Chapter 1

REVIEW ON LINEAR ALGEBRAS

The main mathematical tool in group theory is linear algebras. In this chapter, we will review some fundamental concepts and calculation methods in linear algebras, which are often used in group theory.

1.1 Linear Space and Basis Vector

Let $H(x)$ be the Hamiltonian of a system. Suppose that the eigenvalue E of $H(x)$ is m-degenerate,

$$H(x)\psi_\mu(x) = E\psi_\mu(x), \qquad \mu = 1,\ 2,\ \cdots,\ m, \tag{1.1}$$

where x briefly denotes the set of coordinates for all degrees of freedom. $\psi_\mu(x)$ are linearly independent to one another. Any linear combination of $\psi_\mu(x)$ is an eigenfunction of $H(x)$ with the same eigenvalue

$$\phi(x) = \sum_{\mu=1}^{m} \psi_\mu(x) a_\mu, \qquad H(x)\phi(x) = E\phi(x). \tag{1.2}$$

Conversely, any eigenfunction of $H(x)$ with the eigenvalue E can be expressed as a linear combination of $\psi_\mu(x)$ like (1.2). Two eigenfunctions satisfy the following calculation rule,

$$c\left(\sum_{\mu=1}^{m} \psi_\mu(x)a_\mu + \sum_{\mu=1}^{m} \psi_\mu(x)b_\mu\right) = \sum_{\mu=1}^{m} \psi_\mu(x)\left(c\ a_\mu + c\ b_\mu\right). \tag{1.3}$$

The set of $\phi(x)$ is called a linear space $\mathcal{L}$ of dimension m, generated by m basis vectors of $\psi_\mu(x)$. $\phi(x)$ is an arbitrary vector in $\mathcal{L}$ and a_μ is the μth-component of the vector $\phi(x)$ with respect to the basis vectors $\psi_\mu(x)$.

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

Generally, m objects $\boldsymbol{e}_\mu$ are said to be linearly independent if there do not exist m coefficients c_μ which are not vanishing simultaneously such that

$$\sum_{\mu=1}^{m} \boldsymbol{e}_\mu c_\mu = 0. \tag{1.4}$$

$\boldsymbol{e}_\mu$ satisfy the following linear formulas:

$$\begin{aligned} &\boldsymbol{e}_\mu a_\mu + \boldsymbol{e}_\nu a_\nu = \boldsymbol{e}_\nu a_\nu + \boldsymbol{e}_\mu a_\mu, \\ &c\left(\sum_\mu \boldsymbol{e}_\mu a_\mu + \sum_\mu \boldsymbol{e}_\mu b_\mu\right) = \sum_\mu \boldsymbol{e}_\mu \left(c a_\mu + c b_\mu\right), \end{aligned} \tag{1.5}$$

where c, a_μ, a_ν, and b_μ are arbitrary complex numbers. The m objects $\boldsymbol{e}_\mu$ generate a linear space $\mathcal{L}$ of dimension m, which is the set of all possible complex combinations $\boldsymbol{a}$ of $\boldsymbol{e}_\mu$

$$\boldsymbol{a} = \sum_{\mu=1}^{m} \boldsymbol{e}_\mu a_\mu. \tag{1.6}$$

$\boldsymbol{a}$ is called a vector in $\mathcal{L}$, $\boldsymbol{e}_\mu$ is a basis vector, and a_μ is the μth component of $\boldsymbol{a}$ with respect to the basis vectors $\boldsymbol{e}_\mu$. The space $\mathcal{L}$ is called a real space if the components a_μ of all vectors $\boldsymbol{a}$ in $\mathcal{L}$ are real. A vector is called a null vector if its components are all vanishing. Two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are said to be equal to each other if and only if their components are respectively equal, $a_\mu = b_\mu$. In linear algebras, the concepts of vectors and linear space are independent of the physical content of the objects.

For a given space $\mathcal{L}$ and a given set of basis vectors, vector $\boldsymbol{a}$ is completely described by the m components a_μ. Usually, the m ordered numbers are arranged as a column-matrix $\underline{a}$,

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}. \tag{1.7}$$

The column-matrix $\underline{a}$ is another form to denote vector $\boldsymbol{a}$. Sometimes, we do not distinguish two symbols $\underline{a}$ and $\boldsymbol{a}$.

A basis vector is a special vector where only one component is nonvanishing and to be one,

$$\left(\boldsymbol{e}_\mu\right)_\nu = \delta_{\mu\nu} = \begin{cases} 1 & \text{when} \quad \mu = \nu, \\ 0 & \text{when} \quad \mu \neq \nu, \end{cases} \tag{1.8}$$

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

where $\delta_{\mu\nu}$ is the Kronecker δ function.

n vectors $\boldsymbol{a}^{(1)}$, $\boldsymbol{a}^{(2)}$, $\cdots$, $\boldsymbol{a}^{(n)}$ are linearly dependent if there exists a linear relation

$$\sum_{i=1}^{n} \boldsymbol{a}^{(i)} c_i = 0, \tag{1.9}$$

where n coefficients c_i are not vanishing simultaneously. Otherwise, they are linearly independent. In an m-dimensional space $\mathcal{L}$ the number n of linearly independent vectors is not larger than m.

In $\mathcal{L}$, n linearly independent vectors generate a subspace $\mathcal{L}_1$ of dimension n. A subspace is called a null space $\emptyset$ if it contains only the null vector. The whole space $\mathcal{L}$ and the null space $\emptyset$ are two trivial subspaces. Usually, we only consider nontrivial subspaces.

The sum of two subspaces $\mathcal{L}_1$ and $\mathcal{L}_2$ is a subspace, denoted by $\mathcal{L}_1 + \mathcal{L}_2$, which contains all linear combinations of the vectors belonging to $\mathcal{L}_1$ and $\mathcal{L}_2$. The intersection of two subspaces is a subspace, denoted by $\mathcal{L}_1 \cap \mathcal{L}_2$, which contains all vectors belonging to both subspaces.

$\mathcal{L}$ is said to be the direct sum of two subspaces, $\mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2$, if $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$, and one of the following three equivalent conditions is satisfied.

(1) The intersection of $\mathcal{L}_1$ and $\mathcal{L}_2$ is a null space.

(2) The dimension of $\mathcal{L}$ is equal to the sum of the dimensions of $\mathcal{L}_1$ and $\mathcal{L}_2$.

(3) Each vector in $\mathcal{L}$ can be expressed uniquely as the sum of two vectors, respectively belonging to two subspaces $\mathcal{L}_1$ and $\mathcal{L}_2$.

$\mathcal{L}_2$ is called the complement of $\mathcal{L}_1$ in $\mathcal{L}$ if $\mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2$. $\mathcal{L}_1$ is also the complement of $\mathcal{L}_2$. The complement of $\mathcal{L}_1$ in $\mathcal{L}$ is not unique. A space $\mathcal{L}$ can be decomposed as a direct sum of some subspaces more than two.

1.2 Linear Transformations and Linear Operators

A transformation gives a rule, with which a function changes to another function. An operator is the mathematical symbol for a transformation. An operator $R(x)$ is linear if it satisfies

$$R(x)\left\{c_1\phi_1(x) + c_2\phi_2(x)\right\} = c_1 R(x)\phi_1(x) + c_2 R(x)\phi_2(x), \tag{1.10}$$

where the coefficients c_1 and c_2 are constant. A linear operator describes a linear transformation. The operators used in this textbook are linear if without special indication. The multiplication $R(x)S(x)$ of two operators

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

$R(x)$ and $S(x)$ is an operator, defined as a successive application to the function first with $S(x)$ and then with $R(x)$. Namely, if $S(x)\psi(x) = \phi(x)$, then $R(x)S(x)\psi(x) = R(x)\phi(x)$. Generally, the order of two operators in multiplication cannot be changed, namely $R(x)S(x) \neq S(x)R(x)$.

If an operator $R(x)$ is commutable with the Hamiltonian $H(x)$,

$$[H(x), R(x)] \equiv H(x)R(x) - R(x)H(x) = 0, \tag{1.11}$$

the application of $R(x)$ to the eigenfunction $\psi_\mu(x)$ of $H(x)$ is still an eigenfunction of $H(x)$ with the same eigenvalue,

$$H(x)\left\{R(x)\psi_\mu(x)\right\} = R(x)\left\{H(x)\psi_\mu(x)\right\} = E\left\{R(x)\psi_\mu(x)\right\}. \tag{1.12}$$

Thus, $R(x)\psi_\mu(x)$ belongs to the space $\mathcal{L}$ generated by m eigenfunctions ψ_μ of $H(x)$ with the same eigenvalue E, and can be expressed as Eq. (1.2),

$$R(x)\psi_\mu(x) = \sum_\nu \ \psi_\nu(x)D_{\nu\mu}(R). \tag{1.13}$$

$\mathcal{L}$ is called an invariant space to $R(x)$. The coefficients $D_{\nu\mu}(R)$ are arranged as a matrix $D(R)$ of dimension m, called the matrix of an operator $R(x)$ in the basis functions $\psi_\mu(x)$ of $\mathcal{L}$, or simply called the matrix of $R(x)$. Note that $D(R)$ depends on the operator $R(x)$, but not on x. The action of $R(x)$ to any function $\phi(x)$ in $\mathcal{L}$ can be calculated by the matrix $D(R)$. Namely, if $\phi(x) = \sum_\mu \ \psi_\mu(x)a_\mu$, and $R(x)\phi(x) = \phi_1(x) = \sum_\nu \ \psi_\nu(x)b_\nu$, one has

$$R(x)\phi(x) = \sum_{\mu=1}^{m} \ \left[R(x)\psi_\mu(x)\right]a_\mu = \sum_{\nu\mu} \ \psi_\nu D_{\nu\mu}(R)a_\mu,$$

$$b_\nu = \sum_{\mu=1}^{m} \ D_{\nu\mu}(R)a_\mu. \tag{1.14}$$

Generally, a linear operator R describes a transformation of vectors in a linear space $\mathcal{L}$ satisfying

$$R\left\{c_1\boldsymbol{a} + c_2\boldsymbol{b}\right\} = c_1R\ \boldsymbol{a} + c_2R\ \boldsymbol{b}. \tag{1.15}$$

$\mathcal{L}$ is invariant to R if the application of R to any vector $\boldsymbol{a}$ in $\mathcal{L}$ is still a vector in $\mathcal{L}$,

$$R\boldsymbol{a} = \boldsymbol{b} \in \mathcal{L}, \qquad \forall\ \boldsymbol{a} \in \mathcal{L}. \tag{1.16}$$

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

The matrix $D(R)$ of R in its invariant space $\mathcal{L}$ is calculated from the application of R to the basis vectors $\boldsymbol{e}_\mu$

$$R\,\boldsymbol{e}_\mu = \sum_\nu \boldsymbol{e}_\nu D_{\nu\mu}(R). \tag{1.17}$$

The action of R to any vector $\boldsymbol{a}$ in $\mathcal{L}$ can be calculated by $D(R)$. If

$$\boldsymbol{a} = \sum_\mu \boldsymbol{e}_\mu a_\mu, \qquad R\,\boldsymbol{a} = \boldsymbol{b} = \sum_\nu \boldsymbol{e}_\nu b_\nu, \tag{1.18}$$

then

$$R\,\boldsymbol{a} = \sum_\mu (R\boldsymbol{e}_\mu)\, a_\mu = \sum_{\nu\mu} \boldsymbol{e}_\nu D_{\nu\mu}(R) a_\mu,$$

$$b_\nu = \sum_\mu D_{\nu\mu}(R) a_\mu, \qquad \underline{b} = D(R)\ \underline{a}. \tag{1.19}$$

It is worthy to emphasize the difference between Eqs. (1.17) and (1.19). In Eq. (1.17) a basis vector $\boldsymbol{e}_\mu$ transforms in the operator R to a combination of basis vectors, where the combination index of the basis vectors is the row index ν of $D_{\nu\mu}(R)$. Equation (1.19) is a component equation for vector $\boldsymbol{a}$ transformed by the operator R to another vector $\boldsymbol{b}$, where the combination index of the vector components is the column index μ of $D_{\nu\mu}(R)$. Two equations are consistent because the basis vector $\boldsymbol{e}_\mu$ is a special vector, where only one component is nonvanishing but equal to one,

$$\left(R\boldsymbol{e}_\mu\right)_\rho = \sum_\lambda D_{\rho\lambda}(R) \left(\boldsymbol{e}_\mu\right)_\lambda = D_{\rho\mu}(R) = \sum_\nu \left(\boldsymbol{e}_\nu\right)_\rho D_{\nu\mu}(R). \tag{1.20}$$

1.3 Similarity Transformation

For a given set of basis vectors $\boldsymbol{e}_\mu$ in a linear space $\mathcal{L}$ of dimension m, there is a one-to-one correspondence between vector $\boldsymbol{a}$ and its column-matrix $\underline{a}$, and there is a one-to-one correspondence between an operator R and its matrix $D(R)$. However, the basis vectors in $\mathcal{L}$ are not unique. Any set of m linearly independent vectors can be chosen to be basis vectors. In this section we will discuss how the column-matrix of a vector and the matrix of an operator change when the basis vectors are changed.

Let $\boldsymbol{e}'_\nu$ be m linearly independent vectors with the components $S_{\mu\nu}$ in the original basis vectors $\boldsymbol{e}_\mu$,

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

$$\boldsymbol{e}'_\nu = \sum_\mu \boldsymbol{e}_\mu S_{\mu\nu}, \qquad \underline{e'_\nu} = \underline{S_{\cdot\nu}}. \tag{1.21}$$

Since $\boldsymbol{e}'_\nu$ are linearly independent, S is a nonsingular matrix ($\det S \neq 0$) and has its inverse matrix S^{-1}.

$$\boldsymbol{e}_\mu = \sum_\nu \boldsymbol{e}'_\nu \left(S^{-1}\right)_{\nu\mu}. \tag{1.22}$$

Choosing $\boldsymbol{e}'_\nu$ to be new basis vectors, the components a'_ν of the vector $\boldsymbol{a}$ and the matrix $\overline{D}(R)$ of the operator R can be calculated as follows:

$$\boldsymbol{a} = \sum_\mu \boldsymbol{e}_\mu a_\mu = \sum_{\nu\mu} \boldsymbol{e}'_\nu \left(S^{-1}\right)_{\nu\mu} a_\mu = \sum_\nu \boldsymbol{e}'_\nu a'_\nu,$$

$$a'_\nu = \sum_\mu \left(S^{-1}\right)_{\nu\mu} a_\mu, \qquad \underline{a}' = S^{-1}\,\underline{a}. \tag{1.23}$$

$$\begin{aligned} R\boldsymbol{e}'_\nu &= \sum_\rho \left(R\boldsymbol{e}_\rho\right) S_{\rho\nu} = \sum_{\mu\rho} \boldsymbol{e}_\mu D_{\mu\rho}(R) S_{\rho\nu}, \\ R\boldsymbol{e}'_\nu &= \sum_\rho \boldsymbol{e}'_\rho \overline{D}_{\rho\nu}(R) = \sum_{\mu\rho} \boldsymbol{e}_\mu S_{\mu\rho} \overline{D}_{\rho\nu}(R), \end{aligned}$$

$$\sum_\rho D_{\mu\rho}(R) S_{\rho\nu} = \sum_\rho S_{\mu\rho} \overline{D}_{\rho\nu}(R), \qquad \overline{D}(R) = S^{-1} D(R) S. \tag{1.24}$$

The relation (1.24) between $\overline{D}(R)$ and $D(R)$ is called a similarity transformation and S is the matrix of the similarity transformation. In literature, $\overline{D}(R)$ and $D(R)$ are said to be equivalent to each other if Eq. (1.24) holds. Obviously, the matrix S has the same matrix form with respect to both the original set and the new set of basis vectors. If the new set of basis vectors is the same as the original one except for the order of basis vectors, the similarity transformation is called the simple one. Note that, the similarity transformation for two equivalent matrices is not unique. If X is commutable with $D(R)$ and Y is commutable with $\overline{D}(R)$, both XS and SY satisfy the similarity transformation relation (1.24).

Since $\underline{S_{\cdot\nu}}$ is nothing but the column matrix of new basis vector $\boldsymbol{e}'_\nu$ in the original basis vectors, Eq. (1.24) can be written as

$$D(R)\ \underline{S_{\cdot\nu}} = \sum_\rho \underline{S_{\cdot\rho}}\ \overline{D}_{\rho\nu}(R), \qquad R\boldsymbol{e}'_\nu = \sum_\rho \boldsymbol{e}'_\rho \overline{D}_{\rho\nu}(R). \tag{1.25}$$

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

It is nothing but the definition of the matrix of R in the new basis vectors $\boldsymbol{e}'_\nu$. The different choices of the basis vectors do not change the action of an operator on a vector. If $\boldsymbol{b} = R\boldsymbol{a}$, one has $\underline{b} = D(R)\underline{a}$ in the original basis vectors $\boldsymbol{e}_\mu$. In the new set of basis vectors $\boldsymbol{e}'_\nu$ one has

$$\underline{b}' = S^{-1}\underline{b} = S^{-1}D(R)\underline{a} = S^{-1}D(R)S\underline{a}' = \overline{D}(R)\underline{a}'. \tag{1.26}$$

Let $\mathcal{L}$ be an m-dimensional space, $\mathcal{L}_1$ be its n-dimensional subspace, invariant to the operator R, and $\mathcal{L}_2$ be the complement of $\mathcal{L}_1$. Choose a new set of basis vectors in $\mathcal{L}$ such that the first n basis vectors belong to $\mathcal{L}_1$, and the next $(m-n)$ ones belong to $\mathcal{L}_2$. Arrange the new basis vectors $\boldsymbol{e}'_\nu$ to be the column matrices of S, $\underline{S_{.\nu}} = \underline{e'_\nu}$. Through the similarity transformation S the matrix $D(R)$ of R is changed to $\overline{D}(R)$. Since $\mathcal{L}_1$ is invariant to R, one has

$$R\boldsymbol{e}'_\mu = \sum_{\nu=1}^{n} \boldsymbol{e}'_\nu \overline{D}_{\nu\mu}(R), \qquad 1 \le \mu \le n. \tag{1.27}$$

Namely, the down-left corner of $\overline{D}(R)$ is vanishing,

$$\overline{D}_{\rho\mu}(R) = 0 \qquad \text{when} \ \ \mu \le n < \rho,$$

$$S^{-1}D(R)S = \overline{D}(R) = \begin{pmatrix} D^{(1)}(R) & M \\ 0 & D^{(2)}(R) \end{pmatrix}. \tag{1.28}$$

This matrix $\overline{D}(R)$ is called a ladder one. Furthermore, if $\mathcal{L}_2$ is also invariant to R, one has $M = 0$,

$$\overline{D}(R) = \begin{pmatrix} D^{(1)}(R) & 0 \\ 0 & D^{(2)}(R) \end{pmatrix} = D^{(1)}(R) \oplus D^{(2)}(R). \tag{1.29}$$

This matrix $\overline{D}(R)$ is a direct sum of two submatrices, and is called a block one in the type $[n, (m-n)]$. Generally, a matrix is also called a block one if it can be changed into the direct sum of submatrices by a simple similarity transformation. In order to determine whether or not a matrix is a block one, one may separate the indices of the matrix into two parts and check whether the matrix entries with indices respectively belonging to different parts all are vanishing. If Λ is diagonal and $X\Lambda = \Lambda X$, X is a block matrix.

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

1.4 Eigenvectors and Diagonalization of a Matrix

In quantum mechanics, the eigenequation of a physical operator $R(x)$ is

$$R(x)\psi(x) = \lambda\psi(x).$$

The eigenvalue λ is the possible observed value for the physical quantity. The eigenvalues describe the characteristic of the physical quantity and are independent of the choice of basis functions. In linear algebras, the eigenequation of an operator R is

$$R\,\boldsymbol{a} = \lambda\,\boldsymbol{a}. \tag{1.30}$$

If $\mathcal{L}$ is an m-dimensional space invariant to R and $\boldsymbol{a} \in \mathcal{L}$, one has

$$D(R)\ \underline{a} = \lambda\ \underline{a}, \qquad \sum_{\nu} D_{\mu\nu}(R)a_{\nu} = \lambda a_{\mu}. \tag{1.31}$$

The eigenequation (1.31) is a set of coupled linear homogeneous equations for m variables a_{ν}. The condition for the existence of nonvanishing solution is its coefficient determinant to be vanishing,

$$\det\left[D(R) - \lambda\mathbf{1}\right] = 0. \tag{1.32}$$

Equation (1.32) is called the secular equation for $D(R)$, and its roots are the eigenvalues of $D(R)$. The secular equation is invariant in similarity transformation, so that the eigenvalues are independent of the choice of the basis vectors. It is easy to see from Eq. (1.32) that the sum of the eigenvalues of $D(R)$ is its trace, Tr $D(R)$, and the product of the eigenvalues of $D(R)$ is its determinant, $\det D(R)$ (see Prob. 1).

Equation (1.31) means that the eigenvector generates a one-dimensional subspace which is invariant to R. If in $\mathcal{L}$ there exist m linearly independent eigenvectors $\boldsymbol{a}^{(\nu)}$ of R with the eigenvalues λ_{ν}, respectively, one may choose a new set of basis vectors $\boldsymbol{e}'_{\nu} = \boldsymbol{a}^{(\nu)}$. Then, the matrix of R in the new basis vectors $\boldsymbol{e}'_{\nu}$ is a diagonal one.

$$S^{-1}D(R)S = \Lambda, \qquad \underline{S_{\cdot\nu}} = \underline{a^{(\nu)}}, \qquad \Lambda_{\nu\mu} = \delta_{\nu\mu}\lambda_{\nu}. \tag{1.33}$$

Therefore, the key for diagonalizing an m-dimensional matrix is to find its m linearly independent eigenvectors. Since the eigenequation (1.30) is linearly homogeneous with respect to the eigenvector $\boldsymbol{a}^{(\nu)}$, $\boldsymbol{a}^{(\nu)}$ can be multiplied with a constant c_{ν}. When the eigenvalue is degenerate, its eigenvectors can be made a nonsingular linear combination. This is the reason why the similarity transformation matrix S is not unique. The number of the

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

arbitrary parameters in S is equal to $\sum_\nu n_\nu^2$, where n_ν is the multiplicity of the eigenvalue λ_ν. Those parameters play an important role in calculating a common similarity transformation for a few pairs of equivalent matrices, which is often used in group theory.

Substituting an eigenvalue into Eq. (1.31), one is always able to solve at least one eigenvector. The eigenvectors for different eigenvalues are linearly independent. However, when an eigenvalue is a root of Eq. (1.32) with multiplicity n, it is not certain to obtain n linearly independent eigenvectors with the given eigenvalue by solving Eq. (1.31). The following matrix is the simplest example which has the eigenvalue 1 with multiplicity 2, but only one linearly independent eigenvector,

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad b \neq 0. \tag{1.34}$$

It cannot be diagonalized by a similarity transformation.

The eigenvalues are invariant in a similarity transformation. Therefore, two equivalent matrices can be diagonalized into the same matrix,

$$X^{-1}D(R)X = \Lambda = Y^{-1}\overline{D}(R)Y,$$

such that the similarity transformation related to them is easy to be calculated,

$$\overline{D}(R) = \left(XY^{-1}\right)^{-1} D(R) \left(XY^{-1}\right). \tag{1.35}$$

It can be proved that the sufficient and necessary condition for a matrix $D(R)$ which can be diagonalized by a unitary similarity transformation is that $D(R)^\dagger$ is commutable with $D(R)$ (see Prob. 11). A matrix H is called Hermitian if $H^\dagger = H$. A matrix u is called unitary if $u^\dagger = u^{-1}$. Both a Hermitian matrix and a unitary matrix can be diagonalized by a unitary similarity transformation. A real unitary matrix is called real orthogonal. A real Hermitian matrix is called real symmetric. A real symmetric matrix can be diagonalized by a real orthogonal similarity transformation. Generally, a real orthogonal matrix can be diagonalized by a unitary similarity transformation, but not a real orthogonal one.

1.5 Inner Product of Vectors

There are three types of products of vectors, depending on the products to be a scalar, a vector, or a tensor. The product of two vectors is called the

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

inner product if the product is a scalar. In quantum mechanics the inner product of two wave functions is defined as

$$\langle \phi(x)|\psi(x)\rangle = \int (dx)\phi(x)^*\psi(x), \tag{1.36}$$

where the sign $\int$ denotes an integral for the continuous coordinate and a sum for the discrete coordinate. The form on the left-hand side of Eq. (1.36) is called the Dirac symbol. Generally, the inner product of two vectors $\langle \boldsymbol{a}|\boldsymbol{b}\rangle$ in the linear algebras satisfies

$$\begin{aligned}
&\langle c_1\boldsymbol{a}^{(1)} + c_2\boldsymbol{a}^{(2)}|\boldsymbol{b}\rangle = c_1^*\langle \boldsymbol{a}^{(1)}|\boldsymbol{b}\rangle + c_2^*\langle \boldsymbol{a}^{(2)}|\boldsymbol{b}\rangle, \\
&\langle \boldsymbol{a}|c_1\boldsymbol{b}^{(1)} + c_2\boldsymbol{b}^{(2)}\rangle = c_1\langle \boldsymbol{a}|\boldsymbol{b}^{(1)}\rangle + c_2\langle \boldsymbol{a}|\boldsymbol{b}^{(2)}\rangle, \\
&\langle \boldsymbol{b}|\boldsymbol{a}\rangle = \langle \boldsymbol{a}|\boldsymbol{b}\rangle^*, \qquad \langle \boldsymbol{a}|\boldsymbol{a}\rangle = |\boldsymbol{a}|^2 > 0, \ \text{ if } \boldsymbol{a} \neq 0.
\end{aligned} \tag{1.37}$$

The inner product is linear for the second vector and antilinear for the first vector. The inner product becomes its complex conjugate if changing the order of two factors. The self-inner product of a nonvanishing vector is real positive, called the square module of the vector. Denote by a Hermitian matrix Ω the inner product of two basis vectors

$$\langle \boldsymbol{e}_\mu|\boldsymbol{e}_\nu\rangle = \Omega_{\mu\nu}, \qquad \Omega_{\nu\mu} = \Omega_{\mu\nu}^* = \left(\Omega^\dagger\right)_{\nu\mu}. \tag{1.38}$$

Let $\boldsymbol{a}$ be an nonvanishing eigenvector of Ω with the eigenvalue λ,

$$\boldsymbol{a} = \sum_\nu \boldsymbol{e}_\nu a_\nu \neq 0, \qquad \sum_\nu \Omega_{\mu\nu}a_\nu = \lambda a_\mu,$$

$$\lambda \sum_\mu |a_\mu|^2 = \sum_{\mu\nu} a_\mu^*\Omega_{\mu\nu}a_\nu = \sum_{\mu\nu} a_\mu^*\langle \boldsymbol{e}_\mu|\boldsymbol{e}_\nu\rangle a_\nu = \langle \boldsymbol{a}|\boldsymbol{a}\rangle > 0. \tag{1.39}$$

Hence, the eigenvalue λ of Ω is positive, namely, Ω is positive definite. The inner product of two arbitrary vectors and the matrix entry of an operator can be calculated with Ω,

$$\boldsymbol{a} = \sum_\mu \boldsymbol{e}_\mu a_\mu, \qquad \boldsymbol{b} = \sum_\nu \boldsymbol{e}_\nu b_\nu, \qquad \langle \boldsymbol{a}|\boldsymbol{b}\rangle = \sum_{\mu\nu} a_\mu\Omega_{\mu\nu}b_\nu, \tag{1.40}$$

$$\sum_\rho \left(\Omega^{-1}\right)_{\mu\rho} \langle \boldsymbol{e}_\rho|R\boldsymbol{e}_\nu\rangle = \sum_{\rho\tau} \left(\Omega^{-1}\right)_{\mu\rho} \langle \boldsymbol{e}_\rho|\boldsymbol{e}_\tau\rangle D_{\tau\nu}(R) = D_{\mu\nu}(R). \tag{1.41}$$

A vector is called normalized if its module is one. Two vectors are orthogonal if their inner product is zero. Two nonvanishing orthogonal vectors must be linearly independent.

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

The basis vectors are called orthonormal if

$$\langle \boldsymbol{e}_\mu | \boldsymbol{e}_\nu \rangle = \Omega_{\mu\nu} = \delta_{\mu\nu}. \tag{1.42}$$

In the orthonormal basis vectors the formulas for the inner product become simpler,

$$\langle \boldsymbol{a} | \boldsymbol{b} \rangle = \sum_\mu a_\mu b_\mu, \qquad \langle \boldsymbol{e}_\mu | R \boldsymbol{e}_\nu \rangle = D_{\mu\nu}(R). \tag{1.43}$$

There are a few different definitions for the inner product of vectors. Another inner product is defined to be linear for both factors,

$$\begin{aligned} &\langle c_1 \boldsymbol{a}^{(1)} + c_2 \boldsymbol{a}^{(2)} | \boldsymbol{b} \rangle = c_1 \langle \boldsymbol{a}^{(1)} | \boldsymbol{b} \rangle + c_2 \langle \boldsymbol{a}^{(2)} | \boldsymbol{b} \rangle, \\ &\langle \boldsymbol{a} | c_1 \boldsymbol{b}^{(1)} + c_2 \boldsymbol{b}^{(2)} \rangle = c_1 \langle \boldsymbol{a} | \boldsymbol{b}^{(1)} \rangle + c_2 \langle \boldsymbol{a} | \boldsymbol{b}^{(2)} \rangle, \\ &\langle \boldsymbol{e}_\mu | \boldsymbol{e}_\nu \rangle = \Omega_{\mu\nu} = \Omega_{\nu\mu}, \qquad \det \Omega \neq 0. \end{aligned} \tag{1.44}$$

In this definition, the self-inner product of a vector may not be real.

The inner product of the column matrices has been defined, namely

$$\underline{a}^\dagger \ \underline{b} = \sum_\mu a_\mu^* b_\mu, \qquad \underline{a}^T \ \underline{b} = \sum_\mu a_\mu b_\mu. \tag{1.45}$$

In comparison with Eq. (1.40) the inner product (1.45) means that the basis vectors for the column matrices are orthonormal.

At last, we discuss the concept of the adjoint operator $R^\dagger$ of an operator R. The conjugate matrix $D(R)^\dagger$ of a matrix $D(R)$ was well defined,

$$\left[D(R)^\dagger \right]_{\mu\nu} = D_{\nu\mu}(R)^*, \tag{1.46}$$

$$\left[\underline{a}^\dagger \ D(R) \ \underline{b} \right]^* = \left[D(R) \underline{b} \right]^\dagger \ \underline{a} = \underline{b}^\dagger \ D^\dagger(R) \ \underline{a}. \tag{1.47}$$

The definition for the adjoint operator in quantum mechanics is the generalization of Eq. (1.47),

$$\langle \boldsymbol{a} | R \boldsymbol{b} \rangle^* = \langle R \boldsymbol{b} | \boldsymbol{a} \rangle = \langle \boldsymbol{b} | R^\dagger \boldsymbol{a} \rangle. \tag{1.48}$$

The adjoint relation between two operators is mutual. Note that the matrices of two adjoint operators are not necessary to be conjugate if the basis vectors are not orthonormal. Denote by $D(R)$ and X the matrices of two operators R and $R^\dagger$, respectively

$$R \boldsymbol{e}_\mu = \sum_\rho \boldsymbol{e}_\rho D_{\rho\mu}(R), \qquad R^\dagger \boldsymbol{e}_\nu = \sum_\rho \boldsymbol{e}_\rho X_{\rho\nu},$$

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

$$\begin{gathered}\sum_{\rho} D^*_{\rho\mu}(R)\langle \boldsymbol{e}_\rho|\boldsymbol{e}_\nu\rangle = \langle R\boldsymbol{e}_\mu|\boldsymbol{e}_\nu\rangle = \langle \boldsymbol{e}_\mu|R^\dagger\boldsymbol{e}_\nu\rangle = \sum_{\rho} \langle \boldsymbol{e}_\mu|\boldsymbol{e}_\rho\rangle X_{\rho\nu}, \\ D^\dagger(R)\Omega = \Omega X, \qquad X = \Omega^{-1}D^\dagger(R)\Omega.\end{gathered} \tag{1.49}$$

Conversely, if the basis vectors are orthonormal, two mutual conjugate matrices correspond to two operators adjoint to each other. Namely, in this case, a Hermitian (or an unitary) matrix corresponds to a Hermitian (or a unitary) operator.

1.6 The Direct Product of Matrices

If a quantum system consists of two subsystems, the wave function of the system is expressed as the product of two wave functions of the subsystems, or the combination of the products. Suppose that the two functional spaces $\mathcal{L}_1$ and $\mathcal{L}_2$ for two subsystems are respectively invariant to the operator R,

$$R\psi_\mu = \sum_{\nu=1}^{m} \psi_\nu D^{(1)}_{\nu\mu}(R), \qquad R\phi_i = \sum_{j=1}^{n} \phi_j D^{(2)}_{ji}(R). \tag{1.50}$$

For the composed system, the functional space $\mathcal{L}$ generated by

$$\psi_\mu\phi_i, \qquad 1 \leq \mu \leq m, \ \ 1 \leq i \leq n \tag{1.51}$$

is called the product of two subspaces, $\mathcal{L} = \mathcal{L}_1\mathcal{L}_2$, which is (mn)-dimensional. The space $\mathcal{L}$ is also invariant to R, i.e.,

$$\begin{aligned} R\left(\psi_\mu\phi_i\right) &= \sum_{\nu\ j} \left(\psi_\nu\phi_j\right)\left[D^{(1)}(R) \times D^{(2)}(R)\right]_{\nu j,\mu i}, \\ \left[D^{(1)}(R) \times D^{(2)}(R)\right]_{\nu j,\mu i} &= D^{(1)}_{\nu\mu}(R)D^{(2)}_{ji}(R). \end{aligned} \tag{1.52}$$

The matrix $D^{(1)}(R) \times D^{(2)}(R)$ of R in $\mathcal{L}$, which is (nm)-dimensional, is called the direct product of two submatrices $D^{(1)}(R)$ and $D^{(2)}(R)$. The row (column) of the direct product matrix is denoted by two indices μ and i. The order of indices is usually arranged such that the second index i increases for a given μ, and then the first index μ increases. For example, the direct product of two 2-dimensional matrices X and Y is

$$X \times Y = \begin{pmatrix} X_{11}Y & X_{12}Y \\ X_{21}Y & X_{22}Y \end{pmatrix} = \begin{pmatrix} X_{11}Y_{11} & X_{11}Y_{12} & X_{12}Y_{11} & X_{12}Y_{12} \\ X_{11}Y_{21} & X_{11}Y_{22} & X_{12}Y_{21} & X_{12}Y_{22} \\ X_{21}Y_{11} & X_{21}Y_{12} & X_{22}Y_{11} & X_{22}Y_{12} \\ X_{21}Y_{21} & X_{21}Y_{22} & X_{22}Y_{21} & X_{22}Y_{22} \end{pmatrix}.$$

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

The product of two matrices $D^{(1)}(R) \times D^{(2)}(R)$ and $D^{(1)}(S) \times D^{(2)}(S)$ is

$$\begin{aligned} &\left[D^{(1)}(R) \times D^{(2)}(R)\right]\left[D^{(1)}(S) \times D^{(2)}(S)\right] \\ &\quad = \left[D^{(1)}(R) D^{(1)}(S)\right] \times \left[D^{(2)}(R) D^{(2)}(S)\right]. \end{aligned} \tag{1.53}$$

Thus,

$$\begin{aligned} &\left[D^{(1)}(R) \times D^{(2)}(R)\right]^{-1} = D^{(1)}(R)^{-1} \times D^{(2)}(R)^{-1}, \\ &\left[D^{(1)}(R) \times D^{(2)}(R)\right]^{T} = D^{(1)}(R)^{T} \times D^{(2)}(R)^{T}, \\ &\left[D^{(1)}(R) \times D^{(2)}(R)\right]^{\dagger} = D^{(1)}(R)^{\dagger} \times D^{(2)}(R)^{\dagger}. \end{aligned} \tag{1.54}$$

The trace and the determinant of direct product $D^{(1)}(R) \times D^{(2)}(R)$ are

$$\begin{aligned} &\text{Tr } \left[D^{(1)}(R) \times D^{(2)}(R)\right] = \left[\text{Tr } D^{(1)}(R)\right]\left[\text{Tr } D^{(2)}(R)\right], \\ &\det\left[D^{(1)}(R) \times D^{(2)}(R)\right] = \left[\det\ D^{(1)}(R)\right]^{n}\left[\det\ D^{(2)}(R)\right]^{m}. \end{aligned} \tag{1.55}$$

If two matrices $D^{(1)}(R)$ and $D^{(2)}(R)$ depend on a continuous parameter α, one has

$$\begin{aligned} &\frac{d}{d\alpha}\left[D^{(1)}(R) \times D^{(2)}(R)\right] \\ &\quad = \left[\frac{dD^{(1)}(R)}{d\alpha}\right] \times D^{(2)}(R) + D^{(1)}(R) \times \left[\frac{dD^{(2)}(R)}{d\alpha}\right]. \end{aligned} \tag{1.56}$$

The direct product reduces to the product of a number and a matrix if one of the two factor matrices is one-dimensional. Generally, $D^{(1)}(R) \times D^{(2)}(R)$ is not equal to $D^{(2)}(R) \times D^{(1)}(R)$, but their difference is only a simple similarity transformation. For example, when $n = m = 2$, the similarity transformation matrix for the two direct products is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.57}$$

1.7 Exercises

1. Prove that the sum of the eigenvalues of a matrix is equal to the trace of the matrix, and the product of eigenvalues is equal to the determinant of the matrix.

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

2. Calculate the eigenvalues and eigenvectors of the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

3. Calculate the eigenvalues and eigenvectors of the matrix

$$R = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

4. Calculate the eigenvalues and eigenvectors of the matrix

$$R = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

5. If $\det R \neq 0$, prove that both $R^\dagger R$ and $RR^\dagger$ are positive definite Hermitian matrices.

6. Prove: (1) if $R^\dagger R = \mathbf{1}$, then $RR^\dagger = \mathbf{1}$;

(2) if $R^{-1}R = \mathbf{1}$, then $RR^{-1} = \mathbf{1}$;

(3) if $R^T R = \mathbf{1}$, then $RR^T = \mathbf{1}$.

7. Find the independent real parameters in a 2×2 unitary matrix, a real orthogonal matrix, and a Hermitian matrix, respectively, and give their general expressions.

8. Find the similarity transformation to diagonalize the following matrices:

$$(1) \quad \begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}, \qquad (2) \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

9. Find a similarity transformation matrix M which satisfies

$$M^{-1} \begin{pmatrix} 0 & -\cos\theta & \sin\theta\sin\varphi \\ \cos\theta & 0 & -\sin\theta\cos\varphi \\ -\sin\theta\sin\varphi & \sin\theta\cos\varphi & 0 \end{pmatrix} M = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

10. Find a similarity transformation matrix M which satisfies the following three equations simultaneously

$$M^{-1}\begin{pmatrix}0 & -i & 0\\ i & 0 & 0\\ 0 & 0 & 0\end{pmatrix}M=\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{pmatrix},$$

$$M^{-1}\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & -i\\ 0 & i & 0\end{pmatrix}M=\frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix},$$

$$M^{-1}\begin{pmatrix}0 & 0 & i\\ 0 & 0 & 0\\ -i & 0 & 0\end{pmatrix}M=\frac{i}{\sqrt{2}}\begin{pmatrix}0 & -1 & 0\\ 1 & 0 & -1\\ 0 & 1 & 0\end{pmatrix}.$$

11. Let

$$R=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\qquad S=\frac{1}{2}\begin{pmatrix}-1 & -\sqrt{3}\\ \sqrt{3} & -1\end{pmatrix}.$$

Find the common similarity transformation matrix X satisfying

$$X^{-1}(R\times R)X=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix},$$

$$X^{-1}(S\times S)X=\frac{1}{2}\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & -1 & -\sqrt{3}\\ 0 & 0 & \sqrt{3} & -1\end{pmatrix}.$$

12. Find the similarity transformation matrix X to diagonalize the following three matrices simultaneously,

$$\begin{pmatrix}0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\\1&0&0&0&0&0\\0&1&0&0&0&0\\0&0&1&0&0&0\end{pmatrix},\quad \begin{pmatrix}0&0&0&1&0&0\\0&0&0&0&0&1\\0&0&0&0&1&0\\1&0&0&0&0&0\\0&0&1&0&0&0\\0&1&0&0&0&0\end{pmatrix},\quad \begin{pmatrix}0&0&0&1&1&1\\0&0&0&1&1&1\\0&0&0&1&1&1\\1&1&1&0&0&0\\1&1&1&0&0&0\\1&1&1&0&0&0\end{pmatrix}.$$

13. Show the general form of an $m\times m$ matrix, both unitary and Hermitian.

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html

14. Prove that any unitary matrix R can be diagonalized by a unitary similarity transformation, and any Hermitian matrix R can be diagonalized by a unitary similarity transformation.

15. Prove that R and $R^\dagger$ can be diagonalized by a common unitary similarity transformation if $R^\dagger$ is commutable with R. Further prove that the necessary and sufficient condition for a matrix R to be diagonalized by a unitary similarity transformation is that $R^\dagger$ is commutable with R.

16. Prove that any matrix can be transformed into a direct sum of the standard Jordan forms, each of which is in the form

$$R_{ab} = \begin{cases} \lambda & \text{when} \quad a = b, \\ 0 \quad \text{or} \quad 1 & \text{when} \quad a + 1 = b, \\ 0 & \text{the remaining cases.} \end{cases}$$

GROUP THEORY FOR PHYSICISTS
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/6596.html