# SYMMETRIES

This part of the course is based on refs. [1,2,3,4].

## 1. Groups and Representations

A *group* is a set $G$ on which a multiplication operation $\cdot$ is defined with the following properties:

$(i)$ If $x$ and $y$ are in $G$, $x \cdot y$ is in $G$;

$(ii)$ There is an identity element $e$ in $G$ such that $e \cdot x = x \cdot e = x$ for any $x$ in $G$;

$(iii)$ For every $x$ in $G$, there is an inverse element in $G$ called $x^{-1}$ such that $x \cdot x^{-1} = x^{-1}x = e$;

$(iv)$ For every $x$, $y$ and $z$ in $G$, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

For example, the integers form a group under addition, that is the "multiplication" law for $x \cdot y$ is defined as $(x + y)$. This is called the additive group of integers.

Another example is the possible permutations of three objects. Suppose we have three things, labeled 1, 2 and 3. The elements of the permutation group are the six ways we can reshuffle the objects:

$(\ )$, do nothing, so that $(a, b, c)$ remains $(a, b, c)$;

$(12)$, interchange the objects in positions 1 and 2, so $(a, b, c) \rightarrow (b, a, c)$;

$(23)$, $(a, b, c) \rightarrow (a, c, b)$;

$(13)$, $(a, b, c) \rightarrow (c, b, a)$;

$(123)$, take the object in 1 and put it in 2, etc. $(a, b, c) \rightarrow (c, a, b)$ (this is called a cyclic permutation);

$(321)$, $(a, b, c) \rightarrow (b, c, a)$.

There is a natural multiplication law for such transformations of a physical system:

$(v)$ $A \cdot B$ is the transformation obtained by first making the transformation $B$, then making the transformation $A$.

For example,

$$(23) \cdot (12) = (321)$$

because

$$(12) \text{ takes } (a, b, c) \rightarrow (b, a, c),$$

$$(23) \text{ takes } (b, a, c) \rightarrow (b, c, a).$$

Thus

$$(23) \cdot (12) \text{ takes } (a, b, c) \rightarrow (b, c, a).$$

The permutation group is an example of a *transformation group* on a physical system. It is clear that the reversible transformations on a system form a group under the multiplication law $(v)$. Doing nothing is the identity operator. Undoing a transformation is the inverse transformation. Associativity, $(iv)$, follows because at each step in the product $x \cdot y \cdot z$, there is a definite physical state of the system. This class of groups is very important because it includes any possible symmetry of a physical system.

In quantum mechanics, a transformation of the system is associated with a unitary transformation in the Hilbert space (below we discuss this point in more detail and mention the possibility of antiunitary operators which we ignore in this section). Thus, a transformation of a quantum mechanical system is associated with a mapping of the group into a set of unitary operators. For each $x$ in $G$ there is a $D(x)$ which is a unitary operator. Furthermore, the mapping must preserve the multiplication law (to make physical sense). Thus, for all $x$ and $y$ in $G$,

$$D(x)D(y) = D(x \cdot y). \tag{1.1}$$

A mapping which satisfies (1.1) is called a *representation* of the group $G$. In fact, a representation can involve nonunitary linear operators as long as they satisfy the multiplication law (1.1).

For example, the mapping

$$D(n) = e^{in\theta}$$

is a representation of the additive group of the integers, because

$$e^{in\theta}e^{im\theta} = e^{i(n+m)\theta}.$$

The following mapping is a representation of the permutation group on three objects:

$$\begin{aligned}
D(\ \ ) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad & D(12) &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
D(13) &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad & D(23) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
D(123) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad & D(321) &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.
\end{aligned} \tag{1.2}$$

The multiplication law $(12)(23) = (321)$ is mapped into the matrix multiplication

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

In other words, a group is a multiplication table satisfying $(i)-(iv)$. A representation is a realization of the multiplication by (finite or infinite dimensional) matrices. Group theory makes it possible to determine many properties of any representation from the abstract proprties of the group.

Notice that the definition $(i)-(iv)$ does not require commutativity $(x \cdot y = y \cdot x)$. In fact, while the additive group of integers is commutative, the permutation group is not. A commutative group is called *Abelian*, and a non-commutative group is non-Abelian. The permutations are non-Abelian, since for example $(23)(12) = (123)$.

We will find it convenient to view representations as abstract linear operators and as matrices. The connection is as follows: let $|i\rangle$ be an orthonormal basis in the space on which $D(g)$ acts as a linear operator. Then

$$[D(g)]_{ij} = \langle i|D(g)|j\rangle. \tag{1.3}$$

So

$$D(g)\,|i\rangle = \sum_j |j\rangle\,\langle j|D(g)|i\rangle = |j\rangle\,[D(g)]_{ji}. \tag{1.4}$$

From now on, we will freely translate from one language to the other.

Two representations $D_1$ and $D_2$ are *equivalent* if they are related by a *similarity transformation*,

$$D_2(x) = SD_1(x)S^{-1} \tag{1.5}$$

with a fixed operator $S$ for all $x$ in the group $G$.

A representation $D$ is *reducible* if it is equivalent to a representation $D'$ with a *block-diagonal* form:

$$D'(x) = SD(x)S^{-1} = \begin{pmatrix} D'_1(x) & 0 \\ 0 & D'_2(x) \end{pmatrix}. \tag{1.6}$$

The vector space on which $D'$ acts breaks up into two orthogonal subspaces, each of which is mapped into itself by all the operators $D'(x)$. The representation $D'$ is said to be the *direct sum* of $D'_1$ and $D'_2$:

$$D' = D'_1 \oplus D'_2. \tag{1.7}$$

A representation is *irreducible* if it is not reducible, that is if it cannot be put into block diagonal form by a similarity transformation.

## 2. Lie Groups and Lie Algebras

Compact Lie groups are groups of unitary operators in which the group elements are labeled by a set of continuous parameters with a multiplication law that depends smoothly on the parameters. The term "compact" refers to a global property of the group which we will not discuss in detail. In a certain sense, the volume of the parameter space for a compact group is finite.

Any representation of a compact Lie group is equivalent to a representation by unitary operators. Any group element which can be obtained from the identity by continuous changes in the parameters can be written as

$$\exp(i\alpha_a X_a) \tag{2.1}$$

where $\alpha_a$ $(a = 1, \ldots, N)$ are real parameters and $X_a$ are linearly independent Hermitian operators, and a sum over the repeated index $a$ is implied. The set of all linear combination $\alpha_a X_a$ is a vector space, and $X_a$ are a basis in the space. We will use the term *group generator* to refer (interchangably) to an arbitrary element of the vector space or specifically to the basis vectors $X_a$.

There are several spaces that are relevant here. Do not confuse the space of the group generators (which is $N$ dimensional because $a = 1, \ldots, N$) with the space on which the generators act, which is some yet unspecified Hilbert space. In fact, for the compact Lie groups, we can always take the space on which the generators act to be finite dimensional, so that you can think of the $X_a$ as finite hermitian matrices.

There are two nice things about the generators. The first is simply that they form a vector space. Unlike group elements themselves, the generators can be multiplied by numbers and added to obtain other generators. The second is that they satisfy simple commutation relations which determine (almost) the full structure of the group. Consider the product

$$\exp(i\lambda X_b)\exp(i\lambda X_a)\exp(-i\lambda X_b)\exp(-i\lambda X_a) = 1 + \lambda^2[X_a, X_b] + \cdots. \tag{2.2}$$

Because of the group property, the product of group elements ia another group element and can be written as $\exp(i\beta_c X_c)$. As $\lambda \to 0$, we must have

$$\lambda^2[X_a, X_b] \to i\beta_c X_c, \quad \beta_c = \lambda^2 f_{abc},$$

so

$$[X_a, X_b] = i f_{abc} X_c. \tag{2.3}$$

The constants $f_{abc}$ are called the *structure constants* of the group. Clearly they are determined by the group multiplication law (at least for elements continuously connected to the identity) as follows:

$$\exp(i\alpha_a X_a)\exp(i\beta_b X_b) = \exp(i\delta_c X_c) \tag{2.4}$$

where $\delta_c$ is determined by $\alpha$, $\beta$ and $f$:

$$\delta_c = \alpha_c + \beta_c - \frac{1}{2} f_{abc}\alpha_a\beta_b + \cdots .$$

This identity follows in a straightforward way from the definition of $\exp(A)$. We can obtain $\delta_a$ to any desired order in $\alpha_a$ and $\beta_a$ just by expanding both sides of (2.4) and comparing.

The generators also satisfy the following identity:

$$[X_a, [X_b, X_c]] + \text{ cyclic permutations } = 0. \tag{2.5}$$

This is called the Jacobi identity. It is obvious for a representation since then the $X_a$ are just linear operators, but in fact it is true for the abstract group generators. In terms of the structure constants, (2.5) becomes

$$f_{bcd} f_{ade} + f_{abd} f_{cde} + f_{cad} f_{bde} = 0. \tag{2.6}$$

If we define a set of matrices $T_a$,

$$(T_a)_{bc} = -i f_{abc}, \tag{2.7}$$

then (2.5) and (2.6) can be rewritten as

$$[T_a, T_b] = i f_{abc} T_c. \tag{2.8}$$

In other words, the structure constants themselves generate a representation of the algebra. This representation is called the *adjoint representation.* The *dimension* of a representation is the dimension of the vector space on which it acts. The dimension of the adjoint representation is just the number of generators, which is the number or real parameters necessary to describe a group element.

The generators and the commutation relations define the *Lie algebra* associated with the Lie group. Clearly, every representation of the group defines a representation of the algebra. The generators in the representation, when exponentiated, give the operators of the group representation.

The definitions of equivalence, reducibility and irreducibility can be transferred from the group to the algebra with no change.

The structure constants depend on what basis we choose in the vector space of the generators. To choose an appropriate basis, we use the adjoint representation, (2.7). Consider the trace $\mathrm{tr}(T_a T_b)$. This is a real symmetric matrix, so we can diagonalize it by choosing appropriate real linear combinations of the $X_a$'s and therefore of the $T_a$'s. Suppose we have done this, so that

$$\mathrm{tr}(T_a T_b) = k_a \delta_{ab} \quad \text{(no sum)}. \tag{2.9}$$

We still have the freedom to rescale the generators, so for example, we could choose all the nonzero $k_a$'s to have absolute value 1. But, we cannot change the sign of the $k_a$'s.

We will be concerned almost entirely with algebras in which all the $k_a$'s are positive. For these algebras, which are associated with the *compact semisimple* Lie groups, we can take

$$\mathrm{tr}(T_a T_b) = \lambda \delta_{ab} \tag{2.10}$$

for any convenient positive $\lambda$. In this basis, the structure constants are completely antisymmetric, because we can write $f_{abc} = -i\lambda^{-1}\mathrm{tr}([T_a, T_b]T_c)$ which is completely antisymmetric because of the cyclic property of the trace.

Notice that in this basis, the generators in the adjoint representation are hermitian matrices. In fact, one can show that for the compact Lie groups (as for finite groups) any representation is equivalent to a representation by hermitian operators and all irreducible representations are finite hermitian matrices.

A few more definitions may help the student to appreciate when the structure constants can be useful. An *invariant subalgebra* is some set of generators which goes into itself (or zero) under commutation with any element of the algebra. That is if $X$ is any generator in the invariant subalgebra and $Y$ is any generator in the whole algebra, $[X, Y]$ is a generator in the invariant subalgebra (or it is zero).

An algebra which has no nontrivial invariant subalgebra (the whole algebra and the set of no generators at all are always trivial invariant subalgebras) is called *simple*. A simple algebra generates a simple group.

Particularly annoying are Abelian invariant subalgebras. The generators in an Abelian invariant subalgebra commute with everything. Each of these generators is associated with what we call a $U(1)$ factor of the group. $U(1)$ is the group of phase transformations. At any rate, $U(1)$ factors do not show up in the structure constants at all. If $X_a$ is a $U(1)$ generator, $f_{abc} = 0$ for all $b$ and $c$ and $k_a = 0$. The structure constants do not tell us anything about the Abelian invariant subalgebras.

Algebras without Abelian invariant subalgebras are called *semisimple*. They are built by putting simple algebras together. In these algebras, every generator has a nonzero commutator with some other generator. In such a case the structure constants carry a great deal of information. We can use them to determine the entire structure of the algebra and its representations. From here on, we will be discussing semisimple algebras, and we will deal with representations by unitary operators.

The generators (like the linear operators of the representations they generate) can be thought of as either linear operators or matrices via

$$X_a \left|i\right\rangle = \left|j\right\rangle \left\langle j\right|X_a\left|i\right\rangle = \left|j\right\rangle [X_a]_{ji}. \tag{2.11}$$

Note that in a matrix language, the states form row vectors and the matrix representing a linear operator acts on the right.

In the Hilbert space on which a representation acts, the group elements can be thought of as a transformation on the states. The group element $\exp(i\alpha_a X_a)$ maps or "transforms" the kets as follows:

$$\left|i\right\rangle \to \exp(i\alpha_a X_a)\left|i\right\rangle . \tag{2.12}$$

Then the corresponding bras transform as

$$\left\langle i\right| \to \left\langle i\right| \exp(-i\alpha_a X_a). \tag{2.13}$$

The ket obtained by acting on $\left|i\right\rangle$ with an operaor $O$ must also transform as in (2.12). This implies that any operator $O$ transforms as follows:

$$O \to \exp(i\alpha_a X_a) O \exp(-i\alpha_a X_a). \tag{2.14}$$

Eqs. (2.12), (2.13) and (2.14) imply that all matrix elements are unchanged under the transformation.

To obtain the action of the algebra on these objects, we differentiate with respect to $\alpha$ and divide by $i$. Thus, corresponding to the action of the generator $X_a$ on a ket,

$$X_a \left| i \right\rangle ,$$

is $-X_a$ acting on a bra,

$$- \left\langle i \right| X_a ,$$

and the commutator of $X_a$ with an operator,

$$[X_a, O].$$

Then the invariance of a matrix element $\langle i|O|j\rangle$ is expressed by the statement

$$\langle i|O(X_a|j\rangle) + \langle i|[X_a, O]|j\rangle + (-\langle i|X_a)O|j\rangle = 0.$$

## 3. Quantum Mechanics and Symmetries

First, some good news: quantum field theory is based on the same quantum mechanics that was invented by Schrödinger, Heisenberg, Pauli, Born and others in 1925-6, and has been used ever since in atomic, molecular, nuclear and condensed matter physics. The student is assumed to be familiar with quantum mechanics; this section provides only the briefest of summaries of quantum mechanics, in the generalized version of Dirac.

(i) Physical states are represented by rays in Hilbert space. A Hilbert space is a kind of complex vector space; that is, if $\Phi$ and $\Psi$ are vectors in the space (often called 'state-vectors') then so is $\xi\Phi + \eta\Psi$, for arbitrary complex numbers $\xi, \eta$. It has a norm; for any pair of vectors there is a complex number $(\Phi, \Psi)$ such that

$$(\Phi, \Psi) = (\Psi, \Phi)^*, \tag{3.1}$$

$$(\Phi, \xi_1\Psi_1 + \xi_2\Psi_2) = \xi_1(\Phi, \Psi_1) + \xi_2(\Phi, \Psi_2), \tag{3.2}$$

$$(\eta_1\Phi_1 + \eta_2\Phi_2, \Psi) = \eta_1^*(\Phi_1, \Psi) + \eta_2^*(\Phi_2, \Psi). \tag{3.3}$$

The norm $(\Psi, \Psi)$ also satisfies a positivity condition: $(\Psi, \Psi) \geq 0$, and vanishes if and only if $\Psi = 0$. (There are also certain technical assumptions that allow us to take limits of vectors within Hilbert space.) A *ray* is a set of normalized vectors (i.e. $(\Psi, \Psi) = 1$) with

$\Psi$ and $\Psi'$ belonging to the same ray if $\Psi' = \xi\Psi$, where $\xi$ is an arbitrary complex number with $|\xi| = 1$.

(ii) Observables are represented by Hermitian operators. These are mappings $\Psi \to A\Psi$ of Hilbert space into itself, linear in the sense that

$$A(\xi\Psi + \eta\Phi) = \xi A\Psi + \eta A\Phi \tag{3.4}$$

and satisfying the reality condition $A^\dagger = A$, where for any linear operator $A$ the adjoint $A^\dagger$ is defined by

$$(\Phi, A^\dagger\Psi) \equiv (A\Phi, \Psi) = (\Psi, A\Phi)^*. \tag{3.5}$$

(There are also technical assumptions about the continuity of $A\Psi$ as a function of $\Psi$.) A state represented by the ray $\mathcal{R}$ has a definite value $\alpha$ for the observable represented by an operator $A$ if vectors $\Psi$ belonging to this ray are eigenvectors of $A$ with eigenvalue $\alpha$:

$$A\Psi = \alpha\Psi \quad \text{for} \quad \Psi \text{ in } \mathcal{R}. \tag{3.6}$$

An elementary theorem tells us that for $A$ hermitian, $\alpha$ is real, and eigenvectors with different $\alpha$'s are orthogonal.

(iii) If a system is in a state represented by a ray $\mathcal{R}$, and an experiment is done to test whether it is in any one of the different states represented by the mutually orthogonal rays $\mathcal{R}_1, \mathcal{R}_2, \ldots$ (for instance, by measuring one or more observables) then the probability of finding it in the state represented by $\mathcal{R}_n$ is

$$P(\mathcal{R} \to \mathcal{R}_n) = |(\Psi, \Psi_n)|^2, \tag{3.7}$$

where $\Psi$ and $\Psi_n$ are any vectors belonging to the rays $\mathcal{R}$ and $\mathcal{R}_n$, respectively. (A pair of rays is said to be orthogonal if the state vectors from the two rays have vanishing scalar products.) Another elementary theorem gives a total probability unity:

$$\sum_n P(\mathcal{R} \to \mathcal{R}_n) = 1 \tag{3.8}$$

if the state vectors $\Psi_n$ form a complete set.

A symmetry transformation is a change in our point of view that does not change the results of possible experiments. If an observer $O$ sees a system in a state represented by a ray $\mathcal{R}$ or $\mathcal{R}_1$ or $\mathcal{R}_2 \ldots$, then an equivalent observer $O'$ who looks at the *same* system will observe it in a different state, represented by a ray $\mathcal{R}'$ or $\mathcal{R}'_1$ or $\mathcal{R}'_2 \ldots$, respectively, but the two observers must find the same probabilities

$$P(\mathcal{R} \to \mathcal{R}_n) = P(\mathcal{R}' \to \mathcal{R}'_n). \tag{3.9}$$

(This is only a necessary condition for a ray transformation to be a symmetry.) An important theorem proved by Wigner in the early 1930s tells us that for any such transformation $\mathcal{R} \to \mathcal{R}'$ of rays we may define an operator $U$ on Hilbert space, such that if $\Psi$ is in ray $\mathcal{R}$ then $U\Psi$ is in the ray $\mathcal{R}'$, with $U$ either *unitary* and *linear*

$$(U\Phi, U\Psi) = (\Phi, \Psi), \tag{3.10}$$

$$U(\xi\Phi + \eta\Psi) = \xi U\Phi + \eta U\Psi \tag{3.11}$$

or else *antiunitary* and *antilinear*

$$(U\Phi, U\Psi) = (\Phi, \Psi)^*, \tag{3.12}$$

$$U(\xi\Phi + \eta\Psi) = \xi^* U\Phi + \eta^* U\Psi. \tag{3.13}$$

As already mentioned, the adjoint of a linear operator $L$ is defined by

$$(\Phi, L^\dagger\Psi) \equiv (L\Phi, \Psi). \tag{3.14}$$

This condition cannot be satisfied for an antilinear operator $A$, because in this case the right-hand side of (3.14) would be linear in $\Phi$,

$$(A\xi\Phi, \Psi) = \xi(A\Phi, \Psi), \tag{3.15}$$

while the left-hand side of (3.14) is antilinear in $\Phi$,

$$(\xi\Phi, A^\dagger\Psi) = \xi^*(\Phi, A^\dagger\Psi). \tag{3.16}$$

Instead, the adjoint of an antilinear operator $A$ is defined by

$$(\Phi, A^\dagger\Psi) \equiv (A\Phi, \Psi)^* = (\Psi, A\Phi). \tag{3.17}$$

With this definitions, unitarity and linearity imply

$$(U\Phi, U\Psi) = (U^\dagger U\Phi, \Psi) = (\Phi, \Psi), \tag{3.18}$$

while antiunitarity and antilinearity imply

$$(U\Phi, U\Psi) = (U^\dagger U\Phi, \Psi)^* = (\Phi, \Psi)^*, \tag{3.19}$$

so the conditions for unitarity or antiunitarity both take the form

$$U^\dagger = U^{-1}. \tag{3.20}$$

There is always a trivial symmetry transformation $\mathcal{R} \to \mathcal{R}$, represented by the identity operator $U = 1$. This operator is, of course, unitary and linear. Continuity then demands that any symmetry (like a rotation or translation or Lorentz transformation) that can be made trivial by a continuous change of some parameters (like angles or distances or velocities) must be represented by a linear unitary operator $U$ rather than one that is antilinear and antiunitary. (Symmetries represented by antiunitary antilinear operators are less prominent in physics; they all involve a reversal in the direction of time's flow.)

In particular, a symmetry transformation that is infinitesimally close to being trivial can be represented by a linear unitary operator that is infinitesimally close to the identity:

$$U = 1 + i\epsilon t \tag{3.21}$$

with $\epsilon$ a real infinitesimal. For this to be unitary and linear, $t$ must be Hermitian and linear, so it is a candidate for an observable. Indeed, most (and perhaps all) of the observables of physics, such as angular momentum or momentum, arise in this way from symmetry transformations.

The set of symmetry transformations has certain properties that define it as a *group*. If $T_1$ is a transformation that takes rays $\mathcal{R}_n$ into $\mathcal{R}'_n$, and $T_2$ is another symmetry transformation that takes $\mathcal{R}'_n$ into $\mathcal{R}''_n$, then the result of performing both symmetry transformations is another symmetry transformation, which we write $T_2T_1$, that takes $\mathcal{R}_n$ into $\mathcal{R}''_n$. Also, a symmetry transformation $T$, that takes rays $\mathcal{R}_n$ into $\mathcal{R}'_n$, has an inverse, written $T^{-1}$, which takes $\mathcal{R}'_n$ into $\mathcal{R}_n$, and there is an identity transformation, $T = 1$, which leaves rays unchanged.

The unitary or antiunitary operators $U(T)$ corresponding to these symmetry transformations have properties that mirror this group structure, but with complication due to the fact that, unlike the symmetry transformations themselves, the operators $U(T)$ act on vectors in the Hilbert space rather than on rays. If $T_1$ takes $\mathcal{R}_n$ into $\mathcal{R}'_n$, then acting on vector $\Psi_n$ in the ray $\mathcal{R}_n$, $U(T_1)$ must yield a vector $U(T_1)\Psi_n$ in the ray $\mathcal{R}'_n$, and if $T_2$

takes this ray into $\mathcal{R}''_n$, then acting on $U(T_1)\Psi_n$ it must yield a vector $U(T_2)U(T_1)\Psi_n$ in the ray $\mathcal{R}''_n$. But $U(T_2T_1)\Psi_n$ is also in this ray, so these vectors can differ only by a phase $\phi_n(T_2,T_1)$

$$U(T_2)U(T_1)\Psi_n = e^{i\phi_n(T_2,T_1)}U(T_2T_1)\Psi_n. \tag{3.22}$$

Furthermore, with one significant exception, the linearity (or antilinearity) of $U(T)$ tells us that these phases are independent of the state $\Psi_n$. Here is the proof. Consider any two different vectors $\Psi_A, \Psi_B$, which are not proportional to each other. Then, applying eq. (3.22) to the state $\Psi_{AB} \equiv \Psi_A + \Psi_B$, we have

$$\begin{aligned} e^{i\phi_{AB}}U(T_2T_1)(\Psi_A+\Psi_B) &= U(T_2)U(T_1)(\Psi_A+\Psi_B) \\ &= U(T_2)U(T_1)\Psi_A + U(T_2)U(T_1)\Psi_B \\ &= e^{i\phi_A}U(T_2T_1)\Psi_A + e^{i\phi_B}U(T_2T_1)\Psi_B. \end{aligned} \tag{3.23}$$

Any unitary or antiunitary operator has an inverse (its adjoint) which is also unitary or antiunitary. Multiplying (3.23) on the left with $U^{-1}(T_2T_1)$, we have then

$$e^{\pm i\phi_{AB}}(\Psi_A+\Psi_B) = e^{\pm i\phi_A}\Psi_A + e^{\pm i\phi_B}\Psi_B, \tag{3.24}$$

the upper and lower signs referring to $U(T_2T_1)$ unitary or antiunitary, respectively. Since $\Psi_A$ and $\Psi_B$ are linearly independent, this is only possible if

$$e^{i\phi_{AB}} = e^{i\phi_A} = e^{i\phi_B}. \tag{3.25}$$

So as promised, the phase in (3.22) is independent of the state vector $\Psi_n$, and therefore this can be written as an operator relation

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)}U(T_2,T_1). \tag{3.26}$$

For $\phi = 0$, this would say that $U(T)$ furnish a representation of the group of symmetry transformations. For general phases $\phi(T_2,T_1)$, we have what is called a projective representation, or a representation 'up to a phase'. The structure of the Lie group cannot by itself tell us whether the physical state-vectors furnish an ordinary or a projective representation, but as we shall see, it can tell us whether the group has any intrinsically projective representations at all.

The exception to the argument that led to (3.26) is that it may not be possible to prepare the system in a state represented by $\Psi_A + \Psi_B$. For instance, it is widely believed to be impossible to prepare a system in a superposition of two states whose total angular momentum are integers and half-integers, respectively. In such cases, we say that there is a 'superselection rule' between different classes of states, and the phases $\phi(T_2, T_1)$ *may* depend on which of these classes of states the operators $U(T_2)U(T_1)$ and $U(T_2T_1)$ act upon. We will have more to say about these phases and projective representations below. As we shall see later, any symmetry group with projective representations can always be enlarged (without otherwise changing its physical implications) in such a way that its representations can all be defined as non-projective, with $\phi = 0$. Until then, we will just assume that this has been done, and take $\phi = 0$ in eq. (3.26).

There is a kind of group called *connected Lie group*, of special importance in physics. These are groups of transformations $T(\theta)$ that are described by a finite set of real continuous parameters, say $\theta^a$, with each element of the group connected to the identity by a path within the group. The group multiplication law then takes the form

$$T(\overline{\theta})T(\theta) = T\left(f(\overline{\theta}, \theta)\right) \tag{3.27}$$

with $f^a(\overline{\theta}, \theta)$ a function of the $\overline{\theta}$s and $\theta$s. Taking $\theta^a = 0$ as the coordinates of the identity, we must have

$$f^a(\theta, 0) = f^a(0, \theta) = \theta^a. \tag{3.28}$$

As already mentioned, the transformations of such continuous groups must be represented on the physical Hilbert space by unitary (rather than antiunitary) operators $U(T(\theta))$. For a Lie group, these operators can be represented in at least a finite neighborhood of the identity by a power series

$$U(T(\theta)) = 1 + i\theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \cdots, \tag{3.29}$$

where $t_a$, $t_{bc} = t_{cb}$, etc. are Hermitian operators independent of the $\theta$s. Suppose that the $U(T(\theta))$ form an ordinary (i.e. non-projective) representation of this group of transformations, i.e.

$$U(T(\overline{\theta}))U(T(\theta)) = U(T(f(\overline{\theta}, \theta)). \tag{3.30}$$

Let us see what this condition looks like when expanded in powers of $\theta^a$ and $\overline{\theta}^a$. According to eq. (3.28), the expansion of $f^a(\overline{\theta}, \theta)$ to second order must take the form

$$f^a(\overline{\theta}, \theta) = \theta^a + \overline{\theta}^a + f^a{}_{bc}\overline{\theta}^b\theta_c + \cdots \tag{3.31}$$

with real coefficients $f^a{}_{bc}$. (The presence of any terms of order $\theta^2$ or $\overline{\theta}^2$ would violate eq. (3.28).) Then eq. (3.30) reads

$$\begin{aligned}
&\left[1 + i\overline{\theta}^a t_a + \frac{1}{2}\overline{\theta}^b\overline{\theta}^c t_{bc} + \cdots\right] \times \left[1 + i\theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \cdots\right] \\
&\quad = 1 + i f^a t_a + \frac{1}{2} f^b f^c t_{bc} + \cdots \\
&\quad = 1 + i\left(\theta^a + \overline{\theta}^a + f^a{}_{bc}\overline{\theta}^b\theta^c + \cdots\right) t_a \\
&\quad + \frac{1}{2}(\theta^b + \overline{\theta}^b + \cdots)(\theta^c + \overline{\theta}^c + \cdots) t_{bc} + \cdots
\end{aligned} \tag{3.32}$$

The terms of order $1, \theta, \overline{\theta}, \theta^2$ and $\overline{\theta}^2$ automatically match on both sides of eq. (3.32), but from the $\overline{\theta}\theta$ terms we obtain a non-trivial condition

$$t_{bc} = -t_b t_c - i f^a{}_{bc} t_a. \tag{3.33}$$

This shows that if we are given the structure of the group, i.e. the function $f(\overline{\theta}, \theta)$, and hence its quadratic coefficient $f^a{}_{bc}$, we can calculate the second order terms in $U(T(\theta))$ from the generators $t_a$ appearing in the first order terms. However, there is a consistency condition: the operator $t_{bc}$ must be symmetric in $b$ and $c$ (because it is the second derivative of $U(T(\theta))$ with respect to $\theta^b$ and $\theta^c$) so eq. (3.33) requires that

$$[t_b, t_c] = i C^a{}_{bc} t_a, \tag{3.34}$$

where $C^a{}_{bc}$ are a set of real constants known as *structure constants*

$$C^a{}_{bc} \equiv -f^a{}_{bc} + f^a{}_{cb}. \tag{3.35}$$

Such a set of commutation relations is known as a *Lie algebra*. We will later prove in effect that the commutation relation (3.34) is the single condition needed to assure that this process can be continued: the complete power series for $U(T(\theta))$ may be calculated from an infinite series of relations like (3.33), provided we know the first order terms, the generators $t_a$. This does not necessarily mean that the operators $U(T(\theta))$ are uniquely determined for all $\theta_a$ if we know $t_a$, but it does mean that the $U(T(\theta))$ are uniquely determined in at least a finite neighborhood of the coordinates $\theta^a = 0$ of the identity, in such a way that eq. (3.27) is satisfied if $\theta, \overline{\theta}$ and $f(\theta, \overline{\theta})$ are in this neighborhood. The extension to all $\theta^a$ is discussed later.

There is a special case of some importance, that we will encounter again and again. Suppose that the function $f(\theta, \overline{\theta})$ (perhaps just for some subset of the coordinates $\theta^a$) is simply additive

$$f^a(\theta, \overline{\theta}) = \theta^a + \overline{\theta}^a. \tag{3.36}$$

This is the case for instance for translations in spacetime, or for rotations around any fixed axis (though *not* for both together). Then the coefficients ${f^a}_{bc}$ in eq. (3.31) vanish, and so do the structure constants (3.35). The generators then all commute

$$[t_b, t_c] = 0. \tag{3.37}$$

Such a group is called *Abelian*. In this case, it is easy to calculate $U(T(\theta))$ for all $\theta^a$. From eqs. (3.30) and (3.36), we have $U(T(\overline{\theta}))U(T(\theta)) = U(T(\overline{\theta} + \theta))$, and, in particular, $U(T(\theta))U(T(\theta)) = U(T(2\theta))$. Consequently, for any integer $N$, we have

$$U(T(\theta)) = \left[U(T(\theta/N))\right]^N . \tag{3.38}$$

Letting $N \to \infty$, and keeping only the first order term in $U(T(\theta/N))$, we have then

$$U(T(\theta)) = \lim_{N\to\infty} \left[1 + \frac{i}{N}\theta^a t_a\right]^N \tag{3.39}$$

and hence

$$U(T(\theta)) = \exp(i t_a \theta^a). \tag{3.40}$$

## 4. SU(2)

The simplest non-Abelian Lie algebra consists of three generators $J_a$, $i = 1, 2, 3$, satisfying $f_{abc} = \varepsilon_{abc}$, where $\varepsilon_{abc}$ is the completely antisymmetric tensor ($\varepsilon_{123} = 1$ and $\varepsilon_{abc}$ is antisymmetric in the exchange of any two indices). The commutation relations should be familiar:

$$[J_1, J_2] = iJ_3, \quad [J_2, J_3] = iJ_1, \quad [J_3, J_1] = iJ_2. \tag{4.1}$$

This is the angular momentum algebra. As is well known, the generators of the rotation group satisfy (4.1). We will analyze the representations of the algebra (4.1) in two different ways, both of which generalize to other Lie groups.

The first technique is essentially the operator method. Assume we are given an $N$ dimensional irreducible representation of (4.1). We can diagonalize the hermitian operator $J_3$. Consider the states with the highest value of $J_3$. If there is more than one, label them by integers $\alpha = 1, 2, \ldots$:

$$J_3 \left| j, \alpha \right\rangle = j \left| j, \alpha \right\rangle .$$

Choose $|j, \alpha\rangle$ such that

$$\langle j, \alpha | j, \beta \rangle = \delta_{\alpha\beta}.$$

Define the raising and lowering operators

$$J^{\pm} = (J_1 \pm i J_2)/\sqrt{2},$$

so called because when they act on an eigenstate of $J_3$ they raise or lower the eigenvalue by one unit. The commutation relations for these are

$$[J_3, J^{\pm}] = \pm J^{\pm}, \quad [J^+, J^-] = J_3,$$

so if $J_3 \left| m \right\rangle = m \left| m \right\rangle$, then

$$J_3 J^{\pm} \left| m \right\rangle = J^{\pm} J_3 \left| m \right\rangle \pm J^{\pm} \left| m \right\rangle = (m \pm 1) J^{\pm} \left| m \right\rangle .$$

Since $j$ is the highest $J_3$ value, we must have

$$J^+ \left| j, \alpha \right\rangle = 0.$$

The states $J^- \left| j, \alpha \right\rangle$ have $J_3 = j - 1$. Define

$$J^- \left| j, \alpha \right\rangle = N_j(\alpha) \left| j - 1, \alpha \right\rangle .$$

States with different $\alpha$'s are orthogonal because

$$\begin{aligned} N_j^*(\beta) N_j(\alpha) \langle j-1, \beta | j-1, \alpha \rangle &= \langle j, \beta | J^+ J^- | j, \alpha \rangle = \langle j, \beta | [J^+, J^-] | j, \alpha \rangle \\ &= \langle j, \beta | J_3 | j, \alpha \rangle = j \langle j, \beta | j, \alpha \rangle = j \delta_{\alpha\beta}. \end{aligned}$$

So we can choose $|j-1, \alpha\rangle$ orthonormal, then $|N_j(\alpha)|^2 = j$. Clearly we can choose phases so $N_j(\alpha) = \sqrt{j} = N_j$. Then

$$J^+ \left| j-1, \alpha \right\rangle = \frac{1}{N_j} J^+ J^- \left| j, \alpha \right\rangle = \frac{1}{N_j} [J^+, J^-] \left| j, \alpha \right\rangle = N_j \left| j, \alpha \right\rangle .$$

Note that $J^\pm$ change $j$ without changing $\alpha$.

Now we can look at $J^- |j-1,\alpha\rangle$ and show by an analogous argument that there are orthonormal states $|j-2,\alpha\rangle$ satisfying

$$\begin{aligned} J^- |j-1,\alpha\rangle &= N_{j-1} |j-2,\alpha\rangle , \\ J^+ |j-2,\alpha\rangle &= N_{j-1} |j-1,\alpha\rangle . \end{aligned}$$

Continuing the process we find orthonormal states $|j-k,\alpha\rangle$, such that

$$\begin{aligned} J^- |j-k,\alpha\rangle &= N_{j-k} |j-k-1,\alpha\rangle , \\ J^+ |j-k-1,\alpha\rangle &= N_{j-k} |j-k,\alpha\rangle . \end{aligned}$$

The $N$'s (chosen real) satisfy

$$\begin{aligned} N_{j-k}^2 &= \langle j-k,\alpha|J^+J^-|j-k,\alpha\rangle \\ &= \langle j-k,\alpha|[J^+,J^-]|j-k,\alpha\rangle + \langle j-k,\alpha|J^-J^+|j-k,\alpha\rangle = N_{j-k+1}^2 + j - k. \end{aligned}$$

This recusrsion relation can be solved by addition (do it!), giving

$$N_m = \sqrt{(j+m)(j-m+1)/2}.$$

Eventually, we must come to a $j-l$ such that $J^- |j-l,\alpha\rangle = 0$. But this requires $N_{j-l} = 0$ or $l = 2j$. So, $j = l/1$ for some integer $l$.

The space breaks up to $\alpha$ subspaces of dimension $2j+1$ and the subspace of states that we have not already found by our procedure. But the $J$'s act on these subspaces in block diagonal form,

$$\begin{pmatrix} J & & & & \\ & J & & & \\ & & J & & \\ & & & . & \\ & & & & ? \end{pmatrix} . \tag{4.2}$$

Since we assumed the representation was irreducible, there must be only one $\alpha$ and no other states.

We have explicitly constructed the irreducible representations of (4.1). They are characterized by their highest $J_3$ eigenvalue $j$ which must be an integer or half-odd-integer. They are $2j+1$ dimensional.

Furthermore, the same construction can be used to decompose an arbitrary representation into its irreducible components. Starting with any representation, we obtain (4.2) in which the irreducible representations with the highest $j$ are explicitly separated out. Now we can apply the procedure again to the other states to find the representations with the next highest $j$, and so on until the representation is completely reduced.

The values of $J_3$ are called *weights*, and $j$ is the *highest weight*. This construction, which characterizes the irreducible representations in terms of their highest weight, generalizes to any Lie algebra.

The representation of (4.1) with highest weight $j$ is sometimes called *spin j representation* because it is the representation associated with a particle at rest with spin angular momentum $j$ ($J^2 = j(j+1)\hbar^2$).

When hermitian angular momentum operators (that is $J$'s satisfying (4.1)) exist, the entire Hilbert space is a representation and can be decomposed into states labeled by $|m, j, \alpha\rangle$ in the spin $j$ representation with $J_3 = m$. The $\alpha$ stands for all other labels necessary to characterize the states. We can (and will) choose the state orthonormal in $\alpha$, so the states satisfy

$$\langle m, j, \alpha | m', j', \beta\rangle = \delta_{mm'}\delta_{jj'}\delta_{\alpha\beta}. \tag{4.3}$$

These states form a convenient basis for studying the properties of the system under rotations.

The $\delta$ functions on the RHS of (4.3) are not purely a convention. They follow from the algebraic structure. We know that $m$ and $m'$ must be equal for the matrix element to be nonzero because they are eigenvalues of a hermitian operator. Of course, we know from the familiar analysis in terms of the operator $J^2$ that $j$ and $j'$ are also eigenvalues of a hermitian operator. But we can also show that $j = j'$ for a nonzero matrix element directly with raising and lowering operators. Suppose that $j' > j$ and consider

$$\langle j, j, \alpha | j, j', \beta\rangle = \sqrt{\frac{2}{(j'+j)(j'-j+1)}}\langle j, j, \alpha | J^- | j+1, j', \beta\rangle = 0.$$

Applying $J^\pm$ judiciously, we can show that $j = j'$ for all $m$.

The highest value construction shows us how to break down a representation into simpler irreducible representations. There is another way of looking at it. We can build up an arbitrary representation out of the simplest representation. The simplest nontrivial representation of (4.1) is generated by $\sigma_a/2$ where $\sigma_a$ are the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.4}$$

Notice that the group elements, $\exp(i\alpha_a\sigma_a/2)$, are the Special (determinant 1) Unitary $2\times 2$ matrices. The group is often called $SU(2)$.

Given two representations $D_1$ (dim $n$) and $D_2$ (dim $m$) of a group $G$, we can form another representation *in two ways*:

(1) The direct sum $D_1 \oplus D_2$ has dimension $n+m$. It is formed by the block-diagonal matrices

$$\begin{pmatrix} D_1(g) & 0 \\ 0 & D_2(g) \end{pmatrix}.$$

The generators of the representation $D_1 \oplus D_2$ are

$$X_a^{D_1\oplus D_2} = \begin{pmatrix} X_a^{D_1} & 0 \\ 0 & X_a^{D_2} \end{pmatrix}.$$

The operation of taking the direct sum is the reverse of the process of reducing a representation.

(2) We can also form a representation of dimension $n\times m$ as follows. If $|i\rangle$ $(i=1,\ldots,n)$ is an orthonormal basis in the space on which $D_1$ acts, and $|x\rangle$ $(x=1,\ldots,m)$ is an orthonormal basis for $D_2$, we can identify the product $\{|i\rangle\,|x\rangle\}$ with an orthonormal basis in an $n\times m$ dimensional space, called the *direct product space*. On this space, the *direct product representation* $D_1\otimes D_2$ is

$$(D_1\otimes D_2)(g)\{|i\rangle\,|x\rangle\} = \{D_1(g)\,|i\rangle\,D_2(g)\,|x\rangle\}. \tag{4.5}$$

In the matrix language,

$$[(D_1\otimes D_2)(g)]_{ix,jy} = [D_1(g)]_{ij}[D_2(g)]_{xy}. \tag{4.6}$$

The generators of the direct product representation are the sums (obtained by differentiating the product in (4.6))

$$[X_a^{D_1\otimes D_2}]_{ix,jy} = [X_a^{D_1}]_{ij}\delta_{xy} + \delta_{ij}[X_a^{D_2}]_{xy}. \tag{4.7}$$

Clearly we can form the direct product of an arbitrary number of representations.

The ideas of direct product space and direct product representation are very important. The space occurs whenever a state can be labeled by two or more independent indices. When the indices are associated with representations of a Lie algebra, the states are in a direct product representation. The addition of angular momentum in quantum mechanics is an example of (4.7). Often, the $\delta$ functions in (4.7) are suppressed, since they are just identity matrices in the relevant space. The important thing to remember in dealing with direct products is what matrices act on which space.

Now, getting back to the two dimensional representations of $SU(2)$, we can form the direct product of $n$ of them,

$$[D \otimes \cdots \otimes D]_{i_1 \cdots i_n, j_1 \cdots j_n} = D_{i_1 j_1} \cdots D_{i_n j_n} \tag{4.8}$$

acting on the $n$ component objects $u_{j_1 \cdots j_n}$. This representation is reducible. In particular, (4.8) preserves any symmetry of the $u$ vectors under permutation of the labels 1 to $n$. The product representation breaks up into irreducible components corresponding to vectors with definite symmetry. The component acting on the completely symmetric subspace is $n+1$ dimensional, because the completely symmetric $n$ index object has only $n+1$ independent components. In fact, this representation is the irreducible representation with $j = n/2$. This construction also generalizes to more complicated groups and their algebras. It is the basis of what are called "tensor methods" for dealing with representations.

## 5. Tensor Operators and the Wigner-Eckart Theorem

A *tensor operator* $O_x$ $(x = 1, \ldots, 2s+1)$ is a set of linear operators with simple commutation relations with the angular momentum operators:

$$[J_a, O_x] = O_y [J_a^s]_{yx} \tag{5.1}$$

where $J_a^s$ are the matrix generators of the irreducible spin-$s$ representation. The operators $O_x$ are said to transform according to the spin $s$ representation.

For example, consider a system of a particle in a spherically symmetric potential. Here $J_a$ are the orbital angular momentum operators

$$J_a = L_a = \varepsilon_{abc} r_b p_c .$$

The position vector $r_a$ is a tensor operator because

$$[J_a, r_b] = \varepsilon_{acd} [r_c p_d, r_b] = -i \varepsilon_{abc} r_c = r_c [J_a^1]_{cb} .$$

It transforms according to the spin 1 (adjoint) representation.

The reason that the definition (5.1) is useful is this: Consider the state $O_x\,|m,j,\alpha\rangle$. It satisfies

$$\begin{aligned} J_a\{O_x\,|m,j,\alpha\rangle\} &= [J_a,O_x]\,|m,j,\alpha\rangle + O_x J_a\,|m,j,\alpha\rangle \\ &= O_y\,|m,j,\alpha\rangle\,[J_a^s]_{yx} + O_x\,|m',j,\alpha\rangle\,[J_a^j]_{m'm}. \end{aligned} \tag{5.2}$$

Comparing (5.2) with (4.7), we see that the state $O_x\,|m,j,\alpha\rangle$ transforms according to the tensor product representation $s\otimes j$.

We can choose linear combinations of the $O_x$ as follows, to diagonalize the hermitian matrix $[J_3^s]$:

$$\begin{aligned} O'_\ell &= O_y[S^{-1}]_{y\ell} \quad (\ell=-s,\ldots,s), \\ [SJ_3^sS^{-1}]_{\ell\ell'} &= \ell\delta_{\ell\ell'}. \end{aligned} \tag{5.3}$$

Then

$$\begin{aligned} [J_3,O'_\ell] &= \ell O'_\ell, \\ J_3O'_\ell\,|m,j,\alpha\rangle &= (\ell+m)O'_\ell\,|m,j,\alpha\rangle\,. \end{aligned} \tag{5.4}$$

In the basis defined by (5.3), the operators $O'_\ell$ are analogous to the $J_3$ eigenstates $|\ell,s\rangle$. The state $O'_\ell\,|m,j,\alpha\rangle$ is the analog of the tensor product state $\{|\ell,s\rangle\,|m,j,\alpha\rangle\}$. We can decompose it, just as we can decompose the product state into states which transform irreducibly under the algebra. You have shown in the exercise that $s\otimes j=(s+j)\oplus\cdots\oplus|s-j|$. Thus

$$O'_\ell\,|m,j,\alpha\rangle = \sum_{J=|s-j|}^{(s+j)} \langle \ell+m,J|\ell,s;m,j\rangle\cdot|\ell+m,J\rangle\,. \tag{5.5}$$

The constants $\langle \ell+m,J|\ell,s;m,j\rangle$ are pure group theory. They are called *Clebsch-Gordan coefficients*. You should be familiar with them from your study of addition of angular momenta in quantum mechanics. You can look them up in tables. Or, better still, you can determine them yourself by working out the corresponding decomposition of the tensor product state

$$|\ell,s\rangle\,|m,j\rangle = \sum_{J=|s-j|}^{(s+j)} \langle \ell+m,J|\ell,s;m,j\rangle\cdot|\ell+m,J\rangle\,, \tag{5.6}$$

using the highest weight decomposition algorithm. For example, we know that the state of highest weight is

$$|s,s\rangle\,|j,j\rangle = |s+j,s+j\rangle\,. \tag{5.7}$$

Thus

$$\langle s+j, s+j | s, s; j, j\rangle = 1.$$

Of course, it is equally obvious that the highest weight state in the product $O'_\ell |m, j, \alpha\rangle$ is $O'_s |j, j, \alpha\rangle$. But to go beyond this, it is much more convenient to work with analog tensor product states. Acting on (5.7) with the lowering operator $J^-$ we find

$$\sqrt{\frac{s}{s+j}}\,|s-1, s\rangle\,|j, j\rangle + \sqrt{\frac{j}{s+j}}\,|s, s\rangle\,|j-1, j\rangle = |s+j-1, s+j\rangle\,. \tag{5.8}$$

Thus, for example,

$$\langle s+j-1, s+j | s-1, s; j, j\rangle = \sqrt{\frac{s}{s+j}}.$$

Proceeding in this way we can find all of the coefficients for $J = s+j$. To get them for $J = s+j-1$, we proceed with the highest weight algorithm and throw out all the $J = s+j$ states. The highest weight state remaining is the state orthogonal to (5.8):

$$\sqrt{\frac{j}{s+j}}\,|s-1, s\rangle\,|j, j\rangle - \sqrt{\frac{s}{s+j}}\,|s, s\rangle\,|j-1, j\rangle = |s+j-1, s+j-1\rangle\,. \tag{5.9}$$

Thus, for example,

$$\langle s+j-1, s+j-1 | s, s; j-1, j\rangle = -\sqrt{\frac{s}{s+j}}. \tag{5.10}$$

(5.9) is a definition. The state has been properly normalized, but its phase is ours to choose. The Clebsch-Gordan coefficient (5.10) depends on the choice of the phase. When you use tables of Clebsch-Gordan coefficients, you must be sure that you understand their phase convention.

Now we know by analogy that, for example, the state

$$\sqrt{\frac{j}{s+j}}O'_{s-1}\,|j, j, \alpha\rangle - \sqrt{\frac{s}{s+j}}O'_s\,|j-1, j, \alpha\rangle = |s+j-1, s+j-1\rangle \tag{5.11}$$

is the highest weight state in the $J = s+j-1$ representation. We can confirm this by applying $J^+$ to (5.11) and checking that it annihilates the state. Indeed this must happen because the action of $J^+$ on (5.11) is the same as the action of $J^+$ on (5.9).

Something slightly nontrivial has happened here. We used the orthogonality property of the tensor product,

$$(\langle \ell', s|\,\langle m', j|)(|l, s\rangle\,|m, j\rangle) = \delta_{\ell\ell'}\delta_{mm'}, \tag{5.12}$$

to derive the state (5.9) which is annihilated by $J^+$. We then argued that the analog state (5.11) is also annihilated by $J^+$. And this is true even though the orthogonality property analogous to (5.12) is not true:

$$\langle m', j, \alpha | O'^{\dagger}_{\ell'} O'_{\ell} | m, j, \alpha \rangle \neq \delta_{\ell\ell'} \delta_{mm'} . \tag{5.13}$$

Indeed, group theory does not tell us anything about how the state (5.11) is normalized.

Now, back to (5.5). We can decompose the states $|\ell + m, J\rangle$ which appear on the RHS of (5.5) in terms of our complete set of states $|m, j, \alpha\rangle$ as follows:

$$|\ell + m, J\rangle = \sum_{\beta} k_{\beta\alpha} \, |\ell + m, j, \beta\rangle \, . \tag{5.14}$$

The constants $k_{\beta\alpha}$ depend on $j$, $\alpha$ and on the operator $O$ which appear on the LHS of (5.5) and also on $J$ and $\beta$. *But $k_{\beta\alpha}$ is independent of $\ell$ and $m$.* It is enough to find $k_{\beta\alpha}$ for, say, $\ell + m = J$. This is obvious because $J^{\pm}$ applied to (5.14) changes $\ell + m$ without changing the constants $k_{\beta\alpha}$. The $k_{\beta\alpha}$ are called *reduced matrix elements* and denoted by

$$k_{\beta\alpha} = \langle J, \beta | O | j, \alpha \rangle. \tag{5.15}$$

Finally, we can put all of this together to get what is called the *Wigner-Eckart theorem* for matrix elements of tensor operators:

$$\langle m', J, \beta | O'_{\ell} | m, j, \alpha \rangle = \delta_{m', \ell+m} \langle \ell + m, J | \ell, s; m, j \rangle \cdot \langle J, \beta | O | j, \alpha \rangle. \tag{5.16}$$

According to this theorem, if we know any nonzero matrix element of a tensor operator between states of some given $J, \beta$ and $j, \alpha$, we can compute all the others using group theory. This sounds remarkable, but remember we just arrived at it using nothing more than the properties of the algebra. The point is simply that we can use $J^{\pm}$ to go up and down within representations using pure group theory. Thus, it should not be surprising that by judicious use of the raising and lowering operators we can relate the various matrix elements.

Let us work out an example involving the position operator. Given

$$\langle 1/2, 1/2, \alpha | r_3 | 1/2, 1/2, \beta \rangle = A,$$

we should be able to find

$$\langle 1/2, 1/2, \alpha | r_1 | -1/2, 1/2, \beta \rangle.$$

The first step is to construct the operators $r'_\ell$. This can be done easily by noting that $[J_3, r_3] = 0$, so that we can take

$$r'_0 = r_3 .$$

Then the commutation relations for the spin 1 operator give the rest:

$$[J^\pm, r'_0] = r'_{\pm 1} = \mp(r_1 \pm i r_2)/\sqrt{2}.$$

Thus

$$A = \langle 1/2, 1/2, \alpha | r'_0 | 1/2, 1/2, \beta \rangle,$$

and

$$\begin{aligned}\langle 1/2, 1/2, \alpha | r_1 | -1/2, 1/2, \beta \rangle =& \langle 1/2, 1/2, \alpha | (r'_{-1} - r'_{+1})/\sqrt{2} | -1/2, 1/2, \beta \rangle \\ =& -\frac{1}{\sqrt{2}} \langle 1/2, 1/2, \alpha | r'_{+1} | -1/2, 1/2, \beta \rangle .\end{aligned}$$

Then the theorem (5.16) gives

$$A = \langle 1/2, 1/2 | 0, 1; 1/2, 1/2 \rangle \langle 1/2, \alpha | r | 1/2, \beta \rangle,$$

and the matrix element we want is

$$-\frac{1}{\sqrt{2}} \langle 1/2, 1/2 | 1, 1; -1/2, 1/2 \rangle \langle 1/2, \alpha | r | 1/2, \beta \rangle .$$

From (5.9) and (5.10) (or your favorite table) we see that

$$\begin{aligned}\langle 1/2, 1/2 | 1, 1; -1/2, 1/2 \rangle =& -\sqrt{2/3}, \\ \langle 1/2, 1/2 | 0, 1; 1/2, 1/2 \rangle =& \sqrt{1/3}.\end{aligned}$$

Thus

$$\langle 1/2, 1/2, \alpha | r_1 | -1/2, 1/2, \beta \rangle = A.$$

## 6. The Wigner Theorem

Consider two experimenters $O$ and $O'$ using differently oriented coordinate frames with a common origin. Let $R$ specify their relative orientation – as given by the Euler angles, for example. Let $|jm\rangle$ and $|jm; R\rangle$ be the angular momentum eigenstates prepared by $O$ and $O'$, respectively. Thus $|jm\rangle$ is a state with eigenvalues $j(j+1)$ and $m$ under $J^2$

and $J_z$, while $|jm;R\rangle$ is an eigenstate of $J_{z'}$ with the *same* eigenvalue $m$. Furthermore, let $O$ prepare a one electron state of momentum $p$ travelling along the $z$-axis, and let $O'$ prepare a one-electron state of the same energy travelling along the $z'$ axis. We shall designate these states by $|u\rangle$ and $|u;R\rangle$, respectively. If $O$ carries out an experiment to determine the probability that the state $|u\rangle$ be found to have definite eigenvalues of $J^2$ and $J_z$, he will find that this probability is $|\langle u|jm\rangle|^2$. But $O$ can also carry out a similar experiment on the states prepared by $O'$, and measure the probability $|\langle u;R|jm;R\rangle|^2$. If we assume that the relationship between rotations and translations is the same in the two frames, we must require $|\langle u|jm\rangle| = |\langle u;R|jm;R\rangle|$. This relation must also hold for an arbitrary state $|u\rangle$, provided $|u;R\rangle$ has precisely the same properties when observed by $O'$ as $|u\rangle$ has when observed by $O$. Note the absolute value signs appear because the rays, and not the kets, correspond uniquely to physical states. It is also important to realize that we have *not* assumed rotational invariance of the equations of motion.

The considerations of the preceding paragraph can be generalized to other relationships between pairs of observers. Other important examples are pairs of observers whose coordinate systems are translated with respect to each other, and pairs of observers who use different convention for the origin of time. In relativistic quantum mechanics important consequences emerge when one considers observers who use different conventions for the sign of the electric charge.

We now inquire into the mathematical nature of the relationship between the two complete, though equivalent, descriptions used by $O$ and $O'$. In the case of the example used above, we would like to determine the properties of the operator that perform the mapping $|u\rangle \to |u;R\rangle$ on the Hilbert space. If the vectors, and scalar products, were themselves of physical significance, we would immediately conclude that the correspondence between the two equivalent descriptions is given by a unitary transformation. Because there need only be a one-to-one correspondence between physical states prepared by $O$ and $O'$, however, only one-to-one correspondence between rays is required. We must therefore be prepared for more complicated transformations than the unitary ones. Fortunately, Wigner has proven the remarkably simple result that we need only concern ourselves with one kind of transformations beyond the unitary type. Moreover, we shall see that nonunitary transformations can only be associated with a very restricted class of transformations.

Henceforth we shall use the following notation: $O$ uses the kets $|\alpha\rangle$, $|\beta\rangle$, to describe a pair of states with certain properties, $O'$ associates $|\alpha'\rangle$, $|\beta'\rangle$ with states which, in his description, have identical properties. If, moreover, $c_1\,|\alpha_1\rangle + c_2\,|\alpha_2\rangle$ is a ket describing some state prepared by $O$, then $(c_1 |\alpha_1\rangle + c_2 |\alpha_2\rangle)'$ is the corresponding $O'$-ket. The ray correspondence that we are about to investigate is then

$$|\langle\alpha|\beta\rangle| = |\langle\alpha'|\beta'\rangle|. \tag{6.1}$$

Let $\{|a_n\rangle\}$ be an orthonormal basis prepared by $O$, and let $\{|a'_n\rangle\}$ be the equivalent basis prepared by $O'$. By hypothesis

$$\sum_n |a_n\rangle \langle a_n| = \sum_n |a'_n\rangle \langle a'_n| = 1. \tag{6.2}$$

The phases of these sets can be chosem at our convenience. Furthermore, if $|\alpha\rangle = \sum_n |a_n\rangle \langle a_n|\alpha\rangle$ and $|\alpha'\rangle = \sum_n |a'_n\rangle \langle a'_n|\alpha'\rangle$, we are also free to choose the arbitrary overall phases of $|\alpha\rangle$ and $|\alpha'\rangle$. Only the ray containing $|\alpha\rangle$ must be mapped into the ray containing $|\alpha'\rangle$ in the transformation from the $O$- to the $O'$-description.

The content of this subsection is expressed by the Wigner's theorem: We may choose the arbitrary phases so that *either*

$$\begin{aligned} \langle\alpha|\beta\rangle =& \langle\alpha'|\beta'\rangle, \qquad (I) \\ (c_1 |\alpha_1\rangle + c_2 |\alpha_2\rangle)' =& c_1 |\alpha'_1\rangle + c_2 |\alpha'_2\rangle, \end{aligned} \tag{6.3}$$

or

$$\begin{aligned} \langle\alpha|\beta\rangle =& \langle\beta'|\alpha'\rangle = \langle\alpha'|\beta'\rangle^*, \quad (II), \\ (c_1 |\alpha_1\rangle + c_2 |\alpha_2\rangle)' =& c_1^* |\alpha'_1\rangle + c_2^* |\alpha'_2\rangle. \end{aligned} \tag{6.4}$$

Case (I) corresponds to a linear unitary transformation, while case (II) corresponds to an antilinear antiunitary transformation.

To prove the theorem, we begin with the ket $|\phi_n\rangle = |a_1\rangle + |a_n\rangle$, and show that we can arrange the phases of $\{|a'_n\rangle\}$ so that $|\phi'_n\rangle = |a'_1\rangle + |a'_n\rangle$. From (6.2) we have

$$|\phi'_n\rangle = \sum_m |a'_m\rangle \langle a'_m| (|a_1\rangle + |a_n\rangle)',$$

and from (6.1)

$$\langle a'_m| (|a_1\rangle + |a_n\rangle)' = e^{i\tau_m} [\langle a_m|a_1\rangle + \langle a_m|a_n\rangle] = e^{i\tau_m} (\delta_{m1} + \delta_{mn}),$$

or

$$|\phi'_n\rangle = e^{i\tau_1} |a'_1\rangle + e^{i\tau_n} |a'_n\rangle.$$

We are also free to redefine the phases of $\{|a'_n\rangle\}$ so as to absorb the phase factors $e^{i\tau_n}$. Then we have

$$|\phi'_n\rangle = |a'_1\rangle + |a'_n\rangle\,.$$

Consider the ket

$$|\alpha\rangle = \sum_n c_n\,|a_n\rangle\,,$$

and its image

$$|\alpha'\rangle = \sum_n c'_n\,|a'_n\rangle\,.$$

Of course, $c_n = \langle a_n|\alpha\rangle$ and $c'_n = \langle a'_n|\alpha'\rangle$., and therefore

$$|c_n| = |c'_n|.$$

But

$$|\langle\phi_n|\alpha\rangle| = |c_1 + c_n|, \quad |\langle\phi'_n|\alpha'\rangle| = |c'_1 + c'_n|,$$

and therefore

$$|c_1 + c_n| = |c'_1 + c'_n|.$$

We are still free to choose the phases of $|\alpha'\rangle$ and we do this to make $c_1 = c'_1$. Thus we must solve

$$|c_1 + c_n| = |c_1 + c'_n|,$$

which gives the quadratic equation

$$c_1 c^*_n + c^*_1 c_n = c_1 c'^*_n + c^*_1 c'_n.$$

Multiplying by $c'_n$ gives

$$c^*_1 c'^2_n - c'_n(c_1 c^*_n + c^*_1 c_n) + c_1|c_n|^2 = 0.$$

The roots of this equation are

$$c'_n = \begin{cases} c_n & \text{(i)} \\ (c_1/c^*_1)c^*_n & \text{(ii)} \end{cases}.$$

We can still redefine the phase of $|\alpha\rangle$ itself, and make $c_1$ real. Thus we finally have

$$|\alpha'\rangle = \sum_n c^*_n\,|a'_n\rangle$$

in case (ii). Hence (i) corresponds to the unitary transformation, (ii) to the antiunitary.

Consider another ket $|\beta\rangle = \sum_n d_n |a_n\rangle$. It transforms to either $\sum_n d_n |a'_n\rangle$ or $\sum_n d_n^* |a'_n\rangle$. Assume the unitary case, $|\alpha'\rangle = \sum_n c_n |a'_n\rangle$. Then we see that if $|\beta\rangle$ undergoes the unitary transformation, $\langle\beta'|\alpha'\rangle = \sum_n d_n^* c_n = \langle\beta|\alpha\rangle$, while if $|\beta\rangle$ undergoes an antiunitary transformation, $|\langle\beta'|\alpha'\rangle| = |\sum_n d_n c_n| \neq |\langle\beta|\alpha\rangle|$. Hence in a particular transformation $O \to O'$, *all* vectors are transformed according to unitary or antiunitary transformations. In detail, if

$$|\alpha'\rangle = \sum_n c_n^* |\alpha'_n\rangle \,, \quad |\beta'\rangle = \sum_n d_n^* |\alpha'_n\rangle \,,$$

then

$$\langle\beta'|\alpha'\rangle = \sum_n d_n c_n^* = \langle\alpha|\beta\rangle.$$

This completes the proof of the theorem.

## 7. Time Translations and the Hamiltonian

At $t = t_0$, an observer prepares a variety of states $|a't_0\rangle$, $|a''t_0\rangle$, etc., which are eigenstates of an operator $A$:

$$(A - a')\,|a't_0\rangle = 0.$$

The time evolution of this state is described by a unitary operator $U$ which (for an isolated system) can depend only on $t - t_0$:

$$|a't_0;t\rangle = U(t - t_0)\,|a't_0\rangle\,.$$

Consider a time $t'$ such that $t > t' > t_0$. We could permit the system to evolve from $|a't_0\rangle$ until time $t'$, i.e., into the state represented by $U(t' - t_0)\,|a't_0\rangle$. This state $|a't_0;t'\rangle$ could then be allowed to evolve until $t$, at which time it is $U(t - t')\,|a't_0;t'\rangle$. The latter state is clearly identical to $|a't_0;t\rangle$, and we therefore obtain the *group property* of the unitary operators:

$$U(t - t')U(t' - t_0) = U(t - t_0).$$

Setting $t_0 = t$, and noting that $U(0) = 1$, yields

$$U^{-1}(t - t') = U(t' - t) = U^\dagger(t - t').$$

We are therefore dealing with a continuous one-parameter group of operators, which means that $U$ can only be unitary and not antiunitary.

Consider the case where $t - t_0$ is infinitesimal, and put $t_0 = 0$. $U(\delta t)$ will differ from 1 by an infinitesimal operator:

$$U(\delta t \to 0) \longrightarrow 1 + \delta U.$$

Unitarity then implies

$$1 = UU^\dagger = (1 + \delta U)(1 + \delta U^\dagger) \simeq 1 + (\delta U + \delta U^\dagger) + \mathcal{O}[(\delta U)^2].$$

Hence we require that $\delta U + \delta U^\dagger = 0$, which implies that $\delta U$ is anti-hermitian, or $\delta U = i\times$(hermitian operator). We make the further observation that the group property

$$U(\delta t_1)U(\delta t_2) = U(\delta t_1 + \delta t_2)$$

requires $\delta U(\delta t)$ to be proportional to $\delta t$. Thus in the case of an isolated system, $\delta U$ equals $i\delta t$ multiplied by a time-independent hermitian operator, which has dimension of energy. Conventionally, one writes

$$\delta U = -iH\delta t,$$

where $H$ is the *Hamiltonian operator* of the system. Note that $H$ is the *generator* of infinitesimal time displacements.

Using the group property we have

$$U(t + dt) - U(t) = [U(dt) - 1]U(t) = -iH\ dt\ U(t),$$

or

$$i\frac{\partial}{\partial t}U(t) = HU(t).$$

Because $H$ commutes with itself, this equation can be integrated:

$$U(t) = \exp(-iHt), \tag{7.1}$$

since $U(0) = 1$. In the same way, we have

$$|a' t_0; t + dt\rangle - |a' t_0; t\rangle = [U(dt) - 1]\,|a' t_0; t\rangle = -iH\ dt\,|a' t_0; t\rangle\,,$$

and therefore $|a' t_0; t\rangle$ satisfies the *Schrödinger equation*,

$$\left(i\frac{\partial}{\partial t} - H\right)|a' t_0; t\rangle = 0, \tag{7.2}$$

with the initial condition $|a' t_0; t_0\rangle = |a' t_0\rangle$. Eq. (7.2) has the formal solution

$$|a' t_0; t\rangle = \exp[-iH(t - t_0)]\,|a' t_0\rangle\,. \tag{7.3}$$

Eq. (7.3) assumes a very simple form when the Hamiltonian itself is a member of the complete set of observables that specifies the initial state. When this is the case we shall write the initial state as $|Ea' t_0\rangle$, where $E$ is the eigenvalue of $H$. Eq. (7.3) then reads

$$|Ea' t_0; t\rangle = \exp[-iE(t - t_0)]\,|Ea' t_0\rangle\,. \tag{7.4}$$

These eigenstates of $H$ therefore belong to the same ray for all values of $t$. They represent the same physical states at different times and are therefore stationary. The matrix elements of a time independent observable $B$ between eigenstates of $H$ also have a very simple time dependence:

$$\langle E' a' t_0; t|B|E'' a'' t_0; t\rangle = \exp[-i(E'' - E')(t - t_0)]\langle E' a' t_0|B|E'' a'' t_0\rangle.$$

Note that in our formulation, the linearity of the Schrödinger equation (7.2) is an immediate consequence of the unitarity and continuous nature of time displacements.

Our discussion has been based on a description where the kets move about in Hilbert space as time goes on, but the operators stay fixed. This is only one of many possible and completely equivalent modes of description. It is called the *Schrödinger picture*. Another very important way of discussing time development is with the *Heisenberg picture*. Here we fix the kets and allow the operators to evolve, i.e., the function (or mapping) they perform depends on $t$. The existence of this possibility can be seen by writing an arbitrary matrix element of $A$ at time $t$:

$$\langle b' t_0; t|A|c' t_0; t\rangle = \langle b' t_0|U^\dagger(t, t_0)AU(t, t_0)|c' t_0\rangle. \tag{7.5}$$

Clearly we can calculate equally well with the stationary kets $|b' t_0\rangle$ and the time dependent operator

$$A(t, t_0) = U^\dagger(t, t_0)AU(t, t_0).$$

Consider again the case of an isolated system, and choose $t_0 = 0$. Because of (7.1), (7.5) now reads

$$\langle b't|A|c't\rangle = \langle b'|e^{iHt}Ae^{-iHt}|c'\rangle. \tag{7.6}$$

The Heisenberg picture observable is then defined as

$$A(t) = e^{iHt}Ae^{-iHt}, \qquad A = A(0). \tag{7.7}$$

Let us briefly summarize the differences between the two pictures. In the S-picture, a system evolves in time according to the Schrödinger equation, or more succintly according to the moving ket

$$|a't\rangle_S = U(t)\,|a't\rangle_H\,,$$

where the subscript H(S) stands for Heisenberg (Schrödinger). The ket $|a't\rangle_H$ is just the initial condition which must, in any case, be stipulated in addition to the Schrödinger equation. In the H-picture, the system is described by the time-dependent observable $A(t)$ and a *fixed ket* $|a't\rangle_H$. The equation of motion must now determine $A(t)$. The equation that does this is obtained by differentiating (7.7):

$$i\frac{d}{dt}A(t) = [A(t), H]. \tag{7.8}$$

This is the *Heisenberg equation of motion*, and it replaces the S-equation in the H-picture. It should be noted that a ket $|\alpha\rangle_S$ which was taken to be fixed in the S-picture, moves in the H-picture,

$$|\alpha\rangle_H = U^\dagger(t)\,|\alpha\rangle_S\,.$$

From a practical point of view, the S-picture is often (though not always, by any means) the better to use. For theoretical discussions the H-picture is frequently superior. For example, (7.8) tells us immediately that $\langle A\rangle$ is $t$-independent if $[A, H] = 0$.

## 8. Spatial Translations

Consider two observers whose coordinate frames are rigidly displaced with respect to each other. The first observer has a complete arsenal of measurement devices with which he prepares the states $|a'\rangle$, $|b'\rangle$, etc. The second observer's origin is at $\mathbf{a}$ from the first one's point of view. The second observer also supplies herself with a complete set of states $|a';\mathbf{a}\rangle$ that, as far as she is concerned, have the same properties as the $|a'\rangle$ of the first observer.

Thus $|\langle a'|b'\rangle| = |\langle a';\mathbf{a}|b';\mathbf{a}\rangle|$, and by Wigner's theorem we may choose the arbitrary phases in $|a'\rangle$, $|a';\mathbf{a}\rangle$ so that these two sets of kets are connected by a unitary operator.

The translations of the preparing and measuring apparatus in ordinary three-dimensional space form a group parameterized by the vectors which specify the origin of the coordinate frame before and after displacement. A group element is written as $\mathcal{G}(\mathbf{a}_1, \mathbf{a}_2)$, and it satisfies the composition law

$$\mathcal{G}(\mathbf{a}_1, \mathbf{a}_2)\mathcal{G}(\mathbf{a}_2, \mathbf{a}_3) = \mathcal{G}(\mathbf{a}_1, \mathbf{a}_3).$$

Because we assume the three dimensional space to be Euclidean, only the displacements are of significance, and $\mathcal{G}(\mathbf{a}_1, \mathbf{a}_2)$ can be written as $\mathcal{G}(\mathbf{a}_1 - \mathbf{a}_2)$. The group relations are then

$$\begin{aligned} \mathcal{G}(\mathbf{a})\mathcal{G}(\mathbf{a}') &= \mathcal{G}(\mathbf{a}+\mathbf{a}'), \\ \mathcal{G}(\mathbf{a}) &= \mathcal{G}^{-1}(-\mathbf{a}). \end{aligned}$$

The operations $\mathcal{G}(\mathbf{a})$ discussed up to now work on the apparatus and take place in ordinary 3-space. To each of these there must correspond a unitary operator in the Hilbert space which we call $U(\mathbf{a})$. It is clear that the correspondence $\mathcal{G} \leftrightarrow U$ must preserve the group relations; thus, if $U(\mathbf{a})$, $U(\mathbf{a}')$, correspond to the translations $\mathcal{G}(\mathbf{a})$, $\mathcal{G}(\mathbf{a}')$, respectively, then the translation $\mathcal{G}(\mathbf{a})\mathcal{G}(\mathbf{a}')$ corresponds to $U(\mathbf{a}+\mathbf{a}')$. Hence

$$\begin{aligned} U(\mathbf{a})U(\mathbf{a}') &= U(\mathbf{a}+\mathbf{a}'), \\ U(\mathbf{a})U(-\mathbf{a}) &= U(-\mathbf{a})U(\mathbf{a}) = 1, \end{aligned}$$

or

$$U(-\mathbf{a}) = U^{-1}(\mathbf{a}) = U^{\dagger}(\mathbf{a}).$$

Just as in the time displacement case we can consider infinitesimal transformations

$$U(\delta\mathbf{a}) = 1 + \delta U,$$

and once more the group property and unitarity imply that $\delta U = \delta\mathbf{a}\cdot\mathbf{F}$, where $\mathbf{F}$ is an anti-hermitian operator that does not depend on $\mathbf{a}$. Thus we write

$$U(\delta\mathbf{a}) = 1 - i\ \delta\mathbf{a}\cdot\mathbf{P}. \tag{8.1}$$

The hermitian operator $\mathbf{P}$ is called *the total momentum* of the system.

When we translate an object along the $y$-direction through a distance $a_y$ and then along the $x$-direction through $a_x$, we arrive at the same resultant configuration as when we carry these displacements in the opposite order. Hence $[\mathcal{G}(a_x, 0, 0), \mathcal{G}(0, a_y, 0)] = 0$, and it then follows that the infinitesimal generators must also commute:

$$[P_i, P_j] = 0, \quad i, j = x, y, z. \tag{8.2}$$

The group of spatial translations is then Abelian. As a consequence of of (8.2), we can integrate (8.1) to obtain the form of the unitary operator when the translation is finite:

$$U(\mathbf{a}) = \exp(-i\mathbf{P} \cdot \mathbf{a}). \tag{8.3}$$

The kets belonging to the translated observer are therefore given by

$$|a'; \mathbf{a}\rangle = e^{-i\mathbf{P}\cdot\mathbf{a}} |a'\rangle \,. \tag{8.4}$$

If the compatible set $A$ includes $\mathbf{P}$, then (8.4) is merely a phase change and $|a'\rangle$ and $|a'; \mathbf{a}\rangle$ belong to the same ray and represent the same physical state. To be precise, if both observers prepare states that belong to some particular eigenvalue $\mathbf{P}'$ of $\mathbf{P}$, and some other set of eigenvalues, they are preparing identical states in the sense that one of the observers will not be able to distinguish his own state from that of his colleague. Thus, *eigenstates of* $\mathbf{P}$ *are invariant under translations.* In particular, if $[H, \mathbf{P}] = 0$, there exist states that are simultaneously invariant under translations in space and time. In this situation, the Heisenberg picture momentum operator does not depend on time; in the Schrödinger picture, the expectation value of $\mathbf{P}$ in *any* state that satisfies the Schrödinger equation is conserved. This is the momentum conservation law.

Let the *first* observer measure the expectation value of an observable $X$ in a state $|a', \mathbf{a}\rangle$ prepared by the *displaced* observer:

$$\langle a'; \mathbf{a}|X|a'; \mathbf{a}\rangle = \langle a'|e^{i\mathbf{P}\cdot\mathbf{a}} X e^{-i\mathbf{P}\cdot\mathbf{a}}|a'\rangle. \tag{8.5}$$

If $\mathbf{a}$ is infinitesimal, (8.5) reads

$$\langle a'; \delta\mathbf{a}|X|a'; \delta\mathbf{a}\rangle = \langle a'|X + i\delta\mathbf{a} \cdot [\mathbf{P}, X]|a'\rangle. \tag{8.6}$$

Thus if $[\mathbf{P}, X] = 0$, the result of a measurement of the property $X$ does not depend on the choice of origin of the coordinate frame; *an observable is translation invariant if it* *commutes with* $\mathbf{P}$. (If $X$ commutes with only one of the $P_i$, it is only invariant under translations along the corresponding axis.) It is clear that $H$ must be translation invariant if $\mathbf{P}$ is to be a constant of motion.

We also note from (8.5) that the freedom of mathematical description that existed in the time displacement case exists here too. Thus the observer measuring properties of the displaced system can use his observable $X$ and the displaced kets $|a', \mathbf{a}\rangle$, or he can use the undisplaced kets and the unitary transform $e^{i\mathbf{P}\cdot\mathbf{a}} X e^{-i\mathbf{P}\cdot\mathbf{a}}$ of the observable. In the case of infinitesimal transformation, (8.6) tells us that this new observable is

$$X_{\delta\mathbf{a}} = X + i\delta\mathbf{a} \cdot [\mathbf{P}, X].$$

## 9. Rotations

Let $R_1, R_2, \ldots$ stand for an arbitrary proper rotations of a rigid coordinate frame in ordinary three-dimensional space. By "proper" we mean that the rotation does not include a reflection, i.e., that both the original and final frame are either right- or left-handed. The symbols $R_i$ stand for any convenient set of three real and continuous parameters; the most popular are the Euler angles, and the orientation of an axis plus an angle of rotation about that axis. If two rotations are carried out in the sequence $R_1$ first and then $R_2$, we shall designate the resulting orientation as $R_2R_1$. Note here that $R_2$ is a rotation away from the orientation previously achieved in the first rotation; it is not a rotation referred to some fixed reference frame. The Euler angles that correspond to $R_2R_1$ are complicated functions of those which separately parameterize $R_2$ and $R_1$, but we shall not require this relation. If $\mathcal{O}$ stands for no rotation whatever, then the inverse rotation to $R$, designated by $R^{-1}$, is defined by $RR^{-1} = R^{-1}R = \mathcal{O}$. The set of all proper rotations constitutes *the rotation group*.

We may think of the variously oriented coordinate systems as being attached to the measurement apparatus used in preparing states. The equivalence of the descriptions associated with these differently oriented frames implies that to each rotation $R$ there corresponds a unitary operator on the Hilbert space $D(R)$, with the properties

$$D(R_2)D(R_1) = D(R_2R_1), \tag{9.1}$$

$$D(R) = D^\dagger(R^{-1}). \tag{9.2}$$

The set of objects $\{D(R)\}$ constitutes a group of unitary operators. When $D(R)$ is applied to a ket $|a'\rangle$, it produces a ket

$$|a'; R\rangle = D(R)\,|a'\rangle$$

whose properties are identical to $|a'\rangle$ as far as an observer using a frame rotated by $R$ is concerned. Put more concisely, the configuration $|a'; R\rangle$ is obtained by rotating $|a'\rangle$ through $R$.

Let us begin by considering a subgroup, the rotations about an axis– say the $z$-axis. These rotations are parameterized by one angle $\alpha$, and the unitary operators that perform the correspondings mappings on the Hilbert space are defined by

$$D_z(\alpha_1)D_z(\alpha_2) = D_z(\alpha_1 + \alpha_2),$$

$$D_z(\alpha) = D_z^\dagger(-\alpha).$$

This subgroup is Abelian, and therefore has a very much simpler structure than the full group. In fact, the same argument as we used in the case of translations shows that

$$D_z(\alpha) = e^{-i\alpha J_z}, \tag{9.3}$$

where the hermitian operator $J_z$ is called the $z$-component of the total angular momentum. The analogy with translations is not as strong as one might think, however, because $\alpha$ is a confined parameter: $0 \le \alpha \le 2\pi$, whereas the translations are unrestricted. A group with a confined range of parameters is called *compact.* The spectrum of the generators (e.g. $J_z$) of a compact group is discrete, while that of an open group is, at least in part, continuous.

Consider two successive rotations about *different* axes. It is clear that a rotation through $\alpha'$ about the $y$-axis is given by $D_y(\alpha') = e^{-i\alpha' J_y}$. Previously we knew from geometry that displacements along different axes commute, and therefore concluded that $[P_x, P_y] = 0$, etc. Let us apply a similar argument to rotations. As the physical object that we rotate we take a vector $\mathbf{r}$, which can be thought of as specifying the position of a piece of measurement apparatus. We can describe the rotations conveniently by writing $\mathbf{r}$ as a column vector

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

We also introduce the $3 \times 3$ real antisymmetric matrices

$$I_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad I_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad I_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

One can easily verify that these matrices obey the commutation relations

$$[I_x, I_y] = I_z, \tag{9.4}$$

and cyclic permutations. These matrices can be used to describe rotations about the indicated axis, and by using all three of them in the appropriate combination any proper rotation can be carried out. Thus if we rotate $\mathbf{r}$ through $\phi$ about the $x$ axis,

$$\mathbf{r} \to \mathbf{r}' = e^{\phi I_x}\mathbf{r}. \tag{9.5}$$

The column vector on the RHS of (9.5) gives the components of the rotated vector with respect to the *original* coordinate system. For an infinitesimal rotation, (9.5) simplifies to

$$\mathbf{r} \to (1 + \delta\phi I_x + \cdots)\mathbf{r} = \begin{pmatrix} x \\ y - z\delta\phi \\ z + y\delta\phi \end{pmatrix} + O[(\delta\phi)^2].$$

It should be noted that for positive $\phi$ the sense of the rotation is counter-clockwise convention.

An arbitrary rotation can be obtained by applying the orthogonal matrix

$$\exp(\phi_x I_x)\exp(\phi_y I_y)\exp(\phi_z I_z) \tag{9.6}$$

to $\mathbf{r}$, but the three angles introduced thereby do not constitute a convenient parameterization, and will rarely be used here. It is important to note that (9.6) is characterized by the fact that it leaves the quadratic form $\mathbf{r} \cdot \mathbf{r}$ invariant, and has determinant $+1$.

We shall now carry out two different sequences of rotations on the vector $\mathbf{r}$. The first sequence consists of the rotation $\delta\phi_x$ about the $x$-axis followed by a rotation $\delta\phi_y$ about the *original* $y$-axis; it yields the vector

$$\mathbf{r}' = (1 + \delta\phi_y I_y + \cdots)(1 + \delta\phi_x I_x + \cdots)\mathbf{r}.$$

In the second sequence these rotations are performed in the opposite order, and we obtain instead

$$\mathbf{r}'' = (1 + \delta\phi_x I_x + \cdots)(1 + \delta\phi_y I_y + \cdots)\mathbf{r},$$

which does not equal $\mathbf{r}'$. In fact, because of (9.4),

$$\begin{aligned} \mathbf{r}'' - \mathbf{r}' &= [\delta\phi_x \delta\phi_y (I_x I_y - I_y I_x) + \cdots]\mathbf{r} \\ &= \delta\phi_x \delta\phi_y I_z \mathbf{r} + O(\delta\phi^3). \end{aligned} \tag{9.7}$$

As mentioned above, to each rotation of the measurement apparatus (i.e., rotation of $\mathbf{r}$) there corresponds, in a one-to-one fashion, a unitary mapping in the Hilbert space. Because we have already found that the unitary operators for rotations about a single axis are given by (9.3), we require the analog of (9.7) to be

$$[e^{-i\delta\phi_x J_x}, e^{-i\delta\phi_y J_y}] = (e^{-i\delta\phi_x \delta\phi_y J_z} - 1) + O(\delta\phi^3).$$

This condition determines the angular momentum commutation rules:

$$[J_x, J_y] = iJ_z, \quad \text{etc.} \tag{9.8}$$

This is then the algebra of the infinitesimal generators; it is obtained formally from the $I$'s by the substitution $I_x \to -iJ_x$, but we should remember that the $J$'s work in Hilbert space, whereas the $I$'s work on objects in three dimensions. The argument leading to (9.8) also shows that the angular momentum commutation rules depend only on the properties of rotations, and not at all on the detailed structure of the dynamical system in question.

We shall now construct the general rotation operator $D(R)$ parameterized with the Euler angles. These angles are defined by three successive rotations:

(a) a rotation through $\alpha$ about the $z$-axis, to which there corresponds the unitary operator $e^{-i\alpha J_z}$;

(b) a rotation through $\beta$ about the $y'$-axis, generated by the unitary operator $e^{-i\beta J_{y'}}$;

(c) a rotation through $\gamma$ about the $\zeta$-axis, generated by $e^{-i\gamma J_\zeta}$.

All these rotations are counter-clockwise when one is looking "down" the rotation axis towards the origin. The ranges of the Euler angles are

$$0 \le \alpha \le 2\pi, \quad 0 \le \beta \le \pi, \quad 0 \le \gamma \le 2\pi.$$

Upon putting the three operators just introduced together we obtain the unitary operator for an arbitrary rotation,

$$D(\alpha, \beta, \gamma) = e^{-i\gamma J_\zeta} e^{-i\beta J_{y'}} e^{-i\alpha J_z}.$$

This expression is not convenient since it refers to several coordinate frames. We may easily ammend this if we notice that $J_{y'}$ is related to $J_y$ by a unitary transformation. If we call $\{|a';\alpha\rangle\}$ a complete set prepared in the coordinate system rotated through $\alpha$ about the $z$-axis, and $\{|a'\rangle\}$ the corresponding set as prepared in the original frame, we have $|a';\alpha\rangle = e^{-i\alpha J_z}|a'\rangle$, and $\langle a';\alpha|J_{y'}|a'';\alpha\rangle = \langle a'|J_y|a''\rangle$. Hence

$$J_{y'} = e^{-i\alpha J_z} J_y e^{i\alpha J_z},$$

and of course

$$e^{-i\beta J_{y'}} = e^{-i\alpha J_z} e^{-i\beta J_y} e^{i\alpha J_z}.$$

Repeated application of this result yields

$$J_\zeta = e^{-i\beta J_{y'}} J_z e^{i\beta J_{y'}} = e^{-i\alpha J_z} e^{-i\beta J_y} J_z e^{i\beta J_y} e^{i\alpha J_z}.$$

Upon putting all this together we find the important and simple result

$$D(\alpha,\beta,\gamma) = e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z}.$$

Again, letting $R \equiv (\alpha,\beta,\gamma)$ we therefore have

$$|a';R\rangle = e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z}|a'\rangle .$$

The noncommutative algebra of the generators means that the determination of the matrix elements of $D$ is not a trivial task, as it was for the Abelian groups. There we found that in the representation which diagonalizes the generators (e.g., $\mathbf{P}$), the unitary transformation reduces to a phase factor. Here no representation can diagonalize all the generators because of (9.8). If we choose a representation that diagonalizes $J^2$ and $J_z$, we can make progress, nevertheless. The reason for this is that $J^2$ is a scalar, and hence rotations of its eigenvectors do not change the eigenvalue of $J^2$. Let $|jm\rangle$ be a simultaneous eigenket of $J^2$ and $J_z$. The other quantum numbers are now irrelevant; they belong to observables that commute with $\mathbf{J}$ and we suppress them. Because $[D(R),J^2] = 0$, the matrix of $D$ in this representation is diagonal in $j$. We can therefore confine ourselves to the matrix elements

$$\begin{aligned} D^{(j)}_{mm'}(\alpha,\beta,\gamma) &\equiv \langle jm|jm';R\rangle = \langle jm|D(\alpha,\beta,\gamma)|jm'\rangle \\ &= e^{-i\alpha m} d^{(j)}_{mm'}(\beta) e^{-i\gamma m'}, \end{aligned}$$

where

$$d_{mm'}(\beta) = \langle jm|e^{-i\beta J_y}|jm'\rangle.$$

The matrix $D^{(j)}$ has dimension $(2j+1)$, and is called the $(2j+1)$-dimensional irreducible representation of the rotation group. It is a representation because

$$D^{(j)}_{mm'}(R_1R_2) = \sum_{m''} D^{(j)}_{mm''}(R_1)D^{(j)}_{m''m'}(R_2). \tag{9.9}$$

This can be demonstrated immediately by taking the matrix elements of (9.1). Furthermore, the matrices associated with $R$ and $R^{-1}$ are related by

$$D^{(j)}_{mm'}(R) = D^{(j)}_{m'm}(R^{-1})^*, \tag{9.10}$$

as can be seen from (9.2). As an important case of (9.9) we may note

$$\sum_{m''} D^{(j)}_{mm''}(R)D^{(j)}_{m''m'}(R^{-1}) = \delta_{mm'}, \tag{9.11}$$

since

$$D^{(j)}_{mm'}(\mathcal{O}) = \delta_{mm'},$$

where $\mathcal{O}$ stands for no rotation whatsoever (the identity element in the language of group theory). Eq. (9.11) can be written in a somewhat more concrete fashion as

$$\sum_{m''} D^{(j)}_{mm''}(\alpha,\beta,\gamma)D^{(j)}_{m''m'}(-\gamma,-\beta,-\alpha) = \delta_{mm'}, \tag{9.12}$$

because $D(\alpha,\beta,\gamma) = D^{-1}(-\gamma,-\beta,-\alpha)$. By combining (9.10) and (9.11) we also obtain the unitarity relations

$$\sum_{m''} D^{(j)}_{mm''}(R)D^{(j)}_{m'm''}(R)^* = \sum_{m''} D^{(j)}_{m''m}(R)^*D^{(j)}_{m''m'}(R) = \delta_{mm'}. \tag{9.13}$$

## 10. Parity

Consider two observers with coordinate systems $(x,y,z)$ and $(\xi,\eta,\zeta)$, where $\xi=-x$, $\eta=-y$ and $\zeta=-z$. The second frame is obtained from the first by an inversion through the origin. The matrix which performs the inversion of a vector,

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

has determinant $-1$, whereas the rotations discussed earlier are generated by matrices with determinant 1. This is due to the fact that rotations can be obtained by a continuous transformation from unity, whereas the inversion cannot. (Note that the inversion through a plane, e.g., $x \to -x$, $y \to y$, $z \to z$, is also a discrete transformation, whereas $x \to -x$, $y \to -y$, $z \to z$, is not, as is of course clear since it is a rotation of $\pi$ around the $z$-axis.)

Let $\mathcal{R}$ be the operator in Hilbert space that connects an arbitrary ket $|a'\rangle$ to its reflected counterpart $|a';\mathrm{ref}\rangle$: $\mathcal{R}\,|a'\rangle = |a';\mathrm{ref}\rangle$. Since $\mathcal{R}^2$ must (to within a phase) equal one, we are now dealing with a group of two elements, and so we must first decide whether we are now to use a unitary or antiunitary operator $\mathcal{R}$.

This question is answered once we set up mappings in the Hilbert space that $\mathcal{R}$ is supposed to perform. Let us consider a single spinless particle, and the position and momentum eigenstates $|\mathbf{x}'\rangle$, $|\mathbf{p}'\rangle$. In order to be consistent with our definitions of the observables $\mathbf{x}$ and $\mathbf{p}$, we must have

$$\mathcal{R}\,|\mathbf{x}'\rangle = |-\mathbf{x}'\rangle\,, \quad \mathcal{R}\,|\mathbf{p}'\rangle = |-\mathbf{p}'\rangle\,. \tag{10.1}$$

In writing (10.1) we have adopted a convenient phase convention. Let us now apply (10.1) to $(\mathbf{x}-\mathbf{x}')\,|\mathbf{x}'\rangle = 0$: $(\mathcal{R}\mathbf{x}\mathcal{R}^{-1})\mathcal{R}\,|\mathbf{x}'\rangle = \mathbf{x}'\,|-\mathbf{x}'\rangle = -\mathbf{x}\,|-\mathbf{x}'\rangle$. Hence

$$\mathcal{R}\mathbf{x}\mathcal{R}^{-1} = -\mathbf{x}, \tag{10.2}$$

and the identical argument shows that

$$\mathcal{R}\mathbf{p}\mathcal{R}^{-1} = -\mathbf{p}. \tag{10.3}$$

Vectors that change sign under reflection are called *polar vectors*. From (10.2) and (10.3), on the other hand, the transformation law for orbital angular momentum is

$$\mathcal{R}(\mathbf{x}\times\mathbf{p})\mathcal{R}^{-1} = \mathbf{x}\times\mathbf{p}. \tag{10.4}$$

A vector that is invariant under reflection is called an *axial* or *pseudo-vector.*

To answer the question of whether $\mathcal{R}$ is unitary or not, we must bring in a further physical requirement. For example, we may use the geometric observation that a spatial transformation through $\mathbf{a}$ followed by a reflection is the same as a reflection followed by a translation by $-\mathbf{a}$:

$$\mathcal{R}e^{i\mathbf{a}\cdot\mathbf{P}} = e^{-i\mathbf{a}\cdot\mathbf{P}}\mathcal{R}, \tag{10.5}$$

where $\mathbf{P}$ is the total momentum operator, which of course satisfies

$$\mathcal{R}\,\mathbf{P}\,\mathcal{R}^{-1} \;=\; -\mathbf{P}. \tag{10.6}$$

We note that (10.5) and (10.6) are consistent if $\mathcal{R}$ is unitary, whereas if $\mathcal{R}$ were antiunitary, we would have

$$\mathcal{R}e^{i\mathbf{a}\cdot\mathbf{P}}\mathcal{R}^{-1} = e^{-i\mathbf{a}\cdot(\mathcal{R}\mathbf{P}\mathcal{R}^{-1})},$$

which is inconsistent. Hence $\mathcal{R}$ is a unitary operator.

Let us choose the phases such that $\mathcal{R}^2 = 1$. The eigenvalues $\mathcal{R}'$ of $\mathcal{R}$ are therefore $\mathcal{R}' = \pm 1$. These eigenvalues are called the *parity*. If the Hamiltonian is parity invariant, $\mathcal{R}H\mathcal{R}^{-1} = H$, or

$$[\mathcal{R}, H] = 0.$$

$\mathcal{R}$ is then a constant of motion, and the parity $\mathcal{R}'$ may be chosen as a member of the complete set of quantum numbers.

The behavior of the one-particle orbital angular momentum eigenstates under parity is of considerable importance in many applications. In view of (10.1) we have $\langle\mathbf{x}'|\mathcal{R}|lm\rangle = \langle -\mathbf{x}'|lm\rangle = Y_{lm}(-\mathbf{x}')$. But $Y_{lm}(\pi-\theta, \phi+\pi) = (-1)^l\langle\mathbf{x}'|lm\rangle$, and therefore

$$\langle\mathbf{x}'|lm;\mathrm{ref}\rangle = (-1)^l\langle\mathbf{x}'|lm\rangle.$$

The parity of a one-particle orbital angular momentum eigenstate is therefore $(-1)^l$.

In this discussion we have passed over the question of *intrinsic parity.* The existence of this notion can be recognized, for example, in the two nucleon system. The deuteron is a $^3S$-state, and therefore has the parity $+1$ in its center-of-mass system, if we *define* the parity of the neutron and proton as $+1$ in their respective center-of-mass frames. One could imagine, however, that the deuteron is in the state $^3P_1$, in which case the c.o.m (or intrinsic) parity would be $-1$. If the deuteron is treated as an elementary particle and placed into a state of orbital angular momentum $l$, the parity of the state would be $\pm(-1)^l$, depending on the internal structure of the deuteron. These two possibilities also exist for spinless particles as one realizes by constructing the $n-p$ states $^1S$ and $^3P_0$. The formulas should therefore contain a further factor $\pm 1$ to account for the intrinsic parity of the particle in question. The intrinsic parity of other "elementary" particles, such as the pion, is no longer arbitrary once an intrinsic parity has been assigned to the nucleons, and must be determined by experiment.

We may now draw some important conclusions concerning the matrix elements of operators between states of definite parity. Let $\mathcal{E}$ ($\mathcal{O}$) be an even (odd) observable under reflection, i.e., $\mathcal{R}\mathcal{E}\mathcal{R}^{-1} = \mathcal{E}$, $\mathcal{R}\mathcal{O}\mathcal{R}^{-1} = -\mathcal{O}$. Then

$$\langle\mathcal{R}'|\mathcal{E}|\mathcal{R}''\rangle = \langle\mathcal{R}'|\mathcal{R}^{-1}(\mathcal{R}\mathcal{E}\mathcal{R}^{-1})\mathcal{R}|\mathcal{R}''\rangle = \mathcal{R}'\mathcal{R}''\langle\mathcal{R}'|\mathcal{E}|\mathcal{R}''\rangle,$$

and similarly

$$\langle\mathcal{R}'|\mathcal{O}|\mathcal{R}''\rangle = \langle\mathcal{R}'|\mathcal{R}^{-1}(\mathcal{R}\mathcal{O}\mathcal{R}^{-1})\mathcal{R}|\mathcal{R}''\rangle = -\mathcal{R}'\mathcal{R}''\langle\mathcal{R}'|\mathcal{O}|\mathcal{R}''\rangle.$$

*Hence an observable even under reflections only has nonzero matrix elements between states of the same parity while an odd observable has nonvanishing matrix elements between states of opposite parity.*

The detailed properties of the reflection operator depend on the system to which it pertains. We shall illustrate this by constructing $\mathcal{R}$ in the spin space of a single spin 1/2 particle. Note that the spin is an axial vector,

$$\mathcal{R}\sigma\mathcal{R}^{-1} = \sigma.$$

The only $2 \times 2$ matrix that commutes with all three Pauli matrices is a multiple of the unit matrix. Hence we may choose the phases so that $\mathcal{R} = 1$. A reflection in a plane is a more interesting operation. The reflection in the $x - y$ plane produces the following transformations of the orbital angular momentum:

$$\begin{aligned}
\mathcal{R}_{xy}L_x\mathcal{R}_{xy}^{-1} &= -L_x, \\
\mathcal{R}_{xy}L_y\mathcal{R}_{xy}^{-1} &= -L_y, \\
\mathcal{R}_{xy}L_z\mathcal{R}_{xy}^{-1} &= +L_z,
\end{aligned}$$

where $\mathcal{R}_{xy}$ is the unitary operator associated with this reflection. These relations must hold for any angular momentum operator. In the spin space we therefore require an operator that anti-commutes with $\sigma_x$ and $\sigma_y$, and commutes with $\sigma_z$. These requirements are fulfilled by

$$\mathcal{R}_{xy} = \zeta_z\sigma_z, \tag{10.7}$$

where $\zeta_z$ is a number of modulus 1. Let us briefly verify that (10.7) is consistent with some other facts at our disposal. The reflection through the origin $\mathcal{R}$ is related to reflections in the plane through $\mathcal{R} = \mathcal{R}_{xy}\mathcal{R}_{yz}\mathcal{R}_{zx}$. From (10.7) and $\sigma_x\sigma_y\sigma_z = i$ we obtain $\mathcal{R} = -\zeta_x\zeta_y\zeta_z$. The sequence of reflections $\mathcal{R}_{yz}\mathcal{R}_{zx}$ is equivalent to a rotation through $\pi$ around the $z$ axis, which corresponds to the operator $\exp(-i\pi\sigma_z/2) = -i\sigma_z$, whereas $\mathcal{R}_{yz}\mathcal{R}_{zx} = -\zeta_y\zeta_x\sigma_z$. A consistent set of definitions for these operators can therefore be attained by setting $\zeta_x = \zeta_y = \zeta_z = i$.

*10.1. Space inversion is not a symmetry of Nature*

A giant step in the understanding of weak interactions came in 1956 when T.D. Lee and C.N. Yang pointed out that there was no evidence in favor of parity conservation in weak interactions. The precipitating issue was the $\tau - \theta$ puzzle. The $\tau$ was the $3\pi$ decay of the $K^+$. The analysis begun by Dalitz had shown that the $3\pi$ system had $J^P = 0^-, 2^-, \ldots$. On the other hand, the $\theta^+$ decayed into $\pi^0\pi^+$ and had $J^P = 0^+, 1^-, \ldots$. Measurements showed that the masses and lifetimes of the $\theta$ and $\tau$ were very similar, perhaps equal. The $\theta$ and the $\tau$ seemed to be the same particle, except that they had different values of $J^P$. Despite the overwhelming prejudice that parity must be a good symmetry because it was a symmetry of space itself just as rotational invariance is, Lee and Yang demonstrated that there was no evidence for or against parity conservation in weak interactions. To test for a possible violation of parity it was necessary to observe a dependence of a decay rate (or cross section) on a term that changed sign under parity operation. In nuclear decay, the momenta available are $\mathbf{p}_e$, $\mathbf{p}_\nu$ and $\mathbf{p}_N$. The invariant $\mathbf{p}_e \cdot \mathbf{p}_\nu \times \mathbf{p}_N$ would change sign under parity, but vanishes because the momenta are coplanar. To test for parity in nuclear beta decay required consideration of spin. If the decaying nucleus were oriented, it would be possible to measure the angular dependence of the decay, looking for a term proportional to $\langle \mathbf{J} \rangle \cdot \mathbf{p}_e$. This was achieved by C.S. Wu in collaboration with E. Ambler and co-workers (C.S. Wu *et al.*, Phys. Rev. 105 (1957) 1413).

Wu and her coworkers chose to work with $Co^{60}$, whose ground state has $J^P = 5^+$. It beta-decays through a Gamow-Teller transition with a half-life of 5.2 y, yielding $Ni^{60}$ in the $4^+$ state. The excited Ni state decays through two successive $\gamma$ emissions to $2^+$ and then $0^+$, with $\gamma$ energies 1.173 and 1.332 MeV, respectively. Nuclear polarization was produced through adiabatic demagnetization. The degree of polarization of the Co sample was monitored by observing the anisotropy of the gamma radiation. The polarization of the $Co^{60}$ was transmitted to the $Ni^{60}$, giving a difference between the rates for gamma emission in the polar and equatorial directions, relative to the axis of the applied magnatic field.

The beta-decay rate along the direction of the magnaric field, that is, along the nuclear polarization direction was monitored. Reversing the magnatic field reversed the direction of $\langle \mathbf{J} \rangle$. The counting rate indeed showed a dependence on $\langle \mathbf{J} \rangle \cdot \mathbf{p}_e$. Not only was the rate different for the two magnetic field orientations, but as the sample warmed, the dependence of the rate on the field orientation disappeared at the same speed as the polarization itself disappeared, showing the connection of the decay angular distribution was with the nuclear orientation, not simply with the applied magnetic field.

Word of this *tour de force* spread rapidly and new experiments were undertaken before the results of Wu's team appeared in print. Indeed, two further experiments appeared in rapid succession showing parity violation in the sequence $\pi^+ \to \mu^+ \to e^+$. Rather than beginning with a polarized beam, these experiments exploited the prediction of Lee and Yang that parity violation would lead to polarization of the $\mu$ along its line of flight in the $\pi \to \mu\nu$ decay. The polarization of the $\mu$ is retained when it slows down in matter. A distribution of decay electrons relative to the incident beam direction of the form $1 + a\cos\theta$ is then expected, where $a$ depends on the degree of polarization of the $\mu$. Garwin, Lederman and Weinrich (R.L. Garwin *et al.*, Phys. Rev. 105 (1957) 1415) applied a magnetic field to the region where the muons stopped. This caused the spin of the muon to precess. In this elegant fashion, they demonstrated parity violation, measured its strength and simultaneously measured the magnetic moment of the $\mu^+$ by measuring the rate of precession. At the same time, Friedman and Telegdi (J.I. Friedman and V.L. Telegdi, Phys. Rev. 105 (1957) 1681) also found parity violation by observing the same decay sequence, but working in emulsions and without a magnetic field. The emulsion experiment was started before the others, but took longer to complete because of the laborious scanning procedure.

## 11. Time Reversal

Consider a classical particle subjected to some (unknown) static forces. Let $\mathbf{r}(t_0)$ and $\mathbf{v}(t_0)$ be the position and velocity of the particle at $t = t_0$, and allow it to proceed undisturbed for a time $t_1$ when its position and velocity have become $\mathbf{r}(t_0 + t_1)$, $\mathbf{v}(t_0 + t_1)$. Now (i.e., at $t = t_0 + t_1$) start another identical particle off at $\mathbf{r}(t_0 + t_1) \equiv \mathbf{r}'$ with velocity $-\mathbf{v}(t_0 + t_1) \equiv \mathbf{v}'$. Then at later time $t_0 + 2t_1$, we will either find that the new position and velocity equal $\mathbf{r}(t_0)$ and $-\mathbf{v}(t_0)$, or we will not. In the first case, we say that the basic laws (equation of motion) are invariant under time reversal; this is clearly the case for motion in a force field derivable from a potential. If, on the other hand, we consider the motion of a charged particle in a prescribed magnetic field, there is no symmetry of this kind.

To summarize, the classical equations of motion are said to be invariant under time reversal if one can separate the manifold of all motions into two subsets between which there is one-to-one correspondence that is affected by changing the signs of all momenta at some instant. Let $q_I(t)$ and $p_I(t)$ be the canonical variables of the motion belonging to subset $I$, and $q_{II}(t)$ and $p_{II}(t)$ the variables of subset $II$. The one-to-one coprrespondence just referred to is then

$$\begin{aligned} q_I(t_0+t) &= q_{II}(t_0-t), \\ p_I(t_0+t) &= -\,p_{II}(t_0-t). \end{aligned} \tag{11.1}$$

In the case of a motion in a magnetic field, the correspondence (11.1) does not exist. Instead one has

$$\begin{aligned} q_I(\mathcal{H},t_0+t) &= q_{II}(-\mathcal{H},t_0-t), \\ p_I(\mathcal{H},t_0+t) &= -\,p_{II}(-\mathcal{H},t_0-t). \end{aligned} \tag{11.2}$$

Let $|\psi\rangle$ be a physical state, and $|\psi'\rangle$ the motion-reversed state, both at $t=0$, say. For example, if $|\psi\rangle$ is the momentum eigenstate of a single particle $|\mathbf{p}'\rangle$, then $|\psi'\rangle = |-\mathbf{p}'\rangle$. Here it is important to notice that we have not assumed any invariance because at any instant we can always set up this correspondence. Define the operator $\Theta$ by $\Theta\,|\psi\rangle = |\psi'\rangle$. (In the one-particle state, it is easiest to define it in the $p$-basis.) Clearly $\Theta^2 = 1$ (to within a phase) and $\Theta$ can therefore be unitary or antiunitary. Consider the $p$-basis. We have $\mathbf{p}\,|-\mathbf{p}'\rangle = -\mathbf{p}'\,|-\mathbf{p}'\rangle$; applying $\Theta$ to this equation gives

$$\Theta\mathbf{p}\Theta^{-1}\Theta\,|-\mathbf{p}'\rangle = -\mathbf{p}'\Theta\,|-\mathbf{p}'\rangle\,,$$

or

$$\Theta\mathbf{p}\Theta^{-1} = -\mathbf{p}.$$

Hence, at least for a free particle with Hamiltonian $H = p^2/2m$, $\Theta H = H\Theta$, which is hardly surprising.

Now comes the crucial point: we demand that the correspondence between motion reversed states be maintained for all $t$. (This will of course lead to a condition on $H$ among other things.) We formulate this requirement in complete analogy to (11.1). Equation (11.1) asserts that the configuration attained by permitting the system to evolve for a time $t$ and then reversing all momenta is indistinguishable from the configuration obtained by reversing the momenta at $t_0$ and evolving "backwards" for a period $t$. The analogous condition in quantum mechanics is

$$\Theta e^{-iHt}\,|\psi\rangle = e^{iHt}\Theta\,|\psi\rangle\,,$$

where $t_0$ was put to zero. There are again two choices:

$$\Theta \text{ is unitary}, \quad \{\Theta, H\} = 0,$$
$$\Theta \text{ is antiunitary}, \quad [\Theta, H] = 0.$$

The example of the free particle is already enough to show that $\Theta$ *must be antiunitary.* This important fact was discovered by Wigner in 1931.

Let us now consider the other observables. We certainly want $\Theta$ to leave space coordinates unchanged:

$$\Theta \mathbf{x} \Theta^{-1} = \mathbf{x}. \tag{11.3}$$

It then follows that $\Theta$ anticommutes with $\mathbf{x} \times \mathbf{p}$, and we are therefore justified in requiring

$$\Theta \mathbf{J} \Theta^{-1} = -\mathbf{J} \tag{11.4}$$

for any angular momentum operator.

If we demand invariance under $t$-reversal, we can discard certain kinds of Hamiltonians. For example, in the case of one particle, $H$ cannot contain the term $\mathbf{x} \cdot \sigma$. Of course $\mathbf{x} \cdot \sigma$ could have been discarded by the requirement of parity invariance. Quite generally, the Hamiltonian of spin 1/2 particle in a central field cannot be a true scalar and odd under time reversal. There is such a term, however, in the interaction of two spin 1/2 particles, namely

$$H' = C(\sigma_1 \times \sigma_2) \cdot \mathbf{L}_{\text{rel}},$$

where $\mathbf{L}_{\text{rel}}$ is the relative angular momentum. Clearly $\mathcal{R} H' \mathcal{R}^{-1} = H'$, $\Theta H' \Theta^{-1} = -H'$. Experiments that look for such a term in nucleon-nucleon scattering have been done, but no signal was found so far.

### *11.1. Spin 0 particles*

Because of (11.3), the coordinate eigenstates are unaffected by $\Theta$. If $|\psi'\rangle$ is the time-reversed state of a single spin 0 particle, its wave function is therefore given by $\psi'(\mathbf{x}) \equiv \langle \mathbf{x} | \psi' \rangle = \langle \psi | \mathbf{x} \rangle$, where the last equation follows from the antiunitarity of $\Theta$. Hence

$$\psi'(\mathbf{x}) = \psi^*(\mathbf{x}).$$

This conclusion can be drawn directly from the Schrödinger equation when the Hamiltonian is real. By taking the complex conjugate of $(i\partial_t - H)\psi(\mathbf{x}, t) = 0$ we obtain

$$(i\partial_t + H)\psi^*(\mathbf{x}, t) = 0.$$

Therefore $\psi(\mathbf{x},t)$ and $\psi^*(\mathbf{x},-t)$ satisfy the same equation. From this argument we also conclude that in the $x$-basis the time reversal operator $\Theta$ is simply a complex conjugation operator $K$.

The significance of time reversal is best appreciated when one also considers the initial condition, and not only the evolution of the system in $t$. The scattering problem is ideally suited for this. Consider a particle of momentum $\mathbf{p}$ incident on the potential $V$. The state vector that satisfies outgoing-wave boundary conditions is $\left|\mathbf{p}^{(+)}\right\rangle$:

$$\left|\mathbf{p}^{(+)}\right\rangle = |\mathbf{p}\rangle + \frac{1}{E - H_0 + i\epsilon} V \left|\mathbf{p}^{(+)}\right\rangle.$$

We now assume that $\Theta V \Theta^{-1} = V$, and investigate the time reversed situation, $\Theta\left|\mathbf{p}^{(+)}\right\rangle$. Because $H_0$ commutes with $\Theta$,

$$\Theta\left|\mathbf{p}^{(+)}\right\rangle = |-\mathbf{p}\rangle + \frac{1}{E - H_0 - i\epsilon} V\Theta \left|\mathbf{p}^{(+)}\right\rangle. \tag{11.5}$$

Two things have happened: the "incident" state has momentum $-\mathbf{p}$, and the Green's operator is $(E - H_0 - i\epsilon)^{-1}$, and represents *incoming* spherical waves, because

$$\left\langle \mathbf{x} \left| \frac{1}{E - H_0 - i\epsilon} \right| \mathbf{x}' \right\rangle = -\frac{2m}{4\pi} \frac{e^{-ip|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}$$

in contrast to

$$\left\langle \mathbf{x} \left| \frac{1}{E - H_0 + i\epsilon} \right| \mathbf{x}' \right\rangle = -\frac{2m}{4\pi} \frac{e^{ip|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}.$$

All this is just as it should be. The state $\left|\mathbf{p}^{(+)}\right\rangle$ is a stationary-state idealization of the situation actually attained in the laboratory when we send in a beam with momentum $\mathbf{p}$ and eventually observe it scattered outwards from the target. The motion reversed situation is of course impossible to achieve in practice. Here we start with a spherical wave converging to the target, which, by an incredible feat of coherence, then combines to form a collimated beam of momentum $-\mathbf{p}$ subsequent to the collision. Eq. (11.5) shows that $\Theta\left|\mathbf{p}^{(+)}\right\rangle$ describes precisely this situation.

A more trivial example is a particle in an angular momentum eigenstate, wherein its wave function contains the factor $Y_{lm}(\theta,\phi)$. The time reversed state is $Y^*_{lm}(\theta,\phi) = (-1)^m Y_{l,-m}(\theta,\phi)$. In the time reversed situation the particle travels about in an opposite direction and therefore has its component of angular momentum about *any* axis reversed.

Another simple but important consequence of time reversal is the following. If $\psi_X(\mathbf{x})$ is an energy eigenfunction, and $\Theta H \Theta^{-1} = H$, then $\psi^*_E(\mathbf{x})$ is also an eigenfunction of energy $E$. There are then two possibilities: Either the level is degenerate and $\psi^*_E(\mathbf{x})$ and $\psi_E(\mathbf{x})$ are *really* different (i.e. belong to different rays), or there is no degeneracy and $\psi_E \propto \psi^*_E$. Hence for a time reversal invariant $H$ the nondegenerate eigenfunctions can always be chosen to be real.

*11.2. Spin 1/2 particles*

We begin with a single spin 1/2 particle. The $p$-space and $x$-space kets will be written as $|\mathbf{p}\ m_s\rangle$ and $|\mathbf{x}\ m_s\rangle$, respectively, where $m_s = \pm 1/2$. As the spin is an angular momentum we tentatively set

$$\Theta\,|\mathbf{p}\ m_s\rangle = |-\mathbf{p}\ \ -m_s\rangle\,, \tag{11.6}$$

$$\Theta\,|\mathbf{x}\ m_s\rangle = |\mathbf{x}\ \ -m_s\rangle\,. \tag{11.7}$$

Let us work in the $|\mathbf{x}\ m_s\rangle$ basis. We write $\Theta$ in the form

$$\Theta = UK \tag{11.8}$$

(where $K$ means that one takes the complex conjugates of all the expansion coefficients of arbitrary kets) and seek $U$. For the position kets, $U = 1$, and we need only determine $U$ in the spin-space. According to (11.4) we require

$$UK\sigma KU^{\dagger} = -\sigma = U\sigma^{*}U^{\dagger}.$$

In the Pauli representation of spin,

$$\sigma_x^{*} = \sigma_x, \quad \sigma_y^{*} = -\sigma_y, \quad \sigma_z^{*} = \sigma_z,$$

and therefore we need

$$U\sigma_x U^{\dagger} = -\sigma_x, \quad U\sigma_y U^{\dagger} = \sigma_y, \quad U\sigma_z U^{\dagger} = -\sigma_z,$$

which is satisfied by

$$U = e^{i\delta}\sigma_y. \tag{11.9}$$

Note that (11.9) represents a rotation of $\pi$ about the $y$-axis, which certainly will make a state running clockwise about the $z$-axis run counter-clockwise. We also note from the form of the Pauli matrices that

$$\Theta\,|1/2\rangle = -ie^{i\delta}\,|-1/2\rangle\,, \quad \Theta\,|-1/2\rangle = ie^{i\delta}\,|1/2\rangle\,,$$

and therefore no choice of $\delta$ will satisfy (11.7) with the phase $+1$ throughout. This is because we have held to our convention that $J_{\pm}\,|jm\rangle = \sqrt{j(j+1) - m(m\pm 1)}\,|jm \pm 1\rangle$. The phases in these relations are also at our disposal and could have been chosen in accordance with (11.7). In the same way the $Y_{lm}$'s could be redefined so that $Y_{lm}^{*} = Y_{l,-m}$.

For some purposes the phases that simplify $t$-reversal considerations are convenient, but we shall not bother to go into this because our present convention is perfectly consistent and actually easier to remember for most purposes. With our phase conventions the correct relations are not (11.6) and (11.7), but

$$\Theta \left|\mathbf{p}\ m_s\right\rangle = e^{i\delta}(-i)^{2m_s}\left|-\mathbf{p}\ -m_s\right\rangle, \tag{11.10}$$

$$\Theta \left|\mathbf{x}\ m_s\right\rangle = e^{i\delta}(-i)^{2m_s}\left|\mathbf{x}\ -m_s\right\rangle. \tag{11.11}$$

The operator for a single spin 1/2 particle is then

$$\Theta = e^{i\delta}\sigma_y K. \tag{11.12}$$

The square of this operator turns out to have considerable significance. Its evaluation proceeds as follows: $\Theta^2 = e^{i\delta}\sigma_y e^{-i\delta}\sigma_y^* K^2 = -\sigma_y^2 = -1$. This result does not depend on the phase conventions, nor on the representation. Thus

$$\Theta^2 = \begin{cases} +1 & \text{for a spin-zero particle,} \\ -1 & \text{for a spin-1/2 particle.} \end{cases} \tag{11.13}$$

For a system of $N$ spin 1/2 particles we clearly have

$$\Theta = \Pi_{i=1}^N \sigma_y^{(i)} K$$

(aside from an arbitrary phase), where $\sigma_y^{(i)}$ is the $i$th particle's spin operator. Hence for $N$ spin 1/2 particles

$$\Theta^2 = (-1)^N.$$

Let $\{|E, N\rangle\}$ be the energy eigenstates of a system of $N$ spin 1/2 particles. If the Hamiltonian of this system is invariant under time reversal, $\Theta\,|E, N\rangle$ is also a stationary state of energy $E$. We may then ask whether this is a new state, i.e., one that is linearly independent of $|E, N\rangle$, or simply $|E, N\rangle$ apart from a possible phase. If $\Theta^2 = 1$ one cannot answer this question in general, i.e., the level $E$ may or may not be degenerate. If $\Theta^2 = -1$, however ($N$ odd), it is clear that $\Theta\,|E, N\rangle$ cannot be the same state (i.e., belong to the same ray $|E, N\rangle$) because $\Theta\,|E\rangle = C\,|E\rangle$ implies $\Theta^2\,|E\rangle = |C|^2\,|E\rangle = -\,|E\rangle$ which is inconsistent. (This argument also shows that one cannot associate an eigenvalue with the operator $\Theta$ in general.) Hence if $N$ is odd, $\Theta\,|E, N\rangle$ and $|E, N\rangle$ are linearly independent, and therefore all levels are at least doubly degenerate. In fact, a bit of thought shows that the order of the degeneracy must be even. This type of degeneracy is called *Kramers degeneracy* and is a useful tool for treating complicated problems with few obvious symmetries such as occur in molecular and solid state physics.

## 12. Lorentz Transformations

### *12.1. The Lorentz Group*

The postulates of Special Relativity tell us that the speed of light is the same in all inertial frames. This means that if $x_i$ is the position of a light signal at time $t$ in one frame, and the same light ray is found at $x'_i$ at time $t'$ in another frame, we must have

$$s^2 \equiv c^2t^2 - x_i x_i = c^2t'^2 - x'_i x'_i. \tag{12.1}$$

The set of linear transformations which relate $(x'_i, t')$ to $(x_i, t)$ while preserving the above expressions form a group called the Lorentz group. Choose units such that $c = 1$, and introduce the notation

$$x^\mu,\ \ \mu = 0, 1, 2, 3 \ \text{ with } \ x^0 = t, \ \ (x^1, x^2, x^3) = (x^i) = \vec{x},$$

i.e.

$$x^\mu = (x^0, x^i) = (t, \vec{x}), \quad i = 1, 2, 3.$$

In this compact notation, $s^2$ can be written as

$$s^2 = x^0x^0 - x^ix^i \equiv x^\mu x^\nu g_{\mu\nu}, \tag{12.2}$$

where the metric $g_{\mu\nu} = g_{\nu\mu}$ is zero except for $\mu = \nu$ where $g_{00} = -g_{11} = -g_{22} = -g_{33} = 1$. Repeated indices are summed except when otherwise indicated. Then eq. (12.1) becomes

$$g_{\mu\nu}x^\mu x^\nu = g_{\mu\nu}x'^\mu x'^\nu. \tag{12.3}$$

Now look for a set of linear transformations

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu = \Lambda^\mu{}_0 x^0 + \Lambda^\mu{}_i x^i, \tag{12.4}$$

which preserves $s^2$. The $\Lambda^\mu{}_\nu$ must therefore satisfy

$$g_{\mu\nu}x'^\mu x'^\nu = g_{\mu\nu}\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma x^\rho x^\sigma = g_{\rho\sigma}x^\rho x^\sigma. \tag{12.5}$$

As (12.5) must hold for any $x^\mu$, we conclude that

$$g_{\rho\sigma} = g_{\mu\nu}\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma. \tag{12.6}$$

It is more convenient for certain purposes to use a matrix notation: regard $x^\mu$ as a column vector $\mathbf{x}$ and $g_{\mu\nu}$ as a matrix $\mathbf{g}$. Then

$$s^2 = \mathbf{x}^T \mathbf{g} \mathbf{x}, \tag{12.7}$$

and

$$\mathbf{x}' = \mathbf{L}\mathbf{x}, \tag{12.8}$$

where $\mathbf{L}$ is the matrix equivalent of the $\Lambda^\mu{}_\nu$ coefficients, and $T$ means the transpose. The $\mathbf{L}$'s must obey

$$\mathbf{g} = \mathbf{L}^T \mathbf{g} \mathbf{L}, \tag{12.9}$$

to be Lorentz transformations (LT). Examine the consequences of eq. (12.9). First take its determinant

$$\det \mathbf{g} = \det \mathbf{L}^T \det \mathbf{g} \det \mathbf{L}, \tag{12.10}$$

from which we deduce that

$$\det \mathbf{L} = \pm 1. \tag{12.11}$$

The case $\det \mathbf{L} = +1(-1)$ corresponds to proper (improper) LT's. As an example, the LT given numerically by $\mathbf{L} = \mathbf{g}$ is an improper one; physically it corresponds to $x^0 \to x^0$, $x^i \to -x^i$, i.e. space inversion. Second, take the 00 entry of eq. (12.6):

$$1 = \Lambda^\rho{}_0 g_{\rho\sigma} \Lambda^\sigma{}_0 = (\Lambda^0{}_0)^2 - (\Lambda^i{}_0)^2, \tag{12.12}$$

which shows that

$$|\Lambda^0{}_0| \geq 1. \tag{12.13}$$

When $\Lambda^0{}_0 \geq 1$, the LT's are said to be orthochronous, while $\Lambda^0{}_0 \leq -1$ gives non-orthochronous LT's. It follows then that LT's can be put in four categories:

1) proper orthochronous ($L_+^\uparrow$): $\det \mathbf{L} = +1$, $\Lambda^0{}_0 \geq +1$.
2) proper non-orthochronous ($L_+^\downarrow$): $\det \mathbf{L} = +1$, $\Lambda^0{}_0 \leq -1$.
3) improper orthochronous ($L_-^\uparrow$): $\det \mathbf{L} = -1$, $\Lambda^0{}_0 \geq +1$.
4) improper non-orthochronous ($L_-^\downarrow$): $\det \mathbf{L} = -1$, $\Lambda^0{}_0 \leq -1$.

Let us give a few examples:

a) Rotations: $x'^0 = x^0$, $x'^i = a^{ij} x^j$ with $a^{ij}$ an orthogonal matrix. Then we can write $\mathbf{L}$ in the block form

$$\mathbf{L} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{a} \end{pmatrix}, \tag{12.14}$$

so that $\det \mathbf{L} = \det \mathbf{a}$. We can have $\det \mathbf{a} = \pm 1$, corresponding to proper and improper rotations, with $\mathbf{L}$ belonging to $L_+^\uparrow$ and $L_-^\uparrow$, respectively.

b) Boosts: the transformations

$$\begin{aligned} x'^0 =& x^0 \cosh\eta - x^1 \sinh\eta, \\ x'^1 =& -x^0 \sinh\eta + x^1 \cosh\eta, \\ x'^{2,3} =& x^{2,3}, \end{aligned} \tag{12.15}$$

describe a boost in the 1-direction. Then, in a $2 \times 2$ block form

$$\mathbf{L} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.16}$$

$$\det \mathbf{L} = \cosh^2\eta - \sinh^2\eta = 1, \tag{12.17}$$

$$\Lambda^0{}_0 = \cosh\eta \geq 1. \tag{12.18}$$

This transformation therefore belongs to $L_+^\uparrow$. Note that the identifications

$$\cosh\eta = (1 - v^2)^{-1/2}, \quad \sinh\eta = v(1 - v^2)^{-1/2}, \tag{12.19}$$

where $v$ is the velocity of the boosted frame leads to the more familiar form.

c) Time inversion: defined by $x'^0 = -x^0$, $x'^i = x^i$. It has $\det \mathbf{L} = -1$, $\Lambda^0{}_0 = -1$, and therefore belongs to $L_-^\downarrow$.

d) Full inversion: defined by $x'^\mu = -x^\mu$. It has $\det \mathbf{L} = +1$, $\Lambda^0{}_0 = -1$, and belongs to $L_+^\downarrow$. Full inversion can be obtained as the product of a space and a time inversion.

Any LT can be decomposed as the product of transformations of these four types. Thus, it suffices to concentrate on rotations and boosts. Since there are three rotations and three boosts, one for each space direction, the LT are described in terms of six parameters. We now proceed to build the six corresponding generators. Consider an infinitesimal LT

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu, \tag{12.20}$$

where $\delta^\mu{}_\nu$ is the Kronecker delta which vanishes when $\mu \neq \nu$ and equals $+1$ otherwise, Evaluation of (12.6) yields to $O(\omega)$

$$0 = g_{\nu\rho}\omega^\rho{}_\mu + g_{\mu\rho}\omega^\rho{}_\nu. \tag{12.21}$$

We use the metric $g_{\mu\nu}$ to lower indices, for example

$$x_\mu \equiv g_{\mu\nu} x^\nu = (x^0, -\vec{x}).$$

Eq. (12.21) becomes

$$0 = \omega_{\nu\mu} + \omega_{\mu\nu}, \tag{12.22}$$

that is, $\omega_{\mu\nu}$ is an antisymmetric tensor with (as advertised) $\frac{4\cdot 3}{2} = 6$ independent entries. Introduce the hermitian generators

$$L_{\mu\nu} \equiv i(x_\mu \partial_\nu - x_\nu \partial_\mu), \tag{12.23}$$

where $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$. In terms of these we can write

$$\delta x^\mu = \frac{i}{2} \omega^{\rho\sigma} L_{\rho\sigma} x^\mu = \omega^{\mu\rho} x_\rho. \tag{12.24}$$

It is easy to see that the $L_{\mu\nu}$'s satisfy a Lie algebra

$$[L_{\mu\nu}, L_{\rho\sigma}] = i g_{\nu\rho} L_{\mu\sigma} - i g_{\mu\rho} L_{\nu\sigma} - i g_{\nu\sigma} L_{\mu\rho} + i g_{\mu\sigma} L_{\nu\rho}, \tag{12.25}$$

to be identified with the Lie algebra of $SO(3,1)$. The most general representation of the generators of $SO(3,1)$ that obeys the commutation relations (12.25) is given by

$$M_{\mu\nu} \equiv i(x_\mu \partial_\nu - x_\nu \partial_\mu) + S_{\mu\nu}, \tag{12.26}$$

where the hermitian $S_{\mu\nu}$ satisfy the same Lie algebra as the $L_{\mu\nu}$ and commute with them. The hermitian generators $M_{ij}$ for $i, j = 1, 2, 3$ form an algebra among themselves

$$[M_{ij}, M_{kl}] = -i\delta_{jk} M_{il} + i\delta_{ik} M_{jl} + i\delta_{jl} M_{ik} - i\delta_{il} M_{jk}, \tag{12.27}$$

which is that of the rotation group $SU(2)$. A more familiar expression can be obtained by introducing the new operators

$$J_i \equiv \frac{1}{2} \epsilon_{ijk} M_{jk}. \tag{12.28}$$

Then we find

$$[J_i, J_j] = i\epsilon_{ijk} J_k. \tag{12.29}$$

Define the boost generators

$$K_i \equiv M_{0i}. \tag{12.30}$$

It follows from the Lie algebra that

$$\begin{aligned}
[K_i, K_j] =& - i\epsilon_{ijk}J_k, \\
[J_i, K_j] =& i\epsilon_{ijk}K_k.
\end{aligned} \tag{12.31}$$

Here, both $K_i$ and $J_i$ are hermitian generators. The $K_i$ are non-compact generators. We can disentangle these commutation relations by introducing the new linear combinations

$$N_i \equiv \frac{1}{2}(J_i + iK_i). \tag{12.32}$$

Although not hermitian, $N_i \neq N_i^\dagger$, they have the virtue of yielding simple commutation relations:

$$\begin{aligned}
[N_i, N_j^\dagger] =& 0, \\
[N_i, N_j] =& i\epsilon_{ijk}N_k, \\
[N_i^\dagger, N_j^\dagger] =& i\epsilon_{ijk}N_k^\dagger.
\end{aligned} \tag{12.33}$$

This means that the $N_i$ and the $N_i^\dagger$ independently obey the Lie algebra of $SU(2)$. We can further appeal to its well known representation theory. In particular, we have two Casimir operators (operators that commute with all the generators):

$$\begin{aligned}
&N_i N_i \text{ with eigenvalues } n(n+1), \\
&N_i^\dagger N_i^\dagger \text{ with eigenvalues } m(m+1),
\end{aligned} \tag{12.34}$$

where $m, n = 0, 1/2, 1, 3/2, \ldots$. The representations are then labelled by the pair $(n, m)$ while the states within a representation are further distinguished by the eigenvalues of $N_3$ and $N_3^\dagger$. Observe that the two $SU(2)$'s are not independent as they can be interchanged by the operation of parity, $P$,

$$J_i \to J_i, \qquad K_i \to -K_i,$$

and the operation of hermitian conjugation which changes the sign of $i$ and therefore switches $N_i$ to $N_i^\dagger$. In general, representations of the Lorentz group are neither parity nor (hermitian) conjugation eigenstates. Since $J_i = N_i + N_i^\dagger$, we can identify the spin of the representation with $m + n$. As an example, consider the following representations:

a) $(0, 0)$ with spin zero is the scalar representation, and it has a well defined parity (can appear as scalar or pseudoscalar);
b) $(1/2, 0)$ has spin $\frac{1}{2}$ and represents a left-handed spinor (the handedness is a convention);

c) $(0, 1/2)$ describes a right-handed spinor.

These spinors have two components ("spin up" and "spin down"); they are called Weyl spinors. When parity is relevant, one considers the linear combination $(0, 1/2) \oplus (1/2, 0)$, which yields a Dirac spinor.

Given the spinor representations, we can generate any other representation by multiplying them together. For example,

d) $(1/2, 0) \otimes (0, 1/2) = (1/2, 1/2)$ gives a spin 1 representation with four components. In tensor notation it is written as a 4-vector.

e) $(1/2, 0) \otimes (1/2, 0) = (0, 0) \oplus (1, 0)$. Here the scalar representation is given by the antusymmetric product. The new representation (1,0) is represented by an antisymmetric, self-dual second rank tensor, i.e., a tensor $B_{\mu\nu}$ which obeys

$$B_{\mu\nu} = -B_{\nu\mu}, \quad B_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu}{}^{\rho\sigma}B_{\rho\sigma}. \tag{12.35}$$

The (0,1) representation corresponds to a tensor that is antiself-dual,

$$B_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu}{}^{\rho\sigma}B_{\rho\sigma}. \tag{12.36}$$

For example, Maxwell's field strength tensor $F_{\mu\nu}$ transforms under the Lorentz group as $(0, 1) \oplus (1, 0)$.

### *12.2. The Poincaré Group*

Another fundamental principle is the invariance of the behavior of an isolated system under uniform translation in space and time. Such a transformation is given by

$$x^\mu \to x'^\mu = x^\mu + a^\mu, \tag{12.37}$$

where $a^\mu$ is an arbitrary constant four-vector. Hence the general invariance group is a ten-parameter group called the Poincaré group, under which

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu. \tag{12.38}$$

The translations (12.37) do not commute with the LT's. Indeed consider two successive PG transformations with parameters $(\Lambda_1, a_1)$, $(\Lambda_2, a_2)$,

$$x^\mu \to \Lambda^\mu_{1\nu}x^\nu + a_1^\mu \to \Lambda^\mu_{2\rho}\Lambda^\rho_{1\nu}x^\nu + \Lambda^\mu_{2\rho}a_1^\rho + a_2^\mu, \tag{12.39}$$

and we see that the translation parameters get rotated. Nothing surprising here since this is what four-vectors do for living! Such a coupling of the translation and Lorentz group is called a semi-direct product. Still, the PG transformations form a group. In order to obtain the algebra of the generators, observe that we can write the change in $x$ under a small translation as

$$\delta x^\mu = i\epsilon^\rho P_\rho x^\mu = \epsilon^\mu, \tag{12.40}$$

where $\epsilon^\mu$ are the small parameters and

$$P_\rho = -i\partial_\rho \tag{12.41}$$

are the hermitian generators of the transformation. They clearly commute with one another,

$$[P_\mu, P_\nu] = 0, \tag{12.42}$$

but not with the Lorentz generators:

$$[M_{\mu\nu}, P_\rho] = -ig_{\mu\rho}P_\nu + ig_{\nu\rho}P_\mu. \tag{12.43}$$

The commutation relations (12.42) and (12.43) and those among the $M_{\mu\nu}$ define the Lie algebra of the Poincaré group. The 'length' $P_\mu P^\mu$ of the vector $P_\rho$ is obviously invariant under LT's and in view of (12.42) is therefore a Casimir operator. The other Casimir operator is not so obvious to construct. It is the length of the Pauli-Lubanski four vector defined by

$$W^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}P_\nu M_{\rho\sigma}. \tag{12.44}$$

Use of (12.42) and (12.43) gives

$$[W^\mu, P^\rho] = 0, \tag{12.45}$$

$$[M_{\mu\nu}, W_\rho] = -ig_{\mu\rho}W_\nu + ig_{\nu\rho}W_\mu. \tag{12.46}$$

Its length $W^\mu W_\mu$ is therefore a Casimir invariant. The most general representation of the Poincaré group generators is

$$\begin{aligned} P_\rho =& -i\partial_\rho, \\ M_{\mu\nu} =& i(x_\mu\partial_\nu - x_\nu\partial_\mu) + S_{\mu\nu}, \end{aligned} \tag{12.47}$$

so that

$$W^\mu = -\frac{i}{2}\epsilon^{\mu\nu\rho\sigma}S_{\rho\sigma}\partial_\nu. \tag{12.48}$$

The representations of the PG fall then into three classes:

(1) The eigenvalue of $P_\rho P^\rho \equiv m^2$ is a real positive number. Then the eigenvalue of $W_\rho W^\rho$ is $-m^2 s(s+1)$, where $s$ is the spin, which assumes discrete values $s = 0, 1/2, 1, \ldots$. This representation is labelled by the mass $m$ and the spin $s$. States within the representation are distinguished by the third component of the spin $s_3 = -s, -s+1, \ldots, s-1, s$, and the continuous eigenvalues of $P_i$. Physically a state corresponds to a particle of mass $m$, spin $s$, 3-momentum $P_i$ and spin projection $s_3$. Massive particles of spin $s$ have $2s+1$ degrees of freedom.

(2) The eigenvalue of $P_\rho P^\rho$ is equal to zero, corresponding to a particle of rest nass zero. $W_\rho W^\rho$ is also zero. Since $P^\rho W_\rho = 0$, it follows that $W_\mu$ and $P_\mu$ are proportional. The constant of proportionality is called the helicity, and is equal to $\pm s$, where $s = 0, 1/2, 1, \ldots$ is the spin of the representation. Thus massless particles of spin $s \neq 0$ have two degrees of freedom. They are further distinguished by the three values of their momenta along the $x, y$ and $z$ directions. Examples of particles falling in the category are the photon with spin 1 and two helicity states $\pm 1$, the graviton with two states of polarization $\pm 2$, and – as far as we know – the neutrinos, each with helicity $\pm 1/2$.

(3) $P^\rho P_\rho = 0$, but the spin is continuous. The length of $W$ is minus the square of a positive number. This type of representation describes a particle of zero rest mass with an infinite number of polarization states labelled by a continuous parameter. These do not seem to be realized in nature.

There are other representations of the Poincaré group; however they are not unitary. QM allows for the identification of only the unitary representations with particle states. The Wigner representations are infinite dimensional, corresponding to particles of unbounded momenta. The introduction of field theory will enable us to make use of these representations.