Manifestly Covariant Lagrangians, Classical Particles with Spin, and the Origins of Gauge Invariance

Jacob A. Barandes[1], [*]

[1]*Jefferson Physical Laboratory, Harvard University, Cambridge, MA 02138*
(Dated: July 28, 2023)

In this paper, we review a general technique for converting the standard Lagrangian description of a classical system into a formulation that puts time on an equal footing with the system's degrees of freedom. We show how the resulting framework anticipates key features of special relativity, including the signature of the Minkowski metric tensor and the special role played by theories that are invariant under a generalized notion of Lorentz transformations. We then use this technique to revisit a classification of classical particle-types that mirrors Wigner's classification of quantum particle-types in terms of irreducible representations of the Poincaré group, including the cases of massive particles, massless particles, and tachyons. Along the way, we see gauge invariance naturally emerge in the context of classical massless particles with nonzero spin, as well as study the massless limit of a massive particle and derive a classical-particle version of the Higgs mechanism.

I. INTRODUCTION

The Lagrangian formulation of classical physics provides an elegant and powerful set of techniques for analyzing the behavior of physical systems. For classical fields, it is customary to employ Lagrangians that make the symmetries of special relativity manifest, but textbook treatments of mechanical systems tend to treat time and energy very differently from degrees of freedom and momenta.

In this paper,[1] we cast new light on a technique for resolving this shortcoming. Among its useful features, we show that this framework anticipates key aspects of special relativity, like the signature of the Minkowski metric tensor and the special role played by classical systems that exhibit generalizations of Lorentz invariance.

Extending earlier work, including [2–5], we then present a fully classical version of Wigner's famous classification [6] of quantum particles into general types—massive, massless, and tachyonic. In close parallel with Wigner's construction, which is based on identifying the Hilbert spaces of quantum particles with irreducible representations of the Poincaré group, our classification of classical particle-types consists of identifying their phase spaces with "irreducible" (or, more properly, transitive) group actions of the Poincaré group, so that those phase spaces serve as homogeneous spaces for the Poincaré group. Our classical particles generically possess fixed total spin, but without spin quantization, and therefore correspond to the limit of large spin quantum numbers.

Along the way, and as a case study in how kinematics can determine dynamics, we show that the structure of these phase spaces leads to a simple Lagrangian formulation that can handle both massive and massless particles, and that neatly accommodates spin. In addition,

by paying careful attention to the compactness properties of these phase spaces at fixed energy, we show that physically acceptable massless particles with spin feature a classical point-particle manifestation of gauge invariance that is deeply connected to the gauge invariance of electromagnetism—meaning that this form of gauge invariance is not solely a property of classical field theory or of relativistic quantum mechanics. By studying the relationship between the massive and massless cases through the massless limit, we also derive a classical point-particle version of the Higgs mechanism.

II. THE LAGRANGIAN FORMULATION

We start with a brief review of general classical systems and their standard Lagrangian formulation.[2] Afterward, we will turn to the development of a manifestly covariant approach.

A. Classical Systems

In general, the mathematical description of a classical system consists of a configuration space whose points denote the possible "snapshots" that the system can occupy, together with a list of rules or laws that determine how the system's instantaneous configuration is allowed to evolve.

If q_α are a collection of independent numerical coordinates that label the points in the system's configuration space, with α an index distinguishing the different coordinates, then we call q_α a set of degrees of freedom for the system. We will assume for simplicity that we can cover the entire configuration space with a single such coordinate system, apart from possible regions of measure zero where the coordinates are not well-defined.

---

[*] jacob_barandes@harvard.edu
[1] For a synopsis of the results obtained in this paper, see [1].

[2] For a more extensive introduction, see [7].

A candidate trajectory of the system is an arbitrary smooth path through the system's configuration space, and is conveniently defined by specifying the system's degrees of freedom $q_\alpha(t)$ as functions of a real-valued parameter t called the time. The system's rates of change are then denoted by $\dot{q}_\alpha(t)$, where dots denote derivatives with respect to t:

$$\dot{q}_\alpha(t) \equiv \frac{dq_\alpha(t)}{dt}, \tag{1}$$

$$\ddot{q}_\alpha(t) \equiv \frac{d^2 q_\alpha(t)}{dt^2}, \tag{2}$$

and so forth. Altogether, the system's configuration space, a choice of degrees of freedom q_α, and all the system's candidate trajectories make up the system's kinematics.

The rules or laws that prescribe which *candidate* trajectories are *physical* trajectories that the system can actually follow make up the system's dynamics. In the simplest cases, these rules take the form of first- or second-order differential equations of the form

$$f_\alpha(q, \dot{q}, \ddot{q}) = 0, \tag{3}$$

which are called the system's equations of motion.

As a simple example, consider a Newtonian particle of constant inertial mass m in an inertial reference frame in three spatial dimensions. At the level of kinematics, the particle has a three-dimensional configuration space isomorphic to $\mathbb{R}^3$, and three degrees of freedom q_x, q_y, q_z that make up the particle's position vector $\mathbf{X}$ in Cartesian coordinates:

$$\mathbf{X} \equiv (X, Y, Z) \equiv (q_x, q_y, q_z). \tag{4}$$

At the level of dynamics, we assume a given force vector

$$\mathbf{F} \equiv (F_x, F_y, F_z), \tag{5}$$

in which case the system's equations of motion make up the three components of Newton's second law,

$$\mathbf{F} = m\mathbf{a}, \tag{6}$$

where $\mathbf{a}$ is the system's acceleration vector:

$$\mathbf{a} \equiv \ddot{\mathbf{X}} = (\ddot{X}, \ddot{Y}, \ddot{Z}). \tag{7}$$

B. The Lagrangian Formulation

Returning again to the case of a general classical system, let $L(q, \dot{q}, t)$, assumed to have units of energy, be a function of the system's degrees of freedom q_α, its rates of change $\dot{q}_\alpha$, and the time t. Notice that if we do not specify a candidate trajectory, then q_α, $\dot{q}_\alpha$, and t are all independent variables.

By contrast, if we are given a candidate trajectory $q_\alpha(t)$ from an arbitrary initial time t_A to an arbitrary final time t_B, then the degrees of freedom $q_\alpha(t)$ and their rates of change $\dot{q}_\alpha(t)$ become functions of t, and we can define an integral of $L(q(t), \dot{q}(t), t)$ over time:

$$S[q] \equiv \int_{t_A}^{t_B} dt\, L(q(t), \dot{q}(t), t). \tag{8}$$

The bracketed argument $[q]$ in this notation indicates that $S[q]$ is a functional of the system's candidate trajectory, meaning that $S[q]$ depends on the infinite continuum of real numbers—parametrized by t—that make up the entire candidate trajectory $q_\alpha(t)$.

If we extremize $S[q]$ over all candidate trajectories that share the same initial and final conditions,

$$\delta S[q] = 0,$$
$$\text{with } q_\alpha(t_A) \text{ and } q_\alpha(t_B) \text{ held fixed for all } \alpha, \tag{9}$$

then, as we will review in detail, we obtain the Euler-Lagrange equations,

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\alpha}\right) = 0. \tag{10}$$

For many Lagrangians used in practice, the Euler-Lagrange equations are typically second-order in the time t. If the Euler-Lagrange equations collectively turn out to be equivalent to the system's equations of motion (3), then we respectively call $L = L(q, \dot{q}, t)$ and $S[q]$ a Lagrangian and an action functional for the system, and we say that $S[q] \equiv \int dt\, L$ provides a Lagrangian formulation for the system. (Note that L and $S[q]$ are generally not unique for a given system.)

Deriving the Euler-Lagrange equations from the extremization condition (9), known as Hamilton's principle or the principle of least action, takes just a few steps, and will be an illustrative exercise before we generalize the construction later on. We start by varying the system's candidate trajectory $q_\alpha(t)$ according to

$$q_\alpha(t) \mapsto q_\alpha(t) + \delta q_\alpha(t), \tag{11}$$

where the variations $\delta q_\alpha(t)$ are infinitesimal functions of the time t that are assumed to vanish at the endpoints of the system's trajectory in keeping with (9),

$$\delta q_\alpha(t_A) = 0, \quad \delta q_\alpha(t_B) = 0, \tag{12}$$

but are otherwise arbitrary and independent. Taking a time derivative of the variation rule (11) yields the corresponding variations in the system's rates of change $\dot{q}_\alpha(t)$:

$$\dot{q}_\alpha(t) = \frac{dq_\alpha(t)}{dt} \mapsto \frac{d(q_\alpha(t) + \delta q_\alpha(t))}{dt}$$
$$= \dot{q}_\alpha(t) + \frac{d}{dt}\delta q_\alpha(t). \tag{13}$$

We infer that the induced variation in $\dot{q}_\alpha(t)$ is precisely the time derivative of the variation in $q_\alpha(t)$,

$$\delta\dot{q}_\alpha(t) = \frac{d}{dt}\delta q_\alpha(t), \tag{14}$$

so, loosely speaking, the variation operator δ "commutes" with the time derivative d/dt.

Applying the extremization condition (9), using the chain rule, carrying out an integration by parts, and dropping boundary terms that vanish by the assumption that the variations vanish at the initial and final times, we find

$$
\begin{aligned}
\delta S[q] &\equiv \int dt\, L(q + \delta q, \dot{q} + \delta \dot{q}, t) \\
&\quad - \int dt\, L(q, \dot{q}, t) \\
&= \int dt \sum_\alpha \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right) \\
&= \int dt \sum_\alpha \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \frac{d}{dt} \delta q_\alpha \right) \\
&= \int dt \sum_\alpha \left(\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \right) \delta q_\alpha = 0. \quad (15)
\end{aligned}
$$

Because the infinitesimal variations $\delta q_\alpha(t)$ are assumed to be arbitrary and independent within the domain of integration, we conclude that the factor in parentheses must be zero, so we end up with the Euler-Lagrange equations (10), as claimed.

As an example, consider a Newtonian particle of constant inertial mass m and position vector $\mathbf{X} \equiv (X, Y, Z)$ with kinetic energy

$$
T(\dot{\mathbf{X}}) = \frac{1}{2} m \dot{\mathbf{X}}^2 = \frac{1}{2} m (\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2) \quad (16)
$$

and subject to a conservative force

$$
\mathbf{F} = -\nabla V = \left(-\frac{\partial V}{\partial X}, -\frac{\partial V}{\partial Y}, -\frac{\partial V}{\partial Z} \right), \quad (17)
$$

corresponding to a potential energy $V(\mathbf{X}) = V(X, Y, Z)$. If we choose the Lagrangian

$$
L(\mathbf{X}, \dot{\mathbf{X}}) \equiv T - V = \frac{1}{2} m \dot{\mathbf{X}}^2 - V(\mathbf{X}), \quad (18)
$$

then the Euler-Lagrange equations (10), with $\mathbf{X} = (X, Y, Z) = (q_x, q_y, q_z)$, give

$$
\begin{aligned}
\frac{\partial L}{\partial X_i} &- \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{X}_i} \right) \\
&= -\frac{\partial V}{\partial X_i} - m \ddot{X}_i = 0,
\end{aligned}
$$

which replicate the three components of Newton's second law (6), $\mathbf{F} = m\mathbf{a}$. Notice also that the object's momentum

$$
\mathbf{p} \equiv (p_x, p_y, p_z) \equiv m \dot{\mathbf{X}} \quad (19)
$$

is related to the Lagrangian (18) by

$$
p_i = m \dot{X}_i = \frac{\partial L}{\partial \dot{X}_i}, \quad (20)
$$

and that the object's total mechanical energy

$$
E \equiv T + V \quad (21)
$$

is related to $\mathbf{p}$ and L by

$$
\begin{aligned}
E &= \frac{1}{2} m \dot{\mathbf{X}}^2 + V(\mathbf{X}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{X}) \\
&= \mathbf{p} \cdot \dot{\mathbf{X}} - L. \quad (22)
\end{aligned}
$$

For a generic physical system that may not resemble a Newtonian object, we might not have an obvious choice for defining the system's momenta and energy. The formulas at the end of (20) and at the end of (22) have the virtue of being general and of leading to quantities p_i and E that, as we will see shortly, are respectively conserved if the system's action functional (8) is symmetric under translations in space, $X_i \mapsto X_i + $ (constant), or under translations in time, $t \mapsto t + $ (constant).

Given a generic system with a Lagrangian formulation, we are therefore motivated to define each of the system's canonical momenta p_α in terms of the system's Lagrangian L as the partial derivative of L with respect to the corresponding rate of change $\dot{q}_\alpha$:

$$
p_\alpha \equiv \frac{\partial L}{\partial \dot{q}_\alpha}. \quad (23)
$$

Recalling that the set of points labeled by particular values q_α of a system's degrees of freedom define the system's configuration space, the set of points (q, p) labeled by particular values of the system's canonical variables q_α and p_α define the system's phase space.

If we can solve the definitions (23) for the rates of change $\dot{q}_\alpha$ as functions of the canonical variables q_α and p_α, then the system's Hamiltonian $H(q, p, t)$, which is a function on the system's phase space and roughly describes the system's energy, is defined as

$$
\begin{aligned}
H &\equiv \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} \dot{q}_\alpha - L, \\
&= \sum_\alpha p_\alpha \dot{q}_\alpha - L, \quad (24)
\end{aligned}
$$

which is known as a Legendre transformation of L.

In terms of the canonical momenta (23), we can recast the Euler-Lagrange equations (10) as

$$
\frac{dp_\alpha}{dt} = \frac{\partial L}{\partial q_\alpha}. \quad (25)
$$

One can also use the chain rule together with the Euler-

Lagrange equations to show that

$$
\begin{aligned}
\frac{dH}{dt} &= \sum_\alpha \dot{p}_\alpha \dot{q}_\alpha + \sum_\alpha p_\alpha \ddot{q}_\alpha - \frac{dL}{dt} \\
&= \sum_\alpha \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\alpha}\right)\dot{q}_\alpha + \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha}\ddot{q}_\alpha - \frac{dL}{dt} \\
&= \sum_\alpha \frac{\partial L}{\partial q_\alpha}\dot{q}_\alpha + \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha}\ddot{q}_\alpha - \frac{dL}{dt} \\
&= \underbrace{\left(\sum_\alpha \frac{\partial L}{\partial q_\alpha}\dot{q}_\alpha + \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha}\ddot{q}_\alpha + \frac{\partial L}{\partial t}\right)}_{dL/dt} - \frac{dL}{dt} - \frac{\partial L}{\partial t} \\
&= -\frac{\partial L}{\partial t},
\end{aligned}
$$

from which we conclude that the time derivative of the Hamiltonian (24) is given by

$$
\frac{dH}{dt} = -\frac{\partial L}{\partial t}. \tag{26}
$$

The two equalities (25) and (26) look very similar, apart from an overall minus sign that we will eventually see is not an accident but has an important physical significance.

Moreover, we see right away from (25) that if the Lagrangian is invariant under constant translations along a specific degree of freedom, $q_\alpha \mapsto q_\alpha + $ (constant), so that $\partial L/\partial q_\alpha = 0$, then the corresponding canonical momentum p_α is conserved, $dp_\alpha/dt = 0$. Similarly, we see from (26) that if the Lagrangian is invariant under constant translations in time, $t \mapsto t + $ (constant), so that $\partial L/\partial t = 0$, then the Hamiltonian H is conserved, $dH/dt = 0$. These results are both special cases of Noether's theorem, which establishes a general correspondence between continuous symmetries of a classical system's dynamics and quantities that are conserved when the system follows its equations of motion.

Taking partial derivatives of the Hamiltonian H with respect to the canonical variables q_α and p_α, now treated as independent variables, and regarding $\dot{q}_\alpha$ as a function of the canonical variables, we have

$$
\begin{aligned}
\frac{\partial H}{\partial p_\alpha} &= \frac{\partial}{\partial p_\alpha}\left(\sum_\beta p_\beta \dot{q}_\beta - L\right) \\
&= \dot{q}_\alpha + \sum_\beta p_\beta \frac{\partial \dot{q}_\beta}{\partial p_\alpha} - \sum_\beta \frac{\partial L}{\partial \dot{q}_\beta}\frac{\partial \dot{q}_\beta}{\partial p_\alpha} \\
&= \dot{q}_\alpha + \sum_\beta \left(p_\beta - \frac{\partial L}{\partial \dot{q}_\beta}\right)\frac{\partial \dot{q}_\beta}{\partial p_\alpha} \\
&= \dot{q}_\alpha
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial H}{\partial q_\alpha} &= \frac{\partial}{\partial q_\alpha}\left(\sum_\beta p_\beta \dot{q}_\beta - L\right) \\
&= \sum_\beta p_\beta \frac{\partial \dot{q}_\beta}{\partial q_\alpha} - \frac{\partial L}{\partial q_\alpha} - \sum_\beta \frac{\partial L}{\partial \dot{q}_\beta}\frac{\partial \dot{q}_\beta}{\partial q_\alpha} \\
&= \sum_\beta \left(p_\beta - \frac{\partial L}{\partial \dot{q}_\beta}\right)\frac{\partial \dot{q}_\beta}{\partial q_\alpha} - \frac{\partial L}{\partial q_\alpha} \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial q_\alpha}\right) = \dot{p}_\alpha,
\end{aligned}
$$

where we have used the Euler-Lagrange equations (10) in the last line. Hence, the Euler-Lagrange equations (10) imply the canonical equations of motion:

$$
\left.\begin{aligned}
\dot{q}_\alpha &= \frac{\partial H}{\partial p_\alpha}, \\
\dot{p}_\alpha &= -\frac{\partial H}{\partial q_\alpha}.
\end{aligned}\right\} \tag{27}
$$

By a similar calculation going the other way, one can also show that the canonical equations of motion imply the Euler-Lagrange equations, so the two sets of equations are equivalent. The canonical equations of motion therefore make it possible to encode the system's dynamics in an alternative way, known as the Hamiltonian formulation.

C. The Manifestly Covariant Lagrangian Formulation

The standard Lagrangian formulation of classical physics treats time and energy differently from space and momentum, in tension with the spirit of special relativity. Fortunately, we can recast the Lagrangian formulation in a more elegant way that puts time and degrees of freedom on the same footing, with the result that energy and momentum will naturally also end up on the same footing.[3]

To begin, we turn again to the case of a general classical system with degrees of freedom q_α, Lagrangian $L(q, \dot{q}, t)$, and action functional (8),

$$
S[q] \equiv \int dt\, L(q, \dot{q}, t).
$$

We carry out a smooth, strictly monotonic change of integration variable from t to a new parameter λ:

$$
t \mapsto t(\lambda). \tag{28}
$$

————

[3] For an early example of this formalism, see [8]. See also [9]. For more modern reviews, see [4, 10].

Letting dots now denote derivatives with respect to λ,

$$\dot{f} \equiv \frac{df}{d\lambda}, \qquad (29)$$

we obtain the following differential relationships:

$$dt = d\lambda\,\dot{t}, \qquad \frac{dq_\alpha}{dt} = \frac{\dot{q}_\alpha}{\dot{t}}. \qquad (30)$$

Our action functional then becomes

$$S[q] \equiv \int d\lambda\,\dot{t}\, L(q, \dot{q}/\dot{t}, t). \qquad (31)$$

This formula for the system's action functional is reparametrization invariant, meaning that it would maintain its form if we were to carry out any subsequent smooth, strictly monotonic change of parametrization $\lambda \mapsto \lambda(\lambda')$:

$$S[q] \equiv \int d\lambda' \frac{dt}{d\lambda'} L\left(q, \frac{dq}{d\lambda'}\Big/\frac{dt}{d\lambda'}, t\right). \qquad (32)$$

Reparametrization invariance is an example of a gauge invariance, meaning a redefinition of the system's degrees of freedom that leaves all the system's physically observable features unchanged. A gauge invariance represents a "redundancy" in the mathematical description of a physical system, in the sense that if we were to redefine the system's degrees of freedom according to a gauge invariance, then we would obtain a distinct but mathematically equivalent description of the same system in the same physical state.

A gauge invariance should be distinguished from a dynamical symmetry, which consists of transformations that alter the system's physical state but leave the system's dynamics unchanged. For example, a Newtonian system of particles could have dynamical symmetries under translations or rotations in three-dimensional space, both of which would alter the system's physical state.

We can formally regard the reparametrization-invariant formula (31) for the action functional as describing a system with an additional "degree of freedom" t and a modified Lagrangian

$$\mathscr{L}(q, \dot{q}, t, \dot{t}) \equiv \dot{t}\, L(q, \dot{q}/\dot{t}, t). \qquad (33)$$

Notice that

$$\frac{\partial\mathscr{L}}{\partial\dot{t}} = \frac{\partial}{\partial\dot{t}}(\dot{t}\, L(q, \dot{q}/\dot{t}, t))$$

$$= L + \sum_\alpha \dot{t}\frac{\partial L}{\partial(\dot{q}_\alpha/\dot{t})}\left(-\frac{\dot{q}_\alpha}{\dot{t}^2}\right)$$

$$= L - \sum_\alpha p_\alpha \frac{dq_\alpha}{dt} = -H$$

and

$$\frac{\partial\mathscr{L}}{\partial\dot{q}_\alpha} = \frac{\partial}{\partial\dot{q}_\alpha}(\dot{t}\, L(q, \dot{q}/\dot{t}, t))$$

$$= \dot{t}\frac{\partial L}{\partial(\dot{q}_\alpha/\dot{t})}\frac{1}{\dot{t}} = p_\alpha.$$

Thus, the system's new canonical momenta (23) conjugate to our original degrees of freedom q_α are the same as before, $\mathscr{P}_\alpha = p_\alpha$, whereas the system's canonical momentum $\mathscr{P}_t$ conjugate to t is equal to *minus* the system's original Hamiltonian H:

$$\left.\begin{array}{l} \mathscr{P}_t \equiv \dfrac{\partial\mathscr{L}}{\partial\dot{t}} = -H, \\[2mm] \mathscr{P}_\alpha \equiv \dfrac{\partial\mathscr{L}}{\partial\dot{q}_\alpha} = p_\alpha. \end{array}\right\} \qquad (34)$$

These formulas motivate introducing "upper-index" and "lower-index" versions of our canonical variables by mimicking the analogous rules for the components of the four-vectors that are used in special relativity:

$$\left.\begin{array}{ll} q^t \equiv c\,t, & q_t \equiv -c\,t, \\ q^\alpha \equiv q_\alpha, & \\ p^t \equiv H/c, & p_t \equiv -H/c, \\ p^\alpha \equiv p_\alpha. & \end{array}\right\} \qquad (35)$$

To ensure that we are using the same units for q^t and q^α and also the same units for p^t and p^α, we have introduced an arbitrary constant c with units of energy divided by momentum. (The constant c also has units of distance divided by time, or speed, but not all classical systems possess a notion of distance.) Note also that we have defined $p_t \equiv \mathscr{P}_t/c$.

Applying the extremization condition (9) to the action functional with respect to the new degrees of freedom q^t and q^α, we obtain a new set of Euler-Lagrange equations given by

$$\left.\begin{array}{l} \dfrac{\partial\mathscr{L}}{\partial q^t} - \dfrac{d}{d\lambda}\left(\dfrac{\partial\mathscr{L}}{\partial\dot{q}^t}\right) = 0, \\[3mm] \dfrac{\partial\mathscr{L}}{\partial q^\alpha} - \dfrac{d}{d\lambda}\left(\dfrac{\partial\mathscr{L}}{\partial\dot{q}^\alpha}\right) = 0. \end{array}\right\} \qquad (36)$$

Observe that

$$\frac{\partial\mathscr{L}}{\partial q^\alpha} - \frac{d}{d\lambda}\left(\frac{\partial\mathscr{L}}{\partial\dot{q}^\alpha}\right)$$

$$= \dot{t}\frac{\partial L}{\partial q_\alpha} - \frac{dt}{d\lambda}\frac{d}{dt}\left(\dot{t}\frac{\partial L}{\partial(\dot{q}_\alpha/\dot{t})}\left(\frac{1}{\dot{t}}\right)\right)$$

$$= \dot{t}\left(\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial(dq_\alpha/dt)}\right)\right),$$

so the Euler-Lagrange equations for the degrees of freedom q^α unsurprisingly give us back our original Euler-Lagrange equations (10),

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial(dq_\alpha/dt)}\right) = 0,$$

which, as we recall from (25), can be written more compactly as

$$\frac{dp_\alpha}{dt} = \frac{\partial L}{\partial q_\alpha}.$$

Meanwhile, we also have

$$\frac{\partial \mathscr{L}}{\partial q^t} - \frac{d}{d\lambda}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}^t}\right)$$

$$= \frac{1}{c}\frac{\partial \mathscr{L}}{\partial t} - \frac{1}{c}\frac{d}{d\lambda}\left(\frac{\partial \mathscr{L}}{\partial \dot{t}}\right)$$

$$= \frac{1}{c}\dot{t}\frac{\partial L}{\partial t} - \frac{1}{c}\frac{d}{d\lambda}\left(L + \sum_\alpha \dot{t}\frac{\partial L}{\partial(\dot{q}_\alpha/\dot{t})}\left(-\frac{\dot{q}_\alpha}{\dot{t}^2}\right)\right)$$

$$= \frac{1}{c}\dot{t}\frac{\partial L}{\partial t} - \frac{1}{c}\dot{t}\frac{d}{dt}\left(L - \sum_\alpha p_\alpha\frac{\dot{q}_\alpha}{\dot{t}}\right)$$

$$= \frac{1}{c}\dot{t}\left(\frac{\partial L}{\partial t} + \frac{dH}{dt}\right),$$

so the Euler-Lagrange equation for q^t replicates the equation (26) that relates the total time derivative of the system's original Hamiltonian H to the partial time derivative of the system's original Lagrangian L,

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t}.$$

We can combine these results in terms of the raised-index versions p^t and p^α of the canonical momenta defined in (35) as the symmetric-looking equations

$$\left.\begin{aligned}\frac{dp^t}{dt} &= \frac{\partial L}{\partial q_t}, \\ \frac{dp^\alpha}{dt} &= \frac{\partial L}{\partial q_\alpha},\end{aligned}\right\} \tag{37}$$

or, equivalently, in terms of $\mathscr{L}$ and derivatives with respect to λ as

$$\left.\begin{aligned}\dot{p}^t &\equiv \frac{dp^t}{d\lambda} = \frac{\partial \mathscr{L}}{\partial q_t}, \\ \dot{p}^\alpha &\equiv \frac{dp^\alpha}{d\lambda} = \frac{\partial \mathscr{L}}{\partial q_\alpha}.\end{aligned}\right\} \tag{38}$$

Furthermore, we can write our action functional (31) as

$$S[q] = \int d\lambda\, \dot{t}\, L = \int d\lambda\, \dot{t}\left(\sum_\alpha p_\alpha\frac{dq_\alpha}{dt} - H\right)$$

$$= \int d\lambda\left(\sum_\alpha p_\alpha\frac{dt}{d\lambda}\frac{dq_\alpha}{dt} - (\dot{t}/c)H\right)$$

$$= \int d\lambda\left(\sum_\alpha p_\alpha\dot{q}^\alpha + \dot{q}^t p_t\right).$$

That is, rather remarkably, we can recast our action functional in a form that resembles a Lorentz-invariant dot product, despite the fact that we have not assumed that our system has anything to do with special relativity or four-dimensional spacetime:

$$S[q] = \int d\lambda\, \mathscr{L} = \int d\lambda\left(p_t\dot{q}^t + \sum_\alpha p_\alpha\dot{q}^\alpha\right). \tag{39}$$

We therefore refer to this framework as the manifestly covariant Lagrangian formulation for our classical system.

Introducing a square matrix $\eta \equiv \operatorname{diag}(-1, 1, \dots)$ that naturally generalizes the Minkowski metric tensor from special relativity,

$$\eta \equiv \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{pmatrix}, \tag{40}$$

we can write the system's action functional (39) in matrix form as

$$S[q] = \int d\lambda\, \begin{pmatrix} p^t & p^\alpha \end{pmatrix}\eta\begin{pmatrix} \dot{q}^t \\ \dot{q}^\alpha \end{pmatrix}, \tag{41}$$

where p^α and $\dot{q}^\alpha$ here are notational abbreviations for their whole lists indexed by α. This expression for $S[q]$ immediately suggests the consideration of systems whose action functionals have a symmetry under rigid linear transformations of the form

$$\begin{pmatrix} q^t \\ q^\alpha \end{pmatrix} \mapsto \Lambda\begin{pmatrix} q^t \\ q^\alpha \end{pmatrix}, \quad \begin{pmatrix} p^t \\ p^\alpha \end{pmatrix} \mapsto \Lambda\begin{pmatrix} p^t \\ p^\alpha \end{pmatrix} \tag{42}$$

for constant matrices Λ that preserve the generalized Minkowski metric tensor η in the sense that

$$\Lambda^{\mathrm{T}}\eta\Lambda = \eta. \tag{43}$$

The matrices Λ therefore represent generalizations of Lorentz transformations.

Recall the group $O(N)$ of orthogonal $N \times N$ matrices R, meaning matrices that preserve the $N \times N$ identity matrix $1 \equiv \operatorname{diag}(1, 1, \dots)$,

$$R^{\mathrm{T}}R = R^{\mathrm{T}}1R = 1. \tag{44}$$

Letting N denote the system's original number of degrees of freedom q_α, we see that the set of generalized Lorentz-transformation matrices Λ preserve the $(N+1) \times (N+1)$ matrix $\eta \equiv \operatorname{diag}(-1, 1, \dots)$, so we correspondingly refer to them as making up the group $O(1, N)$.

The formula (39) for the action functional also implies that the new "Hamiltonian" $\mathscr{H}$, defined in line with (24), trivially vanishes, and therefore (at least classically) does not hold any physical meaning:

$$\mathscr{H} \equiv p_t\dot{q}^t + \sum_\alpha p_\alpha q^\alpha - \mathscr{L} = 0. \tag{45}$$

This equation is closely related to the fact that arbitrary changes of parametrization represent a gauge invariance of the system and likewise do not have any physical meaning.

III.   SPACETIME IN SPECIAL RELATIVITY

We now turn to a brief review of special relativity.[4]

---

[4] For a more extensive introduction, see the opening chapters of [11].

A. Spacetime and Four-Vectors

In special relativity, time t and space $\mathbf{x} \equiv (x, y, z)$ join together to form four-dimensional spacetime coordinates,

$$x^\mu \equiv (x^t, x^x, x^y, x^z)^\mu$$
$$\equiv (c\,t, \mathbf{x})^\mu \equiv (c\,t, x, y, z)^\mu, \qquad (46)$$

where c is the speed of light. We will use Greek letters $\alpha, \beta, \ldots, \mu, \nu, \ldots$ for Lorentz indices, which will each run through the four possible values t, x, y, z, and we will use Latin indices $i, j, k, \ldots$ for the spatial values x, y, z, where we will consistently employ Cartesian coordinate systems.

Defining the (3+1)-dimensional Minkowski metric tensor by

$$\eta_{\mu\nu} \equiv \eta^{\mu\nu} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{\mu\nu}, \qquad (47)$$

and employing Einstein summation notation, we can raise and lower indices on the components of four-vectors according to $v_\mu \equiv \eta_{\mu\nu} v^\nu$ and $w^\mu \equiv \eta^{\mu\nu} w_\nu$, with the following results:

$$\left. \begin{aligned} v^t &= -v_t, \\ v^x &= v_x, \\ v^y &= v_y, \\ v^z &= v_z. \end{aligned} \right\} \qquad (48)$$

We let $\Lambda^\mu{}_\nu$ be a 4×4 Lorentz-transformation matrix, meaning that $\Lambda^\mu{}_\nu$ is an element of $O(1,3)$ and therefore preserves the Minkowski metric tensor $\eta_{\mu\nu}$ in the sense that

$$\Lambda^\mu{}_\rho \eta_{\mu\nu} \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}, \qquad (49)$$

or, in matrix notation,

$$\Lambda^{\mathrm{T}} \eta \Lambda = \eta. \qquad (50)$$

Then Lorentz transformations of four-vectors v^μ, meaning linear transformations of the form

$$v^\mu \mapsto \Lambda^\mu{}_\nu v^\nu, \qquad (51)$$

preserve four-dimensional dot products defined by

$$v \cdot w \equiv v_\nu w^\nu = \eta_{\mu\nu} v^\mu w^\nu. \qquad (52)$$

Four-vectors v^μ are classified as timelike, null, or spacelike according to whether the dot product of v^μ with itself is respectively negative, zero, or positive:

$$v^2 \equiv v \cdot v \begin{cases} < 0 & \text{timelike,} \\ = 0 & \text{null,} \\ > 0 & \text{spacelike.} \end{cases} \qquad (53)$$

The Lorentz invariance of the dot product (52) ensures that this classification is invariant and therefore well-defined under Lorentz transformations.

B. The Spacetime Transformation Groups

The collection $O(1,3)$ of all possible Lorentz transformations (51),

$$v^\mu \mapsto \Lambda^\mu{}_\nu v^\nu,$$

is called the Lorentz group.[5] The largest subgroup that excludes parity transformations,

$$\Lambda_{\mathrm{parity}} = \mathrm{diag}(1, -1, -1, -1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad (54)$$

is called the proper Lorentz group and is denoted by $SO(1,3)$, mirroring the notation $SO(N)$ for $N \times N$ rotation matrices R that do not involve parity transformations. The largest subgroup of the Lorentz group that excludes time-reversal transformations,

$$\Lambda_{\mathrm{time\text{-}reversal}} = \mathrm{diag}(-1, 1, 1, 1) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (55)$$

is called the orthochronous Lorentz group and is denoted by $O^+(1,3)$ or $O^\uparrow(1,3)$. The set of all Lorentz transformations that can be reduced smoothly to the identity transformation $\Lambda = 1$ cannot include parity or time-reversal transformations, and is called the proper orthochronous Lorentz group $SO^+(1,3)$ or $SO^\uparrow(1,3)$.

A simple calculation shows that for four-vectors v^μ that are timelike or null, the *sign* of the temporal component v^t is invariant under orthochronous Lorentz transformations $v^\mu \mapsto \Lambda^\mu{}_\nu v^\nu$:

$$v^2 \leq 0 \implies \text{sign of } v^t \text{ is invariant under } O^+(1,3). \qquad (56)$$

As a consequence, future-directed ($v^t > 0$) four-vectors that are timelike or null remain future-directed under orthochronous Lorentz transformations, with a similar statement for past-directed ($v^t < 0$) four-vectors that are timelike or null. These properties ensure that if the displacement between two spacetime points is timelike or null, then their chronological ordering is an invariant fact of nature. By contrast, the temporal components v^t of spacelike four-vectors ($v^2 > 0$) can change sign under orthochronous Lorentz transformations, a behavior that is closely related to the breakdown of simultaneity in special relativity.

We can also consider additive shifts in the four-dimensional coordinates (46) by constants a^μ:

$$x^\mu \mapsto x^\mu + a^\mu. \qquad (57)$$

----

[5] For a comprehensive presentation of the group theory underlying special relativity, see [12].

These transformations make up the spacetime-translation group, which is isomorphic to $\mathbb{R}^4$ but is denoted by $\mathbb{R}^{1,3}$ to emphasize the mathematical and physical distinctions between time and space.

Combining spacetime translations with Lorentz transformations of the spacetime coordinates x^μ gives the Poincaré group:

$$x^\mu \mapsto \Lambda^\mu{}_\nu x^\nu + a^\mu. \tag{58}$$

Like the Lorentz group, the Poincaré group has proper and orthochronous subgroups that are respectively defined by dropping all Lorentz transformations that involve parity or time-reversal transformations.[6]

IV. TRANSITIVE GROUP ACTIONS OF THE POINCARÉ GROUP

The set of all physical transformations $(q, p) \mapsto (q', p')$ that can be carried out on a system's state (q, p) in its phase space are collectively called a group action on the system's phase space. If we include translations in time among these physical transformations, then by starting with a single convenient choice of reference state (q_0, p_0), we can reach every other possible state that the system can occupy. The group action provided by the system's phase space is therefore "irreducible," or, more precisely, transitive, referring to the fact that no proper subset of the system's phase space can be dropped without violating the group action. One then says that the phase space serves as a homogeneous space for the group of physical transformations.

As we will show, the different possible transitive group actions (or homogeneous spaces) of the Poincaré group turn out to provide a complete classification of the phase spaces of the different categories of particles in physics, in parallel with Wigner's method for classifying quantum particle-types by identifying their Hilbert spaces as irreducible representations of the Poincaré group.[7]

A. Systems Singled Out by the Poincaré Group

To start, we note that the Poincaré group (58) naturally singles out classical systems that have three physical degrees of freedom $(q_x, q_y, q_z) = \mathbf{X} \equiv (X, Y, Z)$ and therefore three corresponding canonical momenta $\mathbf{p} = (p_x, p_y, p_z)$. It follows that the system's manifestly

———

[6] As a mathematical aside, the Poincaré group is formally denoted by the semi-direct product $\mathbb{R}^{1,3} \rtimes O(1,3)$, which generalizes the notion of a direct product $G = H_1 \times H_2$ to the case in which the second factor H_2 is not necessarily a normal subgroup of the overall group G.

[7] For alternative classical approaches to this classification problem, see [2–5].

covariant Lagrangian formulation involves four spacetime degrees of freedom

$$\begin{aligned} X^\mu &\equiv (q^t, q^x, q^y, q^z)^\mu \\ &= (cT, X, Y, Z)^\mu \equiv (cT, \mathbf{X})^\mu, \end{aligned} \tag{59}$$

together with a canonical four-momentum

$$\begin{aligned} p^\mu &\equiv (p^t, p^x, p^y, p^z)^\mu \\ &\equiv (E/c, \mathbf{p})^\mu \end{aligned} \tag{60}$$

whose individual components, in lower-index form p_μ, are defined in terms of the system's covariant Lagrangian $\mathscr{L}$ in accordance with (34),

$$p_\mu \equiv \frac{\partial \mathscr{L}}{\partial \dot{X}^\mu}. \tag{61}$$

Here dots denote derivatives with respect to the arbitrary parameter λ,

$$\dot{X}^\mu \equiv \frac{dX^\mu}{d\lambda}, \tag{62}$$

we have identified the system's energy E as

$$E \equiv H \equiv p^t c, \tag{63}$$

and candidate trajectories of the system are now called worldlines.

B. Angular Momentum and Spin

In analogy with the Newtonian definition $\mathbf{L} \equiv \mathbf{X} \times \mathbf{p}$ of an object's orbital angular momentum, whose individual components are

$$\begin{aligned} L_k &= X_i p_j - X_j p_i, \\ &\text{with } (i, j, k) = (x, y, z),\ (z, x, y),\ \text{or}\ (y, z, x), \end{aligned} \tag{64}$$

we will find it convenient to introduce an antisymmetric tensor

$$L^{\mu\nu} \equiv X^\mu p^\nu - X^\nu p^\mu = -L^{\nu\mu} \tag{65}$$

whose spatial components L^{ij} (that is, for i, j each taking the values x, y, z) encode the components of $\mathbf{L}$. We will accordingly refer to $L^{\mu\nu}$ as the system's orbital angular-momentum tensor, although one should keep in mind that its temporal components L^{ti} (for i a spatial index) are not angular momenta. Indeed, if the system's energy (63) is nonzero, $E \equiv p^t c \neq 0$, then we can write these temporal components as

$$\begin{aligned} L^{ti} &= X^t p^i - X^i p^t = cT p^i - X^i E/c \\ &= -\frac{E}{c}\left(X^i - \frac{p^i c^2}{E} T\right). \end{aligned} \tag{66}$$

We will see later that the factor $\mathbf{p}c^2/E$, which has units of distance divided by time, will typically yield the

system's three-dimensional physical propagation velocity $\mathbf{v} \equiv d\mathbf{X}/dt$ through space, so the quantity in parentheses will turn out to be related to the system's *linear* motion.

To be as general as possible, we can also allow the system to possess an *intrinsic* notion of angular momentum, called spin, that does not involve the system's spacetime coordinates X^μ or its four-momentum p^μ, and that can be encoded in an antisymmetric tensor

$$S^{\mu\nu} = -S^{\nu\mu}, \tag{67}$$

called the system's *spin tensor*. The system's *total* angular momentum is then represented by an antisymmetric tensor defined as the sum of the tensors representing the orbital and spin contributions:

$$J^{\mu\nu} \equiv L^{\mu\nu} + S^{\mu\nu} = -J^{\nu\mu}. \tag{68}$$

We will refer to $J^{\mu\nu}$ as the system's total angular-momentum tensor.

We can define the following three-vectors from the independent components of $J^{\mu\nu}$ and $S^{\mu\nu}$:

$$\mathbf{J} \equiv (J_x,\ J_y,\ J_z) \equiv (J^{yz}, J^{zx}, J^{xy}), \tag{69}$$

$$\mathbf{K} \equiv (K_x, K_y, K_z) \equiv (J^{tx}, J^{ty}, J^{tz}), \tag{70}$$

$$\mathbf{S} \equiv (S_x,\ S_y,\ S_z) \equiv (S^{yz}, S^{zx}, S^{xy}), \tag{71}$$

$$\tilde{\mathbf{S}} \equiv (\tilde{S}_x,\ \tilde{S}_y,\ \tilde{S}_z) \equiv (S^{tx}, S^{ty}, S^{tz}). \tag{72}$$

We will call $\mathbf{S}$ the system's spin three-vector and $\tilde{\mathbf{S}}$ its dual spin-three vector.

We can now write the system's total angular-momentum tensor $J^{\mu\nu}$ and its spin tensor $S^{\mu\nu}$ as

$$J^{\mu\nu} \equiv \begin{pmatrix} 0 & K_x & K_y & K_z \\ -K_x & 0 & J_z & -J_y \\ -K_y & -J_z & 0 & J_x \\ -K_z & J_y & -J_x & 0 \end{pmatrix}^{\mu\nu}, \tag{73}$$

$$S^{\mu\nu} \equiv \begin{pmatrix} 0 & \tilde{S}_x & \tilde{S}_y & \tilde{S}_z \\ -\tilde{S}_x & 0 & S_z & -S_y \\ -\tilde{S}_y & -S_z & 0 & S_x \\ -\tilde{S}_z & S_y & -S_x & 0 \end{pmatrix}^{\mu\nu}. \tag{74}$$

Note that if $\mathbf{S} = 0$, then $\mathbf{J} = \mathbf{L} = \mathbf{X} \times \mathbf{p}$ reduces to the usual Newtonian definition (64) of orbital angular momentum.

C. Defining a System by a Transitive Group Action of the Poincaré Group

The state of our system in its phase space is fully determined by knowing the values of the system's spacetime coordinates X^μ, its four-momentum p^μ, and its spin tensor $S^{\mu\nu}$, which together determine the orbital angular-momentum tensor $L^{\mu\nu}$ and the total angular-momentum tensor $J^{\mu\nu}$. We can therefore define a transitive group

action of the Poincaré group on the system's phase space by defining what Poincaré transformations do to the values of X^μ, p^μ, and $S^{\mu\nu}$ that define the system's state (X, p, S).

Specifically, we define the action of Lorentz transformations on the system's state (X, p, S) by generalizing the transformation rule (51) to the statement that every free upper Lorentz index on X^μ, p^μ, and $S^{\mu\nu}$ receives a linear factor of a shared Lorentz-transformation matrix Λ:

$$X^\mu \mapsto \Lambda^\mu{}_\nu X^\nu, \tag{75}$$

$$p^\mu \mapsto \Lambda^\mu{}_\nu p^\nu, \tag{76}$$

$$S^{\mu\nu} \mapsto \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma S^{\rho\sigma} = \Lambda^\mu{}_\rho S^{\rho\sigma} (\Lambda^{\mathrm{T}})_\sigma{}^\nu. \tag{77}$$

It follows from the definitions (65) of $L^{\mu\nu}$ and (68) of $J^{\mu\nu}$ that we have the additional Lorentz-transformation rules

$$L^{\mu\nu} \mapsto \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma L^{\rho\sigma} = \Lambda^\mu{}_\rho L^{\rho\sigma} (\Lambda^{\mathrm{T}})_\sigma{}^\nu, \tag{78}$$

$$J^{\mu\nu} \mapsto \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma J^{\rho\sigma} = \Lambda^\mu{}_\rho J^{\rho\sigma} (\Lambda^{\mathrm{T}})_\sigma{}^\nu. \tag{79}$$

Meanwhile, we define the action of spacetime translations on the system's state (X, p, S) solely as (57) for the spacetime coordinates X^μ, with the system's four-momentum p^μ and spin tensor $S^{\mu\nu}$ unchanged:

$$X^\mu \mapsto X^\mu + a^\mu, \tag{80}$$

$$p^\mu \mapsto p^\mu, \tag{81}$$

$$S^{\mu\nu} \mapsto S^{\mu\nu}. \tag{82}$$

These definitions then determine the additional translation rules

$$L^{\mu\nu} \mapsto L^{\mu\nu} + a^\mu p^\nu - a^\nu p^\mu, \tag{83}$$

$$J^\mu \mapsto J^{\mu\nu} + a^\mu p^\nu - a^\nu p^\mu. \tag{84}$$

We can then construct general Poincaré transformations from combinations of Lorentz transformations and spacetime translations.

One can check that the three-vectors $\mathbf{J}$, $\mathbf{K}$, $\mathbf{S}$, and $\tilde{\mathbf{S}}$ defined in (69)–(72) all indeed transform as three-vectors under proper rotations. One can also show that $\mathbf{K}$ and $\tilde{\mathbf{S}}$ transform as proper vectors (or polar vectors) under parity transformations (54),

$$\left.\begin{array}{r} \mathbf{K} \mapsto -\mathbf{K}, \\ \tilde{\mathbf{S}} \mapsto -\tilde{\mathbf{S}}, \end{array}\right\} \text{(parity)} \tag{85}$$

whereas $\mathbf{J}$ and $\mathbf{S}$ are pseudovectors (or axial vectors), meaning that they do not change sign under parity transformations:

$$\left.\begin{array}{r} \mathbf{J} \mapsto \mathbf{J}, \\ \mathbf{S} \mapsto \mathbf{S}. \end{array}\right\} \text{(parity)} \tag{86}$$

If the system's phase space provides a transitive group action of the Poincaré group, then, by construction, every state (X, p, S) can be reached by starting with an arbitrary choice of reference state

$$(X_0, p_0, S_0) \tag{87}$$

and then acting on it with an appropriate choice of Poincaré transformation (a, Λ):

$$(X, p, S) \equiv (\Lambda X_0 + a, \Lambda p_0, \Lambda S_0 \Lambda^{\mathrm{T}}). \qquad (88)$$

That is,

$$X \equiv \Lambda X_0 + a, \qquad (89)$$
$$p \equiv \Lambda p_0, \qquad (90)$$
$$S \equiv \Lambda S_0 \Lambda^{\mathrm{T}}, \qquad (91)$$

or, displaying indices explicitly,

$$X^\mu \equiv \Lambda^\mu{}_\nu X_0^\nu + a^\mu, \qquad (92)$$
$$p^\mu \equiv \Lambda^\mu{}_\nu p_0^\nu, \qquad (93)$$
$$S^{\mu\nu} \equiv \Lambda^\mu{}_\rho S_0^{\rho\sigma} (\Lambda^{\mathrm{T}})_\sigma{}^\nu. \qquad (94)$$

Without loss of generality, we will always take the reference value of the system's spacetime point to be at the origin:

$$X_0^\mu \equiv 0. \qquad (95)$$

Due to the transformation rule (92), the system's spacetime point X^μ in any other state (X, p, S) can then be identified with the translation-group four-vector a^μ, so we will refer to a^μ as X^μ in our work ahead,

$$X^\mu \equiv a^\mu, \qquad (96)$$

keeping in mind that these variables are independent of the Lorentz-transformation matrix $\Lambda^\mu{}_\nu$. We will choose the reference values p_0^μ and $S_0^{\mu\nu}$ in (87) on a case-by-case basis later.

D. The Pauli-Lubanski Pseudovector

Introducing the totally antisymmetric, four-index Levi-Civita symbol,

$$\epsilon_{\mu\nu\rho\sigma} \equiv \begin{cases} +1 & \text{for } \mu\nu\rho\sigma \text{ an even permutation of } txyz, \\ -1 & \text{for } \mu\nu\rho\sigma \text{ an odd permutation of } txyz, \\ 0 & \text{otherwise} \end{cases}$$
$$= -\epsilon^{\mu\nu\rho\sigma}, \qquad (97)$$

we can form a convenient mathematical object, called the Pauli-Lubanski pseudovector W^μ, by contracting the Lorentz indices of the system's four-momentum p^μ and the total angular-momentum tensor $J^{\mu\nu}$ with the indices of $\epsilon^{\mu\nu\rho\sigma}$:[8]

$$W^\mu \equiv -\frac{1}{2} \epsilon^{\mu\nu\rho\sigma} p_\nu J_{\rho\sigma}. \qquad (98)$$

———

[8] The minus sign in this definition is a reflection of our metric sign conventions.

Decomposing the total angular-momentum tensor as in (68) into its orbital (65) and spin (67) contributions,

$$\begin{aligned} J_{\rho\sigma} &= L_{\rho\sigma} + S_{\rho\sigma} \\ &= X_\rho p_\sigma - X_\sigma p_\rho + S_{\rho\sigma}, \end{aligned}$$

the contributions from the orbital-angular momentum tensor $L_{\rho\sigma}$ cancel out of the definition of W^μ, so we can replace the total angular-momentum tensor $J_{\rho\sigma}$ with just its spin contribution $S_{\rho\sigma}$ in the formula for W^μ:

$$W^\mu = -\frac{1}{2} \epsilon^{\mu\nu\rho\sigma} p_\nu S_{\rho\sigma}. \qquad (99)$$

It follows from a straightforward calculation that we can express the Pauli-Lubanski pseudovector in terms of the spin three-vector $\mathbf{S}$ defined in (71), the dual spin three-vector $\tilde{\mathbf{S}}$ defined in (72), and the components of the system's four-momentum $p^\mu = (E/c, \mathbf{p})^\mu$ as

$$W^\mu = (\mathbf{p} \cdot \mathbf{S}, \ (E/c)\mathbf{S} - \mathbf{p} \times \tilde{\mathbf{S}})^\mu. \qquad (100)$$

The formula (99) makes manifest that the Pauli-Lubanski pseudovector does not involve the spacetime coordinates X^μ, so it is invariant under translation transformations (80)–(84):

$$W^\mu \mapsto W^\mu \quad \text{(spacetime translations)}. \qquad (101)$$

Meanwhile, under Lorentz transformations of p_ν and $S_{\rho\sigma}$, W^μ transforms as

$$W^\mu \mapsto \det(\Lambda) \Lambda^\mu{}_\nu W^\nu, \qquad (102)$$

where $\det(\Lambda)$ is the determinant of $\Lambda^\mu{}_\nu$. Hence, under parity transformations $\Lambda_{\mathrm{parity}}$, for which $\det(\Lambda_{\mathrm{parity}}) = -1$, W^μ transforms oppositely to the way that ordinary four-vectors transform:

$$W^t \mapsto -W^t, \quad W^i = W^i \quad \text{(parity)}. \qquad (103)$$

It is because of this transformation behavior that W^μ is called a pseudovector.

E. Invariant Quantities of a Transitive Group Action of the Poincaré Group

Notice that the quantities $p^2 \equiv p_\mu p^\mu$, $W^2 \equiv W_\mu W^\mu$, and $S^2 \equiv S_{\mu\nu} S^{\mu\nu}$ are invariant under Poincaré transformations, meaning that they are invariant under all Lorentz transformations (whether or not parity and time-reversal transformations are involved) as well as under all spacetime translations. These quantities therefore each have a single, constant value for all states in any phase space that constitutes a transitive group action of the Poincaré group, and so, in particular, have constant values along the system's worldline.[9]

———

[9] Quantities that have fixed values in a transitive group action or in an irreducible representation of a given transformation group are formally called Casimir invariants.

We name these invariant quantities according to

$$p^2 \equiv p_\mu p^\mu \equiv -m^2 c^2, \tag{104}$$

$$W^2 \equiv W_\mu W^\mu \equiv w^2, \tag{105}$$

$$\frac{1}{2} S^2 \equiv \frac{1}{2} S_{\mu\nu} S^{\mu\nu} \equiv s^2. \tag{106}$$

The scalar constant m has units of momentum-squared divided by energy (that is, units of mass), the scalar constant w has units of momentum multiplied by energy multiplied by time, and the scalar constant s has units of energy multiplied by time (that is, units of angular momentum).

Note that w^2 and s^2 having fixed values does not imply any sort of quantization, any more than m^2 being fixed implies quantization. In our classical context, we are essentially working in the limit of large quantum numbers in which w^2 and s^2 are invariant but are otherwise permitted to take on any one of a continuous range of possible real values.

In terms of the spin three-vector $\mathbf{S}$ defined in (71) and the dual spin three-vector $\tilde{\mathbf{S}}$ defined in (72), we can write the invariant quantity s^2 as

$$s^2 \equiv \frac{1}{2} S_{\mu\nu} S^{\mu\nu} = \mathbf{S}^2 - \tilde{\mathbf{S}}^2. \tag{107}$$

We can also contract two copies of the spin tensor $S^{\mu\nu}$ with the Levi-Civita symbol (97) to obtain another quantity with the same units as s^2:

$$\tilde{s}^2 \equiv \frac{1}{8} \epsilon_{\mu\nu\rho\sigma} S^{\mu\nu} S^{\rho\sigma} = \mathbf{S} \cdot \tilde{\mathbf{S}}. \tag{108}$$

This quantity is invariant under spacetime translations and also under *proper orthochronous* Lorentz transformations. However, due to the transformation rules (85) and (86), $\tilde{s}^2$ changes by an overall sign under parity transformations, so it is called a pseudoscalar.

As was true for the scalar invariant quantities m^2, w^2, and s^2, the pseudoscalar quantity $\tilde{s}^2$ cannot change in value under smooth evolution along the system's worldline. To understand why, observe that if $\tilde{s}^2 = 0$, then it is invariant under parity and time-reversal transformations, and therefore has the unique value $\tilde{s}^2 = 0$ for the system's entire phase space. By contrast, if $\tilde{s}^2 \neq 0$, then our transitive group action of the Poincaré group can contain only the values $\pm\tilde{s}^2$, and no smooth evolution can take the system from $\tilde{s}^2 > 0$ to $\tilde{s}^2 < 0$ or vice versa. (In all our examples, ahead, we will end up finding that $\tilde{s}^2 = 0$.)

Classifying the possible systems whose phase spaces provide transitive group actions of the Poincaré group now reduces to selecting mutually consistent values for the invariant quantities m^2, w^2, s^2, and $\tilde{s}^2$, and then choosing a convenient reference state (X_0, p_0, S_0) that is compatible with those fixed values. Note again that the constancy of m^2, w^2, s^2, and $\tilde{s}^2$—including the constancy of the system's invariant spin-squared s^2—is entirely classical and has nothing to do with quantization or quantum theory.

As an aside, observe that the only other candidates for invariant quantities that are derivable from the system's phase-space variables are

$$p_\mu W^\mu = 0,$$
$$p_\mu p_\nu S^{\mu\nu} = 0,$$
$$W_\mu W_\nu S^{\mu\nu} = 0,$$
$$W_\mu p_\nu S^{\mu\nu} = m^2 c^2 \tilde{s}^2,$$
$$\epsilon^{\mu\nu\rho\sigma} W_\mu p_\nu S_{\rho\sigma} = -2w^2.$$

None of these expressions represent fundamentally new quantities independent of m^2, w^2, s^2, and $\tilde{s}^2$, so we do not need to specify values for them as part of the definition of our transitive group action of the Poincaré group.

F. The Generators of the Lorentz Group

Observe that the system's phase space (88) is fully parametrized by the values a^μ and $\Lambda^\mu{}_\nu$ that make up a generic Poincaré transformation (a, Λ), where a^μ encodes the system's spacetime location and $\Lambda^\mu{}_\nu$ encodes the system's motion and angular orientation. Lorentz-transformation matrices are difficult to manipulate directly, due to the constraint $\Lambda^{\mathrm{T}} \eta \Lambda = \eta$ from (50), so we will find it useful to decompose them into simpler ingredients.[10]

We start by considering a Lorentz transformation $\Lambda(\epsilon) = 1 - \epsilon$ that differs only infinitesimally from the identity matrix 1:

$$\Lambda^\alpha{}_\beta(\epsilon) = \delta^\alpha_\beta - \epsilon^\alpha{}_\beta. \tag{109}$$

(The minus sign is conventional.) Here $\epsilon^\alpha{}_\beta$ represents a collection of infinitesimal parameters, and δ^α_β is the four-dimensional Kronecker delta,

$$\delta^\alpha_\beta \equiv \begin{cases} 1 & \text{for } \alpha = \beta, \\ 0 & \text{for } \alpha \neq \beta, \end{cases} \tag{110}$$

which represents the components of the identity matrix. The constraint $\Lambda^{\mathrm{T}} \eta \Lambda = \eta$ then yields the equation

$$(\delta^\alpha_\beta - \epsilon^\alpha{}_\beta) \eta_{\alpha\gamma} (\delta^\gamma_\delta - \epsilon^\gamma{}_\delta) = \eta_{\beta\delta}.$$

Working to first order in ϵ, we see from this equation that the infinitesimal tensor $\epsilon^{\alpha\beta}$ obtained from $\epsilon^\alpha{}_\beta$ by raising its second index using the Minkowski metric tensor is antisymmetric:

$$\epsilon^{\alpha\beta} = -\epsilon^{\beta\alpha}. \tag{111}$$

―――――

[10] For a more extensive review of the mathematical details ahead, see [12].

The tensor $\epsilon^{\alpha\beta}$ therefore has six independent components, with $\epsilon^{yz}, \epsilon^{zx}, \epsilon^{xy}$ respectively parametrizing rotations around the x, y, z axes, and with $\epsilon^{tx}, \epsilon^{ty}, \epsilon^{tz}$ respectively parametrizing Lorentz boosts along the x, y, z directions.

We can write any two-index, antisymmetric Lorentz tensor $A^{\alpha\beta} = -A^{\beta\alpha}$ as

$$
\begin{aligned}
A^{\alpha\beta} &= \frac{1}{2}(A^{\alpha\beta} - A^{\beta\alpha}) \\
&= \frac{1}{2}A^{\mu\nu}(\delta_\mu^\alpha \delta_\nu^\beta - \delta_\mu^\beta \delta_\nu^\alpha).
\end{aligned}
$$

Hence, the tensors defined by

$$
[\sigma_{\mu\nu}]^{\alpha\beta} \equiv -i\delta_\mu^\alpha \delta_\nu^\beta + i\delta_\mu^\beta \delta_\nu^\alpha \tag{112}
$$

form a basis for all two-index, antisymmetric tensors:

$$
A^{\alpha\beta} = \frac{i}{2}A^{\mu\nu}[\sigma_{\mu\nu}]^{\alpha\beta}. \tag{113}
$$

We can therefore write our infinitesimal Lorentz transformation (109) as

$$
\Lambda^\alpha{}_\beta(\epsilon) = \delta_\beta^\alpha - \frac{i}{2}\epsilon^{\mu\nu}[\sigma_{\mu\nu}]^\alpha{}_\beta. \tag{114}
$$

Equivalently, in matrix notation, with the free indices α and β suppressed, we can write

$$
\Lambda(\epsilon) = 1 - \frac{i}{2}\epsilon^{\mu\nu}\sigma_{\mu\nu}. \tag{115}
$$

The tensors $[\sigma_{\mu\nu}]^\alpha{}_\beta$ are called the Lorentz generators and are obtained by lowering the β index in the definition (112) using the Minkowski metric tensor:

$$
[\sigma_{\mu\nu}]^\alpha{}_\beta = -i\delta_\mu^\alpha \eta_{\nu\beta} + i\eta_{\mu\beta}\delta_\nu^\alpha. \tag{116}
$$

We will often suppress the "additional" α, β indices for notational economy.

Note that with our overall sign convention for (116), the Lorentz generators describe *passive* Lorentz transformations, which transform our spacetime coordinate axes. If we instead wished to describe *active* Lorentz transformations, then we could either replace $\sigma_{\mu\nu} \mapsto -\sigma_{\mu\nu}$ or $\epsilon^{\mu\nu} \mapsto -\epsilon^{\mu\nu}$.

By straightforward calculations, one can show that the Lorentz generators satisfy the commutation relations

$$
\begin{aligned}
[\sigma_{\mu\nu}, \sigma_{\rho\sigma}] &\equiv \sigma_{\mu\nu}\sigma_{\rho\sigma} - \sigma_{\rho\sigma}\sigma_{\mu\nu} \\
&= i\eta_{\mu\rho}\sigma_{\nu\sigma} - i\eta_{\mu\sigma}\sigma_{\nu\rho} - i\eta_{\nu\rho}\sigma_{\mu\sigma} + i\eta_{\nu\sigma}\sigma_{\mu\rho},
\end{aligned} \tag{117}
$$

and that the matrix product of two Lorentz generators $\sigma_{\mu\nu}$ and $\sigma_{\rho\sigma}$ on their additional α, β indices, traced over those additional indices, yields

$$
\begin{aligned}
\frac{1}{2}\mathrm{Tr}[\sigma^{\mu\nu}\sigma_{\rho\sigma}] &\equiv \frac{1}{2}[\sigma^{\mu\nu}]^\alpha{}_\beta [\sigma_{\rho\sigma}]^\beta{}_\alpha \\
&= \delta_\rho^\mu \delta_\sigma^\nu - \delta_\sigma^\mu \delta_\rho^\nu \tag{118} \\
&= i[\sigma_{\rho\sigma}]^{\mu\nu}. \tag{119}
\end{aligned}
$$

This last formula implies that antisymmetric tensors $A^{\mu\nu}$ satisfy the identity

$$
\frac{1}{2}\mathrm{Tr}[\sigma^{\mu\nu}A] = iA^{\mu\nu}. \tag{120}
$$

Using this formalism, we can rewrite our system's spin tensor (94) as

$$
\begin{aligned}
S^{\mu\nu} &= -\frac{i}{2}\mathrm{Tr}[\sigma^{\mu\nu}S] \\
&= -\frac{i}{2}\mathrm{Tr}[\sigma^{\mu\nu}\Lambda S_0 \Lambda^{-1}]. \tag{121}
\end{aligned}
$$

G. The Manifestly Covariant Action Functional

In the absence of spin, the system's manifestly covariant action functional takes the form (39):

$$
\begin{aligned}
S_{\text{no spin}}[X, \Lambda] &= \int d\lambda\, \mathscr{L}_{\text{no spin}} \\
&= \int d\lambda\, p_\mu \dot{X}^\mu = \int d\lambda\, (\Lambda p_0)_\mu \dot{X}^\mu. \tag{122}
\end{aligned}
$$

Here $X^\mu(\lambda)$ and $p^\mu(\lambda) \equiv \Lambda^\mu{}_\nu(\lambda)p_0^\nu$ are functions of the worldline parameter λ, and dots, as usual, denote derivatives with respect to λ. We will eventually see that this action functional is capable of accommodating particle types regardless of their mass—and, in particular, works just as well for massless particles as it does for particles with nonzero mass. We will ultimately also need to establish a definite relationship between the system's four-momentum $p^\mu(\lambda)$ and the system's four-velocity $\dot{X}^\mu(\lambda) \equiv dX^\mu(\lambda)/d\lambda$.

In order to include spin in the system's action functional, we will need to develop a framework for taking derivatives of the variable Lorentz-transformation matrix $\Lambda^\mu{}_\nu(\lambda)$ with respect to the worldline parameter λ in a manner that is consistent with the constraint $\Lambda^{\mathrm{T}}\eta\Lambda = \eta$ from (50). To this end, we examine what happens if we shift slightly forward along the system's worldline, so that

$$
\lambda \to \lambda + d\lambda. \tag{123}
$$

Using the fact that successive Lorentz transformations compose,

$$
\Lambda'' = \Lambda'\Lambda, \tag{124}
$$

and recalling the formula (115) for a Lorentz transformation that differs infinitesimally from the identity, with $d\theta^{\mu\nu} \equiv \epsilon^{\mu\nu}$ denoting our Lorentz-boost and angular parameters, it follows that

$$
\begin{aligned}
\Lambda(\lambda + d\lambda) &= \Lambda(d\lambda)\Lambda(\lambda) \\
&= (1 - (i/2)d\theta^{\mu\nu}(\lambda)\sigma_{\mu\nu})\Lambda(\lambda). \tag{125}
\end{aligned}
$$

We can rearrange this formula to obtain the derivative of $\Lambda(\lambda)$ with respect to λ in terms of the rates of change $\dot{\theta}^{\mu\nu}(\lambda) \equiv d\theta^{\mu\nu}(\lambda)/d\lambda$:

$$
\begin{aligned}
\dot{\Lambda}(\lambda) &\equiv \lim_{d\lambda \to 0} \frac{\Lambda(\lambda + d\lambda) - \Lambda(\lambda)}{d\lambda} \\
&= -\frac{i}{2}\dot{\theta}^{\mu\nu}(\lambda)\sigma_{\mu\nu}\Lambda(\lambda).
\end{aligned} \tag{126}
$$

Hence,

$$
\dot{\Lambda}(\lambda)\Lambda^{-1}(\lambda) = -\frac{i}{2}\dot{\theta}^{\mu\nu}(\lambda)\sigma_{\mu\nu},
$$

and so, invoking the identities (113) and (119), we obtain an important formula for the rates of change $\dot{\theta}^{\mu\nu}(\lambda)$ of the Lorentz-transformation parameters:

$$
\dot{\theta}^{\mu\nu}(\lambda) = \frac{i}{2}\mathrm{Tr}[\sigma^{\mu\nu}\dot{\Lambda}(\lambda)\Lambda^{-1}(\lambda)]. \tag{127}
$$

Despite the factor of i, this expression is purely real, due to the additional factor of i in the definition (112) of $\sigma^{\mu\nu}$.

We now look back at the manifestly covariant Lagrangian appearing as the integrand of our action functional (122):

$$
\mathscr{L}_{\text{no spin}} = p_\mu \dot{X}^\mu. \tag{128}
$$

Using the product rule in reverse (that is, "integration by parts" without an actual integration), we can move the derivative from $X^\mu(\lambda)$ to $p_\mu(\lambda)$ at the cost of an overall minus sign and an additive total derivative that does not affect the system's equations of motion. The result is

$$
\mathscr{L}_{\text{no spin}} = -X_\mu \dot{p}^\mu + (\text{total derivative}).
$$

Remembering that the system's four-momentum $p^\mu(\lambda)$ here is fundamentally defined according to (93) in terms of its fixed reference value p_0^μ and the variable Lorentz-transformation matrix $\Lambda^\mu{}_\nu(\lambda)$,

$$
p^\mu(\lambda) \equiv \Lambda^\mu{}_\nu(\lambda)p_0^\nu,
$$

and relabeling indices for later convenience, we have

$$
\mathscr{L}_{\text{no spin}} = -X_\alpha \dot{\Lambda}^\alpha{}_\gamma p_0^\gamma + (\text{total derivative}).
$$

Invoking (126) for the derivative of the Lorentz-transformation matrix yields

$$
\begin{aligned}
\mathscr{L}_{\text{no spin}} &= -X_\alpha\left(-\frac{i}{2}\dot{\theta}^{\mu\nu}[\sigma_{\mu\nu}]^\alpha{}_\beta \Lambda^\beta{}_\gamma\right)p_0^\gamma \\
&\quad + (\text{total derivative}) \\
&= \frac{1}{2}X_\alpha i[\sigma_{\mu\nu}]^\alpha{}_\beta p^\beta \dot{\theta}^{\mu\nu} + (\text{total derivative}).
\end{aligned}
$$

Recalling our formula (116) for the Lorentz generators $[\sigma_{\mu\nu}]^\alpha{}_\beta$, this expression simplifies to

$$
\begin{aligned}
\mathscr{L}_{\text{no spin}} &= \frac{1}{2}X_\alpha(\delta^\alpha_\mu \eta_{\nu\beta} - \eta_{\mu\beta}\delta^\alpha_\nu)p^\beta \dot{\theta}^{\mu\nu} \\
&\quad + (\text{total derivative}) \\
&= \frac{1}{2}(X_\mu p_\nu - X_\nu p_\mu)\dot{\theta}^{\mu\nu} + (\text{total derivative}).
\end{aligned}
$$

The quantity in parentheses is precisely the system's orbital angular-momentum tensor $L_{\mu\nu}$, as defined in (65), so we end up with

$$
\mathscr{L}_{\text{no spin}} = \frac{1}{2}L_{\mu\nu}\dot{\theta}^{\mu\nu} + (\text{total derivative}). \tag{129}
$$

The first term in (129) has precisely the form of a canonical momentum contracted with the rates of change of its corresponding canonical coordinates, where the factor of $1/2$ naturally prevents the implicit summation from double-counting independent terms in the contraction of the two antisymmetric tensors $L_{\mu\nu} = -L_{\nu\mu}$ and $\dot{\theta}^{\mu\nu} = -\dot{\theta}^{\nu\mu}$. It may seem surprising that we have managed to rewrite the system's kinetic Lagrangian $\mathscr{L}_{\text{no spin}} = p_\mu \dot{X}^\mu$ in terms of what looks superficially like purely orbital angular momentum, but remember that the temporal components L^{ti} of the orbital angular-momentum tensor are not angular momenta—indeed, in light of (66), they actually encode linear motion.

Including the system's spin in the dynamics means generalizing the orbital angular-momentum tensor $L_{\mu\nu}$ in (129) to the *total* angular-momentum tensor $J_{\mu\nu}$ defined in (68),

$$
L_{\mu\nu} \mapsto J_{\mu\nu} \equiv L_{\mu\nu} + S_{\mu\nu},
$$

where $S_{\mu\nu}$ is the system's spin tensor. The system's manifestly covariant Lagrangian correspondingly becomes

$$
\begin{aligned}
\mathscr{L}_{\text{no spin}} \mapsto \mathscr{L} &\equiv \frac{1}{2}J_{\mu\nu}\dot{\theta}^{\mu\nu} + (\text{total derivative}) \\
&= \frac{1}{2}L_{\mu\nu}\dot{\theta}^{\mu\nu} + \frac{1}{2}S_{\mu\nu}\dot{\theta}^{\mu\nu} + (\text{total derivative}).
\end{aligned} \tag{130}
$$

At this point, we are free to recombine the first and last terms in $\mathscr{L}$ to get back the expression $p_\mu \dot{X}^\mu$ that we started with. Moreover, by contracting both sides of our formula (127) for $\dot{\theta}^{\mu\nu}$ with the system's spin tensor $S^{\mu\nu}$, and using $(i/2)S_{\mu\nu}[\sigma^{\mu\nu}]^\alpha{}_\beta = S^\alpha{}_\beta$ from (113), we can write the second term in (130) as

$$
\frac{1}{2}S_{\mu\nu}(\lambda)\dot{\theta}^{\mu\nu}(\lambda) = \frac{1}{2}\mathrm{Tr}[S(\lambda)\dot{\Lambda}(\lambda)\Lambda^{-1}(\lambda)]. \tag{131}
$$

Hence, as originally shown in [3, 13–15], the complete action functional for the system is

$$
S[X, \Lambda] = \int d\lambda\, \mathscr{L} = \int d\lambda\left(p_\mu \dot{X}^\mu + \frac{1}{2}\mathrm{Tr}[S\dot{\Lambda}\Lambda^{-1}]\right). \tag{132}
$$

In using the action functional (132), keep in mind that the four-momentum $p^\mu(\lambda)$ and the spin tensor $S^{\mu\nu}(\lambda)$ are given respectively by (93) and (121) in terms of their constant reference values p_0^μ and $S_0^{\mu\nu}$ together with the variable Lorentz-transformation matrix $\Lambda^\mu{}_\nu(\lambda)$:

$$
p^\mu(\lambda) \equiv \Lambda^\mu{}_\nu(\lambda)p_0^\nu, \tag{133}
$$

$$
\begin{aligned}
S^{\mu\nu}(\lambda) &\equiv \Lambda^\mu{}_\rho(\lambda)S_0^{\rho\sigma}(\Lambda^{\mathrm{T}})_\sigma{}^\nu(\lambda) \\
&= -\frac{i}{2}\mathrm{Tr}[\sigma^{\mu\nu}\Lambda(\lambda)S_0\Lambda^{-1}(\lambda)].
\end{aligned} \tag{134}
$$

Consequently, before the equations of motion are imposed, neither $p^\mu(\lambda)$ nor $S^{\mu\nu}(\lambda)$ depend on the spacetime degrees of freedom $X^\mu(\lambda)$.

H.   The Equations of Motion

To obtain the system's equations of motion, we apply the extremization condition (9) by varying the action functional (132) with respect to its fundamental variables X^μ and $\Lambda^\mu{}_\nu$. The spin term $(1/2)\mathrm{Tr}[S\dot\Lambda\Lambda^{-1}]$ does not involve the spacetime coordinates X^μ, so varying the action functional with respect to X^μ yields

$$\delta_X S = \int d\lambda\,(p_\mu \delta \dot X^\mu + 0)$$
$$= \int d\lambda\, p_\mu \frac{d}{d\lambda}\delta X^\mu$$
$$= -\int d\lambda\,\dot p_\mu \delta X^\mu,$$

where we have dropped a boundary term. Setting this variation equal to zero for arbitrary δX^μ leads to the system's first equation of motion, which we see describes conservation of energy-momentum:

$$\dot p^\mu = 0. \qquad (135)$$

Notice that this equation of motion, by itself, does not determine the system's four-velocity $\dot X^\mu \equiv dX^\mu/d\lambda$, or even establish any sort of relationship between p^μ and $\dot X^\mu$. We will return to this issue later.

Varying the action functional with respect to the variable Lorentz-transformation matrix $\Lambda^\mu{}_\nu$ is more complicated, due to its appearance in both terms in the integrand. As our first step, we find

$$\delta_\Lambda S = \int d\lambda \left((\delta p^\mu)\dot X_\mu + \frac{1}{2}\mathrm{Tr}[\delta(S\dot\Lambda\Lambda^{-1})]\right). \quad (136)$$

Invoking our formula (133) for the four-momentum p^μ in terms of its reference value p_0^μ and the Lorentz-transformation matrix $\Lambda^\mu{}_\nu$, the first term in (136) gives

$$(\delta p^\mu)\dot X_\mu = (\delta\Lambda^\mu{}_\nu)p_0^\nu \dot X_\mu$$
$$= (-(i/2)\delta\theta^{\rho\sigma}\sigma_{\rho\sigma}\Lambda)^\mu{}_\nu p_0^\nu \dot X_\mu$$
$$= -\frac{i}{2}\delta\theta^{\rho\sigma}[\sigma_{\rho\sigma}]^\mu{}_\nu p^\nu \dot X_\mu$$
$$= -\frac{i}{2}\delta\theta^{\rho\sigma}(-i\delta_\rho^\mu \eta_{\sigma\nu} + i\eta_{\rho\nu}\delta_\sigma^\mu)p^\nu \dot X_\mu$$
$$= \frac{1}{2}(-\dot X_\rho p_\sigma + \dot X_\sigma p_\rho)\delta\theta^{\rho\sigma}.$$

Meanwhile, using $S^\alpha{}_\beta = (\Lambda S_0\Lambda^{-1})^\alpha{}_\beta$, the second term

in (136) gives

$$\frac{1}{2}\mathrm{Tr}[\delta(S\dot\Lambda\Lambda^{-1})] = \frac{1}{2}\mathrm{Tr}[S_0\delta(\Lambda^{-1}\dot\Lambda)]$$
$$= \frac{1}{2}\mathrm{Tr}[S_0\delta(\Lambda^{-1})\dot\Lambda + S_0\Lambda^{-1}\delta\dot\Lambda]$$
$$= \frac{1}{2}\mathrm{Tr}[S_0\Lambda^{-1}(-(i/2)\delta\dot\theta^{\rho\sigma}\sigma_{\rho\sigma})\Lambda]$$
$$= -\frac{i}{4}\mathrm{Tr}[S_0\Lambda^{-1}\sigma_{\rho\sigma}\Lambda]\delta\dot\theta^{\rho\sigma}$$
$$= \frac{1}{2}S_{\rho\sigma}\frac{d}{d\lambda}\delta\theta^{\rho\sigma},$$

where we have invoked (134) in the last step. Thus, dropping a boundary term, we see that the overall variation (136) in the action functional reduces to

$$\delta_\Lambda S = \int d\lambda\,\frac{1}{2}(-\dot X_\rho p_\sigma + \dot X_\sigma p_\rho - \dot S_{\rho\sigma})\delta\theta^{\rho\sigma}.$$

Setting this variation equal to zero for arbitrary $\delta\theta^{\rho\sigma}$ leads to the system's second equation of motion:

$$\dot S^{\mu\nu} = -\dot X^\mu p^\nu + \dot X^\nu p^\mu. \qquad (137)$$

To provide an interpretation for this equation of motion, we recall again the definition (65) of the tensor $L^{\mu\nu}$ that encodes the system's orbital angular momentum:

$$L^{\mu\nu} \equiv X^\mu p^\nu - X^\nu p^\mu.$$

Because the system's four-momentum p^μ is conserved, (135), we see that the rate of change in $L^{\mu\nu}$ is given by

$$\dot L^{\mu\nu} = \dot X^\mu p^\nu - \dot X^\nu p^\mu, \qquad (138)$$

so we can recast the equation of motion (137) for the spin tensor $S^{\mu\nu}$ as the statement that the system's total angular momentum $J^{\mu\nu} \equiv L^{\mu\nu} + S^{\mu\nu}$ is conserved:

$$\dot J^{\mu\nu} = 0. \qquad (139)$$

Combining $\dot p^\mu = 0$ and $\dot J^\mu = 0$, it follows immediately that the system's Pauli-Lubanski pseudovector (98) is likewise constant in time:

$$\dot W^\mu = 0. \qquad (140)$$

At a deeper level, the system's two equations of motion (135), $\dot p^\mu = 0$, and (139), $\dot J^{\mu\nu} = 0$, are consequences of Noether's theorem together with the fact that the system's action functional (132) has continuous symmetries under spacetime translations and Lorentz transformations.

I.   Self-Consistency Conditions on the Phase Space

Now that we know the system's equations of motion, we will need to ensure that they are consistent with the

invariance of the fixed quantities m^2, w^2, s^2, and $\tilde{s}^2$ from (104)–(108).

For our first check of self-consistency, we note that the invariance of $p^2 \equiv -m^2c^2$ is compatible with the equation of motion (135), $\dot{p}^\mu = 0$:

$$\frac{d}{d\lambda}(p^2) = 2p_\mu \dot{p}^\mu = 0. \tag{141}$$

Similarly, the constancy of $W^2 \equiv w^2$ is compatible with the constancy (140) of the Pauli-Lubanski pseudovector:

$$\frac{d}{d\lambda}(W^2) = 2W_\mu \dot{W}^\mu = 0. \tag{142}$$

By contrast, the constancy of the spin-squared scalar $(1/2)S_{\mu\nu}S^{\mu\nu} \equiv s^2$, combined with the equation of motion (137), $\dot{S}^{\mu\nu} = -\dot{X}^\mu p^\nu + \dot{X}^\nu p^\mu$, requires that

$$\frac{d}{d\lambda}\left(\frac{1}{2}S_{\mu\nu}S^{\mu\nu}\right) = S_{\mu\nu}\dot{S}^{\mu\nu} = 2\dot{X}^\nu p^\mu S_{\mu\nu} = 0. \tag{143}$$

Again, keep in mind that we have not yet established a definite relationship between the system's four-momentum p^μ and its four-velocity $\dot{X}^\mu \equiv dX^\mu/d\lambda$. In particular, it is not clear at this point whether or not p^μ is proportional to $\dot{X}^\mu$, so the condition (143) is not trivial.

Because the condition (143) must hold for all solution trajectories, it imposes an additional requirement on the system's phase space. Specifically, the system's reference four-momentum p_0^μ and its reference spin tensor $S_0^{\mu\nu}$ must satisfy

$$p_{0,\mu}S_0^{\mu\nu} = 0. \tag{144}$$

The tensor-contraction appearing on the left-hand side therefore vanishes in one inertial reference frame, so it remains zero under all Poincaré transformations and therefore represents a Poincaré-invariant statement about the system's phase space:

$$p_\mu S^{\mu\nu} = 0. \tag{145}$$

Notice that this last self-consistency condition, which was also introduced in [14], is phrased in terms of linear and angular momentum, both of which have physical meanings in classical field theory. In particular, the condition $p_\mu S^{\mu\nu} = 0$ is closely related to the momentum-space version of the Lorenz equation

$$\partial_\mu A^\mu = 0 \tag{146}$$

that appears both in the Proca theory of a massive spin-1 bosonic field and as the condition for Lorenz gauge in electromagnetism. Like the Lorenz equation in those field theories, we will eventually see that the condition (145) ends up eliminating unphysical spin states.

To make this connection with field theory more explicit, consider the classical theory of a spin-1 field, and let

$$A^\mu(x) = \varepsilon^\mu \exp(ip \cdot x/\hbar) \tag{147}$$

be a monochromatic plane wave of the field. Here ε^μ is the wave's polarization four-vector, which ultimately encodes the field's quantized spin. Then the Lorenz equation $\partial_\mu A^\mu = 0$ reduces to the statement that

$$p_\mu \varepsilon^\mu = 0, \tag{148}$$

which eliminates one linear combination of polarizations, meaning one unphysical spin state. In contrast with a spin-1 field theory, our classical particle's invariant spin is not quantized, and our condition $p_\mu S^{\mu\nu} = 0$ will turn out to eliminate a continuous infinity of unphysical spin states.

Our final self-consistency condition is that the derivative of the pseudoscalar invariant quantity $(1/8)\epsilon_{\mu\nu\rho\sigma}S^{\mu\nu}S^{\rho\sigma} \equiv \tilde{s}^2$ must vanish:

$$\begin{aligned}\frac{d}{d\lambda}\left(\frac{1}{8}\epsilon_{\mu\nu\rho\sigma}S^{\mu\nu}S^{\rho\sigma}\right) &= \frac{1}{4}\epsilon_{\mu\nu\rho\sigma}\dot{S}^{\mu\nu}S^{\rho\sigma} \\ &= -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\dot{X}^\mu p^\nu S^{\rho\sigma} \\ &= \dot{X}^\mu W_\mu = 0. \end{aligned} \tag{149}$$

We will need to verify in the explicit examples ahead that this condition is indeed satisfied.

J. The Four-Velocity

The self-consistency condition (145), $p_\mu S^{\mu\nu} = 0$, will play an important role in our work ahead. As we will now investigate, its implications include a general set of relationships between the system's four-momentum p^μ and its four-velocity $\dot{X}^\mu$.

Taking a derivative of both sides of $p_\mu S^{\mu\nu} = 0$ with respect to the worldline parameter λ, and invoking the equations of motion (135), $\dot{p}^\mu = 0$, and (137), $\dot{S}^{\mu\nu} = -\dot{X}^\mu p^\nu + \dot{X}^\nu p^\mu$, we obtain

$$p_\mu \dot{S}^{\mu\nu} = -(p \cdot \dot{X})p^\nu + (-m^2c^2)\dot{X}^\nu = 0, \tag{150}$$

which gives us an equation that relates p^μ and $\dot{X}^\mu$:

$$(p \cdot \dot{X})p^\mu = (-m^2c^2)\dot{X}^\mu. \tag{151}$$

Contracting both sides with $\dot{X}_\mu$, we find

$$(p \cdot \dot{X})^2 = m^2c^2(-\dot{X}^2). \tag{152}$$

Taking the square root of this last equation, and substituting the result back into the equation before it, we arrive at the following pair of equations:

$$p \cdot \dot{X} = \pm mc^2\sqrt{-\dot{X}^2/c^2}, \tag{153}$$

$$m\sqrt{-\dot{X}^2/c^2}\,p^\mu = \mp m^2 \dot{X}^\mu. \tag{154}$$

Together with (135), $\dot{p}^\mu = 0$, (139), $\dot{J}^{\mu\nu} = 0$, and (145), $p_\mu S^{\mu\nu} = 0$, these final two equations complete our specification of the system's dynamics.

V. CLASSIFICATION OF THE TRANSITIVE GROUP ACTIONS OF THE ORTHOCHRONOUS POINCARÉ GROUP

We are now ready to apply the foregoing framework to the task of classifying systems whose phase spaces provide transitive group actions of the Poincaré group. For simplicity, we will focus our attention on transitive group actions of the *orthochronous* Poincaré group, putting aside time-reversal transformations (55) until our paper's conclusion.

Notice then that for $m^2 \geq 0$, the system's four-momentum p^μ is either timelike or null, $p^2 \leq 0$, and so (56) implies that the *sign* of p^μ is an invariant property of the system. When we consider transitive group actions having $m^2 \geq 0$, we will assume the positive-energy case $p^t > 0$ on physical grounds. We will address the "negative-energy" case $p^t < 0$ in our conclusion.

A. Massive, Positive-Energy Particles

As our first example, we consider a transitive group action of the orthochronous Poincaré group for which $m > 0$ is real and positive, and for which the system's energy $E = p^t c > 0$ is likewise positive. Then p^μ is a timelike four-vector, so we know from (56) that the sign of p^t is invariant under orthochronous Lorentz transformations, and thus our choice of positive energy is well-defined.

Given that $p^2 = -m^2 c^2$ for $m > 0$ with positive p^t, we can express the system's energy $E = p^t c$ in terms of its three-dimensional momentum $\mathbf{p} = (p_x, p_y, p_z)$ as

$$E = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4}, \qquad (155)$$

a formula known as the system's mass-shell relation because it takes the visual form of a hyperboloid (a "shell") when plotted in terms of the four variables E, p_x, p_y, p_z. Furthermore, there exists a state of the system in which the four-momentum p^μ takes the specific value $(mc, \mathbf{0})^\mu$, which we will choose to be its reference value:

$$p_0^\mu \equiv (mc, \mathbf{0})^\mu = mc\, \delta_t^\mu. \qquad (156)$$

Due to the condition $m > 0$, the four-momentum p^μ cannot vanish, and under our assumption of a strictly monotonic parametrization $X^\mu(\lambda)$, the four-velocity $\dot{X}^\mu$ cannot vanish either, so the relation (154),

$$m\sqrt{-\dot{X}^2/c^2}\, p^\mu = \mp m^2 \dot{X}^\mu,$$

implies that $\dot{X}^2 \neq 0$. We therefore have

$$p^\mu = m\frac{\dot{X}^\mu}{\sqrt{-\dot{X}^2/c^2}},$$

where we have taken the positive sign by choosing our parametrization $X^\mu(\lambda)$ such that $\dot{X}^\mu$ is future-directed.

We therefore learn that the system's four-momentum p^μ is given by

$$p^\mu = mu^\mu, \qquad (157)$$

where u^μ is the system's *normalized* four-velocity:

$$u^\mu \equiv \frac{\dot{X}^\mu}{\sqrt{-\dot{X}^2/c^2}}, \quad u^2 = -c^2. \qquad (158)$$

We can interpret the equation (157) as supplying our *definition* of $\dot{X}^\mu$ (or u^μ) in terms of p^μ and m. Furthermore, because p^μ is parallel to u^μ, we see that the self-consistency condition (149), $\dot{X}^\mu W_\mu = 0$, is satisfied.

As a consequence of (158), we also see that when the system is in its reference state with $p^\mu = p_0^\mu = (mc, \mathbf{0})^\mu$, the four-velocity describes the system at rest, with

$$u_0^\mu = (c, \mathbf{0})^\mu = u_{\text{rest}}^\mu. \qquad (159)$$

For general states, the equation of motion (135) for the system's four-momentum, $\dot{p}^\mu = 0$, tells us that the system's normalized four-velocity is constant:

$$\dot{u}^\mu = 0. \qquad (160)$$

It follows that the system describes a pointlike particle that travels along a straight, timelike path in spacetime.

Defining the particle's three-dimensional velocity $\mathbf{v} = (v_x, v_y, v_z)$ as

$$\mathbf{v} \equiv \frac{d\mathbf{X}}{dt} = \frac{\dot{\mathbf{X}}}{\dot{T}}, \qquad (161)$$

and using (157), $p^\mu = mu^\mu$, together with $E = p^t c$ and the mass-shell relation (155) between E and $\mathbf{p}$, we also obtain an important equation connecting the system's three-dimensional velocity $\mathbf{v}$ and its three-dimensional momentum $\mathbf{p}$:

$$\mathbf{v} = \frac{\mathbf{p}c^2}{E} = \frac{\mathbf{p}}{|\mathbf{p}|}\frac{c}{\sqrt{1 + m^2 c^2/\mathbf{p}^2}}. \qquad (162)$$

We see right away from this equation that the particle's speed $|\mathbf{v}|$ is always slower than the speed of light c:

$$|\mathbf{v}| < c. \qquad (163)$$

Moreover, in the general case in which particle may be in motion, its normalized four-velocity is

$$u^\mu = (\gamma c, \gamma \mathbf{v})^\mu, \qquad (164)$$

where the Lorentz factor γ is defined by

$$\gamma \equiv \frac{1}{\sqrt{1 - \mathbf{v}^2/c^2}} \geq 1. \qquad (165)$$

We next examine the particle's orbital and spin angular momentum. The relation (157), $p^\mu = mu^\mu = $

$m\dot{X}^\mu/\sqrt{-\dot{X}^2/c^2}$, immediately implies that the particle's orbital angular momentum (65) is conserved:

$$\dot{L}^{\mu\nu} = \dot{X}^\mu p^\nu - \dot{X}^\nu p^\mu = 0. \quad (166)$$

Remembering our formula (66) for the temporal components L^{ti} of the orbital angular-momentum tensor,

$$L^{ti} = -\frac{E}{c}\left(X^i - \frac{p^i c^2}{E}T\right),$$

and invoking the constancy of E and p^i from the equation of motion (135) for p^μ, we see that $\dot{L}^{ti} = 0$ gives the relation

$$\frac{\mathbf{p}c^2}{E} = \frac{\dot{\mathbf{X}}}{\dot{T}},$$

which is just our earlier equation (162) connecting the particle's three-dimensional velocity $\mathbf{v}$ to its three-dimensional momentum $\mathbf{p}$.[11]

Combining the conservation equation (166) for the particle's orbital angular-momentum tensor $L^{\mu\nu}$ with the equation of motion (137) for the particle's spin tensor $S^{\mu\nu}$ tells us that the particle's spin is separately conserved:

$$\dot{S}^{\mu\nu} = 0. \quad (167)$$

Furthermore, the condition (144), $p_{0,\mu}S_0^{\mu\nu} = 0$, becomes

$$mc\,S_0^{t\nu} = 0, \quad (168)$$

so only the purely spatial components of the particle's reference spin tensor $S_0^{\mu\nu}$ are nonzero,

$$S_0^{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & S_{0,z} & -S_{0,y} \\ 0 & -S_{0,z} & 0 & S_{0,x} \\ 0 & S_{0,y} & -S_{0,x} & 0 \end{pmatrix}^{\mu\nu}, \quad (169)$$

where the particle's spin three-vector $\mathbf{S} \equiv (S^{yz}, S^{zx}, S^{xy})$ was defined in (71). Thus, the invariant quantity s^2 defined in (107) and characterizing the system's overall spin is non-negative:

$$\begin{aligned} s^2 &= \mathbf{S}^2 - \tilde{\mathbf{S}}^2 \\ &= \mathbf{S}_0^2 = S_{0,x}^2 + S_{0,y}^2 + S_{0,z}^2 \geq 0. \end{aligned} \quad (170)$$

The corresponding reference value W_0^μ of the Pauli-Lubanski pseudovector (99) is then

$$W_0^\mu = (0, mc\,\mathbf{S}_0)^\mu. \quad (171)$$

The Lorentz dot product of W^μ with itself therefore has the non-negative, Lorentz-invariant value

$$W^2 \equiv w^2 = m^2 c^2 s^2 \geq 0. \quad (172)$$

Notice that the reference value of the particle's dual spin three-vector $\tilde{\mathbf{S}} \equiv (S^{tx}, S^{ty}, S^{tz})$, as defined in (72), vanishes in this case:

$$\tilde{\mathbf{S}}_0 = 0. \quad (173)$$

It follows that the pseudoscalar invariant quantity $\tilde{s}^2$ defined in (108) likewise vanishes:

$$\tilde{s}^2 = \mathbf{S} \cdot \tilde{\mathbf{S}} = \mathbf{S}_0 \cdot \tilde{\mathbf{S}}_0 = 0. \quad (174)$$

On physical grounds, a localized system at fixed energy should have a compact (that is, closed and bounded) set of states, because otherwise its Boltzmann entropy under any equitable choice of coarse-graining of the system's fixed-energy phase space would be infinite, and thus the system would exhibit an infinite heat capacity.[12] The compactness of a system's phase space at fixed energy in any one inertial reference frame determines the compactness of the system's phase space in any other inertial reference frame at the correspondingly Lorentz-transformed energy, so it suffices to examine the compactness of our particle's phase space at the fixed reference energy $E_0 = p_0^t c = mc^2$ corresponding to the reference value (156) of the particle's four-momentum. The size of this subset of the particle's phase space is determined by the set of all orthochronous Lorentz transformations that leave the particle's reference four-momentum $p_0^\mu \equiv (mc, \mathbf{0})^\mu$ fixed. This collection of transformations is called the little group of p_0^μ. In the present case, in which $p_0^\mu = (mc, \mathbf{0})^\mu$, this little group consists solely of the group $O(3)$ of three-dimensional rotations and parity transformations, which collectively form a compact set, so we are assured that the particle's phase space at any fixed energy is likewise compact, as required.

To summarize, we see that a transitive group action of the orthochronous Poincaré group for the case of a real and positive $m > 0$ and positive energy $E = p^t c > 0$ describes a massive pointlike particle of inertial mass m, non-negative spin-squared $s^2 = \mathbf{S}_0^2 \geq 0$, non-negative squared Pauli-Lubanski pseudovector $w^2 = m^2 c^2 s^2 \geq 0$, and timelike four-momentum $p^\mu = mu^\mu$. The particle moves along a straight worldline in spacetime characterized by a normalized four-velocity $u^\mu \equiv \dot{X}^\mu/\sqrt{-\dot{X}^2/c^2}$ and a three-dimensional velocity $\mathbf{v} = \mathbf{p}c^2/E$ that is always slower than the speed of light, $|\mathbf{v}| < c$, and the particle has a compact phase space at any fixed value of its energy E.

---

[11] More generally, for a system of multiple particles labeled by $\alpha = 1, 2, \ldots$, the spatial components L^{ti} generalize to the system's center-of-mass-energy $X_{\mathrm{CM}}^i = \sum_\alpha E_\alpha X_{\mathrm{initial},\alpha}^i/E_{\mathrm{total}}$, and so their conservation implies the constancy of $\mathbf{X}_{\mathrm{CM}}$.

[12] For related arguments, see [12, 16].

B.   Massless, Positive-Energy Particles

As our second example, we consider the case of $m = 0$ and positive energy $E = p^t c > 0$. Because the system's four-momentum p^μ is therefore null, $p^2 = 0$, we again have from (56) that the condition $p^t > 0$ is invariant under orthochronous Lorentz transformations, and thus our positivity condition on E is well-defined.

We can use $p^2 = 0$ to express the system's energy $E = p^t c$ in terms of its three-dimensional momentum $\mathbf{p}$ as the mass-shell relation

$$E = |\mathbf{p}|c. \qquad (175)$$

In contrast with the massive case, there is no rest frame in the massless case—that is, we cannot set all three components of $\mathbf{p}$ to be zero, due to the pair of assumptions $p^2 = 0$ and $E = p^t c > 0$. There exist states in the system's phase space in which the four-momentum p^μ has no x or y components, and we take one such value of the four-momentum to be its reference value, for a fixed but arbitrarily chosen value $E_0 > 0$ of the system's energy:

$$p_0^\mu \equiv (E_0/c, 0, 0, E_0/c)^\mu = \frac{E_0}{c}(\delta_t^\mu + \delta_t^z). \qquad (176)$$

The positive-energy condition $E > 0$ implies that the four-momentum p^μ cannot vanish, and under our assumption of a strictly monotonic parametrization $X^\mu(\lambda)$, the four-velocity $\dot{X}^\mu$ also cannot vanish. With $m = 0$, the relation (153) degenerates to

$$p \cdot \dot{X} = 0. \qquad (177)$$

A nonzero four-vector that has vanishing dot product with a null four-vector must be parallel to that null four-vector. We can therefore take the four-velocity $\dot{X}^\mu$ to be a null four-vector that is parallel to the four-momentum p^μ,

$$p^\mu \propto \dot{X}^\mu, \qquad (178)$$

which then ensures that the self-consistency condition (149), $\dot{X}^\mu W_\mu = 0$, is satisfied.

The equation of motion (135), $\dot{p}^\mu = 0$, implies that p^μ is constant along the system's worldline, so we can always choose our parametrization $X^\mu(\lambda)$ to make the proportionality factor in (178) equal to a constant:

$$p^\mu = (\text{const})\dot{X}^\mu. \qquad (179)$$

We then have

$$\ddot{X}^\mu = 0, \qquad (180)$$

so we see that the system describes a pointlike particle that travels along a straight, null path in spacetime.

In addition, invoking the mass-shell relation (175) between the particle's energy E and its three-dimensional

momentum $\mathbf{p}$, we see that the particle's three-dimensional velocity $\mathbf{v}$ is related to its three-dimensional momentum $\mathbf{p}$ according to

$$\mathbf{v} = \frac{d\mathbf{X}}{dt} = \frac{\dot{\mathbf{X}}}{\dot{T}} = \frac{\mathbf{p}c^2}{E} = \frac{\mathbf{p}}{|\mathbf{p}|}c. \qquad (181)$$

Hence, the particle's speed $|\mathbf{v}|$ is always equal to the speed of light c:

$$|\mathbf{v}| = c. \qquad (182)$$

Turning to the particle's spin, we will find a much more nuanced story than in the massive case.

The proportionality relationship $p^\mu \propto \dot{X}^\mu$ from (178) together with the equation of motion (137) for the particle's spin tensor $S^{\mu\nu}$ again imply that the particle's angular momentum (65) and the particle's spin are separately conserved:

$$\dot{L}^{\mu\nu} = \dot{X}^\mu p^\nu - \dot{X}^\nu p^\mu = 0, \qquad (183)$$

$$\dot{S}^{\mu\nu} = 0. \qquad (184)$$

As in the massive case, the conservation law for L^{ti} gives back the formula (181) relating the particle's three-dimensional velocity $\mathbf{v}$ to its three-dimensional momentum $\mathbf{p}$.

However, the condition (144), $p_{0,\mu}S_0^{\mu\nu} = 0$, is more complicated than it was in the massive case:

$$-\frac{E_0}{c} S_0^{t\nu} + \frac{E_0}{c} S_0^{z\nu} = 0. \qquad (185)$$

This equation implies that

$$S_0^{t\nu} = S_0^{z\nu}, \qquad (186)$$

or, equivalently, that the quantities

$$\left. \begin{array}{l} A \equiv S_x + \tilde{S}_y, \\ B \equiv S_y - \tilde{S}_x, \end{array} \right\} \qquad (187)$$

and $\tilde{S}_z$ all vanish in the particle's reference state:

$$\left. \begin{array}{l} A_0 \equiv S_{0,x} + \tilde{S}_{0,y} = 0, \\ B_0 \equiv S_{0,y} - \tilde{S}_{0,x} = 0, \\ \tilde{S}_{0,z} = 0. \end{array} \right\} \qquad (188)$$

The reference value of the system's spin tensor is therefore

$$S_0^{\mu\nu} = \begin{pmatrix} 0 & S_{0,y} & -S_{0,x} & 0 \\ -S_{0,y} & 0 & S_{0,z} & -S_{0,y} \\ S_{0,x} & -S_{0,z} & 0 & S_{0,x} \\ 0 & S_{0,y} & -S_{0,x} & 0 \end{pmatrix}^{\mu\nu}. \qquad (189)$$

In other words, the reference value of the particle's spin three-vector $\mathbf{S} \equiv (S^{yz}, S^{zx}, S^{xy})$, as defined in (71), and the reference value of the particle's dual spin three-vector

$\tilde{\mathbf{S}} \equiv (S^{tx}, S^{ty}, S^{tz})$, as defined in (72), are mutually perpendicular and are related explicitly by

$$\tilde{\mathbf{S}}_0 = \mathbf{S}_0 \times \mathbf{e}_z, \qquad (190)$$

where $\mathbf{e}_z \equiv (0, 0, 1)$ is the usual three-dimensional Cartesian unit vector pointing along the positive z axis. It follows that the pseudoscalar invariant quantity $\tilde{s}^2$ defined in (108) vanishes, as we also saw was true in the massive case:

$$\tilde{s}^2 = \mathbf{S} \cdot \tilde{\mathbf{S}} = \mathbf{S}_0 \cdot \tilde{\mathbf{S}}_0 = 0. \qquad (191)$$

Meanwhile, the invariant quantity s^2 defined in (107) is non-negative, as in the massive case, but is now determined solely by the z component $S_{0,z}$ of the reference value $\mathbf{S}_0$ of the particle's spin three-vector:

$$s^2 = \mathbf{S}^2 - \tilde{\mathbf{S}}^2 = S_{0,z}^2 \geq 0. \qquad (192)$$

In general, the projection of the particle's spin three-vector $\mathbf{S}$ onto the particle's three-dimensional momentum $\mathbf{p} \equiv (p^x, p^y, p^z)$ is called the particle's helicity σ:

$$\sigma \equiv \frac{\mathbf{p}}{|\mathbf{p}|} \cdot \mathbf{S}. \qquad (193)$$

The massless particle's helicity is insensitive to our reference choice of energy E_0, due to the fact that $\mathbf{p}$ appears only as the ratio $\mathbf{p}/|\mathbf{p}|$, and is also invariant under proper rotations, due to the three-dimensional dot product. Thus, σ represents a fundamental feature of the particle in the $m = 0$ case that can only change under parity transformations (54), meaning that σ is a pseudoscalar:

$$\sigma \mapsto -\sigma \quad \text{(parity)}. \qquad (194)$$

It follows from evaluating the definition (193) of the helicity in the massless particle's reference state that the helicity is equal to the z component $S_{0,z}$ of the massless particle's reference spin three-vector $\mathbf{S}_0$, up to a possible sign that changes under parity transformations:

$$\sigma = (\pm)^{\text{parity}} S_{0,z}. \qquad (195)$$

We can therefore use σ to write our expression (192) for the invariant quantity s^2 as

$$s^2 = \sigma^2 \geq 0. \qquad (196)$$

The reference value W_0^μ of the particle's Pauli-Lubanski pseudovector (99) is parallel to the particle's reference four-momentum (176):

$$W_0^\mu = \left(S_{0,z} \frac{E_0}{c}, 0, 0, S_{0,z} \frac{E_0}{c} \right) = S_{0,z} p_0^\mu. \qquad (197)$$

More generally, W^μ is given in terms of the particle's helicity (193) by

$$W^\mu = \sigma p^\mu, \qquad (198)$$

where the pseudovector nature of W^μ is neatly captured by the pseudoscalar nature of σ. As a consequence of the condition $p^2 = 0$, we see that the invariant quantity w^2 defined in (105) vanishes:

$$W^2 \equiv w^2 = 0. \qquad (199)$$

As in the massive case, we will need to examine the compactness of the subset of the particle's phase space at the fixed reference energy $E_0 = p_0^t c$. Again, this subspace is determined by the little group of the particle's reference four-momentum (176), meaning the set of all orthochronous Lorentz transformations that leave $p_0^\mu \equiv (E_0/c, 0, 0, E_0/c)^\mu$ invariant.

Let Λ be a little-group transformation, so that $\Lambda p_0 = p_0$. For now, we will assume that Λ does not involve a parity transformation. As a trick for finding these little-group transformations,[13] let $v^\mu \equiv (1, \mathbf{0})^\mu$ be a purely timelike four-vector. Then

$$(\Lambda v) \cdot p_0 = -(\Lambda v)^t \frac{E_0}{c} + (\Lambda v)^z \frac{E_0}{c}$$
$$\text{also} = (\Lambda v) \cdot (\Lambda p_0) = v \cdot p_0 = -\frac{E_0}{c}.$$

We conclude that

$$(\Lambda v)^t = 1 + (\Lambda v)^z, \qquad (200)$$

and thus that $(\Lambda v)^\mu$ has the form

$$(\Lambda v)^\mu = (1 + \zeta, \alpha, \beta, \zeta)^\mu \qquad (201)$$

for real-valued parameters α, β, and ζ. The normalization condition $(\Lambda v)^2 = v^2 = -1$ implies that these three parameters are related by

$$\zeta = \frac{\alpha^2 + \beta^2}{2}. \qquad (202)$$

The effect of the little-group Lorentz-transformation matrix Λ on $v^\mu \equiv (1, \mathbf{0})^\mu$ fixes Λ up to an overall three-dimensional rotation, and the little-group requirement $\Lambda p_0 = p_0$ further fixes Λ up to a rotation specifically around the z axis. Hence, the most general such proper orthochronous Lorentz-transformation matrix Λ has the form

$$\Lambda(\alpha, \beta, \theta) = L(\alpha, \beta) R(\theta), \qquad (203)$$

where

$$R(\theta) \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (204)$$

---

[13] See, for example, [12].

is a pure rotation by an angle θ around the z axis, and where

$$L(\alpha, \beta) \equiv \begin{pmatrix} 1+\zeta & \alpha & \beta & -\zeta \\ \alpha & 1 & 0 & -\alpha \\ \beta & 0 & 1 & -\beta \\ \zeta & \alpha & \beta & 1-\zeta \end{pmatrix} \qquad (205)$$

is a complicated combination of proper orthochronous Lorentz boosts and rotations satisfying the required condition $\Lambda^{\mathrm{T}} \eta \Lambda = \eta$ from (50).

By straightforward calculations, one can show that

$$R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2), \qquad (206)$$
$$L(\alpha_1, \beta_1)L(\alpha_2, \beta_2) = L(\alpha_1 + \alpha_2, \beta_1 + \beta_2), \qquad (207)$$

so rotations $R(\theta)$ around the z axis and the proper orthochronous Lorentz transformations $L(\alpha, \beta)$ respectively form a pair of commutative subgroups of the particle's little group. Furthermore, we have

$$R(\theta)L(\alpha, \beta)R^{-1}(\theta)$$
$$= L(\alpha \cos\theta + \beta \sin\theta, -\alpha \sin\theta + \beta \cos\theta), \qquad (208)$$

so we see that rotating $L(\alpha, \beta)$ itself around the z axis has the effect of rotating the two-dimensional vector (α, β).

The little group in this case is therefore the group $ISO(2)$ of translations and rotations in the two-dimensional Euclidean plane. The subgroup $SO(2)$ consisting purely of rotations $R(\theta)$ in the two-dimensional plane is compact, but the subgroup $\mathbb{R}^2$ consisting of two-dimensional translations $L(\alpha, \beta)$ is noncompact. The consequence is that the particle's phase space at the fixed reference four-momentum p_0^μ would seem to be noncompact as well, leading to the thermodynamic problems that we discussed earlier, as well as to various issues that arise in the corresponding quantum field theory, such as those that are explored in [17], for example.[14]

The particle's reference spacetime coordinates $X_0^\mu \equiv 0$, four-momentum $p_0^\mu \equiv (E_0/c, 0, 0, E_0/c)^\mu$, helicity $\sigma = \pm S_{0,z}$, and Pauli-Lubanski pseudovector $W_0^\mu = \sigma p_0^\mu$ are all insensitive to the noncompact transformations $L(\alpha, \beta)$. However, the particle's reference spin tensor (189) transforms nontrivially under the action of $L(\alpha, \beta)$:

$$L(\alpha, \beta)S_0 L^{\mathrm{T}}(\alpha, \beta)$$
$$= S_0 + \begin{pmatrix} 0 & -\beta S_{0,z} & \alpha S_{0,z} & 0 \\ \beta S_{0,z} & 0 & 0 & \beta S_{0,z} \\ \alpha S_{0,z} & 0 & 0 & -\alpha S_{0,z} \\ 0 & -\beta S_{0,z} & \alpha S_{0,z} & 0 \end{pmatrix}. \qquad (209)$$

In particular, at the level of the massless particle's reference spin three-vector $\mathbf{S}_0$, we have the transformation

$$\mathbf{S}_0 \mapsto \mathbf{S}_0 + (-\alpha S_{0,z}, -\beta S_{0,z}, 0). \qquad (210)$$

---

[14] For a more optimistic alternative perspective, see [18].

Notice that the discrepant spin components $(-\alpha S_{0,z}, -\beta S_{0,z}, 0)$ here are perpendicular to the particle's reference three-momentum $\mathbf{p}_0 = (0, 0, E_0/c)$, and, furthermore, parametrize *all* possible values of components perpendicular to $\mathbf{p}_0$, contingent on the z component of $\mathbf{S}_0$ being nonzero, $S_{0,z} \neq 0$. The fact that the discrepant spin components are perpendicular to $\mathbf{p}_0$ is ultimately guaranteed by the invariance of the massless particle's helicity $\sigma \equiv (\mathbf{p}/|\mathbf{p}|) \cdot \mathbf{S}$ under all proper Lorentz transformations, together with the invariance of the particle's reference momentum $\mathbf{p}_0$ under little-group transformations.

Moreover, suppose that two states of our massless particle both have four-momentum equal to the reference four-momentum $p_0^\mu = (E/c, 0, 0, E/c)$, but have different values of the spin tensor $S_0^{\mu\nu} \neq S'^{\mu\nu}$. The invariance of the helicity $\sigma \equiv (\mathbf{p}/|\mathbf{p}|) \cdot \mathbf{S}$ over the entire phase space, up to a possible minus sign under parity transformations, then implies that

$$S'_{0,z} = \pm S_{0,z}. \qquad (211)$$

Hence, all states with the reference momentum p_0^μ have spin-z component equal either to $S_{0,z}$ or to $-S_{0,z}$, where these two sets of states have opposite helicity σ. Fixing the helicity, such states can then differ at most in their spin-x and spin-y components.

The real-valued parameters α and β appearing in the little-group transformations (209) and (210) are arbitrary, and parametrize a noncompact set of states at fixed reference momentum p_0^μ. As a consequence, the only way to ensure that the massless particle's phase space at the fixed reference energy $E = p_0^t c$ is compact is to institute an equivalence relation in which we declare that any two states of the form (X_0, p_0, S_0) and (X_0, p_0, S') that have the same helicity σ and that differ solely in their spin components are to be regarded as the *same* physical state:

$$(X_0, p_0, S_0) \cong (X_0, p_0, S') \quad \text{[for fixed } \sigma]. \qquad (212)$$

This equivalence relation immediately generalizes to arbitrary states as

$$(X, p, S) \cong (X, p, S') \quad \text{[for fixed } \sigma], \qquad (213)$$

where the two states have the same spacetime coordinates X^μ, four-momentum p^μ, and helicity σ.

The equivalence relation (213) is a new result, and is another important example of a gauge invariance, distinct from the reparametrization invariance of a manifestly covariant action functional that we introduced earlier. A space with an equivalence relation is known as a quotient space, and so we see that the phase space of a massless $m = 0$ particle with nonzero spin $s^2 \neq 0$ is a quotient space under the gauge invariance (213).

All physical observables must therefore be gauge invariant, as is indeed the case for the particle's spacetime coordinates X^μ, its four-momentum p^μ, its helicity σ,

and its Pauli-Lubanski pseudovector $W^\mu = \sigma p^\mu$. By contrast, the equivalence relation (213), together with the invariance of the helicity σ, as defined in (193), implies that

$$\sigma \equiv \frac{\mathbf{p}}{|\mathbf{p}|} \cdot \mathbf{S} \cong \frac{\mathbf{p}}{|\mathbf{p}|} \cdot \mathbf{S}', \qquad (214)$$

so components of the particle's spin tensor $S^{\mu\nu}$ that are perpendicular to the particle's three-momentum $\mathbf{p}$—such as $S^{yz} = S_x$ and $S^{zx} = S_y$ if $\mathbf{p}$ points along the z direction—are not gauge invariant, and are consequently not physical observables.

Combining the equivalence relation (213) with the relationship $\sigma = \pm S_{0,z}$ from (195) and our condition (211), we see that a massless particle has just *two* physical spin states for each fixed value of the particle's three-momentum $\mathbf{p}$. These two physical spin states corresponding to the two parity-related helicities $\pm\sigma$.

Summarizing our results, we see that a transitive group action of the orthochronous Poincaré group with $m = 0$ and positive energy $E = p^t c > 0$ describes the phase space of a massless particle with null four-momentum p^μ, helicity σ, non-negative spin-squared $s^2 = \sigma^2 \geq 0$, and a null Pauli-Lubanski pseudovector $W^\mu = \sigma p^\mu$. The particle moves at the speed of light c along a null worldline in spacetime with null four-velocity $\dot{X}^\mu$, and the particle's spin tensor $S^{\mu\nu}$ is uniquely defined only up to gauge transformations $S^{\mu\nu} \mapsto S'^{\mu\nu}$ for which $S'^{\mu\nu}$ differs from $S^{\mu\nu}$ solely by components perpendicular to the particle's three-momentum $\mathbf{p}$. This gauge invariance implies that the massless particle has just two physical spin states for each fixed value of $\mathbf{p}$, corresponding to the two helicities $\pm\sigma$, which are related to each other by parity transformations.

As an aside, we note that in the counterpart quantum theory, spin components that are *perpendicular* to the particle's direction of motion correspond to linear polarizations that are *longitudinal*, meaning that they are *parallel* to the particle's direction of motion. Accordingly, spin components that are *parallel* to the particle's direction of motion correspond to *transverse* linear polarizations. So in the quantum version of this story, gauge-invariant observables are those that are insensitive to the particle's longitudinal linear polarizations.

Moving on to connections with classical field theory, recall our earlier analysis of the relationship between the classical-particle condition $p_\mu S^{\mu\nu} = 0$ from (145) and the Lorenz equation $\partial_\mu A^\mu = 0$ from (146) for a spin-1 field theory. In particular, we showed that the Lorenz equation $\partial_\mu A^\mu = 0$, applied to the case (147) of a monochromatic plane wave $A^\mu = \varepsilon^\mu \exp(ip \cdot x/\hbar)$ with polarization four-vector ε^μ, yielded the condition (148), $p_\mu \varepsilon^\mu = 0$, which was precisely analogous to $p_\mu S^{\mu\nu} = 0$, and played the same role of eliminating unphysical spin states.

Similarly, the classical-particle gauge invariance (213) is directly analogous to electromagnetic gauge invariance:

$$A_\mu \cong A_\mu + \partial_\mu f. \qquad (215)$$

Indeed, for the plane-wave case $A^\mu = \varepsilon^\mu \exp(ip \cdot x/\hbar)$, with $f = \alpha \exp(ip \cdot x/\hbar)$, the identification (215) yields the condition that

$$\varepsilon^\mu \cong \varepsilon^\mu + (i\alpha/\hbar)p^\mu, \qquad (216)$$

which is likewise responsible for eliminating unphysical spin states.

Both our classical-particle gauge invariance (213) and electromagnetic gauge invariance (215) have nontrivial implications for the allowed form of interactions between systems, as any such interactions must be insensitive to quantities that are not gauge invariant. For the case of electromagnetism, the gauge potential A_μ cannot directly appear in Lorentz-covariant field equations, but can only appear indirectly through the gauge-invariant Faraday tensor $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. In an analogous way, for our classical massless particle, the spin tensor $S^{\mu\nu}$ cannot directly appear in Lorentz-covariant interaction terms in equations of motion that couple the particle to other systems.

Interaction terms involving the particle's four-momentum p^μ or Pauli-Lubanski pseudovector $W^\mu = \sigma p^\mu$ would both be permitted, although they get weak for small momentum, corresponding in quantum mechanics to large distances. We therefore anticipate that massless particles with classically large total spin $s \gg \hbar$ cannot mediate long-range interactions, and, indeed, a quantum version of our classification of particle-types suggests that long-range interactions are mediated only by massless particles with total spin less than or equal to $2\hbar$.[15]

C. The Massless Limit

It is an enlightening exercise to re-examine the massless case $m = 0$ from the perspective of the massive case $m > 0$ in the limit $m \to 0$. Along the way, we will provide a deeper explanation for the emergence of gauge invariance, as well as derive a classical-particle version of the Higgs mechanism.

To start, notice that our original choice (156) of reference four-momentum in the massive case, $p_0^\mu \equiv (mc, \mathbf{0})^\mu$, does not have an appropriate massless limit. However, our choice of reference four-momentum is entirely arbitrary apart from the condition that $p^2 = -m^2 c^2$ from (104), so we can instead choose it to be

$$\bar{p}^\mu \equiv (\bar{p}^t, 0, 0, \bar{p}^z)^\mu$$
$$= (\sqrt{(\bar{p}^z)^2 + m^2 c^2}, 0, 0, \bar{p}^z)^\mu. \qquad (217)$$

The bars in this notation are solely meant to distinguish this choice of reference momentum from our original choice $p_0^\mu \equiv (mc, \mathbf{0})^\mu$ in (156).

---

[15] Again, for an alternative point of view, see [18].

The massless limit $m \to 0$ of this alternative reference four-momentum replicates the reference four-momentum (176) that we chose for the case of a massless particle:

$$\lim_{m \to 0} \bar{p}^\mu = (E_0/c, 0, 0, E_0/c)^\mu, \quad E_0 \equiv \bar{p}^z c. \qquad (218)$$

Moreover, the choice (217) is related to our original reference four-momentum (156),

$$p_0^\mu = (mc, \mathbf{0})^\mu,$$

by a simple Lorentz boost $\bar{\Lambda}$ along the z direction,

$$\bar{p}^\mu = \bar{\Lambda}^\mu{}_\nu p_0^\nu, \qquad (219)$$

where

$$\bar{\Lambda} \equiv \begin{pmatrix} \dfrac{\bar{p}^t}{mc} & 0 & 0 & \dfrac{\bar{p}^z}{mc} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \dfrac{\bar{p}^z}{mc} & 0 & 0 & \dfrac{\bar{p}^t}{mc} \end{pmatrix}. \qquad (220)$$

It follows that the new reference value $\bar{S}^{\mu\nu}$ of the massive particle's spin tensor is related to its old reference value $S_0^{\mu\nu}$ from (169) according to

$$\bar{S}^{\mu\nu} \equiv (\bar{\Lambda} S_0 \bar{\Lambda}^{\mathrm{T}})^{\mu\nu}$$

$$= \begin{pmatrix} 0 & \dfrac{\bar{p}^z}{mc} S_{0,y} & -\dfrac{\bar{p}^z}{mc} S_{0,x} & 0 \\ -\dfrac{\bar{p}^z}{mc} S_{0,y} & 0 & S_{0,z} & -\dfrac{\bar{p}^t}{mc} S_{0,y} \\ \dfrac{\bar{p}^z}{mc} S_{0,x} & -S_{0,z} & 0 & \dfrac{\bar{p}^t}{mc} S_{0,x} \\ 0 & \dfrac{\bar{p}^t}{mc} S_{0,y} & -\dfrac{\bar{p}^t}{mc} S_{0,x} & 0 \end{pmatrix}^{\mu\nu}. \qquad (221)$$

Both $\bar{p}^t$ and $\bar{p}^z$ approach the finite, nonzero value $E_0/c > 0$ in the massless limit $m \to 0$, so the components of $\bar{S}^{\mu\nu}$ that involve factors of $\bar{p}^t/mc$ or $\bar{p}^z/mc$ diverge in that limit. Furthermore, the particle's spin-squared scalar s^2 continues to have its invariant value (170), which, despite remaining well-defined in the limit $m \to 0$, does not end up agreeing with the corresponding massless particle's spin-squared scalar (192):

$$s^2 = S_{0,x}^2 + S_{0,y}^2 + S_{0,z}^2 \quad \text{(massive)}$$
$$\neq S_{0,z}^2 \quad \text{(massless)}. \qquad (222)$$

Meanwhile, the new reference value $\bar{W}^\mu$ of the particle's Pauli-Lubanski pseudovector is related to its old reference value $W_0^\mu \equiv (0, mc\,\mathbf{S}_0)^\mu$ from (171) according to

$$\bar{W}^\mu = \bar{\Lambda}^\mu{}_\nu W_0^\mu$$
$$= (\bar{p}^z\, S_{0,z}, mc\, S_{0,x}, mc\, S_{0,y}, \bar{p}^t\, S_{0,z})^\mu. \qquad (223)$$

This expression has a well-defined massless limit that precisely agrees with the reference value (197) of the Pauli-Lubanski pseudovector for a massless particle:

$$\lim_{m \to 0} \bar{W}^\mu = \left(S_{0,z} \frac{E_0}{c}, 0, 0, S_{0,z} \frac{E_0}{c} \right)^\mu. \qquad (224)$$

To make contact with the massless case, we can therefore focus our efforts on the spin tensor (221).

An important hint is the discrete discrepancy (222) between the spin-squared scalar s^2 in the massive and massless cases, signaling that the massive case features spin degrees of freedom that need to be removed before taking the massless limit. As we will see, removing these extraneous spin degrees of freedom will require formally enlarging our massive particle's phase space while simultaneously introducing a compensating equivalence relation to ensure that we are not adding any physically new states to the system. This approach corresponds to an analogous construction in quantum field theory whose origins go back to the work of Stueckelberg in [19].

At the conclusion of this procedure, we will be able to isolate and eliminate the extraneous spin degrees of freedom. Moreover, we will find that the equivalence relation that we introduced along the way becomes the gauge invariance (213) in the massless limit.

We begin by redefining the x and y components of the reference value $\bar{\mathbf{S}} = (\bar{S}_x, \bar{S}_y, \bar{S}_z)$ of the massive particle's spin three-vector according to

$$\begin{pmatrix} \bar{S}_x \\ \bar{S}_y \end{pmatrix} \mapsto \frac{mc}{\bar{p}^t} \begin{pmatrix} \bar{S}_x + \bar{p}^t \varphi_x \\ \bar{S}_y + \bar{p}^t \varphi_y \end{pmatrix} = \frac{mc}{\bar{p}^t} \begin{pmatrix} \bar{S}_x \\ \bar{S}_y \end{pmatrix} + mc \begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix}, \qquad (225)$$

where $\varphi_x(\lambda)$ and $\varphi_y(\lambda)$ are arbitrary new functions on the particle's worldline. The particle's spin tensor (221) then takes the form

$$\bar{S}^{\mu\nu} = \begin{pmatrix} 0 & \dfrac{\bar{p}^z}{\bar{p}^t} S_{0,y} & -\dfrac{\bar{p}^z}{\bar{p}^t} S_{0,x} & 0 \\ -\dfrac{\bar{p}^z}{\bar{p}^t} S_{0,y} & 0 & S_{0,z} & -S_{0,y} \\ \dfrac{\bar{p}^z}{\bar{p}^t} S_{0,x} & -S_{0,z} & 0 & S_{0,x} \\ 0 & S_{0,y} & -S_{0,x} & 0 \end{pmatrix}^{\mu\nu}$$
$$+ \begin{pmatrix} 0 & \bar{p}^z \varphi_y & -\bar{p}^z \varphi_x & 0 \\ -\bar{p}^z \varphi_y & 0 & 0 & -\bar{p}^t \varphi_y \\ \bar{p}^z \varphi_x & 0 & 0 & \bar{p}^t \varphi_x \\ 0 & \bar{p}^t \varphi_y & -\bar{p}^t \varphi_x & 0 \end{pmatrix}^{\mu\nu}, \qquad (226)$$

where we have chosen the various factors of m, c, $\bar{p}^t$, and $\bar{p}^z$ in the redefinition (225) to ensure that the two tensors appearing in (226) separately satisfy the fundamental condition $\bar{p}_\mu (\cdots)^{\mu\nu} = 0$ from (145). The particle's spin-squared scalar s^2, as originally defined in (106), now becomes

$$s^2 = \left(1 - \left(\frac{\bar{p}^z}{\bar{p}^t} \right)^2 \right) \left((S_{0,x} + \bar{p}^t \varphi_x)^2 \right.$$
$$\left. + (S_{0,y} + \bar{p}^t \varphi_y)^2 \right) + S_{0,z}^2. \qquad (227)$$

Notice that the particle's spin tensor (226) is invariant under the simultaneous transformations

$$\begin{pmatrix} \bar{S}_x \\ \bar{S}_y \end{pmatrix} \mapsto \begin{pmatrix} \bar{S}_x \\ \bar{S}_y \end{pmatrix} - \bar{p}^t \begin{pmatrix} f_x \\ f_y \end{pmatrix}, \qquad (228)$$

$$\begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix} \mapsto \begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix} + \begin{pmatrix} f_x \\ f_y \end{pmatrix}, \qquad (229)$$

where $f_x(\lambda), f_y(\lambda)$ are arbitrary functions on the particle's worldline. We claim that our massive particle's original phase space, with states denoted by (X, p, S), is equivalent to a *formally enlarged* phase space consisting of states

$$(X, p, S, \varphi) \tag{230}$$

under the equivalence relation

$$(\bar{X}, \bar{p}, \bar{S}, \varphi) \cong (\bar{X}, \bar{p}, \bar{S} - \bar{p}^t f, \varphi + f), \tag{231}$$

suitably generalized from the reference state $(\bar{X}, \bar{p}, \bar{S}, \varphi)$ to general states (X, p, S, φ) of the system. To see why, observe that the specific choice

$$f \equiv -\varphi, \tag{232}$$

or, more explicitly,

$$\begin{pmatrix} f_x \\ f_y \end{pmatrix} \equiv -\begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix} \tag{233}$$

makes clear that the state $(\bar{X}, \bar{p}, \bar{S}, \varphi)$ is equivalent to the state $(\bar{X}, \bar{p}, \bar{S} + \bar{p}^t \varphi, 0)$, which gives us back the state $(\bar{X}, \bar{p}, \bar{S})$ after undoing the redefinition (225) of $\bar{S}^{\mu\nu}$.

The system's redefined spin tensor (226) now has a nice massless limit,

$$\lim_{m \to 0} \bar{S}^{\mu\nu} = \begin{pmatrix} 0 & S_{0,y} & -S_{0,x} & 0 \\ -S_{0,y} & 0 & S_{0,z} & -S_{0,y} \\ S_{0,x} & -S_{0,z} & 0 & S_{0,x} \\ 0 & S_{0,y} & -S_{0,x} & 0 \end{pmatrix}^{\mu\nu}$$
$$+ \frac{E}{c} \begin{pmatrix} 0 & \varphi_y & -\varphi_x & 0 \\ -\varphi_y & 0 & 0 & -\varphi_y \\ \varphi_x & 0 & 0 & \varphi_x \\ 0 & \varphi_y & -\varphi_x & 0 \end{pmatrix}^{\mu\nu}, \tag{234}$$

as does the particle's spin-squared scalar (227),

$$\lim_{m \to 0} s^2 = S_{0,z}^2. \tag{235}$$

Our system fundamentally has the same number of degrees of freedom as it had before we took the massless limit, but we see that the degrees of freedom describing spin components perpendicular to the particle's reference three-momentum $\bar{\mathbf{p}}$ no longer contribute to the particle's spin-squared scalar s^2, which agrees with the spin-squared scalar (192) of the massless case. If we now formally remove the spin degrees of freedom φ_x, φ_y by setting them equal to zero, then the particle's spin tensor (234) reduces to the reference value of the massless spin tensor (189), and our equivalence relation (231) reduces to the gauge invariance (212).

Notice that if we run all the arguments of this section in reverse, then we can convert a massless particle with spin into a massive particle by introducing additional spin degrees of freedom. We therefore obtain a classical version of the celebrated Higgs mechanism.

To see the connection with the field-theoretic Higgs mechanism in more detail, recall the case of a massless spin-1 field, whose corresponding quantum-mechanical boson has two physical spin states. If we spontaneously break the gauge symmetry, then the spin-1 field gains a positive mass along with an additional spin state that is acquired from the Higgs field, so that the corresponding quantum-mechanical boson ends up with the correct three physical spin states for a massive spin-1 particle.

In a similar way, suppose that we start with a massless classical particle with nonzero spin. From our preceding work, we know that the particle has precisely two physical spin states, corresponding to the two helicities $\pm\sigma$. Reversing the logic of this section, we can convert our massless particle into a massive particle by augmenting the particle with spin states from the "Higgs" degrees of freedom φ_x, φ_y.

D. Tachyons

The case $m^2 < 0$ is also interesting. The invariant quantity m is now purely imaginary and is therefore of the form $m = i\mu$ for a real constant μ. The system's four-momentum p^μ is spacelike, $p^2 = \mu^2 c^2 > 0$, so its temporal component p^t does not have a definite sign under orthochronous Lorentz transformations. As a consequence, we cannot impose a positivity condition on the system's energy.

We can use $p^2 = \mu^2 c^2$ to express the system's energy $E = p^t c$ in terms of its three-dimensional momentum $\mathbf{p}$ as the mass-shell relation

$$E = \sqrt{\mathbf{p}^2 c^2 - \mu^2 c^2}. \tag{236}$$

For convenience, we will take the system's reference four-momentum to be purely spacelike and aligned with the z direction:

$$p_0^\mu \equiv (0, 0, 0, \mu c)^\mu = \mu c \, \delta_z^\mu. \tag{237}$$

Once again, the four-momentum p^μ and the four-velocity $\dot{X}^\mu$ are non-vanishing, and so the relation (154),

$$m\sqrt{-\dot{X}^2/c^2}\, p^\mu = \mp m^2 \dot{X}^\mu,$$

becomes

$$\sqrt{-\dot{X}^2/c^2}\, p^\mu = \mp i\mu \dot{X}^\mu. \tag{238}$$

Because the right-hand side is imaginary, this equality implies that $\dot{X}^2 > 0$, so the four-velocity $\dot{X}^\mu$ is likewise spacelike and is related to the four-momentum p^μ by

$$p^\mu = \mu \frac{\dot{X}^\mu}{\sqrt{\dot{X}^2/c^2}}, \tag{239}$$

where we have taken the positive sign by assuming that our parametrization $X^\mu(\lambda)$ points in the positive direction along p^μ. This relation between p^μ and $\dot{X}^\mu$

again ensures that the self-consistency condition (149), $\dot{X}^\mu W_\mu = 0$, is satisfied.

The equation of motion (135) for the system's four-momentum, $\dot{p}^\mu = 0$, then tells us that the system's path has a fixed, spacelike direction in spacetime. A calculation of the system's three-dimensional velocity $\mathbf{v}$ using the mass-shell relation (236) yields the result

$$\mathbf{v} = \frac{d\mathbf{X}}{dt} = \frac{\dot{\mathbf{X}}}{\dot{T}} = \frac{\mathbf{p}c^2}{E} = \frac{\mathbf{p}}{|\mathbf{p}|}\frac{c}{\sqrt{1 - \mu^2 c^2/\mathbf{p}^2}}. \qquad (240)$$

Hence, the system's speed $|\mathbf{v}|$ is always *greater* than the speed of light c:

$$|\mathbf{v}| > c. \qquad (241)$$

Such a system is appropriately called a tachyon, from the Greek for "swift."

By the same reasoning as in the massive and massless cases, a tachyon's orbital and spin angular momenta are separately conserved,

$$\dot{L}^{\mu\nu} = 0, \qquad (242)$$

$$\dot{S}^{\mu\nu} = 0. \qquad (243)$$

The condition (144), $p_{0,\mu}S_0^{\mu\nu} = 0$, now gives

$$\mu c\, S_0^{z\nu} = 0, \qquad (244)$$

so the reference value of the system's spin tensor is

$$S_0^{\mu\nu} = \begin{pmatrix} 0 & \tilde{S}_{0,x} & \tilde{S}_{0,y} & 0 \\ -\tilde{S}_{0,x} & 0 & S_{0,z} & 0 \\ -\tilde{S}_{0,y} & -S_{0,z} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}^{\mu\nu}. \qquad (245)$$

The system's spin-squared scalar (107) and spin-squared pseudoscalar (108) have respective values

$$s^2 = S_{0,z}^2 - \tilde{S}_{0,x}^2 - \tilde{S}_{0,y}^2, \qquad (246)$$

$$\tilde{s}^2 = 0, \qquad (247)$$

and the reference value of the system's Pauli-Lubanski pseudovector (99) is

$$W_0^\mu = \mu c(S_{0,z}, \tilde{S}_{0,y}, -\tilde{S}_{0,x}, 0)^\mu. \qquad (248)$$

We next consider the little group of orthochronous Lorentz transformations that preserve the value of the reference four-momentum (237), $p_0^\mu \equiv (0,0,0,\mu c)^\mu$. As usual, these little-group transformations parametrize the set of all states that share that same four-momentum, so they include rotations around the z axis as well as Lorentz boosts along the x and y directions.

If the system is to have a compact set of states at any fixed four-momentum, then its spin tensor (245) and Pauli-Lubanski pseudovector (248) must be invariant under these noncompact Lorentz transformations. However, we see right away that W_0^μ transforms nontrivially

under Lorentz transformations along the x or y directions if any of its components are nonzero, so our system's phase space at fixed four-momentum can be compact only if all the components of W_0^μ vanish:

$$\left.\begin{array}{c} S_{0,z} = 0, \\ \tilde{S}_{0,x} = 0, \\ \tilde{S}_{0,y} = 0. \end{array}\right\} \qquad (249)$$

The tachyon's spin tensor and Pauli-Lubanski pseudovector therefore vanish identically,

$$\left.\begin{array}{c} S^{\mu\nu} = 0, \\ W^\mu = 0, \end{array}\right\} \qquad (250)$$

so

$$\left.\begin{array}{c} s^2 = 0, \\ W^2 \equiv w^2 = 0, \end{array}\right\} \qquad (251)$$

and we see that a tachyon cannot have any intrinsic spin at all.

E. The Vacuum

Finally, we consider the case in which $p_0^\mu = 0$, meaning that the system's four-momentum vanishes for all the system's possible states:

$$p^\mu = 0. \qquad (252)$$

The system then has no energy or momentum. The kinetic term $p_\mu \dot{X}^\mu$ in the system's action functional (132) vanishes, and we do not get a meaningful equation describing the behavior of $X^\mu(\lambda)$. The system's orbital angular momentum vanishes,

$$L^{\mu\nu} = 0, \qquad (253)$$

and its spin angular momentum is conserved:

$$\dot{S}^{\mu\nu} = 0. \qquad (254)$$

The little group of orthochronous Lorentz transformations that leave $p_0^\mu = 0$ invariant consists of *all* orthochronous Lorentz transformations, and so the only way to obtain a compact phase space at fixed four-momentum is for the spin tensor to vanish for all the system's states:

$$S^{\mu\nu} = 0. \qquad (255)$$

We conclude that our system is entirely devoid of energy, momentum, and angular momentum, and therefore describes an empty vacuum.

VI. CONCLUSION

In this paper, we reviewed a general method for making the standard Lagrangian formulation manifestly covariant. We employed this framework to develop a classical counterpart of Wigner's classification of quantum particle-types in terms of the structure of the orthochronous Poincaré group. We also showed that classical massless particles with spin exhibit a novel manifestation of gauge invariance, and used the massless limit to derive a classical version of the Higgs mechanism.

An interesting way to extend our approach is to consider phase spaces that provide transitive group actions of the *full* Poincaré group, including time-reversal transformations (55). This generalization does not affect our analysis of tachyons or of the vacuum, which do not feature a definite sign for p^t. But in the case of a system with non-negative mass, $m \geq 0$, enlarging the system's phase space so that it provides a transitive action of the full Poincaré group means doubling the phase space to include "negative-energy" states with $p^t < 0$. Because

the four-momentum p^μ is timelike or null when $m \geq 0$, we know from (56) that the sign of p^t is invariant under all *physically realizable* Lorentz transformations, which are smoothly connected with the identity transformation and therefore do not include time-reversal transformations. Hence, a system with $m \geq 0$ cannot evolve from states with $p^t > 0$ to states with $p^t < 0$ or vice versa. We are therefore free to define the physical energy of the additional $p^t < 0$ states to be $E \equiv -p^t c > 0$, and regard them as states not of our original particle, but of its corresponding antiparticle. In this way, we can classically unify particles with their antiparticles.

ACKNOWLEDGMENTS

[1] J. A. Barandes. "Gauge Invariance for Classical Massless Particles with Spin". *Foundations of Physics*, 51(1):7, February 2021. arXiv:1911.02515, doi:10.1007/s10701-021-00415-2.

[2] E. C. G. Sudarshan and N. Mukunda. *Classical Dynamics: A Modern Perspective*. World Scientific, 1974.

[3] A. P. Balachandran, G. Marmo, B.-S. Skagerstam, and A. Stern. *Gauge Symmetries and Fibre Bundles - Applications to Particle Dynamics*. Springer-Verlag Berlin Heidelberg, 1st edition, 1983. arXiv:1702.08910, doi:10.1007/3-540-12724-0.

[4] J.-M. Souriau. *Structure of Dynamical Systems*. Birkhäuser, 1st edition, 1997.

[5] M. Rivas. *Kinematical Theory of Spinning Particles*. Springer, 2002.

[6] E. P. Wigner. "On Unitary Representations of the Inhomogeneous Lorentz Group". *Annals of Mathematics*, 40(1):149–204, 1939. doi:10.2307/1968551.

[7] H. Goldstein, J. L. Safko, and C. P. Poole Jr. *Classical Mechanics*. Pearson, 3rd edition, 2001.

[8] P. A. M. Dirac. "Relativity Quantum Mechanics with an Application to Compton Scattering". *Proceedings of the Royal Society A*, 111(758):405–423, 1926. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1926.0074, doi:10.1098/rspa.1926.0074.

[9] P. A. M. Dirac. "Quantization on Curved Surfaces", in *Lectures on Quantum Mechanics*, pages 44–66. Belfer Graduate School of Science, 1964.

[10] A. Deriglazov and B. Rizzuti. "Reparametrization-invariant formulation of classical mechanics and the Schrödinger equation". *American Journal of Physics*, 79(8):882–885, 2011. URL: https://aapt.scitation.org/doi/10.1119/1.3593270, arXiv:1105.0313, doi:10.1119/1.3593270.

[11] B. Schutz. *A First Course in General Relativity*. Cambridge University Press, 2nd edition, 2009.

[12] S. Weinberg. *The Quantum Theory of Fields*, Volume 1. Cambridge University Press, 1996.

[13] A. J. Hanson and T. Regge. "The relativistic spherical top". *Annals of Physics*, 87(2):498–566, 1974. URL: https://www.sciencedirect.com/science/article/pii/0003491674900463, doi:10.1016/0003-4916(74)90046-3.

[14] B.-S. Skagerstam and A. Stern. "Lagrangian Descriptions of Classical Charged Particles with Spin". *Physica Scripta*, 24:493–497, 1981. URL: https://iopscience.iop.org/article/10.1088/0031-8949/24/3/002, doi:10.1088/0031-8949/24/3/002.

[15] A. Frydryszak. "Lagrangian Models of Particles with Spin: The First Seventy Years", in *From Field Theory to Quantum Groups*, pages 151–172. World Scientific, 1996. URL: https://www.worldscientific.com/doi/abs/10.1142/9789812830425_0009, arXiv:hep-th/9601020, doi:10.1142/9789812830425_0009.

[16] E. P. Wigner. "Invariant Quantum Mechanical Equations of Motion", in *Theoretical Physics*, pages 161–184. International Atomic Energy Agency, Vienna Austria, 1963. URL: https://link.springer.com/chapter/10.1007/978-3-662-09203-3_18, doi:10.1007/978-3-662-09203-3_18.

[17] L. F. Abbott. "Massless particles with continuous spin indices". *Physical Review D*, 13(8):2291–2294, April 1976. URL: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.13.2291, doi:10.1103/PhysRevD.13.2291.

[18] P. Schuster and N. Toro. "On the theory of continuous-spin particles: wavefunctions and soft-factor scattering amplitudes". *Journal of High Energy Physics*, 2013(104), 2013. URL: https://link.springer.com/article/10.1007/JHEP09(2013)104, arXiv:1302.1198, doi:10.1007/JHEP09(2013)104.

[19] E. C. G. Stueckelberg. "Die Wechselwirkungskräfte in der Elektrodynamik und in der Feldtheorie der Kernkraäfte". *Helvetica Physica Acta*, 11(3):225–244, 299–328, 1938. URL: https://www.e-periodica.ch/digbib/view?pid=hpa-001:1938:11::636#227, doi:10.5169/seals-110852.