RELATIVISTIC QUANTUM MECHANICS

This part of the course is based on ref. [1]

The point of view of [1] is that quantum field theory is the way it is because (with certain qualifications) this is the only way to reconcile quantum mechanics with special relativity. Therefore our first task is to study how symmetries like Lorentz invariance appear in quantum setting.

1. Quantum Mechanics

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 Φ and Ψ are vectors in the space (often called 'state-vectors') then so is $\xi\Phi + \eta\Psi$, for arbitrary complex numbers ξ, η. It has a norm; for any pair of vectors there is a complex number (Φ, Ψ) such that

$$(\Phi, \Psi) = (\Psi, \Phi)^*, \tag{1.1}$$

$$(\Phi, \xi_1\Psi_1 + \xi_2\Psi_2) = \xi_1(\Phi, \Psi_1) + \xi_2(\Phi, \Psi_2), \tag{1.2}$$

$$(\eta_1\Phi_1 + \eta_2\Phi_2, \Psi) = \eta_1^*(\Phi_1, \Psi) + \eta_2^*(\Phi_2, \Psi). \tag{1.3}$$

The norm (Ψ, Ψ) 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 Ψ and Ψ' belonging to the same ray if $\Psi' = \xi\Psi$, where ξ 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

$$A(\xi\Psi + \eta\Phi) = \xi A\Psi + \eta A\Phi \tag{1.4}$$

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

$$(\Phi, A^\dagger \Psi) \equiv (A\Phi, \Psi) = (\Psi, A\Phi)^*. \tag{1.5}$$

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

$$A\Psi = \alpha\Psi \quad \text{for} \quad \Psi \text{ in } \mathcal{R}. \tag{1.6}$$

An elementary theorem tells us that for A hermitian, α is real, and eigenvectors with different α'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

$$P(\mathcal{R} \to \mathcal{R}_n) = |(\Psi, \Psi_n)|^2, \tag{1.7}$$

where Ψ and Ψ_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:

$$\sum_n P(\mathcal{R} \to \mathcal{R}_n) = 1 \tag{1.8}$$

if the state vectors Ψ_n form a complete set.

2. Symmetries

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

$$P(\mathcal{R} \to \mathcal{R}_n) = P(\mathcal{R}' \to \mathcal{R}'_n). \tag{2.1}$$

(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 Ψ is in ray $\mathcal{R}$ then $U\Psi$ is in the ray $\mathcal{R}'$, with U either *unitary* and *linear*

$$(U\Phi, U\Psi) = (\Phi, \Psi), \tag{2.2}$$

$$U(\xi\Phi + \eta\Psi) = \xi U\Phi + \eta U\Psi \tag{2.3}$$

or else *antiunitary* and *antilinear*

$$(U\Phi, U\Psi) = (\Phi, \Psi)^*, \tag{2.4}$$

$$U(\xi\Phi + \eta\Psi) = \xi^* U\Phi + \eta^* U\Psi. \tag{2.5}$$

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

$$(\Phi, L^\dagger\Psi) \equiv (L\Phi, \Psi). \tag{2.6}$$

This condition cannot be satisfied for an antilinear operator A, because in this case the right-hand side of (2.6) would be linear in Φ,

$$(A\xi\Phi, \Psi) = \xi(A\Phi, \Psi), \tag{2.7}$$

while the left-hand side of (2.6) is antilinear in Φ,

$$(\xi\Phi, A^\dagger\Psi) = \xi^*(\Phi, A^\dagger\Psi). \tag{2.8}$$

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

$$(\Phi, A^\dagger\Psi) \equiv (A\Phi, \Psi)^* = (\Psi, A\Phi). \tag{2.9}$$

With this definitions, unitarity and linearity imply

$$(U\Phi, U\Psi) = (U^\dagger U\Phi, \Psi) = (\Phi, \Psi), \tag{2.10}$$

while antiunitarity and antilinearity imply

$$(U\Phi, U\Psi) = (U^\dagger U\Phi, \Psi)^* = (\Phi, \Psi)^*, \tag{2.11}$$

so the conditions for unitarity or antiunitarity both take the form

$$U^\dagger = U^{-1}. \tag{2.12}$$

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:

$$U = 1 + i\epsilon t \tag{2.13}$$

with ϵ 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 $T_2 T_1$, 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 Ψ_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 $\phi_n(T_2, T_1)$

$$U(T_2)U(T_1)\Psi_n = e^{i\phi_n(T_2,T_1)}U(T_2T_1)\Psi_n. \tag{2.14}$$

Furthermore, with one significant exception, the linearity (or antilinearity) of $U(T)$ tells us that these phases are independent of the state Ψ_n. Here is the proof. Consider any two different vectors Ψ_A, Ψ_B, which are not proportional to each other. Then, applying eq. (2.14) to the state $\Psi_{AB} \equiv \Psi_A + \Psi_B$, we have

$$\begin{aligned} e^{i\phi_{AB}} U(T_2 T_1)(\Psi_A + \Psi_B) &= U(T_2)U(T_1)(\Psi_A + \Psi_B) \\ &= U(T_2)U(T_1)\Psi_A + U(T_2)U(T_1)\Psi_B \\ &= e^{i\phi_A} U(T_2 T_1)\Psi_A + e^{i\phi_B} U(T_2 T_1)\Psi_B. \end{aligned} \tag{2.15}$$

Any unitary or antiunitary operator has an inverse (its adjoint) which is also unitary or antiunitary. Multiplying (2.15) on the left with $U^{-1}(T_2 T_1)$, we have then

$$e^{\pm i\phi_{AB}}(\Psi_A + \Psi_B) = e^{\pm i\phi_A}\Psi_A + e^{\pm i\phi_B}\Psi_B, \tag{2.16}$$

the upper and lower signs referring to $U(T_2 T_1)$ unitary or antiunitary, respectively. Since Ψ_A and Ψ_B are linearly independent, this is only possible if

$$e^{i\phi_{AB}} = e^{i\phi_A} = e^{i\phi_B}. \tag{2.17}$$

So as promised, the phase in (2.14) is independent of the state vector Ψ_n, and therefore this can be written as an operator relation

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)} U(T_2, T_1). \tag{2.18}$$

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 (2.18) 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 $U(T_2 T_1)$ 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. (2.18).

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 θ^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

$$T(\overline{\theta})T(\theta) = T\left(f(\overline{\theta}, \theta)\right) \tag{2.19}$$

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

$$f^a(\theta, 0) = f^a(0, \theta) = \theta^a. \tag{2.20}$$

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

$$U(T(\theta)) = 1 + i\theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \cdots, \tag{2.21}$$

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

$$U(T(\overline{\theta}))U(T(\theta)) = U(T(f(\overline{\theta}, \theta)). \tag{2.22}$$

Let us see what this condition looks like when expanded in powers of θ^a and $\overline{\theta}^a$. According to eq. (2.20), the expansion of $f^a(\overline{\theta}, \theta)$ to second order must take the form

$$f^a(\overline{\theta}, \theta) = \theta^a + \overline{\theta}^a + f^a{}_{bc}\overline{\theta}^b\theta_c + \cdots \tag{2.23}$$

with real coefficients $f^a{}_{bc}$. (The presence of any terms of order θ^2 or $\overline{\theta}^2$ would violate eq. (2.20).) Then eq. (2.22) reads

$$\begin{aligned}
&\left[1 + i\overline{\theta}^a t_a + \frac{1}{2}\overline{\theta}^b\overline{\theta}^c t_{bc} + \cdots\right] \times \left[1 + i\theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \cdots\right] \\
&= 1 + if^a t_a + \frac{1}{2}f^b f^c t_{bc} + \cdots \\
&= 1 + i\left(\theta^a + \overline{\theta}^a + f^a{}_{bc}\overline{\theta}^b\theta^c + \cdots\right) t_a \\
&+ \frac{1}{2}(\theta^b + \overline{\theta}^b + \cdots)(\theta^c + \overline{\theta}^c + \cdots)t_{bc} + \cdots
\end{aligned} \tag{2.24}$$

The terms of order $1, \theta, \overline{\theta}, \theta^2$ and $\overline{\theta}^2$ automatically match on both sides of eq. (2.24), but from the $\overline{\theta}\theta$ terms we obtain a non-trivial condition

$$t_{bc} = -t_b t_c - i f^a{}_{bc} t_a. \tag{2.25}$$

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 θ^b and θ^c) so eq. (2.25) requires that

$$[t_b, t_c] = i C^a{}_{bc} t_a, \tag{2.26}$$

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

$$C^a{}_{bc} \equiv -f^a{}_{bc} + f^a{}_{cb}. \tag{2.27}$$

Such a set of commutation relations is known as a *Lie algebra*. We will later prove in effect that the commutation relation (2.26) 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 (2.25), 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 θ_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. (2.19) is satisfied if $\theta, \overline{\theta}$ and $f(\theta, \overline{\theta})$ are in this neighborhood. The extension to all θ^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 θ^a) is simply additive

$$f^a(\theta, \overline{\theta}) = \theta^a + \overline{\theta}^a. \tag{2.28}$$

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. (2.23) vanish, and so do the structure constants (2.27). The generators then all commute

$$[t_b, t_c] = 0. \tag{2.29}$$

Such a group is called *Abelian*. In this case, it is easy to calculate $U(T(\theta))$ for all θ^a. From eqs. (2.22) and (2.28), 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

$$U(T(\theta)) = \left[U(T(\theta/N))\right]^N . \tag{2.30}$$

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

$$U(T(\theta)) = \lim_{N\to\infty}\left[1 + \frac{i}{N}\theta^a t_a\right]^N \tag{2.31}$$

and hence

$$U(T(\theta)) = \exp(it_a\theta^a). \tag{2.32}$$

3. Quantum Lorentz Transformations

Einstein's principle of relativity states the equivalence of certain 'inertial' frames of reference. It is distinguished from the Galilean principle of relativity, obeyed by Newtonian mechanics, by the transformation connecting coordinate systems in different inertial frames. If x^μ are the coordinates in one inertial frame (with x^1, x^2, x^3 Cartesian space coordinates, and $x^0 = t$ a time coordinate, the speed of light being set equal to unity) then in any other inertial frame, the coordinate x'^μ must satisfy

$$\eta_{\mu\nu}dx'^\mu dx'^\nu = \eta_{\mu\nu}dx^\mu dx^\nu \tag{3.1}$$

or equivalently

$$\eta_{\mu\nu}\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma} = \eta_{\rho\sigma}. \tag{3.2}$$

Here $\eta_{\mu\nu}$ is the diagonal matrix, with elements

$$\eta_{11} = \eta_{22} = \eta_{33} = +1, \quad \eta_{00} = -1 \tag{3.3}$$

and the summation convention is in force: we sum over any index like μ and ν in eq. (3.2), which appears twice in the same term, once upstairs and once downstairs. These transformations have the special property that the speed of light is the same (in our units, equal to unity) in all inertial frames; a light wave travelling at unit speed satisfies $|d\mathbf{x}/dt| = 1$, or in other words $\eta_{\mu\nu}dx^\mu dx^\nu = d\mathbf{x}^2 - dt^2 = 0$, from which it follows that also $\eta_{\mu\nu}dx'^\mu dx'^\nu = 0$, and hence $|d\mathbf{x}'/dt| = 1$.

Any coordinate transformation $x^\mu \to x'^\mu$ that satisfies eq. (3.2) is *linear*

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \tag{3.4}$$

with a^μ arbitrary constants, and $\Lambda^\mu{}_\nu$ a constant matrix satisfying the conditions

$$\eta_{\mu\nu}\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}. \tag{3.5}$$

For some purposes, it is useful to write the Lorentz transformation condition in a different way. The matrix $\eta_{\mu\nu}$ has an inverse, written $\eta^{\mu\nu}$, which happens to have the same components: it is diagonal with $\eta^{00} = -1$, $\eta^{11} = \eta^{22} = \eta^{33} = +1$. Multiplying eq. (3.5) with $\eta^{\sigma\tau}\Lambda^\kappa{}_\tau$ and inserting parentheses judiciously, we have

$$\eta_{\mu\nu}\Lambda^\mu{}_\rho(\Lambda^\nu{}_\sigma\Lambda^\kappa{}_\tau\eta^{\sigma\tau}) = \Lambda^\kappa{}_\rho = \eta_{\mu\nu}\eta^{\nu\kappa}\Lambda^\mu{}_\rho. \tag{3.6}$$

Multiplying with the inverse of the matrix $\eta_{\mu\nu}\Lambda^\mu{}_\rho$ then gives

$$\Lambda^\nu{}_\sigma\Lambda^\kappa{}_\tau\eta^{\sigma\tau} = \eta^{\nu\kappa}. \tag{3.7}$$

These transformations form a group. If we first perform a Lorentz transformation (3.4), and then a second Lorentz transformation $x'^\mu \to x''^\mu$, with

$$x''^\mu = \overline{\Lambda}^\mu{}_\rho x'^\rho + \overline{a}^\mu = \overline{\Lambda}^\mu{}_\rho(\Lambda^\rho{}_\nu x^\nu + a^\rho) + \overline{a}^\mu \tag{3.8}$$

then the effect is the same as the Lorentz transformation $x^\mu \to x''^\mu$, with

$$x''^\mu = (\overline{\Lambda}^\mu{}_\rho\Lambda^\rho{}_\nu)x^\nu + (\overline{\Lambda}^\mu{}_\rho a^\rho + \overline{a}^\mu). \tag{3.9}$$

(Note that if $\Lambda^\mu{}_\nu$ and $\overline{\Lambda}^\mu{}_\nu$ both satisfy eq. (3.5), then so does $\overline{\Lambda}^\mu{}_\rho\Lambda^\rho{}_\nu$,

$$\eta_{\mu\nu}\overline{\Lambda}^\mu{}_\rho\Lambda^\rho{}_\kappa\overline{\Lambda}^\nu{}_\tau\Lambda^\tau{}_\sigma = \eta_{\rho\tau}\Lambda^\rho{}_\kappa\Lambda^\tau{}_\sigma = \eta_{\kappa\sigma}, \tag{3.10}$$

so this *is* a Lorentz transformation. The bar is used here just to distinguish one Lorentz transformation from the other.) The transformations $T(\Lambda, a)$ induced on physical states therefore satisfy the composition rule

$$T(\overline{\Lambda}, \overline{a})T(\Lambda, a) = T(\overline{\Lambda}\Lambda, \overline{\Lambda}a + \overline{a}). \tag{3.11}$$

Taking the determinant of (3.5) gives

$$(\det \Lambda)^2 = 1 \tag{3.12}$$

so $\Lambda^\mu{}_\nu$ has an inverse, $(\Lambda^{-1})^\nu{}_\rho$ which we see from eq. (3.5) takes the form

$$(\Lambda^{-1})^\rho{}_\nu = \Lambda_\nu{}^\rho \equiv \eta_{\nu\mu}\eta^{\rho\sigma}\Lambda^\mu{}_\sigma. \tag{3.13}$$

The inverse of the transformation $T(\Lambda, a)$ is seen from eq. (3.11) to be $T(\Lambda^{-1}, -\Lambda^{-1}a)$ and, of course, the identity transformation is $T(1, 0)$.

In accordance with the discussion in the previous section, the transformations $T(\Lambda, a)$ induce a unitary linear transformation on vectors in the physical Hilbert space

$$\Psi \to U(\Lambda, a)\Psi. \tag{3.14}$$

The operators U satisfy the composition rule

$$U(\overline{\Lambda}, \overline{a})U(\Lambda, a) = U(\overline{\Lambda}\Lambda, \overline{\Lambda}a + \overline{a}). \tag{3.15}$$

(As already mentioned, to avoid the appearance of a phase factor on the right hand side of eq. (3.15), it is, in general, necessary to enlarge the Lorentz group. The appropriate enlargement is described below.)

The whole group of transformations $T(\Lambda, a)$ is properly known as the *inhomogeneous Lorentz group*, or *Poincaré group*. It has a number of important subgroups. First, those transformations with $a^\mu = 0$ obviously form a subgroup, with

$$T(\overline{\Lambda}, 0)T(\Lambda, 0) = T(\overline{\Lambda}\Lambda, 0), \tag{3.16}$$

known as the *homogeneous Lorentz group*. Also, we note from (3.12) that either $\det \Lambda = +1$ or $\det \Lambda = -1$; those transformations with $\det \Lambda = +1$ obviously form a subgroup of either the homogeneous or the inhomogeneous Lorentz groups. Further, from the 00-components of eqs. (3.5) and (3.7), we have

$$(\Lambda^0{}_0)^2 = 1 + \Lambda^i{}_0\Lambda^i{}_0 = 1 + \Lambda^0{}_i\Lambda^0{}_i, \tag{3.17}$$

with i summed over the values 1, 2 and 3. We see that either $\Lambda^0{}_0 \geq +1$ or $\Lambda^0{}_0 \leq -1$. Those transformations with $\Lambda^0{}_0 \geq +1$ form a subgroup. Note that if $\Lambda^\mu{}_\nu$ and $\overline{\Lambda}^\mu{}_\nu$ are two such Λs, then

$$(\overline{\Lambda}\Lambda)^0{}_0 = \overline{\Lambda}^0{}_0\Lambda^0{}_0 + \overline{\Lambda}^0{}_1\Lambda^1{}_0 + \overline{\Lambda}^0{}_2\Lambda^2{}_0 + \overline{\Lambda}^0{}_3\Lambda^3{}_0. \tag{3.18}$$

But eq. (3.17) shows that the three-vector $(\Lambda^1{}_0, \Lambda^2{}_0, \Lambda^3{}_0)$ has length $\sqrt{(\Lambda^0{}_0)^2 - 1}$ and similarly the three-vector $(\overline{\Lambda}^0{}_1, \overline{\Lambda}^0{}_2, \overline{\Lambda}^0{}_3)$ has length $\sqrt{(\overline{\Lambda}^0{}_0)^2 - 1}$, so that the scalar product of the two three-vectors is bounded by

$$|\overline{\Lambda}^0{}_1\Lambda^1{}_0 + \overline{\Lambda}^0{}_2\Lambda^2{}_0 + \overline{\Lambda}^0{}_3\Lambda^3{}_0| \leq \sqrt{(\Lambda^0{}_0)^2 - 1}\sqrt{(\overline{\Lambda}^0{}_0)^2 - 1}, \tag{3.19}$$

and so

$$(\overline{\Lambda}\Lambda)^0{}_0 \geq \overline{\Lambda}^0{}_0\Lambda^0{}_0 - \sqrt{(\Lambda^0{}_0)^2 - 1}\sqrt{(\overline{\Lambda}^0{}_0)^2 - 1} \geq 1. \tag{3.20}$$

The subgroup of the Lorentz transformations with $\det \Lambda = +1$ and $\Lambda^0{}_0 \geq +1$ is known as the *proper orthochronous Lorentz group.* Since it is not possible by a continuous change of parameters to jump from $\det \Lambda = +1$ to $\det \Lambda = -1$, or from $\Lambda^0{}_0 \geq +1$ to $\Lambda^0{}_0 \leq -1$, any Lorentz transformation that can be obtained from the identity by a continuous change of parameters must have $\det \Lambda$ and $\Lambda^0{}_0$ of the same sign as the identity, and hence must belong to the proper orthochronous Lorentz group.

Any Lorentz transformation is either proper and orthochronous, or may be written as the product of an element of the proper orthochronous Lorentz group with one of the discrete transformations $\mathcal{P}$ or $\mathcal{T}$ or $\mathcal{PT}$, where $\mathcal{P}$ is the space inversion, whose non-zero elements are

$$\mathcal{P}^0{}_0 = 1, \quad \mathcal{P}^1{}_1 = \mathcal{P}^2{}_2 = \mathcal{P}^3{}_3 = -1, \tag{3.21}$$

and $\mathcal{T}$ is the time-reversal matrix, whose non-zero elements are

$$\mathcal{T}^0{}_0 = -1, \quad \mathcal{T}^1{}_1 = \mathcal{T}^2{}_2 = \mathcal{T}^3{}_3 = 1. \tag{3.22}$$

Thus the study of the whole Lorentz group reduces to the study of its proper orthochronous subgroup, plus space inversion and time reversal. We will consider space inversion and time reversal separately later. Until then, we will deal only with the homogeneous or inhomogeneous proper orthochronous Lorentz group.

4. The Poincaré Algebra

As we saw previously, much of the information about any Lie symmetry group is contained in properties of the group elements near the identity. For the inhomogeneous Lorentz group, the identity is the transformation $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu$, $a^\mu = 0$, so we want to study those transformations with

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu, \quad a^\mu = \epsilon^\mu, \tag{4.1}$$

both $\omega^\mu{}_\nu$ and ϵ^μ taken as infinitesimal. The Lorentz condition (3.5) reads here

$$\begin{aligned} \eta_{\rho\sigma} &= \eta_{\mu\nu}(\delta^\mu{}_\rho + \omega^\mu{}_\rho)(\delta^\nu{}_\sigma + \omega^\nu{}_\sigma) \\ &= \eta_{\sigma\rho} + \omega_{\sigma\rho} + \omega_{\rho\sigma} + \mathcal{O}(\omega^2). \end{aligned} \tag{4.2}$$

We are here using the convention (to be used throughout this chapter) that indices may be lowered or raised by contraction with $\eta_{\mu\nu}$ or $\eta^{\mu\nu}$,

$$\begin{aligned} \omega_{\sigma\rho} &\equiv \eta_{\mu\sigma}\omega^\mu{}_\rho, \\ \omega^\mu{}_\rho &\equiv \eta^{\mu\sigma}\omega_{\sigma\rho}. \end{aligned} \tag{4.3}$$

Keeping only the terms of first order in ω in the Lorentz conditions (3.5), we see that now this condition reduces to the antisymmetry of $\omega_{\mu\nu}$,

$$\omega_{\mu\nu} = -\omega_{\nu\mu}. \tag{4.4}$$

An antisymmetric second-rank tensor in four dimensions has $(4 \times 3)/2 = 6$ independent components, so including the four components of ϵ^μ, an inhomogeneous Lorentz transformation is described by $6 + 4 = 10$ parameters.

Since $U(1, 0)$ carries any ray into itself, it must be proportional to the unit operator, and by a choice of phase may be made equal to it. For an infinitesimal Lorentz transformation (4.1), $U(1 + \omega, \epsilon)$ must then equal $\mathbf{1}$ plus terms linear in $\omega_{\rho\sigma}$ and ϵ_ρ. We write this as

$$U(1 + \omega, \epsilon) = 1 + \frac{1}{2} i\omega_{\rho\sigma}J^{\rho\sigma} - i\epsilon_\rho P^\rho + \cdots \quad . \tag{4.5}$$

Here $J^{\rho\sigma}$ and P^ρ are ω- and ϵ-independent operators, and the dots denote higher order terms in ω and/or ϵ. In order for $U(1 + \omega, \epsilon)$ to be unitary, the operators $J^{\rho\sigma}$ and P^ρ must be Hermitian

$$J^{\rho\sigma\dagger} = J^{\rho\sigma}, \qquad P^{\rho\dagger} = P^\rho. \tag{4.6}$$

Since $\omega_{\rho\sigma}$ is antisymmetric, we can take the coefficient $J^{\rho\sigma}$ to be antisymmetric also

$$J^{\rho\sigma} = -J^{\sigma\rho}. \tag{4.7}$$

As we shall see, P^1, P^2 and P^3 are the components of the momentum operators, J^{23}, J^{31} and J^{12} are the components of the angular momentum vector, and P^0 is the energy operator, or *Hamiltonian*.

We now examine the Lorentz transformation properties of $J^{\rho\sigma}$ and P^ρ. We consider the product

$$U(\Lambda, a)U(1+\omega, \epsilon)U^{-1}(\Lambda, a), \tag{4.8}$$

where $\Lambda^\mu{}_\nu$ and a^μ are here the parameters of a new transformation, unrelated to ω and ϵ. According to eq. (3.15), the product $U(\Lambda^{-1}, -\Lambda^{-1}a)U(\Lambda, a)$ equals $U(1,0)$, so $U(\Lambda^{-1}, -\Lambda^{-1}a)$ is the inverse of $U(\Lambda, a)$. It follows then from (3.15) that

$$U(\Lambda, a)U(1+\omega, \epsilon)U^{-1}(\Lambda, a) = U\left(\Lambda(1+\omega)\Lambda^{-1}, \Lambda\epsilon - \Lambda\omega\Lambda^{-1}a\right). \tag{4.9}$$

To first order in ω and ϵ, we have then

$$U(\Lambda, a)\left[\frac{1}{2}\omega_{\rho\sigma}J^{\rho\sigma} - \epsilon_\rho P^\rho\right]U^{-1}(\Lambda, a) = \frac{1}{2}(\Lambda\omega\Lambda^{-1})_{\mu\nu}J^{\mu\nu} - (\Lambda\epsilon - \Lambda\omega\Lambda^{-1}a)_\mu P^\mu. \tag{4.10}$$

Equating coefficients of $\omega_{\rho\sigma}$ and ϵ_ρ on both sides of this equation (and using eq. (3.13)), we find

$$U(\Lambda, a)J^{\rho\sigma}U^{-1}(\Lambda, a) = \Lambda_\mu{}^\rho\Lambda_\nu{}^\sigma(J^{\mu\nu} - a^\mu P^\nu + a^\nu P^\mu), \tag{4.11}$$

$$U(\Lambda, a)P^\rho U^{-1}(\Lambda, a) = \Lambda_\mu{}^\rho P^\mu. \tag{4.12}$$

For homogeneous Lorentz transformations (with $a^\mu = 0$), these transformations rules simply say that $J^{\mu\nu}$ is a tensor and P^μ is a vector. For pure translations (with $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu$), they tell us that P^ρ is translation -invariant but $J^{\rho\sigma}$ is not. In particular, the change of the space-space components of $J^{\rho\sigma}$ under a spatial translation is just the usual change of the angular momentum under a change of the origin relative to which the angular momentum is calculated.

Next, let's apply rules (4.11), (4.12) to a transformation that is itself infinitesimal, i.e. $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu$ and $a^\mu = \epsilon^\mu$, with infinitesimal $\omega^\mu{}_\nu$ and ϵ^μ unrelated to the previous ω and ϵ. Using eq. (4.5), and keeping only terms of first order in $\omega^\mu{}_\nu$ and ϵ^μ, eqs. (4.11) and (4.12) now become

$$i\left[\frac{1}{2}\omega_{\mu\nu}J^{\mu\nu} - \epsilon_\mu P^\mu, J^{\rho\sigma}\right] = \omega_\mu{}^\rho J^{\mu\sigma} + \omega_\nu{}^\sigma J^{\rho\nu} - \epsilon^\rho P^\sigma + \epsilon^\sigma P^\rho, \tag{4.13}$$

$$i\left[\frac{1}{2}\omega_{\mu\nu}J^{\mu\nu} - \epsilon_\mu P^\mu, P^\rho\right] = \omega_\mu{}^\rho P^\mu. \tag{4.14}$$

Equating coefficients of $\omega_{\mu\nu}$ and ϵ_μ on both sides of these equations, we find the commutation relations

$$i[J^{\mu\nu}, J^{\rho\sigma}] = \eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} - \eta^{\sigma\mu}J^{\rho\nu} + \eta^{\sigma\nu}J^{\rho\mu}, \tag{4.15}$$

$$i[P^\mu, J^{\rho\sigma}] = \eta^{\mu\rho} P^\sigma - \eta^{\mu\sigma} P^\rho, \tag{4.16}$$

$$i[P^\mu, P^\rho] = 0. \tag{4.17}$$

This is the Lie algebra of the Poincaré group.

In quantum mechanics, a special role is played by those operators that are *conserved*, i.e. that commute with the energy operator $H = P^0$. Inspection of (4.16) and (4.17) shows that these are the momentum three-vector

$$\mathbf{P} = \{P^1, P^2, P^3\} \tag{4.18}$$

and the angular-momentum three-vector

$$\mathbf{J} = \{J^{23}, J^{31}, J^{12}\} \tag{4.19}$$

and, of course, the energy P^0 itself. The remaining generators form what is called the 'boost' three-vector

$$\mathbf{K} = \{J^{10}, J^{20}, J^{30}\}. \tag{4.20}$$

These are *not* conserved, which is why we do not use the eigenvalues of $\mathbf{K}$ to label physical states. In a three-dimensional notation, the commutation relations (4.15), (4.16), (4.17) may be written

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \tag{4.21}$$

$$[J_i, K_j] = i\epsilon_{ijk} K_k, \tag{4.22}$$

$$[K_i, K_j] = -i\epsilon_{ijk} J_k, \tag{4.23}$$

$$[J_i, P_j] = i\epsilon_{ijk} P_k, \tag{4.24}$$

$$[K_i, P_j] = iH\delta_{ij}, \tag{4.25}$$

$$[J_i, H] = [P_i, H] = [H, H] = 0, \tag{4.26}$$

$$[K_i, H] = iP_i, \tag{4.27}$$

where i, j, k, etc. run over the values 1, 2, and 3, and ϵ_{ijk} is the totally antisymmetric quantity with $\epsilon_{123} = +1$. The commutation relation (4.21) will be recognized as that of the angular-momentum operator.

The pure translations $T(1, a)$ form a subgroup of the inhomogeneous Lorentz group with a group multiplication rule given by (3.9) as

$$T(1, \overline{a})T(1, a) = T(1, \overline{a} + a). \tag{4.28}$$

This is additive in the same sense as (2.28), so by using (4.5) and repeating the same arguments that led to (2.32), we find that finite translations are represented on the physical Hilbert space by

$$U(1, a) = \exp(-iP^{\mu} a_{\mu}). \tag{4.29}$$

In exactly the same way, we can show that a rotation R_{θ} by an angle $|\theta|$ around the direction of θ is represented on the physical Hilbert space by

$$U(R_{\theta}, 0) = \exp(i\mathbf{J} \cdot \theta). \tag{4.30}$$

It is interesting to compare the Poincaré algebra with the Lie algebra of the symmetry group of Newtonian mechanics, the Galilean group. We could derive this algebra by starting with the transformation rules of the Galilean group and then following the same procedure that we used here to derive the Poincaré algebra. However, since we already have eqs. (4.21)-(4.27), it is easier to obtain the Galilean algebra as a low-velocity limit of the Poincaré algebra, by what is known as an *Inönü-Wigner contraction.* For a system of particles of typical mass m and typical velocity v, the momentum and angular-momentum operators are expected to be of order $\mathbf{J} \sim 1$, $\mathbf{P} \sim mv$. On the other hand, the energy operator is $H = M + W$ with a total mass M and non-mass energy W (kinetic plus potential) of order $M \sim m$, $W \sim mv^2$. Inspection of eqs. (4.21)-(4.27) shows that these commutation relations have a limit for $v \ll 1$ of the form

$$\begin{aligned}
&[J_i, J_j] = i\epsilon_{ijk} J_k, \quad [J_i, K_j] = i\epsilon_{ijk} K_k, \quad [K_i, K_j] = 0, \\
&[J_i, P_j] = i\epsilon_{ijk} P_k, \quad [K_i, P_j] = iM\delta_{ij}, \\
&[J_i, W] = [P_i, W] = 0, \quad [K_i, W] = iP_i, \\
&[J_i, M] = [P_i, M] = [K_i, M] = [W, M] = 0,
\end{aligned} \tag{4.31}$$

with $\mathbf{K}$ of order $1/v$. Note that the product of a translation $\mathbf{x} \to \mathbf{x} + \mathbf{a}$ and a 'boost' $\mathbf{x} \to \mathbf{x} + \mathbf{v}t$ should be the transformation $\mathbf{x} \to \mathbf{x} + \mathbf{v}t + \mathbf{a}$, but this is not true for the action of these operators on Hilbert space:

$$\exp(i\mathbf{K} \cdot \mathbf{v}) \exp(-i\mathbf{P} \cdot \mathbf{a}) = \exp(iM\mathbf{a} \cdot \mathbf{v}/2) \exp(i(\mathbf{K} \cdot \mathbf{v} - \mathbf{P} \cdot \mathbf{a})). \tag{4.32}$$

The appearance of the phase factor $\exp(iM\mathbf{a}\cdot\mathbf{v}/2)$ shows that this is a projective representation, with a superselection rule forbidding the superposition of states of different mass. In this respect, the mathematics of the Poincaré group is simpler than that of the Galilean group. However, there is nothing to prevent us from formally enlarging the Galilean group, by adding one more generator to its Lie algebra, which commutes with all the other generators, and whose eigenvalues are the masses of the various states. In this case, physical states provide an ordinary rather that a projective representation of the extended symmetry group. The difference appears to be a mere matter of notation, except that with this reinterpretation of the Galilean group there is no need for a mass superselection rule.

5. One-Particle States

We now consider the classification of one-particle states according to their transformation under the inhomogeneous Lorentz group.

The components of the energy-momentum four-vector all commute with each other, so it is natural to express physical state-vectors in terms of eigenvectors of the four-momentum. Introducing a label σ to denote all other degrees of freedom, we thus consider state-vectors $\Psi_{p,\sigma}$ with

$$P^{\mu}\Psi_{p,\sigma} = p^{\mu}\Psi_{p,\sigma}. \tag{5.1}$$

For general states, consisting for instance of several unbounded particles, the label σ would have to be allowed to include continuous as well as discrete labels. We take as part of the definition of a *one*-particle state, that the label σ is purely discrete, and will limit ourselves here to that case. (However, a specific bound state of two or more particles, such as the lowest state of the hydrogen atom, is to be considered a one-particle state. It is not an *elementary* particle, but the distinction between composite and elementary particles is of no relevance here.)

Eqs. (5.1) and (4.29) tell us how the states $\Psi_{p,\sigma}$ transform under translations:

$$U(1,a)\Psi_{p,\sigma} = e^{-ip\cdot a}\Psi_{p,\sigma}. \tag{5.2}$$

We must now consider how these state transform under homogeneous Lorentz transformations.

Using (4.12), we see that the effect of operating on $\Psi_{p,\sigma}$ with a quantum homogeneous Lorentz transformation $U(\Lambda, 0) \equiv U(\Lambda)$ is to produce an eigenvector of the four-momentum with eigenvalue Λp

$$\begin{aligned} P^\mu U(\Lambda)\Psi_{p,\sigma} &= U(\Lambda)\left[U^{-1}(\Lambda)P^\mu U(\Lambda)\right]\Psi_{p,\sigma} = U(\Lambda)(\Lambda_\rho^{-1\mu}P^\rho)\Psi_{p,\sigma} \\ &= \Lambda^\mu{}_\rho p^\rho U(\Lambda)\Psi_{p,\sigma}. \end{aligned} \tag{5.3}$$

Hence $U(\Lambda)\Psi_{p,\sigma}$ must be a linear combination of the state-vectors $\Psi_{\Lambda p,\sigma'}$:

$$U(\Lambda)\Psi_{p,\sigma} = \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda, p)\Psi_{\Lambda p,\sigma'}. \tag{5.4}$$

In general, it may be possible by using suitable linear combinations of the $\Psi_{p,\sigma}$ to choose the σ labels in such a way that the matrix $C_{\sigma'\sigma}(\Lambda, p)$ is block-diagonal; in other words, so that the $\Psi_{p,\sigma}$ with σ within any one block by themselves furnish a representation of the inhomogeneous Lorentz group. It is natural to identify the states of a specific particle type with the components of a representation of the inhomogeneous Lorentz group which is irreducible, in the sense that it cannot be further decomposed in this way. Our task now is to work out the structure of the coefficients $C_{\sigma'\sigma}(\Lambda, p)$ in irreducible representations of the inhomogeneous Lorentz group.

For this purpose, note that there are only two functions of p^μ that are left invariant by all proper orthochronous Lorentz transformations $\Lambda^\mu{}_\nu$. First, the invariant square $p^2 \equiv \eta_{\mu\nu}p^\mu p^\nu$:

$$\eta_{\rho\sigma}p'^\rho p'^\sigma = \eta_{\rho\sigma}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu p^\mu p^\nu = \eta_{\mu\nu}p^\mu p^\nu, \tag{5.5}$$

where we used (3.5). Second, for $p^2 \leq 0$, also the sign of p^0. To see this, consider

$$p'^0 = \Lambda^0{}_\rho p^\rho = \Lambda^0{}_0 p^0 + \Lambda^0{}_i p^i. \tag{5.6}$$

For $p^2 \leq 0$, the three-vector (p^1, p^2, p^3) has length $\leq |p^0|$. From (3.17), we learn that the three-vector $(\Lambda^1{}_0, \Lambda^2{}_0, \Lambda^3{}_0)$ has length $\sqrt{(\Lambda^0{}_0)^2 - 1}$. Then, for example, for $p^0 \geq 0$ and proper Lorentz transformations, $\Lambda^0{}_0 \geq 1$, we have

$$p'^0 \geq p^0(\Lambda^0{}_0 - \sqrt{(\Lambda^0{}_0)^2 - 1}) \geq 0. \tag{5.7}$$

Hence, for each value of p^2, and (for $p^2 \leq 0$) for each sign of p^0, we can choose a 'standard' four-momentum, say k^μ, and express any p^μ of this class as

$$p^\mu = L^\mu{}_\nu(p)k^\nu, \tag{5.8}$$

where $L^\mu{}_\nu$ is some standard Lorentz transformation that depends on p^μ, and also implicitly on our choice of the standard k^μ. We can then *define* the states $\Psi_{p,\sigma}$ of momentum p by

$$\Psi_{p,\sigma} \equiv N(p) U(L(p)) \Psi_{k,\sigma}, \tag{5.9}$$

where $N(p)$ is a numerical normalization factor, to be chosen later. Up to this point, we said nothing about how the different σ labels are related for different momenta; Eq. (5.9) now fills that gap.

Operating on (5.9) with an arbitrary homogeneous Lorentz transformation $U(\Lambda)$, we now find

$$\begin{aligned} U(\Lambda)\Psi_{p,\sigma} &= N(p) U(\Lambda L(p))\Psi_{k,\sigma} \\ &= N(p) U(L(\Lambda p)) U(L^{-1}(\Lambda p)\Lambda L(p))\Psi_{k,\sigma}. \end{aligned} \tag{5.10}$$

The point of this last step is that the Lorentz transformation $L^{-1}(\Lambda p)\Lambda L(p)$ takes k to $L(p)k = p$, and then to Λp, and then back to k, so it belongs to the subgroup of homogeneous Lorentz group consisting of Lorentz transformations $W^\mu{}_\nu$ that leave k^μ invariant:

$$W^\mu{}_\nu k^\nu = k^\mu. \tag{5.11}$$

This subgroup is called the *little* group. For any W satisfying (5.11), we have

$$U(W)\Psi_{k,\sigma} = \sum_{\sigma'} D_{\sigma'\sigma}(W)\Psi_{k,\sigma'}. \tag{5.12}$$

The coefficients $D(W)$ furnish a representation of the little group; that is, for any elements $W, \overline{W}$ we have

$$\begin{aligned} \sum_{\sigma'} D_{\sigma'\sigma}(\overline{W}W)\Psi_{k,\sigma'} &= U(\overline{W}W)\Psi_{k,\sigma} = U(\overline{W})U(W)\Psi_{k,\sigma} \\ &= U(\overline{W})\sum_{\sigma''} D_{\sigma''\sigma}(W)\Psi_{k,\sigma''} = \sum_{\sigma'\sigma''} D_{\sigma''\sigma}(W) D_{\sigma'\sigma''}(\overline{W})\Psi_{k,\sigma'} \end{aligned} \tag{5.13}$$

and so

$$D_{\sigma'\sigma}(\overline{W}W) = \sum_{\sigma''} D_{\sigma'\sigma''}(\overline{W}) D_{\sigma''\sigma}(W). \tag{5.14}$$

In particular, we may apply eq. (5.12) to the little group transformation

$$W(\Lambda, p) \equiv L^{-1}(\Lambda p)\Lambda L(p) \tag{5.15}$$

and then (5.10) takes the form

$$U(\Lambda)\Psi_{p,\sigma} = N(p)\sum_{\sigma'} D_{\sigma'\sigma}(W(\Lambda,p))U(L(\Lambda p))\Psi_{k,\sigma'}\,, \tag{5.16}$$

or, recalling the definition (5.9):

$$U(\Lambda)\Psi_{p,\sigma} = \left(\frac{N(p)}{N(\Lambda p)}\right)\sum_{\sigma'} D_{\sigma'\sigma}(W(\Lambda,p))\Psi_{\Lambda p,\sigma'}\,. \tag{5.17}$$

Apart from the question of normalization, the problem of determining the coefficients $C_{\sigma'\sigma}$ in the transformation rule (5.4) has been reduced to the problem of finding the representations of the little group. This approach, of deriving representations of a group like the inhomogeneous Lorentz group from the representations of a little group, is called the method of induced representations.

The table below gives a convenient choice of the standard momentum k^μ and the corresponding little group for the various classes of four-momenta.

	Standard k^μ	Little Group
(a) $p^2 = -M^2 < 0,\ p^0 > 0$	$(0,0,0,M)$	$SO(3)$
(b) $p^2 = -M^2 < 0,\ p^0 < 0$	$(0,0,0,-M)$	$SO(3)$
(c) $p^2 = 0,\ p^0 > 0$	$(0,0,\kappa,\kappa)$	$ISO(2)$
(d) $p^2 = 0,\ p^0 < 0$	$(0,0,\kappa,-\kappa)$	$ISO(2)$
(e) $p^2 = N^2 > 0$	$(0,0,N,0)$	$SO(2,1)$
(f) $p^\mu = 0$	$(0,0,0,0)$	$SO(3,1)$

Here κ is an arbitrary positive energy, say 1 eV. The little groups are mostly pretty obvious: $SO(3)$ is the ordinary rotation group in three dimensions (excluding space-inversions), because rotations are the only proper orthochronous Lorentz transformations that leave at rest a particle with zero momentum, while $SO(2,1)$ and $SO(3,1)$ are the Lorentz groups in $(2+1)$- and $(3+1)$-dimensions, respectively. The group $ISO(2)$ is the group of Euclidean geometry, consisting of rotations and translations in two dimensions.

Of these six classes of four-momenta, only (a), (c) and (f) have any known interpretations in terms of physical states. Not much needs to be said here about case (f) – $p^\mu = 0$;

it describes the vacuum, which is simply left invariant under $U(\Lambda)$. In what follows, we will consider only cases (a) and (c), which cover particles of mass $M > 0$ and zero mass, respectively.

This is a good place to pause, and say something about the normalization of these states. By the usual orthonormalization procedure of quantum mechanics, we may choose the states with standard momentum k^μ to be orthonormal, in the sense that

$$(\Psi_{k',\sigma'}, \Psi_{k,\sigma}) = \delta^3(\mathbf{k}' - \mathbf{k})\delta_{\sigma\sigma'}. \tag{5.18}$$

(The delta function appears here because $\Psi_{k',\sigma'}$ and $\Psi_{k,\sigma}$ are eigenstates of a Hermitian operator with eigenvalues $\mathbf{k}'$ and $\mathbf{k}$, respectively.) This has the immediate consequence that the representation of the little group in eqs. (5.12) and (5.17) must be unitary

$$D^\dagger(W) = D^{-1}(W). \tag{5.19}$$

Now, what about scalar products for arbitrary momenta? Using the unitarity of the operator $U(\Lambda)$ in eqs. (5.9) and (5.17), we find for the scalar product:

$$\begin{aligned}(\Psi_{p',\sigma'}, \Psi_{p,\sigma}) &= N(p)\left(U^{-1}(L(p))\Psi_{p',\sigma'}, \Psi_{k,\sigma}\right) \\ &= N(p)N^*(p')D\left(W(L^{-1}(p), p')\right)_{\sigma\sigma'} \delta^3(\mathbf{k}' - \mathbf{k})\end{aligned} \tag{5.20}$$

where $k' \equiv L^{-1}(p)p'$. Since also $k \equiv L^{-1}(p)p$, the delta function $\delta^3(\mathbf{k}' - \mathbf{k})$ is proportional to $\delta^3(\mathbf{p}' - \mathbf{p})$. For $p' = p$, the little-group transformation here is trivial, $W(L^{-1}(p), p) = 1$, and so the scalar product is

$$(\Psi_{p',\sigma'}, \Psi_{p,\sigma}) = |N(p)|^2 \delta_{\sigma\sigma'}\delta^3(\mathbf{k}' - \mathbf{k}). \tag{5.21}$$

It remains to work out the proportionality factor relating $\delta^3(\mathbf{k}' - \mathbf{k})$ and $\delta^3(\mathbf{p}' - \mathbf{p})$. Note that the Lorentz-invariant integral of an arbitrary scalar function $f(p)$ over four-momenta with $-p^2 = M^2 \geq 0$ and $p^0 > 0$ (i.e. cases (a) or (c)) may be written

$$\begin{aligned}&\int d^4p\ \delta(p^2 + M^2)\theta(p^0)f(p) \\ &= \int d^3\mathbf{p}\ dp^0 \delta((p^0)^2 - \mathbf{p}^2 - M^2)\theta(p^0)f(\mathbf{p}, p^0) \\ &= \int d^3\mathbf{p} \frac{f(\mathbf{p}, \sqrt{\mathbf{p}^2 + M^2})}{2\sqrt{\mathbf{p}^2 + M^2}}\end{aligned} \tag{5.22}$$

($\theta(p^0)$ is the step function: $\theta(x) = 1$ for $x \geq 0$, $\theta(x) = 0$ for $x < 0$). We see that when integrating on the 'mass shell' $p^2 + M^2 = 0$, the invariant volume element is

$$d^3\mathbf{p}/\sqrt{\mathbf{p}^2 + M^2}. \tag{5.23}$$

The delta function is defined by

$$\begin{aligned} F(\mathbf{p}) &= \int F(\mathbf{p}')\delta^3(\mathbf{p} - \mathbf{p}')d^3\mathbf{p}' \\ &= \int F(\mathbf{p}')\left[\sqrt{\mathbf{p}'^2 + M^2}\delta^3(\mathbf{p} - \mathbf{p}')\right]\frac{d^3\mathbf{p}'}{\sqrt{\mathbf{p}'^2 + M^2}}, \end{aligned} \tag{5.24}$$

so we see that the invariant delta function is

$$\sqrt{\mathbf{p}'^2 + M^2}\delta^3(\mathbf{p} - \mathbf{p}') = p^0\delta^3(\mathbf{p} - \mathbf{p}'). \tag{5.25}$$

Since p' and p are related to k' and k respectively by a Lorentz transformation $L(p)$, we have then

$$p^0\delta^3(\mathbf{p} - \mathbf{p}') = k^0\delta^3(\mathbf{k} - \mathbf{k}') \tag{5.26}$$

and therefore

$$(\Psi_{p',\sigma'}, \Psi_{p,\sigma}) = |N(p)|^2\delta_{\sigma\sigma'}\left(\frac{p^0}{k^0}\right)\delta^3(\mathbf{p}' - \mathbf{p}). \tag{5.27}$$

The normalization factor $N(p)$ is sometimes chosen to be just $N(p) = 1$, but then we would need to keep track of the p^0/k^0 factor in scalar products. Instead, we will here adopt the more usual convention that

$$N(p) = \sqrt{k^0/p^0} \tag{5.28}$$

for which

$$(\Psi_{p',\sigma'}, \Psi_{p,\sigma}) = \delta_{\sigma\sigma'}\delta^3(\mathbf{p}' - \mathbf{p}). \tag{5.29}$$

We now consider the two cases of physical interest: particles of mass $M > 0$, and particles of zero mass.

Mass Positive-Definite

The little group here is the three-dimension rotation group. Its unitary representations can be broken up into a direct sum of irreducible unitary representations $D^{(j)}_{\sigma'\sigma}(R)$ of dimensionality $2j+1$, with $j = 0, \frac{1}{2}, 1, \cdots$. These can be built up from the standard matrices for infinitesimal rotations $R_{ik} = \delta_{ik} + \Theta_{ik}$, with $\Theta_{ik} = -\Theta_{ki}$ infinitesimal:

$$D^{(j)}_{\sigma'\sigma}(1+\Theta) = \delta_{\sigma'\sigma} + \frac{i}{2}\Theta_{ik}(J^{(j)}_{ik})_{\sigma'\sigma}, \tag{5.30}$$

$$\begin{aligned}(J^{(j)}_{23} \pm iJ^{(j)}_{31})_{\sigma'\sigma} &= (J^{(j)}_{1} \pm iJ^{(j)}_{2})_{\sigma'\sigma} \\ &= \delta_{\sigma',\sigma\pm1}\sqrt{(j\mp\sigma)(j\pm\sigma+1)},\end{aligned} \tag{5.31}$$

$$(J^{(j)}_{12})_{\sigma'\sigma} \;=\; (J^{(j)}_{3})_{\sigma'\sigma} \;=\; \sigma\delta_{\sigma'\sigma}, \tag{5.32}$$

with σ running over the values $j, j-1, \cdots, -j$. For a particle of mass $M > 0$ and spin j, eq. (5.17) now becomes

$$U(\Lambda)\Psi_{p,\sigma} = \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma'} D^{(j)}_{\sigma'\sigma}(W(\Lambda,p))\Psi_{\Lambda p,\sigma'}, \tag{5.33}$$

with the little-group element $W(\Lambda,p)$ (the Wigner rotation) given by eq. (5.15):

$$W(\Lambda,p) = L^{-1}(\Lambda p)\Lambda L(p). \tag{5.34}$$

To calculate this rotation, we need to choose a 'standard boost' $L(p)$ which carries the four-momentum from $k^\mu = (0,0,0,M)$ to p^μ. This is conveniently chosen as

$$\begin{aligned} L^i{}_k(p) &= \delta_{ik} + (\gamma-1)\hat{p}_i\hat{p}_k, \\ L^i{}_0(p) &= L^0{}_i(p) = \hat{p}_i\sqrt{\gamma^2-1}, \\ L^0{}_0(p) &= \gamma, \end{aligned} \tag{5.35}$$

where

$$\hat{p}_i \equiv p_i/|\mathbf{p}|, \qquad \gamma \equiv \sqrt{\mathbf{p}^2 + M^2}/M. \tag{5.36}$$

It is very important that when $\Lambda^\mu{}_\nu$ is an arbitrary three-dimensional rotation $\mathcal{R}$, the Wigner rotation $W(\Lambda,p)$ is the same as $\mathcal{R}$ for all p. To see this, note that the boost (5.35) may be expressed as

$$L(p) = R(\hat{\mathbf{p}})B(|\mathbf{p}|)R^{-1}(\hat{\mathbf{p}}), \tag{5.37}$$

where $R(\hat{\mathbf{p}})$ is a rotation (to be defined in a standard way below) that takes the three-axis into the direction of $\mathbf{p}$, and

$$B(|\mathbf{p}|) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & \sqrt{\gamma^2-1} \\ 0 & 0 & \sqrt{\gamma^2-1} & \gamma \end{pmatrix}. \tag{5.38}$$

Then for an arbitrary rotation $\mathcal{R}$

$$W(\mathcal{R},p) = R(\mathcal{R}\hat{p})B^{-1}(|\mathbf{p}|)R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}})B(|\mathbf{p}|)R^{-1}(\hat{\mathbf{p}}). \tag{5.39}$$

But the rotation $R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}})$ takes the three-axis into the direction $\hat{\mathbf{p}}$, and then into the direction $\mathcal{R}\hat{\mathbf{p}}$, and then back to the three-axis, so it must be just a rotation by some angle θ around the three-axis

$$R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}}) = R(\theta) \equiv \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.40}$$

Since $R(\theta)$ commutes with $B(|\mathbf{p}|)$, this now gives

$$W(\mathcal{R},p) = R(\mathcal{R}\hat{p})B^{-1}(|\mathbf{p}|)R(\theta)B(|\mathbf{p}|)R^{-1}(\hat{\mathbf{p}}) = R(\mathcal{R}\hat{p})R(\theta)R^{-1}(\hat{\mathbf{p}}) \tag{5.41}$$

and hence

$$W(\mathcal{R},p) = \mathcal{R} \tag{5.42}$$

as was to be shown. Thus the states of a moving massive particle (and, by extension, multi-particle states) have the same transformation under rotations as in non-relativistic quantum mechanics. This is another piece of good news – the whole apparatus of spherical harmonics, Clebsch-Gordan coefficients, etc. can be carried over wholesale from non-relativistic to relativistic quantum mechanics.

Mass Zero

First, we have to work out the structure of the little group. Consider an arbitrary little group element $W^\mu{}_\nu$, with $W^\mu{}_\nu k^\nu = k^\mu$, where k^μ is the standard four-momentum for this case, $k^\mu = (0,0,1,1)$. Acting on a time-like four-vector $t^\mu = (0,0,0,1)$, such a Lorentz transformation must yield a four-vector Wt whose length and scalar product with $Wk = k$ are the same as those for t:

$$\begin{aligned} (Wt)^\mu (Wt)_\mu &= t^\mu t_\mu = -1, \\ (Wt)^\mu k_\mu &= t^\mu k_\mu = -1. \end{aligned} \tag{5.43}$$

Any four-vector that satisfies the second condition may be written

$$(Wt)^{\mu} = (\alpha, \beta, \zeta, 1 + \zeta) \tag{5.44}$$

and the first condition yields the relation

$$\zeta = (\alpha^2 + \beta^2)/2. \tag{5.45}$$

It follows that the effect of $W^{\mu}{}_{\nu}$ on t^{ν} is the same as that of the Lorentz transformation

$$S^{\mu}{}_{\nu}(\alpha, \beta) = \begin{pmatrix} 1 & 0 & -\alpha & \alpha \\ 0 & 1 & -\beta & \beta \\ \alpha & \beta & 1 - \zeta & \zeta \\ \alpha & \beta & -\zeta & 1 + \zeta \end{pmatrix}. \tag{5.46}$$

This does not mean that W equals $S(\alpha, \beta)$, but it does mean that $S^{-1}(\alpha, \beta)W$ is a Lorentz transformation that leaves the time-like four-vector $(0, 0, 0, 1)$ invariant, and is therefore a pure rotation. Also, $S^{\mu}{}_{\nu}$ like $W^{\mu}{}_{\nu}$ leaves the light-like four-vector $(0, 0, 1, 1)$ invariant, so $S^{-1}(\alpha, \beta)W$ must be a rotation by some angle θ around the three-axis

$$S^{-1}(\alpha, \beta)W = R(\theta), \tag{5.47}$$

where

$$R^{\mu}{}_{\nu}(\theta) \equiv \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.48}$$

The most general element of the little group is therefore of the form

$$W(\theta, \alpha, \beta) = S(\alpha, \beta)R(\theta). \tag{5.49}$$

What group is this? We note that the transformations with $\theta = 0$ or with $\alpha = \beta = 0$ form subgroups:

$$S(\overline{\alpha}, \overline{\beta})S(\alpha, \beta) = S(\overline{\alpha} + \alpha, \overline{\beta} + \beta), \tag{5.50}$$

$$R(\overline{\theta})R(\theta) = R(\overline{\theta} + \theta). \tag{5.51}$$

These subgroups are *Abelian* – that is, their elements all commute with each other. Furthermore, the subgroup with $\theta = 0$ is *invariant*, in the sense that its elements are transformed into other elements of the same subgroup by any member of the group

$$R(\theta)S(\alpha, \beta)R^{-1}(\theta) = S(\alpha\cos\theta + \beta\sin\theta, -\alpha\sin\theta + \beta\cos\theta). \tag{5.52}$$

From eqs. (5.50)-(5.52) we can work out the product of any group elements. The student will recognize these multiplication rules as those of the subgroup $ISO(2)$, consisting of translations (by a vector (α, β)) and rotations (by an angle θ) in two dimensions.

Groups that do *not* have invariant subgroups have certain simple properties, and for this reason are called *semi-simple*. As we have seen, the little group $ISO(2)$ like the inhomogeneous Lorentz group is *not* semi-simple, and this leads to interesting complications. First, let's take a look at the Lie algebra of $ISO(2)$. For θ, α, β infinitesimal, the general group element is

$$\begin{aligned} W(\theta,\alpha,\beta)^\mu{}_\nu &= \delta^\mu{}_\nu + \omega^\mu{}_\nu, \\ \omega^\mu{}_\nu &= \begin{pmatrix} 0 & \theta & -\alpha & \alpha \\ -\theta & 0 & -\beta & \beta \\ \alpha & \beta & 0 & 0 \\ -\alpha & -\beta & 0 & 0 \end{pmatrix}. \end{aligned} \tag{5.53}$$

From (4.5), we see that the corresponding Hilbert space operator is

$$U(W(\theta,\alpha,\beta)) = 1 + i\alpha A + i\beta B + i\theta J_3, \tag{5.54}$$

where A and B are the Hermitian operators

$$\begin{aligned} A &= -J^{13} + J^{10} = J_2 + K_1, \\ B &= -J^{23} + J^{20} = -J_1 + K_2, \end{aligned} \tag{5.55}$$

and, as before, $J_3 = J_{12}$. Either from (4.21)-(4.23), or directly from eqs. (5.50)-(5.52), we see that these generators have the commutators

$$[J_3, A] = +iB, \tag{5.56}$$

$$[J_3, B] = -iA, \tag{5.57}$$

$$[A, B] = 0. \tag{5.58}$$

Since A and B are commuting Hermitian operators they (like the momentum generators of the inhomogeneous Lorentz group) can be simultaneously diagonalized by the states $\Psi_{k,a,b}$

$$\begin{aligned} A\Psi_{k,a,b} &= a\Psi_{k,a,b}, \\ B\Psi_{k,a,b} &= b\Psi_{k,a,b}. \end{aligned} \tag{5.59}$$

The problem is that if we find one such set of non-zero eigenvalues of A, B, then we find a whole continuum. From eq. (5.52), we have

$$\begin{aligned} U[R(\theta)]AU^{-1}[R(\theta)] &= A\cos\theta - B\sin\theta, \\ U[R(\theta)]BU^{-1}[R(\theta)] &= A\sin\theta + B\cos\theta, \end{aligned} \tag{5.60}$$

and so, for arbitrary θ,

$$\begin{aligned} A\Psi^{\theta}_{k,a,b} &= (a\cos\theta - b\sin\theta)\Psi^{\theta}_{k,a,b}, \\ B\Psi^{\theta}_{k,a,b} &= (a\sin\theta + b\cos\theta)\Psi^{\theta}_{k,a,b}, \end{aligned} \tag{5.61}$$

where

$$\Psi^{\theta}_{k,a,b} \equiv U^{-1}(R(\theta))\Psi_{k,a,b}. \tag{5.62}$$

Massless particles are not observed to have any continuous degree of freedom like θ; to avoid such a continuum of states, we must require that physical states (now called $\Psi_{k,\sigma}$) are eigenvectors of A and B with $a = b = 0$:

$$A\Psi_{k,\sigma} = B\Psi_{k,\sigma} = 0. \tag{5.63}$$

These states are then distinguished by the eigenvalue of the remaining generator

$$J_3\Psi_{k,\sigma} = \sigma\Psi_{k,\sigma}. \tag{5.64}$$

Since the momentum $\mathbf{k}$ is in the three-direction, σ gives the component of angular momentum in the direction of motion, or *helicity*.

We are now in position to calculate the Lorentz transformation properties of general massless particle states. First note that by use of the general arguments of the previous sections, eq. (5.54) generalizes for finite α and β to

$$U(S(\alpha,\beta)) = \exp(i\alpha A + i\beta B) \tag{5.65}$$

and for finite θ to

$$U(R(\theta)) = \exp(iJ_3\theta). \tag{5.66}$$

An arbitrary element W of the little group can be put in the form (5.49), so that

$$U(W)\Psi_{k,\sigma} = \exp(i\alpha A + i\beta B)\exp(iJ_3\theta)\Psi_{k,\sigma} \tag{5.67}$$

and therefore eq. (5.12) gives

$$D_{\sigma'\sigma}(W) = \exp(i\theta\sigma)\delta_{\sigma'\sigma}, \tag{5.68}$$

where θ is the angle defined by expressing W as in eq. (5.49). The Lorentz transformation rule for a massless particle of arbitrary helicity is now given by eqs. (5.17) and (5.28) as

$$U(\Lambda)\Psi_{p,\sigma} = \sqrt{\frac{(\Lambda p)^0}{p^0}} \exp(i\sigma\theta(\Lambda, p))\Psi_{\Lambda p,\sigma} \tag{5.69}$$

with $\theta(\Lambda, p)$ defined by

$$W(\Lambda, p) \equiv L^{-1}(\Lambda p)\Lambda L(p) \equiv S(\alpha(\Lambda, p), \beta(\Lambda, p))R(\theta(\Lambda, p)). \tag{5.70}$$

The electromagnetic gauge invariance arises from the part of the little group parameterized by α and β.

At this point, we have not yet encountered any reason that would forbid the helicity σ of a massless particle from being an arbitrary real number. There are however topological considerations that restrict the allowed values of σ to integers and half-integers, just as for massive particles.

To calculate the little group element (5.70) for a given Λ and p, (and also to enable us to calculate the effect of space or time inversion on these states in the next section) we need to fix a convention for the standard Lorentz transformation that takes us from $k^\mu = (0, 0, \kappa, \kappa)$ to p^μ. This may be conveniently chosen to have the form

$$L(p) = R(\hat{\mathbf{p}})B(|\mathbf{p}|/\kappa) \tag{5.71}$$

where $B(u)$ is a pure boost along the three-direction:

$$B(u) \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (u^2+1)/2u & (u^2-1)/2u \\ 0 & 0 & (u^2-1)/2u & (u^2+1)/2u \end{pmatrix} \tag{5.72}$$

and $R(\hat{\mathbf{p}})$ is a pure rotation that carries the three-axis into the direction of the unit vector $\hat{\mathbf{p}}$. For instance, suppose we take $\hat{\mathbf{p}}$ to have polar and azimuthal angles θ and ϕ:

$$\hat{\mathbf{p}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta). \tag{5.73}$$

Then we can take $R(\hat{\mathbf{p}})$ as a rotation by an angle θ around the two-axis, which takes $(0,0,1)$ into $(\sin\theta, 0, \cos\theta)$, followed by a rotation by angle ϕ around the three-axis:

$$U(R(\hat{\mathbf{p}})) = \exp(i\phi J_3)\exp(i\theta J_2), \tag{5.74}$$

where $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$. (We give $U(R(\hat{\mathbf{p}}))$ rather than $R(\hat{\mathbf{p}})$, together with a specification of the range of ϕ and θ, because shifting θ or ϕ by 2π would give the same rotation $R(\hat{\mathbf{p}})$, but a different sign for $U(R(\hat{\mathbf{p}}))$ when acting on half-integer spin states.) Since (5.74) is a rotation, and does take the three-axis into the direction (5.73), any other choice of such $R(\hat{\mathbf{p}})$ would differ from this one by at most an initial rotation around the three-axis, corresponding to a mere redefinition of the phase of the one-particle states.

Note that the helicity is Lorentz-invariant; a massless particle of a given helicity σ looks the same (aside from its momentum) in all inertial frames. Indeed, we would be justified in thinking of massless particles of each different helicity as different species of particles. However, particles of opposite helicities are related by the symmetry of space inversion. Thus, because electromagnetic and gravitational forces obey space inversion symmetry, the massless particles of helicity ± 1 associated with electromagnetic phenomena are both called *photons*, and the massless particles of helicity ± 2 that are believed to be associated with gravitation are both called *gravitons*. On the other hand, the supposedly massless particles of helicity $\pm 1/2$ that are emitted in nuclear beta decay have no interaction (apart from gravitation) that respect the symmetry of space inversion, so these particles are given different names: *neutrinos* for helicity $+1/2$, and *antineutrinos* for helicity $-1/2$.

Even though the helicity of a massless particle is Lorentz-invariant, the state itself is not. In particular, because of the helicity-dependent phase factor $\exp(i\sigma\theta)$ in eq. (5.69), a state formed as a linear superposition of one-particle states with opposite helicities will be changed by a Lorentz transformation into a different superposition. For instance, a general one-photon state of four-momenta p may be written

$$\Psi_{p;\alpha} = \alpha_+ \Psi_{p,+1} + \alpha_- \Psi_{p,-1}, \tag{5.75}$$

where

$$|\alpha_+|^2 + |\alpha_-|^2 = 1. \tag{5.76}$$

The generic case is one of *elliptic polarization*, with $|\alpha_\pm|$ both non-zero and unequal. *Circular polarization* is the limiting case where either α_+ or α_- vanishes, and *linear polarization* is the opposite extreme, with $|\alpha_+| = |\alpha_-|$. The overall phase of α_+ and α_- has no physical significance, and for linear polarization may be adjusted so that $\alpha_- = \alpha_+^*$, but the relative phase is still important. Indeed, for linear polarizations with $\alpha_- = \alpha_+^*$, the phase of α_+ may be identified as the angle between the plane of polarization and some fixed reference direction perpendicular to $\hat{\mathbf{p}}$. Eq. (5.69) shows that under a Lorentz transformation $\Lambda^\mu{}_\nu$, this angle rotates by an amount $\theta(\Lambda, p)$. Plane polarized gravitons can be defined in a similar way, and here eq. (5.69) has the consequence that a Lorentz transformation Λ rotates the plane of polarization by an angle $2\theta(\Lambda, p)$.

6. Space Inversion and Time-Reversal

We saw previously that any homogeneous Lorentz transformation is either proper and orthochronous (i.e. $\det \Lambda = +1$ and $\Lambda^0{}_0 \geq +1$) or else equal to a proper orthochronous transformation times either $\mathcal{P}$ or $\mathcal{PT}$, where $\mathcal{P}$ and $\mathcal{T}$ are the space inversion and time-reversal transformations

$$\mathcal{P}^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{T}^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{6.1}$$

It used to be thought self-evident that the fundamental multiplication rule of the Poincaré group

$$U(\overline{\Lambda}, \overline{a})U(\Lambda, a) = U(\overline{\Lambda}\Lambda, \overline{\Lambda}a + \overline{a}) \tag{6.2}$$

would be valid even if Λ and/or $\overline{\Lambda}$ involved factors of $\mathcal{P}$ or $\mathcal{T}$ or $\mathcal{PT}$. In particular, it was believed that there are operators corresponding to $\mathcal{P}$ and $\mathcal{T}$ themselves:

$$\mathrm{P} \equiv U(\mathcal{P}, 0), \quad \mathrm{T} \equiv U(\mathcal{T}, 0) \tag{6.3}$$

such that

$$\mathrm{P}U(\Lambda, a)\mathrm{P}^{-1} = U(\mathcal{P}\Lambda\mathcal{P}^{-1}, \mathcal{P}a), \tag{6.4}$$

$$\mathrm{T}U(\Lambda, a)\mathrm{T}^{-1} = U(\mathcal{T}\Lambda\mathcal{T}^{-1}, \mathcal{T}a), \tag{6.5}$$

for any proper orthochronous Lorentz transformation $\Lambda^\mu{}_\nu$ and translation a^μ. These transformation rules incorporate most of what is meant when we say that P or T are 'conserved'.

In 1956-57 it became understood that this is true for P only in the approximation in which one ignores the effects of weak interactions, such as those that produce nuclear beta decay. Time-reversal survived for a while, but in 1964 there appeared indirect evidence that these properties of T are also only approximately satisfied. In what follows, we will make believe that operators P and T satisfying eqs. (6.4) and (6.5) actually exist, but it should be kept in mind that this is only an approximation.

Let us apply eqs. (6.4) and (6.5) in the case of an infinitesimal transformation, i.e.

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu, \quad a^\mu = \epsilon^\mu \tag{6.6}$$

with $\omega_{\mu\nu} = -\omega_{\nu\mu}$ and ϵ_μ both infinitesimal. Using (4.5), and equating coefficients of $\omega_{\rho\sigma}$ and ϵ_ρ in eqs. (6.4) and (6.5), we obtain the P and T transformation properties of the Poincaré generators

$$\mathrm{P}iJ^{\rho\sigma}\mathrm{P}^{-1} = i\mathcal{P}_\mu{}^\rho\mathcal{P}_\nu{}^\sigma J^{\mu\nu}, \tag{6.7}$$

$$\mathrm{P}iP^\rho\mathrm{P}^{-1} = i\mathcal{P}_\mu{}^\rho P^\mu, \tag{6.8}$$

$$\mathrm{T}iJ^{\rho\sigma}\mathrm{T}^{-1} = i\mathcal{T}_\mu{}^\rho\mathcal{T}_\nu{}^\sigma J^{\mu\nu}, \tag{6.9}$$

$$\mathrm{T}iP^\rho\mathrm{T}^{-1} = i\mathcal{T}_\mu{}^\rho P^\mu. \tag{6.10}$$

This is much like eqs. (4.11) and (4.12), except that we have not cancelled factors of i on both sides of these equations, because at this point we have not yet decided whether P and T are linear and unitary or antilinear and antiunitary.

The decision is an easy one. Setting $\rho = 0$ in eq. (6.8) gives

$$\mathrm{P}iH\mathrm{P}^{-1} = iH, \tag{6.11}$$

where $H \equiv P^0$ is the energy operator. If P were antiunitary and antilinear then it would anticommute with i, so $\mathrm{P}H\mathrm{P}^{-1} = -H$. But then for any state Ψ of energy $E > 0$, there would have to be another state $\mathrm{P}^{-1}\Psi$ of energy $-E < 0$. There are no states of negative energy (energy less than that of the vacuum), so we are forced to choose the other alternative: P *is linear and unitary and commutes rather than anticommutes with* H.

On the other hand, setting $\rho = 0$ in eq. (6.10) yields

$$\mathrm{T}iH\mathrm{T}^{-1} = -iH. \tag{6.12}$$

If we supposed that T is unitary and linear then we could simply cancel the is, and find $\mathrm{T}H\mathrm{T}^{-1} = -H$, with the again disastrous conclusion that for any state Ψ of energy E

there is another state $\mathrm{T}^{-1}\Psi$ of energy $-E$. To avoid this, we are forced here to conclude that T is *antilinear and antiunitary.*

Now that we have decided that P is linear and T is antilinear, we can conveniently rewrite eqs. (6.7)-(6.10) in terms of the generators (4.18)-(4.20) in a three-dimensional notation

$$\mathrm{P}\mathbf{J}\mathrm{P}^{-1} = +\mathbf{J}, \tag{6.13}$$

$$\mathrm{P}\mathbf{K}\mathrm{P}^{-1} = -\mathbf{K}, \tag{6.14}$$

$$\mathrm{P}\mathbf{P}\mathrm{P}^{-1} = -\mathbf{P}, \tag{6.15}$$

$$\mathrm{T}\mathbf{J}\mathrm{T}^{-1} = -\mathbf{J}, \tag{6.16}$$

$$\mathrm{T}\mathbf{K}\mathrm{T}^{-1} = +\mathbf{K}, \tag{6.17}$$

$$\mathrm{T}\mathbf{P}\mathrm{T}^{-1} = -\mathbf{P}, \tag{6.18}$$

and, as shown before,

$$\mathrm{P}H\mathrm{P}^{-1} = \mathrm{T}H\mathrm{T}^{-1} = H. \tag{6.19}$$

It is physically sensible that P should preserve the sign of $\mathbf{J}$, because at least the orbital part is a vector product $\mathbf{r}\times\mathbf{p}$ of two vectors, both of which change sign under an inversion of the spatial coordinate system. On the other hand, T reverses $\mathbf{J}$, because after time-reversal an observer will see all bodies spinning in the opposite direction. Note by the way that eq. (6.16) is consistent with the angular momentum commutation relations $\mathbf{J}\times\mathbf{J} = i\mathbf{J}$, because T reverses not only $\mathbf{J}$ but also i. The student can easily check that eqs. (6.13)-(6.19) are consistent with all the commutation relations (4.21)-(4.27).

Let us now consider what P and T do to one-particle states:

P : $M > 0$

The one-particle states $\Psi_{k,\sigma}$ are defined as eigenvectors of $\mathbf{P}$, H and J_3 with eigenvalues 0, M and σ, respectively. From eqs. (6.13), (6.15) and (6.19), we see that the same must be true of the state $\mathrm{P}\Psi_{k,\sigma}$ and therefore (barring degeneracies) these states can only differ by a phase

$$\mathrm{P}\Psi_{k,\sigma} = \eta_\sigma \Psi_{k,\sigma} \tag{6.20}$$

with a phase factor ($|\eta| = 1$) that may or may not depend on the spin σ. To see that η_σ is σ-independent, we note from (5.12), (5.30) and (5.31) that

$$(J_1 \pm iJ_2)\Psi_{k,\sigma} = \sqrt{(j\mp\sigma)(j\pm\sigma+1)}\Psi_{k,\sigma\pm1}, \tag{6.21}$$

where j is the particle's spin. Operating on both sides with P, we find

$$\eta_\sigma = \eta_{\sigma\pm 1} \tag{6.22}$$

and so η_σ is actually independent of σ. We therefore write

$$\mathrm{P}\Psi_{k,\sigma} = \eta\Psi_{k,\sigma} \tag{6.23}$$

with η a phase, known as the *intrinsic parity*, that depends only on the species of particle on which P acts.

To get to finite momentum states, we must apply the unitary operator $U(L(p))$ corresponding to the 'boost' (5.35):

$$\Psi_{p,\sigma} = \sqrt{M/p^0}U(L(p))\Psi_{k,\sigma}. \tag{6.24}$$

We note that

$$\begin{aligned} \mathcal{P}L(p)\mathcal{P}^{-1} &= L(\mathcal{P}p), \\ \mathcal{P}p &= (-\mathbf{p}, \sqrt{\mathbf{p}^2 + M^2}) \end{aligned} \tag{6.25}$$

so using eqs. (6.4) and (6.23), we have

$$\mathrm{P}\Psi_{p,\sigma} = \sqrt{M/p^0}U(L(\mathcal{P}p))\eta\Psi_{k,\sigma} \tag{6.26}$$

or in other words

$$\mathrm{P}\Psi_{p,\sigma} = \eta\Psi_{\mathcal{P}p,\sigma} \tag{6.27}$$

$\mathrm{T} : M > 0$

From eqs. (6.16), (6.18) and (6.19), we see that the effect of T on the zero-momentum one-particle state $\Psi_{k,\sigma}$ is to yield a state with

$$\begin{aligned} \mathbf{P}(\mathrm{T}\Psi_{k,\sigma}) =& 0, \\ H(\mathrm{T}\Psi_{k,\sigma}) =& M(\mathrm{T}\Psi_{k,\sigma}), \\ J_3(\mathrm{T}\Psi_{k,\sigma}) =& -\sigma(\mathrm{T}\Psi_{k,\sigma}), \end{aligned} \tag{6.28}$$

and so

$$\mathrm{T}\Psi_{k,\sigma} = \zeta_\sigma\Psi_{k,-\sigma}, \tag{6.29}$$

where ζ_σ is a phase factor. Applying the operator T to (6.21), and recalling that T anticommutes with $\mathbf{J}$ *and* i, we find

$$(-J_1 \pm iJ_2)\zeta_\sigma\Psi_{k,-\sigma} = \sqrt{(j \mp \sigma)(j \pm \sigma + 1)}\Psi_{k,-\sigma\mp 1}. \tag{6.30}$$

Using eq. (6.21) again on the left, we see that the square-root factors cancel, and so

$$-\zeta_\sigma = \zeta_{\sigma\pm 1}. \tag{6.31}$$

We write the solution as $\zeta_\sigma = \zeta(-)^{j-\sigma}$, with ζ another phase that depends only on the species of particle:

$$\mathrm{T}\Psi_{k,\sigma} = \zeta(-)^{j-\sigma}\Psi_{k,-\sigma}. \tag{6.32}$$

However, unlike the 'intrinsic parity' η, the time-reversal phase ζ has no physical significance. This is because we can redefine the one-particle state by a change of phase

$$\Psi_{k,\sigma} \to \Psi'_{k,\sigma} = \zeta^{1/2}\Psi_{k,\sigma}, \tag{6.33}$$

so that the phase ζ is eliminated from the transformation rule

$$\mathrm{T}\Psi'_{k,\sigma} = \zeta^{*1/2}\mathrm{T}\Psi_{k,\sigma} = \zeta^{*1/2}\zeta(-)^{j-\sigma}\Psi_{k,-\sigma} = (-)^{j-\sigma}\Psi'_{k,-\sigma}. \tag{6.34}$$

In what follows we will keep the arbitrary phase in (6.32), just to keep open our options in choosing the phase of one-particle states, but it should be kept in mind that this phase is of no real importance.

To deal with states of finite momentum, we again apply the 'boost' (5.35). Note that

$$\begin{aligned} \mathcal{T}L(p)\mathcal{T}^{-1} &= L(\mathcal{P}p), \\ \mathcal{P}p &= (-\mathbf{p}, \sqrt{\mathbf{p}^2 + M^2}). \end{aligned} \tag{6.35}$$

(That is, changing the sign of each element of $L^\mu{}_\nu$ with an odd number of *time*-indices is the same as changing the signs of elements with an odd number of *space*-indices.) Using eqs. (6.5) and (5.9), we have then

$$\mathrm{T}\Psi_{p,\sigma} = \zeta(-)^{j-\sigma}\Psi_{\mathcal{P}p,-\sigma}. \tag{6.36}$$

$\mathrm{P} : M = 0$

Acting on a state $\Psi_{k,\sigma}$, that is defined as an eigenvector of P^μ with eigenvalue $k^\mu = (0, 0, \kappa, \kappa)$ and an eigenvector of J_3 with eigenvalue σ, the parity operator P yields a state with four-momentum $(\mathcal{P}k)^\mu = (0, 0, -\kappa, \kappa)$ and J_3 equal to σ. Thus it takes a state of helicity (the component of spin along the direction of motion) σ into one of helicity $-\sigma$. As mentioned earlier, this shows that the existence of space-inverion symmetry requires that any species of massless particles with non-zero helicity must be accompanied with another of opposite helicity. Because P does not leave the standard momentum invariant, it is convenient to consider instead the operator $U(R_2^{-1})\mathrm{P}$, where R_2 is a rotation that also takes k to $\mathcal{P}k$, conveniently chosen as a rotation by -180^o around the two-axis

$$U(R_2) = \exp(-i\pi J_2). \tag{6.37}$$

Since $U(R_2^{-1})$ reverses the sign of J_3, we have

$$U(R_2^{-1})\mathrm{P}\Psi_{k,\sigma} = \eta_\sigma \Psi_{k,-\sigma} \tag{6.38}$$

with η_σ a phase factor. Now $R_2^{-1}\mathcal{P}$ commutes with the Lorentz 'boost' (5.72), and $\mathcal{P}$ commutes with the rotation $R(\hat{p})$ which takes the three-direction into the direction of $\mathbf{p}$, so by operating on (5.9) with P, we find for a general four-momentum p^μ

$$\begin{aligned} \mathrm{P}\Psi_{p,\sigma} &= \sqrt{\frac{\kappa}{p^0}} U\left(R(\hat{p})R_2 B\left(\frac{|\mathbf{p}|}{\kappa}\right)\right) U(R_2^{-1})\mathrm{P}\Psi_{k,\sigma} \\ &= \sqrt{\frac{\kappa}{p^0}} \eta_\sigma U\left(R(\hat{p})R_2 B\left(\frac{|\mathbf{p}|}{\kappa}\right)\right) \Psi_{k-\sigma}. \end{aligned} \tag{6.39}$$

Note that $R(\hat{\mathbf{p}})R_2$ is a rotation that takes the three-axis into the direction of $-\hat{\mathbf{p}}$, but $U(R(\hat{\mathbf{p}})R_2)$ is not quire equal to $U(R(-\hat{\mathbf{p}}))$. According to (5.74),

$$U(R(-\hat{\mathbf{p}})) = \exp\left(i(\phi \pm \pi)J_3\right) \exp\left(i(\pi - \theta)J_2\right) \tag{6.40}$$

with azimuthal angle chosen as $\phi + \pi$ or $\phi - \pi$ according to whether $0 \le \phi \le \pi$ or $\pi \le \phi \le 2\pi$, so that it remains in the range of 0 to 2π. Then

$$\begin{aligned} U^{-1}\left(R(-\hat{\mathbf{p}})\right) U\left(R(\hat{\mathbf{p}})R_2\right) &= \exp\left(-i(\pi - \theta)J_2\right) \\ &\times \exp\left(-i(\phi \pm \pi)J_3\right) \exp(i\phi J_3) \exp(i\theta J_2) \exp(-i\pi J_2) \\ &= \exp\left(-i(\pi - \theta)J_2\right) \exp\left(\mp i\pi J_3\right) \exp\left(-i(\pi - \theta)J_2\right). \end{aligned} \tag{6.41}$$

But a rotation of $\pm 180^o$ around the three-axis reverses the sign of J_2, so

$$U\left(R(\hat{\mathbf{p}})R_2\right) = U\left(R(-\hat{\mathbf{p}})\right) \exp(\pm i\pi J_3). \tag{6.42}$$

Also, $R(-\hat{\mathbf{p}})B(|\mathbf{p}|/k)$ is just the standard boost $L(\mathcal{P}p)$ in the direction $\mathcal{P}p = (-\mathbf{p}, p^0)$. We have then finally

$$\mathrm{P}\Psi_{p,\sigma} = \eta_\sigma \exp(\mp i\pi\sigma)\Psi_{\mathcal{P}p,-\sigma} \tag{6.43}$$

with the phase $-\pi\sigma$ or $+\pi\sigma$ according to whether the two-component of $\mathbf{p}$ is positive or negative, respectively. This peculiar change of sign in the operation of parity for massless particles of half-integer spin is due to the convention adopted in eq. (5.74) for the rotation used to define massless particle states of arbitrary momentum. Because the rotation group is not simply connected, some discontinuity of this sort is unavoidable.

$\mathrm{T} : M = 0$

Acting on a state $\Psi_{k,\sigma}$, which has values $k^\mu = (0, 0, \kappa, \kappa)$ and σ for P^μ and J_3, the time-reversal operator T yields a state which has values $(\mathcal{P}k)^\mu = (0, 0, -\kappa, \kappa)$ and $-\sigma$ for P^μ and J_3. This does not change the helicity $\mathbf{J} \cdot \hat{k}$, and by itself has nothing to say about whether massless particles of one helicity σ are accompanied with others of helicity $-\sigma$. Because T like P does not leave the standard four-momentum k invariant, it is convenient to consider the generator $U(R_2^{-1})\mathrm{T}$, where R_2 is the rotation (6.37), which also takes k to $\mathcal{P}k$. This commutes with J_3, so

$$U(R_2^{-1})\mathrm{T}\Psi_{k,\sigma} = \zeta_\sigma \Psi_{k,\sigma} \tag{6.44}$$

with ζ_σ another phase. Since $R_2^{-1}\mathcal{T}$ commutes with the 'boost' (5.72), and $\mathcal{T}$ commutes with the rotation $R(\hat{p})$, operating with T on the state (5.9) gives

$$\mathrm{T}\Psi_{p,\sigma} = \sqrt{\frac{\kappa}{p^0}} U \left[R(\hat{p}) R_2 B \left(\frac{|\mathbf{p}|}{\kappa} \right) \right] \zeta_\sigma \Psi_{k,\sigma}. \tag{6.45}$$

Using (6.42), this yields finally

$$\mathrm{T}\Psi_{p,\sigma} = \zeta_\sigma \exp(\pm i\pi\sigma) \Psi_{\mathcal{P}p,\sigma}. \tag{6.46}$$

Again, the top or bottom sign applies according to whether the two-component of $\mathbf{p}$ is positive or negative, respectively.

* * *

It is interesting that the square T^2 of the time-reversal operator has a very simple action on both massive and massless one-particle states. Using eq. (6.36), and recalling that T is antiunitary, we see that for massive one-particle states:

$$\mathrm{T}^2\Psi_{p,\sigma} = \mathrm{T}\zeta(-)^{j-\sigma}\Psi_{\mathcal{P}p,-\sigma} = \zeta^*(-)^{j-\sigma}\zeta(-)^{j+\sigma}\Psi_{p,\sigma} \tag{6.47}$$

or in other words

$$\mathrm{T}^2\Psi_{p,\sigma} = (-)^{2j}\Psi_{p,\sigma}. \tag{6.48}$$

We get the same result for massless particles. If the two-component of $\mathbf{p}$ is positive then the two component of $\mathcal{P}\mathbf{p}$ is negative, and vice-versa, so eq. (6.46) gives

$$\begin{aligned} \mathrm{T}^2\Psi_{p,\sigma} &= \mathrm{T}\zeta_\sigma \exp(\pm i\pi\sigma)\Psi_{\mathcal{P}p,\sigma} = \zeta_\sigma^* \exp(\mp i\pi\sigma)\zeta_\sigma \exp(\mp i\pi\sigma)\Psi_{p,\sigma} \\ &= \exp(\mp 2i\pi\sigma)\Psi_{p,\sigma}. \end{aligned} \tag{6.49}$$

As long as σ is an integer or half-integer, this can be written

$$\mathrm{T}^2\Psi_{p,\sigma} = (-)^{2|\sigma|}\Psi_{p,\sigma}. \tag{6.50}$$

By the 'spin' of a massless particle, we usually mean the absolute value of the helicity, so eq. (6.50) is the same as eq. (6.48).

This result has an interesting consequence. When T^2 acts on any state Ψ of a system of non-interacting particles, either massive or massless, it yields a factor $(-)^{2j}$ or $(-)^{2|\sigma|}$ for each particle. Hence if the state contains an odd number of particles of half-integer spin or helicity (plus any number of particles of integer spin or helicity), we get an overall change of sign

$$\mathrm{T}^2\Psi = -\Psi. \tag{6.51}$$

If we now 'turn on' various interactions, this result will be preserved, provided these interactions respect invariance under time-reversal, even if they do not respect invariance under rotations. (For instance, these arguments will apply even if our system is subjected to arbitrary static gravitational and electric fields.) Now, suppose that Ψ is an eigenstate of the Hamiltonian. Since T commutes with the Hamiltonian, $\mathrm{T}\Psi$ will also be an eigenstate of the Hamiltonian. Is it the same state? If so, then $\mathrm{T}\Psi$ can differ from Ψ only by a phase

$$\mathrm{T}\Psi = \zeta\Psi, \tag{6.52}$$

but then

$$\mathrm{T}^2\Psi = \mathrm{T}(\zeta\Psi) = \zeta^*\mathrm{T}\Psi = |\zeta|^2\Psi = \Psi, \tag{6.53}$$

in contradiction with eq. (6.51). We see that any energy eigenstate Ψ satisfying eq. (6.51) must be degenerate with another eigenstate of the same energy. This is known as the 'Kramers degeneracy'. Of course, this conclusion is trivial if the system is in a rotationally invariant environment, because the total angular-momentum j of any state of this system would have to be half-integer, and there would therefore be $2j + 1 = 2, 4, \cdots$ degenerate states. The surprising result is that at least a two-fold degeneracy persists even if rotational invariance is perturbed by external fields, such as electrostatic fields, as long as these fields are invariant under T. In particular, if any particle had an electric or gravitational dipole moment then the degeneracy among its $2j + 1$ spin states would be entirely removed in a static electric or gravitational field, so such dipole moments are forbidden by time-reversal invariance.

For the sake of completeness, it should be mentioned that P and T can have more complicated effects on multiplets of particles with the same mass. No physically relevant examples are known.

References

[1] S. Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, 1995), chapter 2.