilde{\boldsymbol{\psi}}^{(k)}, \boldsymbol{\phi}^{(j+1)} \rangle$$

In accordance with the assumption of mathematical induction, Eq. (1.3.17) is tenable for $r = k$. Hence, we have

$$\langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \boldsymbol{\phi}^{(j)} \rangle \begin{cases} = (-1)^{k+1} a \neq 0; & \text{for } k+1+j = m-1 \\ = 0; & \text{for } k+1+j \neq m-1 \end{cases}$$

i.e. there exists a set of basic eigenvectors and Jordan form eigenvectors $\{\boldsymbol{\psi}^{(0)}, \ldots, \boldsymbol{\psi}^{(k)}, \tilde{\boldsymbol{\psi}}^{(k+1)}, \ldots, \tilde{\boldsymbol{\psi}}^{(m-1)}\}$ and $\{\boldsymbol{\phi}^{(0)}, \boldsymbol{\phi}^{(1)}, \ldots, \boldsymbol{\phi}^{(m-1)}\}$ corresponding to μ, $-\mu$ such that Eq. (1.3.17) is true for $i \leq k+1$. Hence, the theorem is proven.

Theorem 1.16 merely indicates that there exists an adjoint symplectic orthonormal vector set $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(m-1)}, \boldsymbol{\phi}^{(m-1)}, \ldots, \boldsymbol{\phi}^{(1)}, \boldsymbol{\phi}^{(0)}\}$ formed by the eigenvectors corresponding to eigenvalues $\pm\mu \neq 0$. If we take

$$\tilde{\boldsymbol{\phi}}^{(j)} = \frac{(-1)^{m-1-j}}{a} \boldsymbol{\phi}^{(j)} \quad (j = 0, 1, \ldots, m-1) \tag{1.3.18}$$

then $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(m-1)}, \tilde{\boldsymbol{\phi}}^{(m-1)}, \tilde{\boldsymbol{\phi}}^{(m-2)}, \ldots, \tilde{\boldsymbol{\phi}}^{(0)}\}$ forms a normal adjoint symplectic orthonormal vector set. It should be noted that the Jordan part corresponding to eigenvalue μ remains as the form of Eq. (1.1.16), while the Jordan part corresponding to eigenvalue $-\mu$ should be rewritten as

$$\boldsymbol{D}_{-\mu} = \begin{bmatrix} -\mu & 0 & 0 & \cdots & 0 \\ -1 & -\mu & 0 & \cdots & 0 \\ 0 & -1 & -\mu & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -\mu \end{bmatrix} \tag{1.3.19}$$

Then, we have

$$\left.\begin{aligned} \boldsymbol{H}\tilde{\boldsymbol{\phi}}^{(m-1)} &= -\mu\tilde{\boldsymbol{\phi}}^{(m-1)} - \tilde{\boldsymbol{\phi}}^{(m-2)} \\ \boldsymbol{H}\tilde{\boldsymbol{\phi}}^{(m-2)} &= -\mu\tilde{\boldsymbol{\phi}}^{(m-2)} - \tilde{\boldsymbol{\phi}}^{(m-3)} \\ &\cdots\cdots\cdots \\ \boldsymbol{H}\tilde{\boldsymbol{\phi}}^{(1)} &= -\mu\tilde{\boldsymbol{\phi}}^{(1)} - \tilde{\boldsymbol{\phi}}^{(0)} \\ \boldsymbol{H}\tilde{\boldsymbol{\phi}}^{(0)} &= -\mu\tilde{\boldsymbol{\phi}}^{(0)} \end{aligned}\right\} \tag{1.3.19$'$}$$

In such a way, a Hamiltonian matrix can be ensured when the Hamiltonian matrix $\boldsymbol{H}$ is transformed into a block diagonal matrix. Hence, we establish symplectic orthonormalization which is the basis of the expansion theorem.

Subsequently, we assume that the Jordan form eigenvectors corresponding to the (α)-set eigenvalues in Eq. (1.3.15) is again determined by Eq. (1.1.16), while the Jordan form eigenvectors corresponding to the (β)-set eigenvalues is determined by Eq. (1.3.19).

The discussion above concerns the adjoint symplectic orthonormality between the eigenvectors of the adjoint nonzero eigenvalues. From Theorem 1.14, there must be an even multiple root if a Hamiltonian matrix $\boldsymbol{H}$ has a zero eigenvalue. There usually exists the Jordan form for zero eigenvalue, and their solutions for physical problems have special interpretation in physics. The various physical problems will be introduced in due course in the following chapters.

Due to the special property of zero eigenvalue $\mu = -\mu = 0$, the corresponding basic eigenvectors and Jordan form eigenvectors form an adjoint symplectic orthonormal vector set. In order to discuss their nature of symplectic orthogonality, the following lemma is introduced.

Lemma. Let a Hamiltonian matrix $\boldsymbol{H}$ has a zero eigenvalue, and $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(2m-1)}\}$ be a set of arbitrary basic eigenvectors and Jordan form eigenvectors corresponding to the zero eigenvalue. Then for any $1 \leq p \leq 2m-1, 0 \leq q \leq 2m-2$, we have

$$\langle \boldsymbol{\psi}^{(p)}, \boldsymbol{\psi}^{(q)} \rangle = -\langle \boldsymbol{\psi}^{(p-1)}, \boldsymbol{\psi}^{(q+1)} \rangle \tag{1.3.20}$$

and for even $p+q$, we have

$$\langle \boldsymbol{\psi}^{(p)}, \boldsymbol{\psi}^{(q)} \rangle = 0 \tag{1.3.21}$$

Proof. Firstly, the eigenvectors $\boldsymbol{\psi}^{(p)}, \boldsymbol{\psi}^{(q+1)}$ corresponding to the zero eigenvalue satisfies

$$\langle \boldsymbol{\psi}^{(p)}, \boldsymbol{H}\boldsymbol{\psi}^{(q+1)} \rangle = \langle \boldsymbol{\psi}^{(p)}, \boldsymbol{\psi}^{(q)} \rangle$$

and

$$\langle \boldsymbol{\psi}^{(q+1)}, \boldsymbol{H}\boldsymbol{\psi}^{(p)} \rangle = \langle \boldsymbol{\psi}^{(q+1)}, \boldsymbol{\psi}^{(p-1)} \rangle = -\langle \boldsymbol{\psi}^{(p-1)}, \boldsymbol{\psi}^{(q+1)} \rangle$$

in accordance with Eq. (1.1.19). As $\boldsymbol{H}$ is a Hamiltonian matrix, the left-hand-sides of the two expressions above are equal. Hence, we have Eq. (1.3.20).

Next, we can assume without the loss of generality that $p = q + 2k$ for even $p + q$ where k is a nonnegative integer. Repeatedly applying Eq. (1.3.20) and using Eq. (1.3.3) yield

$$\begin{aligned}\langle \boldsymbol{\psi}^{(q+2k)}, \boldsymbol{\psi}^{(q)} \rangle &= -\langle \boldsymbol{\psi}^{(q+2k-1)}, \boldsymbol{\psi}^{(q+1)} \rangle \\ &= \cdots\cdots = (-1)^k \langle \boldsymbol{\psi}^{(q+k)}, \boldsymbol{\psi}^{(q+k)} \rangle = 0\end{aligned}$$

Hence the proposition is proven.

Theorem 1.17. If there exists a zero eigenvalue with multiplicity $2m$ for a Hamiltonian matrix $\boldsymbol{H}$, then there also exists a set of basic eigenvectors and Jordan form eigenvectors $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(2m-1)}\}$ corresponding to the zero eigenvalue such that these eigenvectors have adjoint symplectic orthonormality as follows

$$\langle \boldsymbol{\psi}^{(i)}, \boldsymbol{\psi}^{(j)} \rangle \begin{cases} = (-1)^i a \neq 0 & \text{for } i + j = 2m - 1 \\ = 0 & \text{for } i + j \neq 2m - 1 \end{cases} \tag{1.3.22}$$

Proof. A proof for i by mathematical induction is established.

(1) For $i = 0$, let $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(2m-1)}\}$ be a set of basic eigenvectors and Jordan form eigenvectors corresponding to a zero eigenvalue.

Firstly, using Eq. (1.1.19) for any $j \leq 2m - 2$ yields

$$\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{H}\boldsymbol{\psi}^{(j+1)} \rangle = \langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(j)} \rangle$$

On the other hand, as $\boldsymbol{H}$ is a Hamiltonian matrix, we have

$$\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{H}\boldsymbol{\psi}^{(j+1)} \rangle = \langle \boldsymbol{\psi}^{(j+1)}, \boldsymbol{H}\boldsymbol{\psi}^{(0)} \rangle = 0$$

hence

$$\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(j)} \rangle = 0 \quad (j \leq 2m - 2)$$

Next, $\boldsymbol{\psi}^{(0)}$ is symplectic orthogonal to all eigenvectors of nonzero eigenvalues as well as the Jordan form eigenvectors in accordance to Theorem 1.15, hence $\boldsymbol{\psi}^{(0)}$ must be symplectic adjoint with $\boldsymbol{\psi}^{(2m-1)}$. Otherwise it will be symplectic orthogonal to all eigenvectors and this conclusion is contradictory. Hence we have

$$\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(2m-1)} \rangle = a \neq 0$$

Hence Eq. (1.3.22) is true for $i = 0$.

(2) Assume that there exists a set of basic eigenvectors and Jordan form eigenvectors $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(2m-1)}\}$ corresponding to a zero eigenvalue

such that Eq. (1.3.22) is true for $i \leq k (k \geq 0)$. Now consider the case of $i \leq k+1$.

(a) For odd k, we denote

$$t = -\frac{1}{2a}\langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\psi}^{(2m-1)} \rangle$$

and take

$$\tilde{\boldsymbol{\psi}}^{(k+1+p)} = \boldsymbol{\psi}^{(k+1+p)} + t\boldsymbol{\psi}^{(p)} \quad (p = 0, 1, \ldots, (2m-k-2))$$

Obviously, $\{\boldsymbol{\psi}^{(0)}, \ldots, \boldsymbol{\psi}^{(k)}, \tilde{\boldsymbol{\psi}}^{(k+1)}, \ldots, \tilde{\boldsymbol{\psi}}^{(2m-1)}\}$ is also a set of basic eigenvectors and Jordan form eigenvectors corresponding to the zero eigenvalue. Furthermore, it still holds for $i \leq k (k \geq 0)$ with respect to Eq. (1.3.22). But here we have

$$\begin{aligned} \langle \tilde{\boldsymbol{\psi}}^{(k+1)}, \tilde{\boldsymbol{\psi}}^{(2m-1)} \rangle &= \langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\psi}^{(2m-1)} \rangle + t\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(2m-1)} \rangle \\ &\quad + t\langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\psi}^{(2m-k-2)} \rangle + t^2\langle \boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(2m-k-2)} \rangle \\ &= -2ta + ta + (-1)t\langle \boldsymbol{\psi}^{(k)}, \boldsymbol{\psi}^{(2m-k-1)} \rangle = 0 \end{aligned}$$

(b) For even k, from the lemma in Eq. (1.3.21), we have

$$\langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\psi}^{(2m-1)} \rangle = 0$$

Combining (a) and (b) yields a set of basic eigenvectors and Jordan form eigenvectors $\{\boldsymbol{\psi}^{(0)}, \ldots, \boldsymbol{\psi}^{(2m-1)}\}$ corresponding to the zero eigenvalue such that Eq. (1.3.22) is true for $i \leq k$ $(k \geq 0)$. Furthermore, they satisfy $\langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\psi}^{(2m-1)} \rangle = 0$. From Eq. (1.3.20), for any $j \leq 2m-2$, we have

$$\langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\psi}^{(j)} \rangle = -\langle \boldsymbol{\psi}^{(k)}, \boldsymbol{\psi}^{(j+1)} \rangle$$

According to the assumption of mathematical induction, Eq. (1.3.22) is tenable for $i \leq k$ $(k \geq 0)$. Hence we have

$$\langle \boldsymbol{\psi}^{(k+1)}, \boldsymbol{\psi}^{(j)} \rangle \begin{cases} = (-1)^{k+1}a \neq 0; & \text{when } k+1+j = 2m-1 \\ = 0; & \text{when } k+1+j \neq 2m-1 \end{cases}$$

Therefore, Eq. (1.3.22) is tenable for $i \leq k+1$ with respect to $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(2m-1)}\}$.

In other words, there exists a set of eigenvectors $\{\boldsymbol{\psi}^{(0)}, \boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(2m-1)}\}$ corresponding to the zero eigenvalue such that Eq. (1.3.22) is tenable for $i \leq k+1$. Therefore according to mathematical induction, the theorem is proven.

Theorems 1.14 to 1.17 indicate that there exists an adjoint symplectic orthonormal basis composed of the basic eigenvectors and Jordan form eigenvectors of the Hamiltonian matrix $\boldsymbol{H}$ in a $2n$-dimensional symplectic space. Through normalization, a normal adjoint symplectic orthonormal basis can be formed. The matrix formed by the column vectors is indeed a symplectic matrix. The properties of Hamiltonian matrix above are verified by a specific example as follows.

Example 1.3. Construct a normal adjoint symplectic orthonormal basis composed of the basic eigenvectors and Jordan form eigenvectors of the following Hamiltonian matrix

$$\boldsymbol{H} = \begin{bmatrix} 1 & -2 & 0 & 2 & 1 & 3 \\ 0 & -1 & 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & 3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution: First determine the eigenvalue of the Hamiltonian matrix $\boldsymbol{H}$. From

$$|\mu \boldsymbol{I} - \boldsymbol{H}| = \mu^2(\mu - 1)^2(\mu + 1)^2$$

hence $\mu = 0, \pm 1$ are the eigenvalues of Hamiltonian matrix $\boldsymbol{H}$. All of them are eigenvalues with multiplicity 2.

Then we solve the eigenvectors. For the eigenvectors corresponding to eigenvalue $\mu = 0$, from

$$\boldsymbol{H}\boldsymbol{\psi} = 0$$

we obtain

$$\boldsymbol{\psi}_0^{(0)} = \left\{ 1, 1, 0, 0, 1, 0 \right\}^{\mathrm{T}}$$

As there is only a single Jordan form chain corresponding to the zero eigenvalue, a Jordan form solution exists. From

$$\boldsymbol{H}\boldsymbol{\psi} = \boldsymbol{\psi}_0^{(0)}$$

we obtain

$$\boldsymbol{\psi}_0^{(1)} = \left\{ 0, 0, 1, 0, 1, 0 \right\}^{\mathrm{T}}$$

Obviously, $\boldsymbol{\psi}_0^{(0)}$ is symplectic adjoint with $\boldsymbol{\psi}_0^{(1)}$

$$\langle \boldsymbol{\psi}_0^{(0)}, \boldsymbol{\psi}_0^{(1)} \rangle = 1$$

For the eigenvectors corresponding to eigenvalue $\mu = 1$, from

$$\boldsymbol{H}\boldsymbol{\psi} = \boldsymbol{\psi}$$

we obtain

$$\boldsymbol{\psi}_1^{(0)} = \left\{ 1, 0, 0, 0, 0, 0 \right\}^{\mathrm{T}}$$

Similarly, there exists the Jordan form eigenvectors corresponding to eigenvalue $\mu = 1$. From

$$\boldsymbol{H}\boldsymbol{\psi} = \boldsymbol{\psi} + \boldsymbol{\psi}_1^{(0)}$$

we obtain

$$\boldsymbol{\psi}_1^{(1)} = \left\{ 0,\ \frac{1}{2},\ 0,\ 0,\ \frac{1}{2},\ \frac{1}{2} \right\}^{\mathrm{T}}$$

For the eigenvectors corresponding to eigenvalue $\mu = -1$, from

$$\boldsymbol{H}\boldsymbol{\psi} = -\boldsymbol{\psi}$$

we obtain

$$\boldsymbol{\psi}_{-1}^{(0)} = \left\{ 0\ 0, 1, 0, 0, 0 \right\}^{\mathrm{T}}$$

Similarly, there exists the Jordan form eigenvectors corresponding to eigenvalue $\mu = -1$. From

$$\boldsymbol{H}\boldsymbol{\psi} = -\boldsymbol{\psi} + \boldsymbol{\psi}_{-1}^{(0)}$$

we obtain

$$\boldsymbol{\psi}_{-1}^{(1)} = \left\{ -\frac{1}{4},\ 0,\ 0,\ \frac{1}{2},\ -\frac{1}{2},\ 0 \right\}^{\mathrm{T}}$$

As there are only two vectors for the Jordan form chains corresponding to eigenvalue $\mu = \pm 1$, hence $\boldsymbol{\psi}_1^{(0)}$ and $\boldsymbol{\psi}_{-1}^{(1)}$, $\boldsymbol{\psi}_{-1}^{(0)}$ and $\boldsymbol{\psi}_1^{(1)}$ must be symplectic adjoint

$$a = \langle \boldsymbol{\psi}_{-1}^{(0)}, \boldsymbol{\psi}_1^{(1)} \rangle = -\langle \boldsymbol{\psi}_{-1}^{(1)}, \boldsymbol{\psi}_1^{(0)} \rangle = \frac{1}{2}$$

Here $\boldsymbol{\psi}_1^{(1)}$ and $\boldsymbol{\psi}_{-1}^{(1)}$ are not symplectic orthogonal. Referring to the proof of Theorem 1.16, denote

$$t = -2\langle \boldsymbol{\psi}_1^{(1)}, \boldsymbol{\psi}_{-1}^{(1)} \rangle = \frac{1}{2}$$

and take

$$\tilde{\boldsymbol{\psi}}_1^{(1)} = \boldsymbol{\psi}_1^{(1)} + t\boldsymbol{\psi}_1^{(0)} = \left\{ \frac{1}{2},\ \frac{1}{2},\ 0,\ 0,\ \frac{1}{2},\ \frac{1}{2} \right\}^{\mathrm{T}}$$

then the eigenvectors $\boldsymbol{\psi}_{-1}^{(1)}$ and $\tilde{\boldsymbol{\psi}}_1^{(1)}$ are symplectic orthogonal. According to Theorems 1.15 and 1.16, the other eigenvectors satisfy symplectic orthogonality. In this way an adjoint symplectic orthonormal basis formed by the eigenvectors of Hamiltonian matrix $\boldsymbol{H}$ is

$$\boldsymbol{\psi}_{-1}^{(0)}, \quad \boldsymbol{\psi}_{-1}^{(1)}, \quad \boldsymbol{\psi}_0^{(0)}; \quad \tilde{\boldsymbol{\psi}}_1^{(1)}, \quad \boldsymbol{\psi}_1^{(0)}, \quad \boldsymbol{\psi}_0^{(1)}$$

In addition, from Eq. (1.3.18) and via symplectic normalization, a normal adjoint symplectic orthonormal basis

$$\boldsymbol{\psi}_{-1}^{(0)}, \quad \boldsymbol{\psi}_{-1}^{(1)}, \quad \boldsymbol{\psi}_0^{(0)}; \quad \hat{\boldsymbol{\psi}}_1^{(1)}, \quad \hat{\boldsymbol{\psi}}_1^{(0)}, \quad \boldsymbol{\psi}_0^{(1)}$$

can be composed where

$$\begin{aligned} \hat{\boldsymbol{\psi}}_1^{(0)} &= -2\boldsymbol{\psi}_1^{(0)} = \{-2, 0, 0, 0, 0, 0\}^{\mathrm{T}} \\ \hat{\boldsymbol{\psi}}_1^{(1)} &= 2\tilde{\boldsymbol{\psi}}_1^{(1)} = \{1, 1, 0, 0, 1, 1\}^{\mathrm{T}} \end{aligned}$$

It should be clearly stated here that the arrangement of $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r\}$ for all adjoint symplectic orthonormal vector sets expressed as Eq. (1.3.8) is adopted in this book. It is absolutely valid to have another arrangement as $\{\boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_r\}$, but in this case the definition of unit symplectic matrix Eq. (1.3.6) should be rewritten as

$$\boldsymbol{J}'_{2n} = \begin{bmatrix} \boldsymbol{J}'_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{J}'_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{J}'_2 \end{bmatrix} \quad \text{where } \boldsymbol{J}'_2 = \boldsymbol{J}_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{1.3.23}$$

Furthermore, other relevant definitions such as symplectic matrix, Hamiltonian matrix, etc. should be modified accordingly. This arrangement is comparatively more convenient for numerical analysis. The details are omitted here.

Table 1.1. The correlation between Euclidean space and symplectic space.

Euclidean space	**Symplectic space**
inner product $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ — {length}	symplectic inner product $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle$ — {area}
unit matrix $\boldsymbol{I}$	unit symplectic matrix $\boldsymbol{J}$
orthogonality $(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}(= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{I}\boldsymbol{y}) = 0$	symplectic orthogonality $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{J}\boldsymbol{y} = 0$
(normal) orthogonal basis	(normal) adjoint symplectic orthonormal basis
orthogonal matrix $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = (\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{I}\boldsymbol{Q} =)\boldsymbol{I}$	symplectic matrix $\boldsymbol{S}^{\mathrm{T}}\boldsymbol{J}\boldsymbol{S} = \boldsymbol{J}$
symmetry transformation $(\boldsymbol{\alpha}, \tilde{\boldsymbol{A}}\boldsymbol{\beta}) = (\boldsymbol{\beta}, \tilde{\boldsymbol{A}}\boldsymbol{\alpha})$	Hamiltonian transformation $\langle \boldsymbol{\alpha}, \tilde{\boldsymbol{H}}\boldsymbol{\beta} \rangle = \langle \boldsymbol{\beta}, \tilde{\boldsymbol{H}}\boldsymbol{\alpha} \rangle$
symmetric matrix $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}(= \boldsymbol{I}\boldsymbol{A}\boldsymbol{I})$	Hamiltonian matrix $\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{JHJ}$
The eigenvalues of real symmetric matrix are real	If μ is an eigenvalue of a Hamiltonian matrix, $-\mu$ is also its eigenvalue
The eigenvectors corresponding to distinct eigenvalues of a real symmetric matrix are orthogonal	The eigenvectors corresponding to non-symplectic-adjoint eigenvalues of a Hamiltonian matrix are symplectic orthogonal
The eigenvectors of a real symmetric matrix can form a normal orthogonal basis	The eigenvectors of a Hamiltonian matrix can form a normal adjoint symplectic orthonormal basis

In this section, the basic concepts of a finite-dimensional symplectic space are elaborated and some fundamental properties are briefly introduced. It is of certainty that the many concepts and properties can be directly generalized to an infinite-dimensional symplectic space. Towards the end of this section, the correlation between a Euclidean space and a symplectic space is presented in order to better describe the relevant concepts and properties of a symplectic space for the benefit of readers.

1.4. Legengre's Transformation

Legendre's transformation in the scope of mathematics is introduced in this section. It is the key to realize a transformation from the Lagrange system to the Hamiltonian system.

Consider Legendre's transformation in two variables[6]. Let $f = f(x, y)$, then

$$\mathrm{d}f = u\mathrm{d}x + v\mathrm{d}y \tag{1.4.1}$$

where

$$u = \frac{\partial f}{\partial x}, \quad v = \frac{\partial f}{\partial y} \tag{1.4.2}$$

Here we choose x, y as the independent variables. In reality, we may choose any two of x, y, u, v as the independent variables to suit the problem under consideration. If we choose u, y as the independent variables, we obtain from Eq. (1.4.2):

$$x = x(u, y), \quad v = v(u, y) \tag{1.4.3}$$

while the function f can be expressed in terms of u, y as:

$$\bar{f}(u, y) = f[x(u, y), y] \tag{1.4.4}$$

then

$$\left.\begin{aligned} \frac{\partial \bar{f}}{\partial y} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial y} + \frac{\partial f}{\partial y} = u\frac{\partial x}{\partial y} + v \\ \frac{\partial \bar{f}}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} = u\frac{\partial x}{\partial u} = \frac{\partial}{\partial u}(ux) - x \end{aligned}\right\} \tag{1.4.5}$$

Equation (1.4.5) can be expressed as

$$\left.\begin{aligned} v &= -\frac{\partial}{\partial y}\left(ux - \bar{f}\right) = -\frac{\partial g}{\partial y} \\ x &= \frac{\partial}{\partial u}\left(ux - \bar{f}\right) = \frac{\partial g}{\partial u} \end{aligned}\right\} \tag{1.4.6}$$

where $g(u, y) = ux - \bar{f} = x\partial f/\partial x - f$. It shows that when the independent variables change from x, y to u, y and using function $\bar{f}$, then x, v cannot be directly expressed in terms of the partial derivative of $\bar{f}$ with respect to u and y as in Eq. (1.4.2). Instead, we should use function g which is equal to the variable to be eliminated x multiplied by the partial derivative of the former function with respect to this variable $u = \partial f/\partial x$ and minus the original function f. Hence, x, v can then be expressed in terms of the partial derivative of g with respect to u and y. This is the basic principle of Legendre's transformation.

The discussion above is merely for the Legendre's transformation on variable x. Of course, we may also perform Legendre's transformation on two variables x, y simultaneously, i.e. we may choose u, v as the independent variables. In a similar way, using (1.4.2) yields:

$$x = x(u, v), \quad y = y(u, v) \tag{1.4.7}$$

and function f can be alternatively expressed in terms of u, v, as

$$\tilde{f}(u,v) = f[x(u,v), y(u,v)] \tag{1.4.8}$$

Introducing transformation function

$$\tilde{g}(u,v) = ux + vy - \tilde{f}(u,v) \tag{1.4.9}$$

it is obvious that following relation exists

$$\left.\begin{aligned} \frac{\partial \tilde{g}}{\partial u} &= x + u\frac{\partial x}{\partial u} + v\frac{\partial y}{\partial u} - \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} - \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} = x \\ \frac{\partial \tilde{g}}{\partial v} &= u\frac{\partial x}{\partial v} + y + v\frac{\partial y}{\partial v} - \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} - \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} = y \end{aligned}\right\} \tag{1.4.10}$$

i.e. x, y can be expressed in terms of the partial derivative of $\tilde{g}$ with respect to u and v.

In this section, we introduce the Legendre's transformation for two variables. The approach can be directly generalized from two variables to multiple variables. The details are omitted here.

1.5. The Hamiltonian Principle and the Hamiltonian Canonical Equations

"The nature always chooses the simplest and most possible way." This is a famous principle of Fermat. The minimum action principle in classical mechanics is originated from the Hamiltonian principle. It is often described in terms of the n-dimensional general displacement $q_i (i = 1, 2, \ldots, n)$ with finite degrees of freedom or in terms of vector $\boldsymbol{q}$. Using $\dot{q}_i$ to indicate differentiation with respect to time, the **Lagrange function** of a dynamic system (kinetic energy minus potential energy) is

$$\mathscr{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \quad \text{or } \mathscr{L}(q_1, q_2, \ldots, q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n) \tag{1.5.1}$$

The Hamiltonian principle states that the actual path of a conservative system from the initial point $(\boldsymbol{q}_0, t_0)$ to the terminal point $(\boldsymbol{q}_\mathrm{e}, t_\mathrm{e})$ is such that the action A is a stationary value,

$$A = \int_{t_0}^{t_\mathrm{e}} \mathscr{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})\mathrm{d}t; \quad \delta A = 0 \tag{1.5.2}$$

Performing variation of Eq. (1.5.2) and integrating by parts yield

$$\delta A = \int_{t_0}^{t_\mathrm{e}} \left[\frac{\partial \mathscr{L}}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathscr{L}}{\partial \dot{\boldsymbol{q}}} \right) \right]^\mathrm{T} \cdot \delta \boldsymbol{q} \mathrm{d}t = 0 \tag{1.5.3}$$

As $\delta \boldsymbol{q}$ is arbitrary, the **Lagrange equation**:

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathscr{L}}{\partial \dot{\boldsymbol{q}}} \right) = \frac{\partial \mathscr{L}}{\partial \boldsymbol{q}} \tag{1.5.4}$$

is derived. Hence, the Hamiltonian principle (1.5.2) corresponds to the Lagrange equation (1.5.4), which is a system of second-order ordinary differential equations. The expression above only includes one class of variables like displacement, and therefore it is a variational principle with a single class of variables.

The Hamiltonian canonical system has already been developed in classical analytical mechanics. It transforms a class of independent variables $\dot{\boldsymbol{q}}$ (generalized velocity) of Lagrange function $\mathscr{L}$ into $\boldsymbol{p}$ (generalized momentum, i.e. dual variable) by Legendre's transformation

$$\boldsymbol{p} = \frac{\partial \mathscr{L}}{\partial \dot{\boldsymbol{q}}} \tag{1.5.5}$$

From Eq. (1.5.5), we may solve $\dot{\boldsymbol{q}}$ such that $\dot{\boldsymbol{q}}$ is a function of $\boldsymbol{p}, \boldsymbol{q}$, or

$$\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}}(\boldsymbol{p}, \boldsymbol{q}) \tag{1.5.6}$$

According to the principle of Legendre's transformation, we introduce a transformation function, i.e. the **Hamiltonian function** (kinetic energy plus potential energy)

$$\mathscr{H}(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{p}^\mathrm{T} \dot{\boldsymbol{q}} - \mathscr{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}(\boldsymbol{p}, \boldsymbol{q})) \tag{1.5.7}$$

Hence, Eq. (1.4.6) yields

$$\frac{\partial \mathscr{L}}{\partial \boldsymbol{q}} = -\frac{\partial \mathscr{H}}{\partial \boldsymbol{q}}; \quad \dot{\boldsymbol{q}} = \frac{\partial \mathscr{H}}{\partial \boldsymbol{p}} \tag{1.5.8}$$

On the other hand, from Eq. (1.5.4)

$$\frac{\partial \mathscr{L}}{\partial \boldsymbol{q}} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathscr{L}}{\partial \dot{\boldsymbol{q}}} \right) = \dot{\boldsymbol{p}} \tag{1.5.9}$$

we obtain

$$\dot{\boldsymbol{q}} = \frac{\partial \mathscr{H}}{\partial \boldsymbol{p}}; \quad \dot{\boldsymbol{q}} = -\frac{\partial \mathscr{H}}{\partial \boldsymbol{q}} \tag{1.5.10}$$

Equation (1.5.10) are the **Hamiltonian canonical equations**, in which there are two classes of variables: the generalized displacement $\boldsymbol{q}$ and the generalized momentum $\boldsymbol{p}$. The variational principle corresponding to the Hamiltonian equations (1.5.10) is

$$\delta \int_{t_0}^{t_e} [\boldsymbol{p}^{\mathrm{T}}\dot{\boldsymbol{q}} - \mathscr{H}(\boldsymbol{q},\boldsymbol{p})]\,\mathrm{d}t = 0 \tag{1.5.11}$$

where $\boldsymbol{q}$ and $\boldsymbol{p}$ should be regarded as two unrelated variables with independent variation. Equation (1.5.10) can be obtained directly when the variational formula (1.5.11) is expanded.

The process of transforming from the variational principle with a single class of variables in Eq. (1.5.2) to the variational principle with two classes of variables in Eq. (1.5.11) bears a classical feature realized through Legendre's transformation.

1.6. The Reciprocal Theorems

An elastic system is a system without energy dissipation. Hence the strain energy stored during the process of deformation must be equal to the work done by the external force during the process. If applied in a quasi-static process, the work done by the external force only depends on the state of displacement state at the particular moment. If a linear infinitesimal deformation problem is considered, the principle of superposition is applicable. A series of reciprocal theorems can therefore be derived by combining the principles of superposition and energy conservation.

1.6.1. *The Reciprocal Theorem for Work*[7]

Consider a linear system under the action of two sets of forces F_A and F_B at two different positions as illustrated in Fig. 1.1. Because it is an elastic system, the work done by the external forces during the process of deformation is independent of the sequence of loading. It only depends on the final state of the external forces. Now consider the outcome of two different loading routes.

For the first loading route, the action of F_A occurs first in a quasi-static progressive manner and it is followed by the action of F_B. F_B is not in action when F_A is applied. Hence the only work done is by F_A with magnitude 0.5 $F_A u_A$, where u_A is the displacement at A caused by F_A.

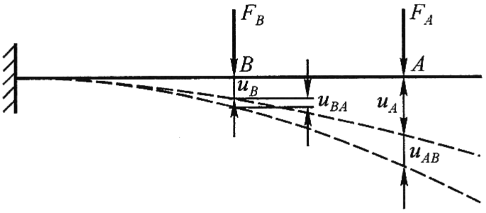

Fig. 1.1. Two different load processes.

After applying F_A, the action of force F_B on the system begins. The work done by F_B is 0.5 $F_B u_B$ where u_B is the displacement at B caused by F_B. During the process of loading F_B, a displacement u_{AB} at A (i.e. the displacement at A caused by F_B) occurs. As F_A applied on, the system remains, the additional work done by F_A is $F_A u_{AB}$. Hence the total work done by the external forces is

$$W_1 = \frac{1}{2}F_A u_A + \frac{1}{2}F_B u_B + F_A u_{AB} \tag{1.6.1}$$

For the second loading route, F_B acts first on the system and follows by F_A. Following the same analytical procedure, we obtain the total work done by the external forces as

$$W_2 = \frac{1}{2}F_B u_B + \frac{1}{2}F_A u_A + F_B u_{BA} \tag{1.6.2}$$

where u_{BA} is the displacement at B caused by F_A.

Because the final state due to different loading sequences does not change, the ultimate deformation of structure for the two varying loading sequences are identical. Therefore, we have $W_1 = W_2$ and

$$F_A u_{AB} = F_B u_{BA} \tag{1.6.3}$$

Hence we establish **the reciprocal theorem for work: in a system undergoing arbitrary linear elastic deformation, the work done by a first set of external forces due to the displacement caused by a second set of external forces is equal to the work done by the second set of external forces due to the displacement caused by the first set of external forces.** The reciprocal theorem for work in elastic mechanics and structural mechanics is very useful. From this theorem it is deduced the reciprocal theorem for displacement and the reciprocal theorem for reaction.

1.6.2. *The Reciprocal Theorem for Displacement*

Supposing $F_A = F_B = 1$ in Eq. (1.6.3), we obtain

$$u_{AB} = u_{BA} \tag{1.6.4}$$

Hence, we establish **the reciprocal theorem for displacement: in a system undergoing linear elastic deformation, the displacement at B caused by unit force acting at A is equal to the displacement at A caused by unit force acting at B.**

1.6.3. *The Reciprocal Theorem for Reaction*

Consider the settlement problem of supports as shown in Fig. 1.2. Let s_A and s_B are the settlements at supports A and B, respectively. For a linear elastic system, the state of system is independent of the order of settlement of supports. It only depends on the final settlement shape. Now we consider settlement in two different cases.

In the first case, s_A takes place first and follows by s_B. During the settlement of s_A, there are reactions F_A and F_{BA} at A and B, respectively. During the process, there is no settlement at support B and therefore the only work done is by the reaction F_A at A, which is 0.5 $F_A s_A$. Then a settlement s_B at support B takes place and it generates the reactions F_B and F_{AB} at B and A, respectively. There is no work done at support A because A does not move. As reaction F_{BA} at B exists during the process, the work done is $F_{BA} s_B$. In addition, the work done by the gradually

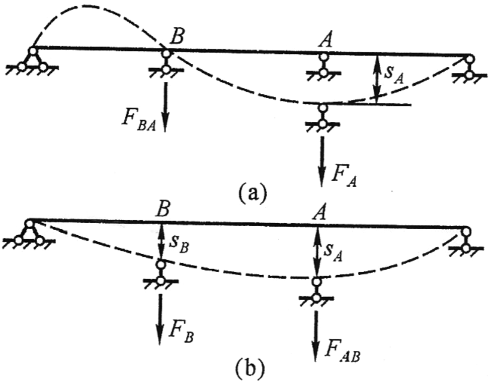

Fig. 1.2. Support settlement.

increasing reaction F_B is 0.5 $F_B s_B$. Hence the total work is

$$W_1 = \frac{1}{2} F_A s_A + \frac{1}{2} F_B s_B + F_{BA} s_B \tag{1.6.5}$$

In the second case, s_B takes place first and follows by s_A. Following the same procedure, we obtain the total work as

$$W_2 = \frac{1}{2} F_B s_B + \frac{1}{2} F_A s_A + F_{AB} s_A \tag{1.6.6}$$

Because the final settlements states are identical, hence $W_1 = W_2$ and

$$F_{BA} s_B = F_{AB} s_A \tag{1.6.7}$$

Further, assuming $s_A = s_B = 1$ yields

$$F_{AB} = F_{BA} \tag{1.6.8}$$

Hence, we establish **the reciprocal theorem for reaction: in a system undergoing linear elastic deformation, the reaction at B caused by unit displacement at A is equal to the reaction at A caused by unit displacement at B.**

1.6.4. *The Reciprocal Theorem for Displacement and Negative Reaction*

Suppose there is a unit external force[a] acting at point A of an elastic system, while there is unit support displacement at support B as illustrated in Fig. 1.3. As the strain energy of a linear elastic system depends only on the final states of loads and support displacements and is independent of orders they are applied, we may consider the following two cases with different orders of application.

---

[a] The magnitude of unit external force and unit displacement is 1. These quantities are omitted in the following equations.

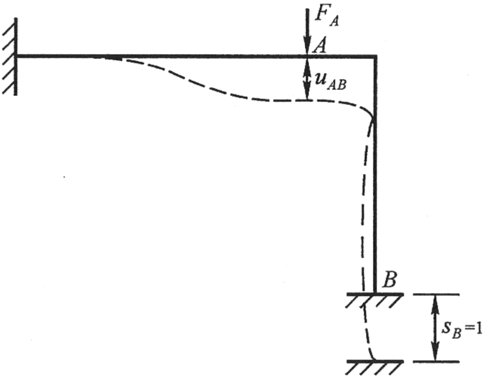

Fig. 1.3. The reciprocal for displacement and negative reaction.

In the first case, we apply a unit displacement $s_B = 1$ at support B and the displacement at A is u_{AB}. There is no work at A because the unit force at A has not been applied. The only work $0.5F_B s_B$ is done by the reaction F_B at B. Then we apply a unit external force $F_A = 1$ at A, which results in a reaction F_{BA} at B. Again the reaction at B does no work because there is no support displacement during this process. The only work $u_A \times 1/2$ is due to the unit external force where u_A is the displacement at A resulting from the unit force at A. The work done by external forces corresponding to this order of application is

$$W_1 = \frac{1}{2}u_A + \frac{1}{2}F_B \tag{1.6.9}$$

In the second case, we apply a unit external force at A first and the work done by this external force is $u_A \times 1/2$. Then we apply a unit displacement at support B. As then there exists a reaction F_{BA} at B, the work done by the external forces during the latter process is $u_{AB} \times 1 + F_{BA} \times 1 + F_B \times 1/2$. Hence the work done by the external forces corresponding to this order of application sums up to

$$W_2 = \frac{1}{2}u_A + \frac{1}{2}F_B + u_{AB} + F_{BA} \tag{1.6.10}$$

These two orders of application correspond the identical final state of external influences and structural deformation. Therefore $W_1 = W_2$ and

$$F_{BA} = -u_{AB} \tag{1.6.11}$$

As a result, we establish **the reciprocal theorem for displacement and negative reaction: in an arbitrary system undergoing linear**

elastic deformation, the reaction at support B caused by a unit force acting at A is equal to the negative value of the displacement at A caused by a unit displacement at support B.

When the hybrid method (i.e. partially the unknown quantities of some displacements and partially the unknown quantities of redundant internal force) is adopted to solve a system of structural mechanics, the symmetry or antisymmetry of the canonical equation is reflected in the reciprocal theorems for displacement, for reaction and for displacement and negative reaction. The Hamiltonian system corresponds to the adoption of the hybrid method, and the properties of Hamiltonian matrix reflects the reciprocal theorems for displacement, for reaction, and for displacement and negative reaction.

The chapter discusses some basic concepts, such as symplectic space, Hamiltonian system, etc. from various aspects of mathematics and analytical mechanics. We will introduce in detail the solutions of symplectic Hamiltonian systems through some elasticity problems in the following chapters.

References

1. Rongchang Shi, *Matrix Analysis*, Beijing: Beijing Institute of Technology Press, 1996 (in Chinese).
2. Yizhong Lan and Chunlai Zhao, *Theory of Linear Algebra*, Beijing: Peking University Press, 1998 (in Chinese).
3. Department of Mathematical Mechanics of Peking University, *Advanced Algebra*, Beijing: People's Education Press, 1978 (in Chinese).
4. J. Koszul and Yiming Zou, *Theory of Symplectic Geometry*, Beijing: Science Press, 1986 (in Chinese).
5. MengZhao Qin, Symplectic geometry and computing Hamiltonian mechanics, *Mechanics and Practice*, 1990, 12(6): 1–20 (in Chinese).
6. Yanbai Zhou, *Lectures on Theoretical Mechanics*, Beijing: People's Education Press, 1979 (in Chinese).
7. Lingxi Qian, *Statically Indeterminate Structural Analysis*, Shanghai: Shanghai Scientific and Technical Publishers, 1951 (in Chinese).