9 Poincaré and Lorentz Symmetries

In this chapter, we examine the symmetry structure underlying relativistic quantum field theory. The invariance of the Lagrangian under the Poincaré group, which comprises Lorentz transformations and spacetime translations, is the foundation of all relativistic field theories. For a comprehensive treatment of group theory in physics, see Tung [29]; a detailed discussion of the Lorentz group and relativistic symmetry can be found in Sexl and Urbantke [30].

Roadmap. This chapter has two connected, but logically distinct, goals. First we study the Lorentz and Poincaré groups as spacetime symmetry groups, derive their Lie algebras, and identify the Casimir operators that classify one-particle states by mass and spin. Then we switch from states to fields: we construct the finite-dimensional Lorentz representations carried by spinor fields, introduce the sigma-matrix technology that links Weyl spinors to four-vectors, and finally translate the two-component language into the four-component Dirac and Majorana formalisms. The main thread is therefore

$$\begin{aligned}&\text{Lorentz algebra} \longrightarrow \text{Poincaré Casimirs}\\ &\longrightarrow \text{Weyl representations} \longrightarrow \text{covariant fermion Lagrangians.}\end{aligned}$$

9.1 The Lorentz group and its Lie algebra

Definition and fundamental properties

Under an **active** Lorentz transformation $\Lambda^\mu{}_\nu$, the spacetime coordinates x^μ transform as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu\,, \tag{9.1}$$

where there is an implicit sum over the repeated index ν. One can regard $\Lambda^\mu{}_\nu$ as the matrix element located in row μ and column ν of the 4×4 matrix Λ. The condition that the spacetime interval $g_{\mu\nu}x^\mu x^\nu$ is invariant under Lorentz transformations implies that

$$\boxed{\Lambda^\mu{}_\rho\, g_{\mu\nu}\, \Lambda^\nu{}_\sigma = g_{\rho\sigma}}\,. \tag{9.2}$$

In matrix notation, this reads $\Lambda^\top g \Lambda = g$, where $g = \mathrm{diag}(+1,-1,-1,-1)$ is the Minkowski metric.

The set of all Λ that satisfy eq. (9.2) forms a Lie group $O(1,3)$. Equation (9.2) implies that Λ has the following two properties:

Proposition 9.1 (Properties of Lorentz transformations). For any Λ satisfying (9.2):

(i) $\det \Lambda = \pm 1$
(ii) $|\Lambda^0{}_0| \geq 1$

Proof. (i) Taking the determinant of both sides of (9.2):

$$\det\big(\Lambda^\top g \Lambda\big) = \det(g) \quad \Longrightarrow \quad (\det \Lambda)^2 \det(g) = \det(g)$$

Since $\det(g) = -1 \neq 0$, we can divide to obtain $(\det \Lambda)^2 = 1$, hence $\det \Lambda = \pm 1$.

(ii) The $(0,0)$ component of (9.2) with $\rho = \sigma = 0$ gives:

$$\Lambda^\mu{}_0\, g_{\mu\nu}\, \Lambda^\nu{}_0 = g_{00} = 1$$

Expanding:

$$(\Lambda^0{}_0)^2 g_{00} + \sum_{i=1}^{3} (\Lambda^i{}_0)^2 g_{ii} = (\Lambda^0{}_0)^2 - \sum_{i=1}^{3} (\Lambda^i{}_0)^2 = 1$$

Therefore $(\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\Lambda^i{}_0)^2 \geq 1$, which implies $|\Lambda^0{}_0| \geq 1$. □

These two properties divide the Lorentz group into four disconnected classes, characterized by a pair of signs $(\mathrm{sgn}(\det \Lambda), \mathrm{sgn}(\Lambda^0{}_0))$:

Class	$\det \Lambda$	$\Lambda^0{}_0$	Description
$(+,+)$	$+1$	$\geq +1$	Proper orthochronous (connected to $\mathbb{1}$)
$(+,-)$	$+1$	≤ -1	Proper non-orthochronous
$(-,+)$	-1	$\geq +1$	Improper orthochronous
$(-,-)$	-1	≤ -1	Improper non-orthochronous

The **proper orthochronous** Lorentz transformations $(+,+)$ are continuously connected to the identity and form a subgroup denoted $SO^+(1,3)$. If $\Lambda \in SO^+(1,3)$, we can generate all elements of the other classes by introducing:

- Space inversion (parity): $\Lambda_P = \mathrm{diag}(1,-1,-1,-1)$
- Time inversion: $\Lambda_T = \mathrm{diag}(-1,1,1,1)$
- Spacetime inversion: $\Lambda_P \Lambda_T = -\mathbb{1}_4$

Then $\{\Lambda, \Lambda_P\Lambda, \Lambda_T\Lambda, \Lambda_P\Lambda_T\Lambda\}$ spans the full Lorentz group.

Remark. [Physical significance] Infinitesimal Lorentz transformations must be proper and orthochronous, since only these are continuously connected to the identity. Space inversion and time inversion are *discrete* symmetries that may or may not be respected by physical interactions (famously, parity is violated in weak interactions).

Infinitesimal Lorentz transformations

The most general proper orthochronous Lorentz transformation matrix Λ is characterized by:

- A rotation angle θ about an axis $\hat{\boldsymbol{n}}$, encoded in $\vec{\theta} \equiv \theta\hat{\boldsymbol{n}}$
- A boost vector $\vec{\zeta} \equiv \hat{\boldsymbol{\beta}} \tanh^{-1}\beta$, where $\hat{\boldsymbol{\beta}}$ is the unit vector along the direction of motion and $\beta = |\vec{v}|/c$

The 4×4 matrix Λ is given by

$$\Lambda = \exp\left(-\tfrac{i}{2}\theta_{\rho\sigma}s^{\rho\sigma}\right) = \exp\left(-i\vec{\theta}\cdot\vec{s} - i\vec{\zeta}\cdot\vec{k}\right), \tag{9.3}$$

where $\theta_{\rho\sigma} = -\theta_{\sigma\rho}$, and the $s^{\rho\sigma} = -s^{\sigma\rho}$ are six independent 4×4 antisymmetric matrices that satisfy the commutation relations of the Lie algebra $\mathfrak{so}(1,3)$:

$$[s^{\alpha\beta}, s^{\rho\sigma}] = i\left(g^{\beta\rho}s^{\alpha\sigma} - g^{\alpha\rho}s^{\beta\sigma} - g^{\beta\sigma}s^{\alpha\rho} + g^{\alpha\sigma}s^{\beta\rho}\right). \tag{9.4}$$

The matrix elements of $s^{\rho\sigma}$ are given explicitly by:

$$(s^{\rho\sigma})^{\mu}{}_{\nu} = i\left(\delta^{\sigma}_{\nu}g^{\mu\rho} - \delta^{\rho}_{\nu}g^{\mu\sigma}\right). \tag{9.5}$$

Derivation of the commutation relations. We compute $[s^{\alpha\beta}, s^{\rho\sigma}]$ using the explicit form (9.5). The matrix product gives:

$$\begin{aligned}(s^{\alpha\beta}s^{\rho\sigma})^{\mu}{}_{\lambda} &= (s^{\alpha\beta})^{\mu}{}_{\nu}(s^{\rho\sigma})^{\nu}{}_{\lambda} \\ &= i(\delta^{\beta}_{\nu}g^{\mu\alpha} - \delta^{\alpha}_{\nu}g^{\mu\beta})\cdot i(\delta^{\sigma}_{\lambda}g^{\nu\rho} - \delta^{\rho}_{\lambda}g^{\nu\sigma}) \\ &= -(\delta^{\beta}_{\nu}g^{\mu\alpha} - \delta^{\alpha}_{\nu}g^{\mu\beta}) \\ &\qquad \times(\delta^{\sigma}_{\lambda}g^{\nu\rho} - \delta^{\rho}_{\lambda}g^{\nu\sigma})\end{aligned} \tag{9.6}$$

Expanding and using $\delta^{\beta}_{\nu}g^{\nu\rho} = g^{\beta\rho}$:

$$\begin{aligned}(s^{\alpha\beta}s^{\rho\sigma})^{\mu}{}_{\lambda} = -(&g^{\beta\rho}g^{\mu\alpha}\delta^{\sigma}_{\lambda} - g^{\beta\sigma}g^{\mu\alpha}\delta^{\rho}_{\lambda} \\ &- g^{\alpha\rho}g^{\mu\beta}\delta^{\sigma}_{\lambda} + g^{\alpha\sigma}g^{\mu\beta}\delta^{\rho}_{\lambda})\end{aligned} \tag{9.7}$$

Computing the commutator by antisymmetrizing in $(\alpha\beta) \leftrightarrow (\rho\sigma)$:

$$
\begin{aligned}
[s^{\alpha\beta}, s^{\rho\sigma}]^{\mu}{}_{\lambda} = -g^{\beta\rho}&(g^{\mu\alpha}\delta^{\sigma}_{\lambda} - g^{\mu\sigma}\delta^{\alpha}_{\lambda}) \\
+ g^{\alpha\rho}&(g^{\mu\beta}\delta^{\sigma}_{\lambda} - g^{\mu\sigma}\delta^{\beta}_{\lambda}) \\
+ g^{\beta\sigma}&(g^{\mu\alpha}\delta^{\rho}_{\lambda} - g^{\mu\rho}\delta^{\alpha}_{\lambda}) \\
- g^{\alpha\sigma}&(g^{\mu\beta}\delta^{\rho}_{\lambda} - g^{\mu\rho}\delta^{\beta}_{\lambda})
\end{aligned}
\tag{9.8}
$$

Recognizing the structure of (9.5) in each term:

$$
[s^{\alpha\beta}, s^{\rho\sigma}] = i(g^{\beta\rho}s^{\alpha\sigma} - g^{\alpha\rho}s^{\beta\sigma} - g^{\beta\sigma}s^{\alpha\rho} + g^{\alpha\sigma}s^{\beta\rho})
$$

□

In (9.3), we have defined:

$$
\begin{aligned}
\theta^{i} &\equiv \tfrac{1}{2}\epsilon^{ijk}\theta^{jk}, \quad \zeta^{i} \equiv \theta^{i0} = -\theta^{0i}, \\
s^{i} &\equiv \tfrac{1}{2}\epsilon^{ijk}s^{jk}, \quad k^{i} \equiv s^{0i} = -s^{i0},
\end{aligned}
\tag{9.9}
$$

where indices $i, j, k \in \{1, 2, 3\}$ and $\epsilon^{123} = +1$. The s^{i} generate infinitesimal three-dimensional rotations in space, and the k^{i} generate infinitesimal Lorentz boosts.

An infinitesimal orthochronous Lorentz transformation is given by:

$$
\begin{aligned}
\Lambda^{\mu}{}_{\nu} &\simeq \delta^{\mu}_{\nu} - \tfrac{i}{2}\theta_{\rho\sigma}(s^{\rho\sigma})^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} + \theta^{\mu}{}_{\nu} \\
&\simeq \left(\mathbb{1}_4 - i\vec{\theta}\cdot\vec{s} - i\vec{\zeta}\cdot\vec{k}\right)^{\mu}{}_{\nu}.
\end{aligned}
\tag{9.10}
$$

Verification of infinitesimal form. Using (9.5):

$$
\begin{aligned}
-\frac{i}{2}\theta_{\rho\sigma}(s^{\rho\sigma})^{\mu}{}_{\nu} &= -\frac{i}{2}\theta_{\rho\sigma}\cdot i(\delta^{\sigma}_{\nu}g^{\mu\rho} - \delta^{\rho}_{\nu}g^{\mu\sigma}) \\
&= \frac{1}{2}\theta_{\rho\sigma}(\delta^{\sigma}_{\nu}g^{\mu\rho} - \delta^{\rho}_{\nu}g^{\mu\sigma}) \\
&= \frac{1}{2}(\theta_{\rho\nu}g^{\mu\rho} - \theta_{\sigma\nu}g^{\mu\sigma}) = \frac{1}{2}(\theta^{\mu}{}_{\nu} + \theta^{\mu}{}_{\nu}) = \theta^{\mu}{}_{\nu}
\end{aligned}
\tag{9.11}
$$

where we used the antisymmetry $\theta_{\sigma\nu} = -\theta_{\nu\sigma}$ and $\theta^{\mu}{}_{\nu} = g^{\mu\rho}\theta_{\rho\nu} = -\theta_{\nu}{}^{\mu}$. □

Physical generators: rotations and boosts

For a field of spin s, the general transformation law under a proper orthochronous Lorentz transformation reads:

$$
\psi'_{\alpha}(x') = \exp\left(-\tfrac{i}{2}\theta_{\mu\nu}S^{\mu\nu}\right)_{\alpha}{}^{\beta}\psi_{\beta}(x),
\tag{9.12}
$$

where the $S^{\mu\nu}$ are (finite-dimensional) irreducible matrix representations of the Lorentz algebra satisfying (9.4), and α, β label the components of the representation space.

In analogy with s^i and k^i defined in (9.9), we identify:

$$\boxed{S^i \equiv \frac{1}{2}\epsilon^{ijk}S^{jk}, \qquad K^i \equiv S^{0i}}, \tag{9.13}$$

where $i, j, k \in \{1,2,3\}$. The inverse relation reads $S^{jk} = \epsilon^{jki}S^i$. Since $S^{\mu\nu}$ is antisymmetric, it can be arranged as a 4×4 *matrix of matrices* (each entry $S^{\mu\nu}$ being itself a matrix acting on the representation space):

$$\left(S^{\mu\nu}\right) = \begin{pmatrix} 0 & K^1 & K^2 & K^3 \\ -K^1 & 0 & S^3 & -S^2 \\ -K^2 & -S^3 & 0 & S^1 \\ -K^3 & S^2 & -S^1 & 0 \end{pmatrix}, \tag{9.14}$$

where the row/column index runs over $\mu, \nu = 0, 1, 2, 3$. Thus the boosts K^i occupy the time–space entries (mixing x^0 with x^i), while the rotations S^i occupy the purely spatial 3×3 block.

Using (9.4), it follows that S^i and K^i satisfy:

$$\boxed{\begin{aligned} [S^i, S^j] &= i\epsilon^{ijk}S^k \\ [S^i, K^j] &= i\epsilon^{ijk}K^k \\ [K^i, K^j] &= -i\epsilon^{ijk}S^k \end{aligned}}. \tag{9.15}$$

Derivation of rotation-boost commutators. **For** $[S^i, S^j]$**:** Using $S^i = \frac{1}{2}\epsilon^{imn}S^{mn}$:

$$[S^i, S^j] = \frac{1}{4}\epsilon^{imn}\epsilon^{jpq}[S^{mn}, S^{pq}] \tag{9.16}$$

From (9.4) with only spatial indices (where $g^{mn} = -\delta^{mn}$):

$$[S^{mn}, S^{pq}] = i(-\delta^{np}S^{mq} + \delta^{mp}S^{nq} + \delta^{nq}S^{mp} - \delta^{mq}S^{np})$$

Using the identity $\epsilon^{imn}\epsilon^{jpq}\delta^{np} = \delta^{ij}\delta^{mq} - \delta^{iq}\delta^{mj}$ and careful index manipulation yields $[S^i, S^j] = i\epsilon^{ijk}S^k$.

For $[S^i, K^j]$**:** Using $S^i = \frac{1}{2}\epsilon^{imn}S^{mn}$ and $K^j = S^{0j}$:

$$[S^i, K^j] = \frac{1}{2}\epsilon^{imn}[S^{mn}, S^{0j}] \tag{9.17}$$

From (9.4):

$$[S^{mn}, S^{0j}] = i(g^{n0}S^{mj} - g^{m0}S^{nj} - g^{nj}S^{m0} + g^{mj}S^{n0})$$

Since $g^{n0} = g^{m0} = 0$ for spatial indices m, n:

$$[S^{mn}, S^{0j}] = i(\delta^{nj}S^{m0} - \delta^{mj}S^{n0}) = i(-\delta^{nj}K^m + \delta^{mj}K^n)$$

Therefore:

$$\begin{aligned}[S^i, K^j] &= \frac{i}{2}\epsilon^{imn}(-\delta^{nj}K^m + \delta^{mj}K^n) \\ &= \frac{i}{2}(-\epsilon^{imj}K^m + \epsilon^{ijn}K^n) = i\epsilon^{ijk}K^k\end{aligned}$$

For $[K^i, K^j]$: Using $K^i = S^{0i}$ and $K^j = S^{0j}$:

$$[K^i, K^j] = [S^{0i}, S^{0j}] = i(g^{i0}S^{0j} - g^{00}S^{ij} - g^{ij}S^{00} + g^{0j}S^{0i}) \tag{9.18}$$

Since $g^{i0} = g^{0j} = 0$, $g^{00} = 1$, and $S^{00} = 0$:

$$[K^i, K^j] = -iS^{ij} = -i\epsilon^{ijk}S^k$$

The crucial minus sign reflects the Lorentzian signature of spacetime. □

Remark. [Physical interpretation] The commutation relations (9.15) encode fundamental physics:

- $[S^i, S^j] = i\epsilon^{ijk}S^k$: This is the familiar $\mathfrak{su}(2)$ algebra of angular momentum.
- $[S^i, K^j] = i\epsilon^{ijk}K^k$: The boosts K^i transform as a 3-vector under rotations.
- $[K^i, K^j] = -i\epsilon^{ijk}S^k$: Two boosts in different directions do not commute—their commutator yields a rotation. This is the origin of *Thomas precession*.

The $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ structure (on the complexification)

Box 9.1 A subtlety: real vs. complex Lie algebra

The decomposition we are about to derive is a statement about the *complexified* Lorentz algebra, not about the real one. As a *real* Lie algebra, the Lorentz algebra $\mathfrak{so}(1,3)$ is *not* isomorphic to $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ (the latter being the compact algebra of $SO(4)$, with different signature). The correct real isomorphism is instead

$$\mathfrak{so}(1,3)_{\mathbb{R}} \simeq \mathfrak{sl}(2,\mathbb{C})_{\mathbb{R}},$$

viewed as a six-dimensional real algebra. Only after **complexification**—i.e. allowing complex linear combinations of the generators—does one obtain

$$\mathfrak{so}(1,3)_{\mathbb{C}} \;\simeq\; \mathfrak{sl}(2,\mathbb{C})\oplus\mathfrak{sl}(2,\mathbb{C}) \;\simeq\; \mathfrak{su}(2)_{\mathbb{C}}\oplus\mathfrak{su}(2)_{\mathbb{C}}\,.$$

The reason is visible in the very definition (9.19) below: the combinations $\vec{S}_\pm = \frac{1}{2}(\vec{S}\pm i\vec{K})$ require the imaginary unit i as a coefficient, so they do *not* belong to the real span of $\{S^i, K^i\}$. Equivalently, the boost generators K^i are anti-Hermitian (non-compact directions), whereas in $\mathfrak{su}(2)\oplus\mathfrak{su}(2)$ all generators would be Hermitian. Working with $\vec{S}_\pm$ is nevertheless extremely convenient, because finite-dimensional irreducible representations of a real Lie algebra are in one-to-one correspondence with those of its complexification.

It is convenient to define the following *complex* linear combinations of the generators:

$$\boxed{\vec{S}_+ \equiv \frac{1}{2}(\vec{S}+i\vec{K}), \qquad \vec{S}_- \equiv \frac{1}{2}(\vec{S}-i\vec{K})}\,, \tag{9.19}$$

with inverse relations $\vec{S} = \vec{S}_+ + \vec{S}_-$ and $\vec{K} = -i(\vec{S}_+ - \vec{S}_-)$. Note that $\vec{S}_\pm$ live in $\mathfrak{so}(1,3)_{\mathbb{C}}$, not in the real algebra.

Then, equations (9.15) decouple and yield two independent $\mathfrak{su}(2)$ Lie algebras (in the complexified sense discussed above):

$$\boxed{\begin{aligned} &[S_+^i, S_+^j] = i\epsilon^{ijk} S_+^k \\ &[S_-^i, S_-^j] = i\epsilon^{ijk} S_-^k \\ &[S_+^i, S_-^j] = 0 \end{aligned}}\,. \tag{9.20}$$

Derivation of decoupled algebra. We compute:

$$\begin{aligned} 4[S_+^i, S_+^j] &= [S^i + iK^i, S^j + iK^j] \\ &= [S^i, S^j] + i[S^i, K^j] + i[K^i, S^j] - [K^i, K^j] \\ &= i\epsilon^{ijk}S^k + i\cdot i\epsilon^{ijk}K^k + i\cdot i\epsilon^{ijk}K^k - (-i\epsilon^{ijk}S^k) \\ &= 2i\epsilon^{ijk}S^k - 2\epsilon^{ijk}K^k = 2i\epsilon^{ijk}(S^k + iK^k) = 4i\epsilon^{ijk}S_+^k \end{aligned} \tag{9.21}$$

Similarly, $4[S_-^i, S_-^j] = [S^i - iK^i, S^j - iK^j] = 4i\epsilon^{ijk}S_-^k$.

For the mixed commutator:

$$\begin{aligned} 4[S_+^i, S_-^j] &= [S^i + iK^i, S^j - iK^j] \\ &= [S^i, S^j] - i[S^i, K^j] + i[K^i, S^j] + [K^i, K^j] \\ &= i\epsilon^{ijk}S^k + \epsilon^{ijk}K^k - \epsilon^{ijk}K^k - i\epsilon^{ijk}S^k = 0 \end{aligned} \tag{9.22}$$

□

The finite-dimensional irreducible representations of $\mathfrak{su}(2)$ are well known: these are the $(2s+1)\times(2s+1)$ representation matrices corresponding to spin s, where $s = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. Hence, the irreducible representations of the Lorentz group can be characterized by two numbers (s_+, s_-), where $s_\pm$ is nonnegative and either an integer or a half-integer. The eigenvalues of $\vec{S}_\pm^2$ are given by $s_\pm(s_\pm + 1)$. The dimension of the representation (s_+, s_-) is $(2s_+ + 1)(2s_- + 1)$.

Remark. [Common representations]

(s_+, s_-)	Dimension	Name	Examples
$(0,0)$	1	Scalar	Higgs boson
$(\frac{1}{2},0)$	2	Left-handed Weyl spinor	ν_L, e_L
$(0,\frac{1}{2})$	2	Right-handed Weyl spinor	e_R
$(\frac{1}{2},\frac{1}{2})$	4	Four-vector	Photon A^μ
$(1,0)\oplus(0,1)$	6	Antisymmetric tensor	$F^{\mu\nu}$

9.2 The Poincaré group and its Lie algebra

Definition of the Poincaré group

The **Poincaré group** is a semidirect product of the group of spacetime translations and the Lorentz group. Under a Poincaré transformation, the spacetime coordinates transform as:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \,, \tag{9.23}$$

where Λ satisfies (9.2) and a^μ is a constant four-vector representing the translation.

Under a Poincaré transformation, a field of spin s transforms according to (9.12) after identifying x' with the Poincaré transformed spacetime coordinate. It is convenient to rewrite this transformation law as:

$$\psi'_\alpha(x) = \exp\Big(-\tfrac{i}{2}\theta_{\mu\nu}S^{\mu\nu}\Big)_\alpha{}^\beta \psi_\beta\big(\Lambda^{-1}(x-a)\big)\,, \tag{9.24}$$

where we have used $x = \Lambda^{-1}(x'-a)$ and redefined the dummy variable x' by removing the prime.

The Poincaré algebra

Under an infinitesimal Poincaré transformation, we expand (9.24) about $\Lambda = \mathbb{1}_4$ and $a = 0$ to obtain:

$$\psi'_\alpha(x) \simeq \left[\mathbb{1} + i a_\mu P^\mu - \tfrac{i}{2}\theta_{\mu\nu}(L^{\mu\nu} + S^{\mu\nu})\right]_\alpha{}^\beta \psi_\beta(x)\,, \tag{9.25}$$

where P^μ and $L^{\mu\nu}$ are the differential operators:

$$\boxed{P^\mu \equiv i\partial^\mu, \qquad L^{\mu\nu} \equiv i(x^\mu\partial^\nu - x^\nu\partial^\mu)}\,, \tag{9.26}$$

and the **total angular momentum tensor** is $J^{\mu\nu} \equiv L^{\mu\nu} + S^{\mu\nu}$.

The four generators of spacetime translations (P^μ) and the six generators of Lorentz transformations ($J^{\mu\nu}$, $\mu < \nu$) satisfy the following commutation relations of the **Poincaré algebra**:

$$\boxed{\begin{aligned} [P^\mu, P^\nu] &= 0 \\ [J^{\mu\nu}, P^\rho] &= i(g^{\nu\rho}P^\mu - g^{\mu\rho}P^\nu) \\ [J^{\mu\nu}, J^{\rho\sigma}] &= i(g^{\nu\rho}J^{\mu\sigma} - g^{\mu\rho}J^{\nu\sigma} \\ &\qquad - g^{\nu\sigma}J^{\mu\rho} + g^{\mu\sigma}J^{\nu\rho}) \end{aligned}}\,. \tag{9.27}$$

Remark. [Physical interpretation of $[J^{\mu\nu}, P^\rho]$] Before turning to the algebraic derivation, we motivate the structure of the most informative commutator, $[J^{\mu\nu}, P^\rho] = i(g^{\nu\rho}P^\mu - g^{\mu\rho}P^\nu)$. It is not a convention: it is forced by the requirement that P^μ transform as a four-vector under Lorentz transformations. The argument compares two ways of writing the infinitesimal action of a Lorentz transformation on P^ρ.

Classical side. A four-vector transforms as $P^\rho \longrightarrow \Lambda^\rho{}_\nu P^\nu$. For an infinitesimal transformation $\Lambda^\rho{}_\nu = \delta^\rho_\nu + \theta^\rho{}_\nu + O(\theta^2)$ with $\theta_{\mu\nu} = -\theta_{\nu\mu}$,

$$\Lambda^\rho{}_\nu P^\nu = P^\rho + \theta^\rho{}_\nu P^\nu + O(\theta^2)\,. \qquad (\star) \tag{9.28}$$

For an active boost, the rapidity vector used later is related to the antisymmetric parameter by $\zeta^i = \theta_{0i} = \theta^0{}_i$. Hence

$$P'^0 = P^0 + \vec{\zeta}\cdot\vec{P} + O(\zeta^2), \qquad P'^i = P^i + \zeta^i P^0 + O(\zeta^2),$$

which is the infinitesimal form of an active boost of the physical four-momentum.

Quantum side. In quantum theory the same transformation is realised by the unitary operator $U(\Lambda) = \exp\!\left(-\tfrac{i}{2}\theta_{\mu\nu}J^{\mu\nu}\right)$, and operators transform by conjugation:

$$U^\dagger(\Lambda)\, P^\rho\, U(\Lambda) = P^\rho + \tfrac{i}{2}\theta_{\mu\nu}\,[J^{\mu\nu}, P^\rho] + O(\theta^2)\,. \qquad (\star\star) \tag{9.29}$$

The commutator $[J^{\mu\nu}, P^\rho]$ *is* therefore the generator of Lorentz transformations acting on P^ρ.

Matching. Setting $(\star) = (\star\star)$ and rewriting the right-hand side of $(\star)$ in manifestly antisymmetric form (using $\theta_{\mu\nu} = -\theta_{\nu\mu}$),

$$\theta_{\mu\nu} g^{\rho\mu} P^\nu = \tfrac{1}{2}\theta_{\mu\nu}\left(g^{\rho\mu}P^\nu - g^{\rho\nu}P^\mu\right),$$

and matching coefficients of the arbitrary antisymmetric $\theta_{\mu\nu}$ fixes the commutator uniquely. The structure is dictated by Lorentz covariance, not by choice; the algebraic proof below verifies the same result by direct computation with the differential operator representation $L^{\mu\nu} = i(x^\mu\partial^\nu - x^\nu\partial^\mu)$.

Component examples.

- $[J^{ij}, P^k] = i(\delta^{ik}P^j - \delta^{jk}P^i)$: rotations mix the spatial components of $\vec{P}$ exactly as a 3-vector. Equivalently, $[J^i, P^k] = i\epsilon^{ik\ell}P^\ell$ for $J^i = \frac{1}{2}\epsilon^{ijk}J^{jk}$.
- $[J^{0i}, P^0] = -iP^i$ and $[J^{0i}, P^j] = -i\delta^{ij}P^0$: boosts mix energy with momentum — with the active boost convention used below, this corresponds infinitesimally to $E' = E + \vec{\zeta}\cdot\vec{P} + O(\zeta^2)$, or finitely to $E' = \gamma(E + \vec{v}\cdot\vec{p})$.

Derivation of the Poincaré algebra. **For** $[P^\mu, P^\nu] = 0$**:** Since $P^\mu = i\partial^\mu$, and partial derivatives commute:

$$[P^\mu, P^\nu] = [i\partial^\mu, i\partial^\nu] = -[\partial^\mu, \partial^\nu] = 0$$

For $[J^{\mu\nu}, P^\rho]$**:** We use the orbital part $L^{\mu\nu} = i(x^\mu\partial^\nu - x^\nu\partial^\mu)$ and $P^\rho = i\partial^\rho$. Since $[S^{\mu\nu}, P^\rho] = 0$ (the spin generators act on internal indices, not coordinates), we have $[J^{\mu\nu}, P^\rho] = [L^{\mu\nu}, P^\rho]$.

Acting on a test function $f(x)$:

$$\begin{aligned}[L^{\mu\nu}, P^\rho]f &= L^{\mu\nu}(P^\rho f) - P^\rho(L^{\mu\nu}f)\\ &= i(x^\mu\partial^\nu - x^\nu\partial^\mu)(i\partial^\rho f)\\ &\quad - i\partial^\rho\left[i(x^\mu\partial^\nu - x^\nu\partial^\mu)f\right]\end{aligned} \tag{9.30}$$

Using $\partial^\rho(x^\mu\partial^\nu f) = g^{\rho\mu}\partial^\nu f + x^\mu\partial^\rho\partial^\nu f$ and the commutativity of partial derivatives:

$$\begin{aligned}[L^{\mu\nu}, P^\rho]f &= -(x^\mu\partial^\nu\partial^\rho f - x^\nu\partial^\mu\partial^\rho f)\\ &\quad + (g^{\rho\mu}\partial^\nu f + x^\mu\partial^\rho\partial^\nu f - g^{\rho\nu}\partial^\mu f - x^\nu\partial^\rho\partial^\mu f)\\ &= g^{\mu\rho}\partial^\nu f - g^{\nu\rho}\partial^\mu f = \tfrac{1}{i}(g^{\mu\rho}P^\nu - g^{\nu\rho}P^\mu)f\end{aligned} \tag{9.31}$$

Therefore $[J^{\mu\nu}, P^\rho] = i(g^{\nu\rho}P^\mu - g^{\mu\rho}P^\nu)$.

For $[J^{\mu\nu}, J^{\rho\sigma}]$**:** This is the Lorentz algebra (9.4), which holds for any representation including $J^{\mu\nu} = L^{\mu\nu} + S^{\mu\nu}$. □

The Pauli-Lubanski vector and Casimir operators

It is convenient to introduce the **Pauli-Lubanski pseudovector**:

$$\boxed{w^\mu \equiv -\tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma}J_{\nu\rho}P_\sigma = \left(\vec{J}\cdot\vec{P};\, P^0\vec{J}+\vec{K}\times\vec{P}\right)}, \tag{9.32}$$

where $\epsilon^{0123} = +1$, $J^i \equiv \frac{1}{2}\epsilon^{ijk}J^{jk}$, and $K^i \equiv J^{0i}$.

Proposition 9.2 (Properties of the Pauli-Lubanski vector).

(i) $w_\mu P^\mu = 0$
(ii) $[w^\mu, P^\nu] = 0$
(iii) $[w^\mu, J^{\rho\sigma}] = i(g^{\mu\rho}w^\sigma - g^{\mu\sigma}w^\rho)$

Proof. (i) $w_\mu P^\mu = -\frac{1}{2}\epsilon^{\mu\nu\rho\sigma}J_{\nu\rho}P_\sigma P_\mu = 0$ since $\epsilon^{\mu\nu\rho\sigma}$ is antisymmetric in $\mu \leftrightarrow \sigma$ while $P_\sigma P_\mu$ is symmetric.

(ii) Using $[J_{\nu\rho}, P^\alpha] = i(g_\rho{}^\alpha P_\nu - g_\nu{}^\alpha P_\rho)$ and $[P_\sigma, P^\alpha] = 0$:

$$\begin{aligned}[w^\mu, P^\alpha] &= -\frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\left([J_{\nu\rho}, P^\alpha]P_\sigma + J_{\nu\rho}[P_\sigma, P^\alpha]\right)\\ &= -\frac{i}{2}\epsilon^{\mu\nu\rho\sigma}(g_\rho{}^\alpha P_\nu - g_\nu{}^\alpha P_\rho)P_\sigma = 0\end{aligned} \tag{9.33}$$

The last step follows because $\epsilon^{\mu\nu\rho\sigma}P_\nu P_\sigma = 0$ (antisymmetric times symmetric).

(iii) This follows from the fact that w^μ transforms as a four-vector under Lorentz transformations. □

The Poincaré algebra possesses two independent **Casimir operators**:

$$\boxed{C_1 = P^2 \equiv P_\mu P^\mu, \qquad C_2 = w^2 \equiv w_\mu w^\mu}, \tag{9.34}$$

which commute with all generators P^μ and $J^{\mu\nu}$.

Proof that P^2 and w^2 are Casimir operators.
$[P^2, P^\mu] = 0$ trivially since $[P_\nu, P^\mu] = 0$. For Lorentz generators:

$$\begin{aligned}[P^2, J^{\mu\nu}] &= P_\rho[P^\rho, J^{\mu\nu}] + [P_\rho, J^{\mu\nu}]P^\rho\\ &= -iP_\rho(g^{\rho\nu}P^\mu - g^{\rho\mu}P^\nu) + \text{h.c.} = 0\end{aligned}$$

Similarly, $[w^2, P^\mu] = 0$ follows from $[w^\nu, P^\mu] = 0$, and $[w^2, J^{\mu\nu}] = 0$ follows from the vector transformation property of w^μ. □

Physical interpretation: mass and spin

The representations of the Poincaré group corresponding to particle states of nonnegative energy P^0 with definite mass and spin can be labeled by the eigenvalues of the Casimir operators P^2 and w^2.

The eigenvalue of P^2 is m^2, where m is the **mass**. To see the physical interpretation of w^2, we consider two cases:

Massive particles ($m \neq 0$): We evaluate w^2 in the particle rest frame (since w^2 is a Lorentz scalar). In this frame, $P^\mu = (m, 0, 0, 0)$ and:

$$w^\mu = (0; m\vec{S}),$$

where S^i is defined in (9.13). Hence:

$$w^2 = -m^2 \vec{S}^2 \quad \text{with eigenvalues} \quad -m^2 s(s+1), \tag{9.35}$$

where $s = 0, \frac{1}{2}, 1, \ldots$ is the **spin**. Massive positive-energy states are labeled by (m, s).

Massless particles ($m = 0$): We cannot use the rest frame. For the ordinary finite-helicity massless representations one has $P^2 = 0$, $w_\mu P^\mu = 0$, and $w^2 = 0$. In a frame where $P = (P^0; 0, 0, P^0)$ with $P^0 > 0$, these conditions imply $w = (w^0; 0, 0, w^0)$, excluding the continuous-spin representations. That is, in any Lorentz frame:

$$w^\mu = h P^\mu, \tag{9.36}$$

where h is called the **helicity operator**. From (9.32):

$$h = \frac{w^0}{P^0} = \frac{\vec{J} \cdot \vec{P}}{P^0} = \frac{\vec{S} \cdot \vec{P}}{P^0}. \tag{9.37}$$

The last equality follows because $\vec{L} \cdot \vec{P} = (\vec{x} \times \vec{P}) \cdot \vec{P} = 0$.

Eigenvalues of h are called the helicity λ. For massless states, $\vec{S} \cdot \hat{P}$ (where $\hat{P} = \vec{P}/|\vec{P}|$) corresponds to the projection of spin along the direction of motion, with spectrum $\lambda = 0, \pm\frac{1}{2}, \pm 1, \ldots$. Under a CPT transformation, $\lambda \to -\lambda$, so both $\pm|\lambda|$ helicity states must appear in any CPT-invariant theory.

9.3 Spin-1/2 representations of the Lorentz group

The previous section used the Poincaré Casimirs to label physical one-particle states. We now change perspective: a relativistic field must also carry a finite-dimensional representation of the Lorentz group on its spinor or tensor indices. These finite-dimensional (non-unitary) representations acting on the field indices are distinct from the Hilbert-space unitary representations used to classify particles; they tell the field components how to mix when spacetime is Lorentz transformed. For spin-$\frac{1}{2}$ fields the relevant finite-dimensional representations are the two inequivalent Weyl representations.

The $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ representations

The simplest nontrivial irreducible representations of the Lorentz algebra are the two-dimensional representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$. The corresponding generators are:

$$\boxed{\begin{aligned} (\tfrac{1}{2},0): \quad & \vec{S}_+ = \tfrac{1}{2}(\vec{S}+i\vec{K}) = \tfrac{1}{2}\vec{\sigma}, \quad \vec{S}_- = 0 \\ \implies \quad & \vec{S} = \tfrac{1}{2}\vec{\sigma}, \quad \vec{K} = -\tfrac{i}{2}\vec{\sigma} \end{aligned}} . \tag{9.38}$$

$$\boxed{\begin{aligned} (0,\tfrac{1}{2}): \quad & \vec{S}_+ = 0, \quad \vec{S}_- = \tfrac{1}{2}(\vec{S}-i\vec{K}) = \tfrac{1}{2}\vec{\sigma} \\ \implies \quad & \vec{S} = \tfrac{1}{2}\vec{\sigma}, \quad \vec{K} = +\tfrac{i}{2}\vec{\sigma} \end{aligned}} . \tag{9.39}$$

Here, $\vec{\sigma} = (\sigma^1,\sigma^2,\sigma^3)$ are the Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.40}$$

We also define $\sigma^0 = \mathbb{1}_2$ and the four-component notation:

$$\sigma^\mu = (\mathbb{1}_2, \vec{\sigma}), \qquad \bar{\sigma}^\mu = (\mathbb{1}_2, -\vec{\sigma})\,. \tag{9.41}$$

Lorentz transformation matrices

There are two kinds of two-component spinors, distinguished by their transformation character under boosts. Under an infinitesimal three-dimensional rotation parameterized by $\vec{\epsilon} = (\epsilon_1,\epsilon_2,\epsilon_3)$, a four-vector transforms (in the active convention of (9.3)) as $E \longrightarrow E$, $\vec{p} \longrightarrow \vec{p}+\vec{\epsilon}\times\vec{p}$, while under an infinitesimal boost parameterized by $\vec{\eta} = (\eta_1,\eta_2,\eta_3)$, it transforms as $E \longrightarrow E+\vec{\eta}\cdot\vec{p}$, $\vec{p} \longrightarrow \vec{p}+\vec{\eta}E$.

Infinitesimal transformations of two-component spinors

Under such Lorentz transformations, the two types of Weyl spinors transform as (**active convention**, consistent with (9.3)):

$$\boxed{\begin{aligned} (\tfrac{1}{2},0): \quad & \chi \to \chi' = \left(1-\tfrac{i}{2}\vec{\epsilon}\cdot\vec{\sigma}-\tfrac{1}{2}\vec{\eta}\cdot\vec{\sigma}\right)\chi \\ (0,\tfrac{1}{2}): \quad & \psi \to \psi' = \left(1-\tfrac{i}{2}\vec{\epsilon}\cdot\vec{\sigma}+\tfrac{1}{2}\vec{\eta}\cdot\vec{\sigma}\right)\psi \end{aligned}} . \tag{9.42}$$

The $(\frac{1}{2},0)$ spinor χ (left-handed, L-type) and the $(0,\frac{1}{2})$ spinor ψ (right-handed, R-type) transform *identically* under rotations (they both have spin-$\frac{1}{2}$), but *differently* under boosts—the sign of the $\vec{\eta}\cdot\vec{\sigma}$ term distinguishes them.

Box 9.2 Active vs. passive convention: a frequent source of sign confusion

We adopt the *active* convention throughout: a rotation parameter $\vec{\theta} = \theta\,\hat{z}$ rotates the *physical system* by $+\theta$ about z, and a boost parameter $\vec{\zeta} = \zeta\,\hat{z}$ ($\zeta > 0$) boosts the system in the $+z$ direction, giving the particle positive z-momentum. A four-vector therefore transforms as

$$\vec{p} \;\longrightarrow\; \vec{p} + \vec{\epsilon}\times\vec{p} \quad (\text{rotation})\,, \qquad E \;\longrightarrow\; E + \vec{\eta}\cdot\vec{p} \quad (\text{boost})\,,$$

consistently with the statement at (9.42) above.
The spinor matrix M is obtained by inserting the $(\frac{1}{2},0)$ representation matrices $\vec{S} = \vec{\sigma}/2$, $\vec{K} = -i\vec{\sigma}/2$ from (9.38) into the representation exponential $D = \exp\left(-i\vec{\theta}\cdot\vec{S} - i\vec{\zeta}\cdot\vec{K}\right)$:

$$M = \exp\left(-\tfrac{i}{2}\vec{\theta}\cdot\vec{\sigma} - \tfrac{1}{2}\vec{\zeta}\cdot\vec{\sigma}\right).$$

The minus signs on *both* terms are a direct consequence of this substitution. The physical content — that this M produces the *active* boost — is verified by the sigma-matrix identity $P' = MPM^\dagger$ (established below in (9.57)), which yields $E' = E + \vec{\eta}\cdot\vec{p}$ and $\vec{p}\,' = \vec{p} + \vec{\eta}\,E$.
Sign subtlety. With the definitions in (9.9), the four-vector exponential (9.3) and the spinor construction give the same active boost. For a boost along $+z$,

$$\Lambda^0{}_3 = \Lambda^3{}_0 = +\sinh\zeta,$$

so a rest momentum $(m,0,0,0)$ is sent to $(m\cosh\zeta, 0, 0, m\sinh\zeta)$. Many textbooks write instead the passive coordinate transformation, or equivalently the inverse active transformation, with off-diagonal entries $-\sinh\zeta$. This is a convention change $\vec{\zeta} \to -\vec{\zeta}$, not a different covering map.
Passive convention. Several classical references (Aitchison–Hey vol. 1, parts of Weinberg) adopt the opposite overall sign: Λ is viewed as the transformation rotating the *frame*, so $\vec{p} \to \vec{p} - \vec{\epsilon}\times\vec{p}$ and the spinor law reads $\chi' = (1 + \frac{i}{2}\vec{\epsilon}\cdot\vec{\sigma} + \frac{1}{2}\vec{\eta}\cdot\vec{\sigma})\chi$. The two conventions are related by $\vec{\epsilon} \leftrightarrow -\vec{\epsilon}$, $\vec{\eta} \leftrightarrow -\vec{\eta}$ and produce identical physical predictions.

Finite transformations with explicit spinor indices: M and $(M^{-1})^\dagger$

Before writing the finite versions of (9.42), we install the index conventions that will be used throughout the rest of the chapter — conventions that immediately become necessary as soon as we have to track which representation each Weyl field belongs to.

Spinor indices: bar and dotted notation. When we attach explicit spinor indices to a Weyl field, we follow two parallel conventions:

- Left-handed $(\frac{1}{2},0)$ spinors carry *undotted* indices $\alpha,\beta,\ldots$ and are written without decoration: $\chi_\alpha, \eta_\alpha, \psi_\alpha, \ldots$
- Right-handed $(0,\frac{1}{2})$ spinors carry *dotted* indices $\dot\alpha,\dot\beta,\ldots$ and are written with an overline: $\bar\chi^{\dot\alpha}, \bar\eta^{\dot\alpha}, \bar\psi^{\dot\alpha}, \ldots$.

The bar is not decorative: it records the fact that complex conjugation maps the $(\frac{1}{2},0)$ representation into the inequivalent $(0,\frac{1}{2})$ representation. Operationally, for any L-type spinor χ_α we *define* its R-type partner by

$$\boxed{\bar\chi_{\dot\alpha} \;\equiv\; \left(\chi_\alpha\right)^*, \qquad \bar\chi^{\dot\alpha} \;=\; \epsilon^{\dot\alpha\dot\beta}\,\bar\chi_{\dot\beta}}\,, \tag{9.43}$$

where $\epsilon^{\dot\alpha\dot\beta}$ is the antisymmetric spinor metric used to raise dotted indices, with $\epsilon^{\dot 1\dot 2} = +1$; its undotted analogue $\epsilon^{\alpha\beta}$ raises undotted indices in the same way (the full conventions, including how the metric is used to form Lorentz-invariant bilinears, are collected in (9.97) below). Whenever an independent R-type field is introduced without an explicit L-type "parent" (such as the ψ of (9.42) above, written below with explicit dotted index as $\bar\psi^{\dot\alpha}$), it is understood to be of this form for some implicit underlying L-type field.

Cross-references with older notations. Older texts — notably Aitchison's Lorentz chapter, early Weinberg, and many original QFT references — denote $\bar\chi^{\dot\alpha}$ by $\chi^{\dagger\dot\alpha}$ ("dagger notation"), with the dagger emphasising the Hermitian-conjugate origin of the dotted-index partner. The two notations are interchangeable; we have chosen the bar form throughout this chapter and Chapter 10 to align with modern SUSY conventions (Wess–Bagger, Srednicki, Aitchison's SUSY chapter).

A forward warning. The same overline symbol will reappear later in the four-component Dirac formalism as the Dirac adjoint $\bar\Psi \equiv \Psi^\dagger\gamma^0$, which is a *distinct* operation acting on a four-component spinor (see (9.141) and the dedicated remark accompanying it). The two usages are disambiguated by the type of the parent field — lowercase Greek two-component spinors versus uppercase Dirac four-spinors — but it is worth flagging the collision once, at the moment the bar notation is introduced, rather than discovering it 800 lines later.

The Weyl transformation matrices M and $(M^{-1})^\dagger$. For the $(\frac{1}{2},0)$ representation, we adopt the active convention (cf. box 9.2) for the L-type Weyl spinor. The finite transformation matrix is:

$$\boxed{M = \exp\!\left(-\tfrac{i}{2}\vec\theta\cdot\vec\sigma - \tfrac{1}{2}\vec\zeta\cdot\vec\sigma\right) \simeq \mathbb{1}_2 - \tfrac{i}{2}\vec\theta\cdot\vec\sigma - \tfrac{1}{2}\vec\zeta\cdot\vec\sigma}\,, \tag{9.44}$$

where $\vec{\theta}$ is the rotation angle vector and $\vec{\zeta} = \hat{\boldsymbol{\beta}} \tanh^{-1}\beta$ is the rapidity vector; at first order these are the $\vec{\epsilon}$ and $\vec{\eta}$ of (9.42).

Remark. [The rapidity vector and the non-relativistic limit] The choice $\vec{\zeta} = \hat{\boldsymbol{\beta}} \tanh^{-1}\beta$ is not arbitrary: it is the unique boost parameter that is *additive* for collinear boosts, in the same sense in which the rotation angle $\vec{\theta} = \theta\hat{n}$ is additive for rotations about a fixed axis. The velocity itself fails this property: the relativistic composition law gives $\vec{v}_1 \oplus \vec{v}_2 \neq \vec{v}_1 + \vec{v}_2$, whereas collinear rapidities simply satisfy $\zeta_1 + \zeta_2$.

Setting $\beta = \tanh\zeta$ gives $\gamma = \cosh\zeta$ and $\gamma\beta = \sinh\zeta$, so an active boost along x with rapidity ζ takes the form

$$\Lambda(M) = \begin{pmatrix} \cosh\zeta & +\sinh\zeta & 0 & 0 \\ +\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

formally a “hyperbolic rotation” (the standard textbook form has $-\sinh\zeta$ on the off-diagonal, corresponding to the passive convention; see box 9.2 for the relation between the two). Promoting $\zeta \to \vec{\zeta} = \hat{\boldsymbol{\beta}} \tanh^{-1}\beta$ extends this to an arbitrary boost direction; the corresponding spinor matrix is $M = \exp\left(-\frac{1}{2}\vec{\zeta}\cdot\vec{\sigma}\right)$ as in (9.44), and the four-vector transformation $\Lambda(M)$ is determined by the sigma identity (9.57).

Non-relativistic limit. For $|\vec{v}| \ll c$, the Taylor expansion $\tanh^{-1}\beta = \beta + \frac{\beta^3}{3} + O(\beta^5)$ gives

$$\vec{\zeta} \simeq \hat{\boldsymbol{\beta}}\beta = \frac{\vec{v}}{c} \equiv \vec{\eta},$$

recovering the infinitesimal parameter of (9.42). Analogously $\vec{\theta} \simeq \vec{\epsilon}$ at first order. Summarising:

finite parameter	infinitesimal	type
$\vec{\theta} = \theta\hat{n}$	$\vec{\epsilon}$	rotation
$\vec{\zeta} = \hat{\boldsymbol{\beta}} \tanh^{-1}\beta$	$\vec{\eta} = \vec{v}/c$	boost

> The finite parameters $\vec\theta, \vec\zeta$ encode the full Lie-group exponential; the infinitesimal $\vec\epsilon, \vec\eta$ are their linearisations and label tangent vectors at the identity. The non-linearity of $\tanh^{-1}\beta$ encodes all the relativistic content of boost composition.

A two-component $(\frac{1}{2},0)$ spinor field $\chi_\alpha(x)$ (left-handed Weyl spinor) transforms as:

$$\chi'_\alpha(x') = M_\alpha{}^\beta \chi_\beta(x), \qquad \alpha,\beta \in \{1,2\}. \tag{9.45}$$

By definition, M carries undotted spinor indices: $M_\alpha{}^\beta$.

For the $(0,\frac{1}{2})$ representation, the transformation matrix is:

$$\boxed{(M^{-1})^\dagger = \exp\left(-\tfrac{i}{2}\vec\theta\cdot\vec\sigma + \tfrac{1}{2}\vec\zeta\cdot\vec\sigma\right) \simeq \mathbb{1}_2 - \tfrac{i}{2}\vec\theta\cdot\vec\sigma + \tfrac{1}{2}\vec\zeta\cdot\vec\sigma}. \tag{9.46}$$

A two-component $(0,\frac{1}{2})$ spinor field $\bar\psi^{\dot\alpha}(x)$ (right-handed Weyl spinor, see (9.43)) transforms as:

$$\bar\psi'^{\dot\alpha}(x') = \left[(M^{-1})^\dagger\right]^{\dot\alpha}{}_{\dot\beta}\,\bar\psi^{\dot\beta}(x). \tag{9.47}$$

The matrix $(M^{-1})^\dagger$ carries dotted spinor indices.

Verification of the relation between M and $(M^{-1})^\dagger$. From (9.44), the inverse (at first order) is obtained by reversing the transformation parameters:

$$M^{-1} \simeq \mathbb{1}_2 + \tfrac{i}{2}\vec\theta\cdot\vec\sigma + \tfrac{1}{2}\vec\zeta\cdot\vec\sigma$$

Using the hermiticity of the Pauli matrices ($\vec\sigma^\dagger = \vec\sigma$):

$$M^\dagger \simeq \mathbb{1}_2 + \tfrac{i}{2}\vec\theta\cdot\vec\sigma - \tfrac{1}{2}\vec\zeta\cdot\vec\sigma$$

Hence $(M^{-1})^\dagger = (M^\dagger)^{-1} \simeq \mathbb{1}_2 - \frac{i}{2}\vec\theta\cdot\vec\sigma + \frac{1}{2}\vec\zeta\cdot\vec\sigma$, confirming (9.46). □

> **Remark.** [Index conventions] We adopt the convention that:
>
> - Undotted indices $\alpha,\beta,\dots$ label $(\frac{1}{2},0)$ spinors and are contracted as ${}^\alpha{}_\alpha$ (northwest to southeast)
> - Dotted indices $\dot\alpha,\dot\beta,\dots$ label $(0,\frac{1}{2})$ spinors and are contracted as ${}_{\dot\alpha}{}^{\dot\alpha}$ (south-west to northeast)

The homomorphism $SL(2,\mathbb{C}) \to SO^+(1,3)$

We now justify why the two Weyl matrices M and $(M^{-1})^\dagger$ of (9.44)–(9.46) are the *correct* objects to call "Lorentz transformations of spinors": they are the two inequivalent two-dimensional irreducible representations of the *universal covering group* of the restricted Lorentz group, $SL(2,\mathbb{C})$.

A direct calculation from (9.44) gives $\det M = 1$ (the exponential of a traceless matrix has unit determinant), so the Weyl matrices live in the special linear group

$$SL(2,\mathbb{C}) \equiv \{ M \in \mathrm{Mat}(2,\mathbb{C}) \,:\, \det M = 1 \}.$$

This is a 6-real-dimensional Lie group, exactly the same dimension as the restricted Lorentz group $SO^+(1,3)$. The connection between the two is more than a counting coincidence: there is an explicit two-to-one group homomorphism $SL(2,\mathbb{C}) \to SO^+(1,3)$, with $SL(2,\mathbb{C})$ playing the role of the *universal covering group*. The construction (following Müller-Kirsten [31], Section 1.3.3) proceeds in three steps.

Step 1 — The four-vector / Hermitian matrix correspondence. Define the map

$$\rho \,:\, \mathbb{M}_4 \longrightarrow \mathbb{H}(2,\mathbb{C}), \qquad x^\mu \longmapsto X \equiv x^\mu \sigma_\mu = \begin{pmatrix} x^0 - x^3 & -x^1 + ix^2 \\ -x^1 - ix^2 & x^0 + x^3 \end{pmatrix}, \tag{9.48}$$

sending each Minkowski four-vector to a 2×2 Hermitian matrix (the Hermiticity is immediate from $\sigma_\mu^\dagger = \sigma_\mu$ and the reality of x^μ). The map ρ is a real-linear bijection: any $X \in \mathbb{H}(2,\mathbb{C})$ has the form (9.48) for a unique real four-vector

$$x^\mu = \tfrac{1}{2}\mathrm{Tr}\big[X\,\bar{\sigma}^\mu\big], \tag{9.49}$$

as one verifies using the trace identity $\mathrm{Tr}[\sigma_\mu \bar{\sigma}^\nu] = 2\,\delta_\mu^\nu$. A direct computation yields the key identity

$$\det X = (x^0)^2 - |\vec{x}|^2 = x_\mu x^\mu, \tag{9.50}$$

i.e. *the determinant of the Hermitian matrix equals the Minkowski square of the four-vector*. This is the algebraic engine of the whole construction.

Step 2 — The adjoint action of $SL(2,\mathbb{C})$ on $\mathbb{H}(2,\mathbb{C})$. For any $M \in SL(2,\mathbb{C})$ and $X \in \mathbb{H}(2,\mathbb{C})$ define

$$\mathrm{ad}(M)\,X \equiv M\,X\,M^\dagger. \tag{9.51}$$

Box 9.3 Why this is an "adjoint" action

The transformation $X \mapsto MXM^\dagger$ is called the *adjoint action* of $SL(2,\mathbb{C})$ on Hermitian matrices. The name carries two layers of physical meaning, both worth unpacking.

The quantum-mechanical analogy. Every Hermitian observable $\hat{O}$ on a quantum Hilbert space transforms under a symmetry $\hat{U}$ as

$$\hat{O} \longmapsto \hat{U}\hat{O}\hat{U}^\dagger,$$

the same rule by which density matrices transform: $\rho \mapsto U\rho U^\dagger$. The matrix $X = x^\mu \sigma_\mu$ *is* literally a 2×2 Hermitian matrix encoding the four-vector x^μ, so it behaves like a quantum observable on a two-dimensional state space, and $MXM^\dagger$ is the natural extension of the QM operator rule to the symmetry group at hand. The conjugation by $\dagger$ in place of M^{-1} is not an exotic device — it is what quantum mechanics has been doing all along to transformations of observables.

From group adjoint to algebra adjoint: the universal pattern. For any Lie group G acting on its own Lie algebra $\mathfrak{g}$, the group adjoint action is the conjugation

$$\mathrm{Ad}(g)\,X \;=\; g\,X\,g^{-1}, \qquad g \in G,\;\; X \in \mathfrak{g}. \tag{9.52}$$

Differentiating this at the identity along the one-parameter subgroup $g(t) = e^{tA}$, with $A \in \mathfrak{g}$, gives

$$\left.\frac{d}{dt}\right|_{t=0} \big(e^{tA}\,X\,e^{-tA}\big) \;=\; AX - XA \;=\; [A, X] \;\equiv\; \mathrm{ad}(A)\,X\,.$$

This is the *algebra adjoint* representation: $\mathrm{ad}(A) : \mathfrak{g} \longrightarrow \mathfrak{g}$ is the linear map $X \mapsto [A, X]$, obtained by differentiating the group conjugation gXg^{-1} at the identity. The relation is universal across Lie theory:

$$\boxed{\text{group adjoint } \mathrm{Ad}(g) \xrightarrow{\;d/dt|_{t=0}\;} \text{algebra adjoint } \mathrm{ad}(A) = [A, \cdot\,]\,.}$$

In quantum mechanics this is the statement that infinitesimal symmetry transformations of an observable are generated by commutators with the Hermitian generator: $\delta_A \hat{O} = -i\,[A, \hat{O}]$ for unitary $U = e^{-iA}$.

Specialising to $SL(2,\mathbb{C})$ on Hermitian matrices. We repeat the differentiation for our slightly modified action $MXM^\dagger$. Writing $M = \mathbb{1}_2 + tA + O(t^2)$ with $A \in \mathfrak{sl}(2,\mathbb{C})$:

$$M\,X\,M^\dagger \;=\; X \;+\; t\big(AX + XA^\dagger\big) \;+\; O(t^2)\,.$$

The generator is $AX + XA^\dagger$ — not a pure commutator, because of the $\dagger$. To compare with the universal pattern above, split A into anti-Hermitian and Hermitian parts:

$$A \;=\; A_- \;+\; A_+\,, \qquad A_-^\dagger = -A_- \text{ (rotations)}, \quad A_+^\dagger = +A_+ \text{ (boosts)}.$$

Then

$$AX + XA^\dagger \;=\; \underbrace{[A_-, X]}_{\text{rotation: Lie bracket}} \;+\; \underbrace{\{A_+, X\}}_{\text{boost: anticommutator}}\,.$$

The **rotation part recovers the standard algebra adjoint** $\mathrm{ad}(A_-)\,X = [A_-, X]$ exactly as for $\mathrm{Ad}(g)$ in (9.52): on the rotation subgroup, our modified action is literally the universal Lie-group adjoint. The **boost part is new**: a symmetric (anticommutator) product, reflecting the fact that boosts in finite-dimensional spinor representations are Hermitian rather than anti-Hermitian, and therefore *rescale* X rather than rotate it.

Physical pictures.

- *Rotations $\longrightarrow$ the familiar $SU(2) \to SO(3)$ story.* For pure rotations $M = U \in SU(2)$ is unitary, $U^\dagger = U^{-1}$, and $\mathrm{ad}(U)\,X = UXU^{-1}$ literally is the standard group adjoint. The infinitesimal action on $\vec{x}\cdot\vec{\sigma}$ is the spin-$\frac{1}{2}$ commutator $\mathrm{ad}(\vec{\sigma})\,\vec{x}\cdot\vec{\sigma} \sim \vec{\sigma}\times\vec{x}$ that underlies Larmor precession, Stern–Gerlach splitting, and every angular-momentum coupling calculation. We recover the 2:1 covering $SU(2) \to SO(3)$ as the spatial sector of the present construction.
- *Boosts $\longrightarrow$ finite-dimensional non-unitarity.* Boosts in $SL(2,\mathbb{C})$ are realised by Hermitian matrices $M = e^{-\vec{\zeta}\cdot\vec{\sigma}/2}$, not unitary ones. Finite-dimensional Lorentz representations *cannot* be unitary — the same obstruction that, in Wigner's one-particle classification, forces massive states into infinite-dimensional Hilbert-space representations. The $\dagger$ in $MXM^\dagger$ is what accommodates this non-unitarity at the level of the finite-dimensional matrix realisation: it rescales the diagonal entries of X by $e^{\pm\zeta}$, mixing energy and momentum (cf. the explicit boost calculation below (9.57)). The Jordan-like anticommutator $\{A_+, X\}$ is the infinitesimal echo of this rescaling.
- *Why $\dagger$ and not -1? Hermiticity = physical content.* The matrix $X = x^\mu\sigma_\mu$ carries a real four-vector precisely because X is Hermitian. Conjugation by M^{-1} would not preserve Hermiticity for non-unitary M: $(MXM^{-1})^\dagger = (M^{-1})^\dagger XM^\dagger \neq MXM^{-1}$. The substitution $M^{-1} \to M^\dagger$ is the same fix

quantum mechanics imposes to keep density matrices density matrices, and observables observables.

Summary. The map $X \mapsto MXM^\dagger$ is the quantum-operator rule applied to a Hermitian matrix realisation of a four-vector. Differentiating at the identity (the universal recipe $\mathrm{Ad}(g) \to \mathrm{ad}(A) = [A, \cdot]$) recovers the ordinary algebra adjoint $[A_-, X]$ on the rotation sector, supplemented by a Jordan-like anticommutator $\{A_+, X\}$ on the boost sector. The first is universal Lie theory; the second is the algebraic signature of the non-compactness of the Lorentz group — the same fact that distinguishes special relativity from purely spatial rotations.

Two properties are immediate:

(i) $\mathrm{ad}(M)X$ is again Hermitian: $(MXM^\dagger)^\dagger = MX^\dagger M^\dagger = MXM^\dagger$ since $X^\dagger = X$;
(ii) $\det\big(MXM^\dagger\big) = |\det M|^2 \det X = \det X$, using $\det M = 1$.

By (9.50), the second property says that the four-vector $x'^\mu \equiv \rho^{-1}(MXM^\dagger)$ has the *same* Minkowski square as x^μ:

$$x'_\mu x'^\mu \;=\; x_\mu x^\mu .$$

The map $x^\mu \mapsto x'^\mu$ is real-linear (since ρ and $\mathrm{ad}(M)$ are), preserves the quadratic form, and is continuously connected to the identity (taking $M = \mathbb{1}_2$). It is therefore an element of the restricted Lorentz group $SO^+(1,3)$. We may write

$$\boxed{x'^\mu \;=\; \Lambda^\mu{}_\nu(M)\, x^\nu, \qquad \Lambda^\mu{}_\nu(M) \;=\; \tfrac{1}{2}\mathrm{Tr}\big[\bar\sigma^\mu M \sigma_\nu M^\dagger\big]}\,. \tag{9.53}$$

This is the *explicit* formula for the homomorphism $SL(2,\mathbb{C}) \to SO^+(1,3)$. The right-hand side is obtained by combining (9.49) with $X' = MXM^\dagger$ from (9.51); the component-level version of the same statement is the sigma-matrix transformation law $M\sigma_\mu M^\dagger = \sigma_\nu \Lambda^\nu{}_\mu$, established below as eqs. (9.57)–(9.58).

Step 3 — Homomorphism, image, and kernel. We now prove three propositions that establish (9.53) as a two-to-one group homomorphism onto $SO^+(1,3)$.

(a) Λ is a group homomorphism: for all $M_1, M_2 \in SL(2,\mathbb{C})$,

$$\Lambda(M_1 M_2) \;=\; \Lambda(M_1)\,\Lambda(M_2), \qquad \Lambda(M^{-1}) \;=\; \Lambda(M)^{-1}.$$

Proof. Apply $\mathrm{ad}(M_2)$ followed by $\mathrm{ad}(M_1)$ to an arbitrary $X \in \mathbb{H}(2,\mathbb{C})$:

$$\mathrm{ad}(M_1)\,\mathrm{ad}(M_2)\,X \;=\; M_1\big(M_2 X M_2^\dagger\big)M_1^\dagger \;=\; (M_1M_2)\,X\,(M_1M_2)^\dagger \;=\; \mathrm{ad}(M_1M_2)\,X,$$

using $(M_1M_2)^\dagger = M_2^\dagger M_1^\dagger$. Reading both sides through ρ^{-1} converts the equality of adjoint actions on $\mathbb{H}(2,\mathbb{C})$ into the equality of the induced Lorentz transformations on $\mathbb{M}_4$, hence $\Lambda(M_1)\Lambda(M_2) = \Lambda(M_1M_2)$. Setting $M_2 = M^{-1}$ in this identity and noting $\Lambda(\mathbb{1}_2) = \mathbb{1}$ gives $\Lambda(M^{-1}) = \Lambda(M)^{-1}$. □

(b) $\Lambda(M)$ **lies in the restricted Lorentz group:**

$$\det\Lambda(M) \;=\; +1\,, \qquad \Lambda^0{}_0(M) \;\geq\; 1\,.$$

Proof. We already know that $\Lambda(M)$ preserves the quadratic form $x_\mu x^\mu$ (Step 2), so $\Lambda(M) \in O(1,3)$ and $\det\Lambda(M) \in \{+1,-1\}$, $\Lambda^0{}_0(M) \in (-\infty,-1] \cup [1,+\infty)$. Both functions $M \mapsto \det\Lambda(M)$ and $M \mapsto \Lambda^0{}_0(M)$ are continuous (polynomial in the entries of $M, M^\dagger$ by (9.53)). The group $SL(2,\mathbb{C})$ is connected (it is homeomorphic to $\mathbb{R}^3 \times S^3$, see paragraph after (9.54) below), and at $M = \mathbb{1}_2$ both quantities take the value $+1$. By the intermediate value theorem, they cannot jump to a value in the disconnected complementary branches, so $\det\Lambda(M) = +1$ and $\Lambda^0{}_0(M) \geq +1$ throughout $SL(2,\mathbb{C})$.

A direct algebraic version of the second statement: from (9.53) with $\mu = \nu = 0$ (and $\sigma_0 = \bar\sigma^0 = \mathbb{1}_2$),

$$\Lambda^0{}_0(M) \;=\; \tfrac{1}{2}\,\mathrm{Tr}\big[MM^\dagger\big]\,.$$

The singular value decomposition $M = UDV^\dagger$ with $U, V \in U(2)$ and $D = \mathrm{diag}(s_1, s_2)$, $s_i > 0$, gives $\mathrm{Tr}[MM^\dagger] = s_1^2 + s_2^2$, while $\det M = 1$ forces $s_1 s_2 = 1$. By AM–GM,

$$s_1^2 + s_2^2 \;\geq\; 2\,s_1 s_2 \;=\; 2\,,$$

so $\Lambda^0{}_0(M) = \frac{1}{2}(s_1^2 + s_2^2) \geq 1$, with equality iff $s_1 = s_2 = 1$, i.e. M unitary — a pure rotation. □

(c) The kernel of Λ **is** $\{\pm\mathbb{1}_2\}$**:**

$$\ker\Lambda \;=\; \{\,M \in SL(2,\mathbb{C}) \;:\; \Lambda(M) = \mathbb{1}\,\} \;=\; \{\,\pm\mathbb{1}_2\,\}.$$

Proof. Suppose $\Lambda(M) = \mathbb{1}$. By (9.58) (which expresses (9.53) at the level of basis sigma matrices),

$$M\sigma_\mu M^\dagger \;=\; \sigma_\nu\Lambda^\nu{}_\mu \;=\; \sigma_\mu \quad \text{for all } \mu = 0,1,2,3\,.$$

Taking $\mu = 0$: $M\mathbb{1}_2M^\dagger = MM^\dagger = \mathbb{1}_2$, i.e. $M^\dagger = M^{-1}$ — M is *unitary*. For $\mu = i$ the same equation then reads $M\sigma_i M^{-1} = \sigma_i$, that is, M commutes with every Pauli matrix:

$$[M, \sigma_i] \;=\; 0\,, \quad i = 1,2,3\,.$$

A 2×2 matrix that commutes with all three Pauli matrices commutes in particular with the basis $\{\mathbb{1}_2, \sigma_1, \sigma_2, \sigma_3\}$ of $\mathrm{Mat}(2,\mathbb{C})$, hence belongs to the centre of $\mathrm{Mat}(2,\mathbb{C})$, which by Schur's lemma is $\mathbb{C}\,\mathbb{1}_2$. Thus $M = c\,\mathbb{1}_2$ for some $c \in \mathbb{C}$. The constraint $\det M = c^2 = 1$ forces $c = \pm 1$, i.e. $M = \pm\mathbb{1}_2$. Conversely $\Lambda(\pm\mathbb{1}_2)$ acts on X as $X \mapsto (\pm\mathbb{1}_2)X(\pm\mathbb{1}_2) = X$, so both $\pm\mathbb{1}_2$ lie in $\ker\Lambda$. □The first isomorphism theorem then yields

$$\boxed{SO^+(1,3) \;\cong\; SL(2,\mathbb{C})/\mathbb{Z}_2\,,} \tag{9.54}$$

with the $\mathbb{Z}_2$ kernel generated by $-\mathbb{1}_2$. Since $SL(2,\mathbb{C})$ is simply connected (as a 6-dimensional real Lie group it is homeomorphic to $\mathbb{R}^3 \times S^3$), (9.54) identifies $SL(2,\mathbb{C})$ as the *universal cover* of $SO^+(1,3)$.

Inverse construction $M(\Lambda)$. The two-valuedness can be made explicit: given any $\Lambda \in SO^+(1,3)$, the element M that maps to it under (9.53) is

$$M(\Lambda) \;=\; \pm\frac{1}{\sqrt{\det\left[\Lambda^\mu{}_\nu\sigma_\mu\bar{\sigma}^\nu\right]}}\,\Lambda^\mu{}_\nu\sigma_\mu\bar{\sigma}^\nu, \tag{9.55}$$

unique up to the overall sign $\pm$ [31].

Physical consequence. The two-valuedness of the cover is the precise mathematical reason why *spinor* fields exist: an irreducible representation of $SL(2,\mathbb{C})$ that does not factor through $SO^+(1,3)$ assigns inequivalent transformations to M and $-M$, and hence corresponds to a representation of $SO^+(1,3)$ only up to sign. A 2π rotation in $SO^+(1,3)$ lifts to $-\mathbb{1}_2$ in $SL(2,\mathbb{C})$, so a spinor returns to itself only after a 4π rotation. This is the algebraic origin of the fermion sign and of all related effects (Berry phase, spin-statistics, etc.). The two fundamental two-dimensional irreducible representations of $SL(2,\mathbb{C})$ are precisely the $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ Weyl spinor representations introduced in (9.42); their direct sum is the Dirac representation.

Conclusion: identifying M and $(M^{-1})^\dagger$. Equation (9.53) singles out, for each Lorentz transformation Λ, exactly two elements $\pm M \in SL(2,\mathbb{C})$ that act on Hermitian matrices as $X \mapsto MXM^\dagger$. The two non-equivalent 2-dimensional irreducible representations of this $SL(2,\mathbb{C})$ act on Weyl spinors as $\chi \mapsto M\chi$ (the $(\frac{1}{2},0)$ representation) and $\bar{\psi} \mapsto (M^{-1})^\dagger\bar{\psi}$ (the $(0,\frac{1}{2})$ representation): one verifies that M and $(M^{-1})^\dagger$ are inequivalent precisely because $\Lambda_L = \Lambda_R^{-1\dagger}$ differs from Λ_R by the sign of the boost (cf. the discussion below (9.46)). Therefore the matrices M and $(M^{-1})^\dagger$ introduced earlier are not ad hoc but are forced by group theory: they are the *unique* two-dimensional spinor representations of the universal cover of the Lorentz group, and any Weyl-spinor field theory is built from these representations and their complex conjugates. The deeper physical consequence — that the sigma matrices act as intertwining operators between the two Weyl sectors, and that this intertwining is what makes the Dirac equation inevitable — is examined at the end of the Sigma matrices subsection below.

Sigma matrices

Sigma matrices and four-vectors: explicit realisation of ρ

We now make the abstract construction of the homomorphism $SL(2,\mathbb{C}) \to SO^+(1,3)$ entirely concrete. Applying the map ρ of (9.48) to a generic four-momentum p^μ gives

$$P \;\equiv\; \rho(p) \;=\; p^\mu\sigma_\mu = p^0\mathbb{1}_2 - \vec{p}\cdot\vec{\sigma} = \begin{pmatrix} p^0 - p^3 & -p^1 + ip^2 \\ -p^1 - ip^2 & p^0 + p^3 \end{pmatrix}, \tag{9.56}$$

a Hermitian 2×2 matrix. The inverse map (9.49) recovers $p^\mu = \frac{1}{2}\mathrm{Tr}[P\,\bar{\sigma}^\mu]$, and the determinant identity (9.50) reads

$$\det P = (p^0)^2 - |\vec{p}|^2 = p_\mu p^\mu\,,$$

as one verifies directly from the matrix on the right-hand side of (9.56). This concretely realises the abstract correspondence $\mathbb{M}_4 \leftrightarrow \mathbb{H}(2,\mathbb{C})$ on momentum space.

Proposition 9.3 (Lorentz transformation of $p^\mu\sigma_\mu$). The Lorentz transformation $p'^\mu = \Lambda^\mu{}_\nu(M)\,p^\nu$ generated by $M \in SL(2,\mathbb{C})$ through the homomorphism (9.53) is realised at the level of Hermitian matrices by the adjoint action $\mathrm{ad}(M)$:

$$p'^\mu\sigma_\mu = M\,(p^\nu\sigma_\nu)\,M^\dagger\,. \tag{9.57}$$

Equivalently, at the level of the basis sigma matrices,

$$M\,\sigma_\mu\,M^\dagger = \sigma_\nu\,\Lambda^\nu{}_\mu(M)\,, \tag{9.58}$$

and, with the inverse transformation and a raised Lorentz index,

$$M^{-1}\sigma^\mu(M^{-1})^\dagger = \Lambda^\mu{}_\nu(M)\,\sigma^\nu\,. \tag{9.59}$$

Proof. The Hermiticity-preservation and determinant-preservation steps required to conclude that $\mathrm{ad}(M)$ induces *some* Lorentz transformation were already carried out in the construction of the homomorphism (Step 2 of the previous subsubsection, using $\det M = 1$). All that remains is to identify the specific $\Lambda \in SO^+(1,3)$ produced by a given M.

Step A: from (9.57) to (9.58). Expand p'^{μ} as the linear combination $p'^{\mu}\sigma_{\mu}$ in (9.57) and use the assumed transformation law $p'^{\mu} = \Lambda^{\mu}{}_{\nu} p^{\nu}$ on the left:

$$\Lambda^{\mu}{}_{\nu}\, p^{\nu}\, \sigma_{\mu} \;=\; M\,(p^{\nu}\sigma_{\nu})\,M^{\dagger} \;=\; p^{\nu}\, M\sigma_{\nu}M^{\dagger}\,.$$

Since this holds for every four-vector p^{ν}, the coefficients of p^{ν} on the two sides must match:

$$M\sigma_{\nu}M^{\dagger} = \sigma_{\mu}\Lambda^{\mu}{}_{\nu},$$

which is (9.58).

To obtain the raised-index inverse form (9.59), we proceed in two substeps.

(A1) Apply (9.58) to M^{-1}. Since the identity holds for *every* $M \in SL(2,\mathbb{C})$, replacing $M \to M^{-1}$ gives

$$M^{-1}\,\sigma_{\nu}\,(M^{-1})^{\dagger} \;=\; \sigma_{\mu}\,\Lambda^{\mu}{}_{\nu}(M^{-1}) \;=\; \sigma_{\mu}\,(\Lambda^{-1})^{\mu}{}_{\nu},$$

where we used the group-homomorphism property $\Lambda(M^{-1}) = \Lambda(M)^{-1}$ (Step 3, (a) of the previous subsubsection).

(A2) Raise the free index ν with $g^{\nu\alpha}$. The left-hand side becomes simply

$$g^{\nu\alpha}\big[M^{-1}\sigma_{\nu}(M^{-1})^{\dagger}\big] \;=\; M^{-1}\,\sigma^{\alpha}\,(M^{-1})^{\dagger}.$$

For the right-hand side we use the standard Lorentz-inverse identity, which follows from the metric-preservation condition $\Lambda^{\mu}{}_{\rho}\,\Lambda^{\nu}{}_{\sigma}\, g_{\mu\nu} = g_{\rho\sigma}$:

$$\boxed{(\Lambda^{-1})^{\mu}{}_{\nu} \;=\; \Lambda_{\nu}{}^{\mu} \;\equiv\; g_{\nu\rho}\, g^{\mu\sigma}\, \Lambda^{\rho}{}_{\sigma}}\,. \tag{9.60}$$

Plugging this in and contracting with $g^{\nu\alpha}$:

$$g^{\nu\alpha}\,\sigma_{\mu}\,(\Lambda^{-1})^{\mu}{}_{\nu} \;=\; \sigma_{\mu}\,g^{\nu\alpha}\,g_{\nu\rho}\,g^{\mu\sigma}\,\Lambda^{\rho}{}_{\sigma} \;=\; \sigma_{\mu}\,\delta^{\alpha}_{\rho}\,g^{\mu\sigma}\,\Lambda^{\rho}{}_{\sigma} \;=\; \underbrace{\sigma_{\mu}\,g^{\mu\sigma}}_{=\sigma^{\sigma}}\,\Lambda^{\alpha}{}_{\sigma} \;=\; \Lambda^{\alpha}{}_{\sigma}\,\sigma^{\sigma}.$$

Equating the two sides and relabelling $\alpha \to \mu$, $\sigma \to \nu$:

$$M^{-1}\,\sigma^{\mu}\,(M^{-1})^{\dagger} \;=\; \Lambda^{\mu}{}_{\nu}(M)\,\sigma^{\nu},$$

which is (9.59).

Step B: explicit verification on a boost. The abstract argument fixes the structure but not the specific value of $\Lambda^{\mu}{}_{\nu}$ in terms of the $(\vec{\theta},\vec{\zeta})$ that parametrise M in (9.44). We pin this down on a representative example. Take an active boost along the z-axis with rapidity $\zeta > 0$: from (9.44) with $\vec{\theta} = 0$ and $\vec{\zeta} = \zeta\hat{z}$,

$$M = e^{-\zeta\sigma^3/2} = \begin{pmatrix} e^{-\zeta/2} & 0 \\ 0 & e^{+\zeta/2} \end{pmatrix}, \qquad M = M^{\dagger}\,. \tag{9.61}$$

Computing $MPM^\dagger$ explicitly:

$$
\begin{aligned}
MPM^\dagger &= \begin{pmatrix} e^{-\zeta/2} & 0 \\ 0 & e^{+\zeta/2} \end{pmatrix} \begin{pmatrix} p^0 - p^3 & -p^1 + ip^2 \\ -p^1 - ip^2 & p^0 + p^3 \end{pmatrix} \begin{pmatrix} e^{-\zeta/2} & 0 \\ 0 & e^{+\zeta/2} \end{pmatrix} \\
&= \begin{pmatrix} e^{-\zeta}(p^0 - p^3) & -p^1 + ip^2 \\ -p^1 - ip^2 & e^{+\zeta}(p^0 + p^3) \end{pmatrix}.
\end{aligned}
\tag{9.62}
$$

Reading off the components from the canonical form $p'^\mu \sigma_\mu = \begin{pmatrix} p'^0 - p'^3 & -p'^1 + ip'^2 \\ -p'^1 - ip'^2 & p'^0 + p'^3 \end{pmatrix}$:

$$
\begin{aligned}
p'^0 + p'^3 = e^{+\zeta}(p^0 + p^3)\,, \qquad p'^0 - p'^3 = e^{-\zeta}(p^0 - p^3)\,, \\
p'^1 = p^1\,, \qquad p'^2 = p^2\,.
\end{aligned}
\tag{9.63}
$$

Adding and subtracting the first two equations:

$$
p'^0 = \cosh\zeta\, p^0 + \sinh\zeta\, p^3\,, \qquad p'^3 = \sinh\zeta\, p^0 + \cosh\zeta\, p^3\,,
\tag{9.64}
$$

which is the active Lorentz boost along $+z$: a particle at rest acquires positive z-momentum $p'^3 = m\sinh\zeta > 0$, as required by our active convention (cf. box 9.2). The corresponding $\Lambda(M)$ has $\Lambda^0{}_3 = +\sinh\zeta$; this differs from the standard textbook form $\Lambda^0{}_3 = -\sinh\zeta$ (which is the passive convention) by the sign of $\sinh\zeta$, as explained in box 9.2. The general case can be assembled by composing such examples (rotations and boosts in arbitrary directions) or, equivalently, by extracting $\Lambda(M)$ directly from the trace formula (9.53). □

Remark. [The two halves of one story] The previous subsubsection established the existence and structural properties of the homomorphism $SL(2,\mathbb{C}) \to SO^+(1,3)$ in abstract terms (hermiticity of X, determinant invariance, kernel $\pm\mathbb{1}_2$). The present subsubsection *realises* the same homomorphism on a concrete object — the Hermitian matrix $P = p^\mu \sigma_\mu$ associated to a four-vector — and identifies the linear transformation of P under the adjoint action with an explicit Lorentz transformation in spacetime. Equations (9.57), (9.58), (9.59) are therefore the component-level statements of which (9.53) is the trace-extracted summary.

Reading off the spinor index structure of σ^μ and $\bar\sigma^\mu$. The transformation laws (9.58)–(9.59) are matrix identities in $\mathrm{Mat}(2,\mathbb{C})$, but each 2×2 slot of σ^μ carries a definite spinor index assignment that is *dictated* by which group element (M or $(M^{-1})^\dagger$) acts on which side. Let us read this off step by step.

(i) The two Weyl reps and their index conventions. From (9.45) the $(\frac{1}{2},0)$ representation acts on *undotted* spinors by

$$\chi'_\alpha = M_\alpha{}^\beta \chi_\beta,$$

so M carries the index structure $M_\alpha{}^\beta$. From (9.47) the $(0,\frac{1}{2})$ representation acts on *dotted* spinors by

$$\bar\psi'^{\dot\alpha} = \left[(M^{-1})^\dagger\right]^{\dot\alpha}{}_{\dot\beta}\bar\psi^{\dot\beta},$$

so $(M^{-1})^\dagger$ carries dotted indices $\left[(M^{-1})^\dagger\right]^{\dot\alpha}{}_{\dot\beta}$. Crucially, $M^\dagger$ itself (which appears on the right of σ^μ in (9.58)) is obtained from $(M^{-1})^\dagger$ by inversion and so it carries the *transposed* dotted index pattern:

$$\left[M^\dagger\right]^{\dot\beta}{}_{\dot\alpha}.$$

(ii) Forcing the index placement of σ^μ. Restoring spinor indices in (9.58), the left-hand side $M\sigma^\mu M^\dagger$ contracts $M_\alpha{}^\beta$ on the left and $[M^\dagger]^{\dot\beta}{}_{\dot\alpha}$ on the right, so σ^μ must be contractable with an upper undotted index from M and an upper dotted index from $M^\dagger$. The only consistent placement is

$$\boxed{\sigma^\mu_{\alpha\dot\alpha}\,,}$$

with one lower undotted index and one lower dotted index. The identity then reads, with all spinor indices explicit:

$$M_\alpha{}^\beta\,\sigma^\mu_{\beta\dot\beta}\left[M^\dagger\right]^{\dot\beta}{}_{\dot\alpha} = \Lambda^\mu{}_\nu(M)\,\sigma^\nu_{\alpha\dot\alpha}. \tag{9.65}$$

The Lorentz index μ on σ^μ and the spinor indices $\alpha,\dot\alpha$ are now both unambiguous: σ^μ is a $(\frac{1}{2},\frac{1}{2})$ object — the unique linear map intertwining $(\frac{1}{2},0)\otimes(0,\frac{1}{2})$ with the four-vector representation $(\frac{1}{2},\frac{1}{2})$.

(iii) Dual story for $\bar\sigma^\mu$. The barred matrices are the index-raised partners of the unbarred ones. Starting from (9.59) and raising the two spinor indices with the antisymmetric metrics gives the equivalent barred transformation law

$$\left[(M^{-1})^\dagger\right]^{\dot\alpha}{}_{\dot\beta}\,\bar\sigma^{\mu\dot\beta\beta}\left[M^{-1}\right]_\beta{}^\alpha = \Lambda^\mu{}_\nu(M)\,\bar\sigma^{\nu\dot\alpha\alpha}.$$

Thus $\bar\sigma^\mu$ contracts with the lower dotted and undotted index slots of the inverse transformations. The natural placement is therefore reversed:

$$\boxed{\bar\sigma^{\mu\dot\alpha\alpha}\,,}$$

with both spinor indices *upper* (first dotted, then undotted, by convention). The two patterns are related by the spinor metric: $\bar\sigma^{\mu\dot\alpha\alpha} = \epsilon^{\alpha\beta}\epsilon^{\dot\alpha\dot\beta}\sigma^\mu_{\beta\dot\beta}$, as recorded in (9.69) below.

(iv) Summary. The spinor index structure

$$\sigma^{\mu}_{\alpha\dot{\alpha}} \quad \text{and} \quad \bar{\sigma}^{\mu\,\dot{\alpha}\alpha}, \tag{9.66}$$

with undotted $\alpha, \beta, \dots$ labelling $(\frac{1}{2}, 0)$ and dotted $\dot{\alpha}, \dot{\beta}, \dots$ labelling $(0, \frac{1}{2})$, is therefore not an arbitrary convention: it is the *unique* placement compatible with the transformation laws (9.58) and (9.59).

Sigma matrices with covariant indices

The sigma matrices have been defined with upper (contravariant) Lorentz indices. They are related to the sigma matrices with lower (covariant) indices in the usual way:

$$\boxed{\sigma_\mu = g_{\mu\nu}\sigma^\nu = (\mathbb{1}_2, -\vec{\sigma}), \qquad \bar{\sigma}_\mu = g_{\mu\nu}\bar{\sigma}^\nu = (\mathbb{1}_2, \vec{\sigma})}\,. \tag{9.67}$$

Both σ^μ and $\bar{\sigma}^\mu$ are hermitian matrices, i.e., $(\sigma^\mu)^\dagger = \sigma^\mu$ and $(\bar{\sigma}^\mu)^\dagger = \bar{\sigma}^\mu$, or equivalently:

$$(\sigma^{\mu}_{\alpha\dot{\beta}})^* = \sigma^{\mu}_{\beta\dot{\alpha}}, \qquad (\bar{\sigma}^{\mu\,\dot{\alpha}\beta})^* = \bar{\sigma}^{\mu\,\dot{\beta}\alpha}\,. \tag{9.68}$$

Key sigma matrix identities

Some useful relations between σ^μ and $\bar{\sigma}^\mu$ involving the ϵ tensor are:

$$\boxed{\begin{aligned} \sigma^{\mu}_{\alpha\dot{\alpha}} &= \epsilon_{\alpha\beta}\epsilon_{\dot{\alpha}\dot{\beta}}\bar{\sigma}^{\mu\,\dot{\beta}\beta}, & \bar{\sigma}^{\mu\,\dot{\alpha}\alpha} &= \epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\sigma^{\mu}_{\beta\dot{\beta}} \\ \epsilon^{\alpha\beta}\sigma^{\mu}_{\beta\dot{\alpha}} &= \epsilon_{\dot{\alpha}\dot{\beta}}\bar{\sigma}^{\mu\,\dot{\beta}\alpha}, & \epsilon^{\dot{\alpha}\dot{\beta}}\sigma^{\mu}_{\alpha\dot{\beta}} &= \epsilon_{\alpha\beta}\bar{\sigma}^{\mu\,\dot{\alpha}\beta} \end{aligned}}\,. \tag{9.69}$$

Proof of (9.69).
Throughout we adopt the convention $\epsilon^{12} = -\epsilon_{12} = +1$ (and analogously for dotted indices). Direct enumeration of the four index combinations gives the *spinor-metric identity*

$$\epsilon^{\alpha\beta}\epsilon_{\beta\gamma} = \delta^{\alpha}_{\gamma}, \qquad \epsilon^{\dot{\alpha}\dot{\beta}}\epsilon_{\dot{\beta}\dot{\gamma}} = \delta^{\dot{\alpha}}_{\dot{\gamma}}\,. \tag{$\dagger$}$$

(i) The second relation is the defining one. The identity

$$\bar{\sigma}^{\mu\,\dot{\alpha}\alpha} = \epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}}\sigma^{\mu}_{\beta\dot{\beta}} \tag{$*$}$$

is the statement that $\bar{\sigma}^\mu$ is obtained from σ^μ by raising both spinor indices with the spinor metric. To check that this prescription, applied to $\sigma^\mu = (\mathbb{1}_2, \vec{\sigma})$, indeed reproduces $\bar{\sigma}^\mu = (\mathbb{1}_2, -\vec{\sigma})$, treat both sides of ($*$) as 2×2 matrices in which σ^μ has row index β (undotted) and column index $\dot{\beta}$ (dotted), and $\bar{\sigma}^\mu$ has row index $\dot{\alpha}$

(dotted) and column index α (undotted). The sum on the right reorganises into the matrix product

$$\bar{\sigma}^{\mu} = \epsilon\,(\sigma^{\mu})^{\top}\,\epsilon^{\top}, \qquad \text{with} \quad \epsilon \equiv i\sigma^{2} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Using $\epsilon^{\top} = -\epsilon$ and $\epsilon^{2} = -\mathbb{1}_{2}$:

$$\begin{aligned} \mu = 0: &\quad \epsilon\,\mathbb{1}_2\,\epsilon^{\top} = -\epsilon^{2} = \mathbb{1}_2 = \bar{\sigma}^{0}, \\ \mu = 1,3: &\quad (\sigma^{\mu})^{\top} = \sigma^{\mu}, \;\; \epsilon\sigma^{\mu}\epsilon^{\top} = -\sigma^{\mu} = \bar{\sigma}^{\mu} \;\; (\text{using } \epsilon\sigma^{1,3}\epsilon = \sigma^{1,3}), \\ \mu = 2: &\quad (\sigma^{2})^{\top} = -\sigma^{2}, \;\; \epsilon(-\sigma^{2})\epsilon^{\top} = -\sigma^{2} = \bar{\sigma}^{2} \;\; (\text{using } \epsilon\sigma^{2}\epsilon^{\top} = \sigma^{2}). \end{aligned}$$

This verifies (∗).

(ii) The first relation is its inverse. Multiply both sides of (∗) by $\epsilon_{\rho\alpha}\epsilon_{\dot{\rho}\dot{\alpha}}$ and apply the spinor-metric identity (†) twice:

$$\epsilon_{\rho\alpha}\epsilon_{\dot{\rho}\dot{\alpha}}\,\bar{\sigma}^{\mu\,\dot{\alpha}\alpha} = \underbrace{\epsilon_{\rho\alpha}\epsilon^{\alpha\beta}}_{\delta_{\rho}^{\beta}}\,\underbrace{\epsilon_{\dot{\rho}\dot{\alpha}}\epsilon^{\dot{\alpha}\dot{\beta}}}_{\delta_{\dot{\rho}}^{\dot{\beta}}}\,\sigma^{\mu}_{\beta\dot{\beta}} = \sigma^{\mu}_{\rho\dot{\rho}}\,.$$

Relabelling $\rho \to \alpha$, $\dot{\rho} \to \dot{\alpha}$ reproduces the first relation in (9.69).

(iii) The mixed relations follow by contracting with a single ϵ. Multiply (∗) by $\epsilon_{\dot{\rho}\dot{\alpha}}$ only:

$$\epsilon_{\dot{\rho}\dot{\alpha}}\,\bar{\sigma}^{\mu\,\dot{\alpha}\alpha} = \epsilon^{\alpha\beta}\,\underbrace{\epsilon_{\dot{\rho}\dot{\alpha}}\epsilon^{\dot{\alpha}\dot{\beta}}}_{\delta_{\dot{\rho}}^{\dot{\beta}}}\,\sigma^{\mu}_{\beta\dot{\beta}} = \epsilon^{\alpha\beta}\,\sigma^{\mu}_{\beta\dot{\rho}}\,.$$

Relabelling $\dot{\rho} \to \dot{\alpha}$, this is the third relation in (9.69). Multiplying instead by $\epsilon_{\rho\alpha}$ yields, by an identical argument,

$$\epsilon_{\rho\alpha}\,\bar{\sigma}^{\mu\,\dot{\alpha}\alpha} = \epsilon^{\dot{\alpha}\dot{\beta}}\,\sigma^{\mu}_{\rho\dot{\beta}},$$

which after relabelling $\rho \to \alpha$ is the fourth relation. □

The following identities are essential for calculations:

$$\sigma^{\mu}_{\alpha\dot{\alpha}}\bar{\sigma}^{\dot{\beta}\beta}_{\mu} = 2\delta^{\beta}_{\alpha}\delta^{\dot{\beta}}_{\dot{\alpha}} \tag{9.70}$$

$$\sigma^{\mu}_{\alpha\dot{\alpha}}\sigma_{\mu\,\beta\dot{\beta}} = 2\epsilon_{\alpha\beta}\epsilon_{\dot{\alpha}\dot{\beta}} \tag{9.71}$$

$$\bar{\sigma}^{\mu\,\dot{\alpha}\alpha}\bar{\sigma}^{\dot{\beta}\beta}_{\mu} = 2\epsilon^{\alpha\beta}\epsilon^{\dot{\alpha}\dot{\beta}} \tag{9.72}$$

$$[\sigma^{\mu}\bar{\sigma}^{\nu} + \sigma^{\nu}\bar{\sigma}^{\mu}]_{\alpha}{}^{\beta} = 2g^{\mu\nu}\delta^{\beta}_{\alpha} \tag{9.73}$$

$$[\bar{\sigma}^{\mu}\sigma^{\nu} + \bar{\sigma}^{\nu}\sigma^{\mu}]^{\dot{\alpha}}{}_{\dot{\beta}} = 2g^{\mu\nu}\delta^{\dot{\alpha}}_{\dot{\beta}}\,. \tag{9.74}$$

Remark. [Triple products of sigma matrices] The following identities are useful for simplifying products of three sigma matrices:

$$\begin{aligned} \sigma^\mu\bar\sigma^\nu\sigma^\rho &= g^{\mu\nu}\sigma^\rho - g^{\mu\rho}\sigma^\nu + g^{\nu\rho}\sigma^\mu \\ &\quad + i\epsilon^{\mu\nu\rho\kappa}\sigma_\kappa \end{aligned} \tag{9.75}$$

$$\begin{aligned} \bar\sigma^\mu\sigma^\nu\bar\sigma^\rho &= g^{\mu\nu}\bar\sigma^\rho - g^{\mu\rho}\bar\sigma^\nu + g^{\nu\rho}\bar\sigma^\mu \\ &\quad - i\epsilon^{\mu\nu\rho\kappa}\bar\sigma_\kappa\,, \end{aligned} \tag{9.76}$$

where $\epsilon^{0123} = +1$.

Trace formulas

Computations of cross sections and decay rates generally require traces of alternating products of σ and $\bar\sigma$ matrices:

$$\boxed{\mathrm{Tr}[\sigma^\mu\bar\sigma^\nu] = \mathrm{Tr}[\bar\sigma^\mu\sigma^\nu] = 2g^{\mu\nu}}\,. \tag{9.77}$$

$$\begin{aligned} \mathrm{Tr}[\sigma^\mu\bar\sigma^\nu\sigma^\rho\bar\sigma^\kappa] &= 2\big(g^{\mu\nu}g^{\rho\kappa} - g^{\mu\rho}g^{\nu\kappa} \\ &\qquad + g^{\mu\kappa}g^{\nu\rho} + i\epsilon^{\mu\nu\rho\kappa}\big)\,. \end{aligned} \tag{9.78}$$

$$\begin{aligned} \mathrm{Tr}[\bar\sigma^\mu\sigma^\nu\bar\sigma^\rho\sigma^\kappa] &= 2\big(g^{\mu\nu}g^{\rho\kappa} - g^{\mu\rho}g^{\nu\kappa} \\ &\qquad + g^{\mu\kappa}g^{\nu\rho} - i\epsilon^{\mu\nu\rho\kappa}\big)\,. \end{aligned} \tag{9.79}$$

Remark. [Computing higher traces] Traces involving a larger even number of σ and $\bar\sigma$ matrices can be systematically obtained from eqs. (9.77)–(9.79) by repeated use of eqs. (9.73) and (9.74) and the cyclic property of the trace. Traces involving an odd number of σ and $\bar\sigma$ matrices cannot arise, since there is no way to connect the spinor indices consistently.

Proof of (9.73). For $\mu = \nu = 0$: $[\sigma^0\bar\sigma^0 + \sigma^0\bar\sigma^0]_\alpha{}^\beta = 2\delta_\alpha^\beta = 2g^{00}\delta_\alpha^\beta$.
For $\mu = 0$, $\nu = i$: $[\sigma^0\bar\sigma^i + \sigma^i\bar\sigma^0]_\alpha{}^\beta = (-\sigma^i + \sigma^i)_\alpha{}^\beta = 0 = 2g^{0i}\delta_\alpha^\beta$.
For $\mu = i$, $\nu = j$: Using $\sigma^i\bar\sigma^j = -\sigma^i\sigma^j = -(\delta^{ij}\mathbb{1}_2 + i\epsilon^{ijk}\sigma^k)$:

$$[\sigma^i\bar\sigma^j + \sigma^j\bar\sigma^i]_\alpha{}^\beta = -2\delta^{ij}\delta_\alpha^\beta = 2g^{ij}\delta_\alpha^\beta$$

□

Remark. [Preview: sigma matrices and the Dirac equation] Having assembled the full sigma-matrix machinery — definitions (9.56), the transformation laws (9.57)–(9.59), the index structure $\sigma^{\mu}_{\alpha\dot{\alpha}}$ and $\bar{\sigma}^{\mu\dot{\alpha}\alpha}$, the anticommutators (9.73)–(9.74), and the trace formulas (9.77) — we can now extract the central physical consequence: the matrices σ^{μ} and $\bar{\sigma}^{\mu}$ are precisely the intertwining operators between the two Weyl representations, and this intertwining is what *forces* the Dirac equation.

Statement. If $\bar{\psi}$ is a $(0,\frac{1}{2})$ spinor (R-type), then $\sigma^{\mu}p_{\mu}\bar{\psi}$ transforms as a $(\frac{1}{2},0)$ spinor (L-type). Conversely, if χ is a $(\frac{1}{2},0)$ spinor, then $\bar{\sigma}^{\mu}p_{\mu}\chi$ transforms as a $(0,\frac{1}{2})$ spinor. Schematically,

$$\chi \underset{\sigma^{\mu}p_{\mu}}{\overset{\bar{\sigma}^{\mu}p_{\mu}}{\rightleftarrows}} \bar{\psi}.$$

The bottom arrow $\sigma^{\mu}p_{\mu}$ sends R→L; the top arrow $\bar{\sigma}^{\mu}p_{\mu}$ sends L→R.

Origin of the intertwining. The two Weyl matrices M and $(M^{-1})^{\dagger}$ in (9.44)–(9.46) agree on rotations but differ by the *sign of the boost*. From this sign flip,

$$M^{\dagger} = M^{-1} \text{ (rotations)}, \qquad M^{\dagger} \neq M^{-1} \text{ (boosts)},$$

follows the identity $M\,(\sigma_{\nu}p^{\nu})\,M^{\dagger} = \sigma_{\nu}p'^{\nu}$ already established in (9.57), which is precisely the statement that σ^{μ} transforms as a four-vector when sandwiched between L-type and R-type representations. The corresponding $\bar{\sigma}$ identity is obtained by conjugating this relation and raising or lowering the spinor indices with ϵ. Combining these sigma-matrix identities with the transformation laws of χ and $\bar{\psi}$ proves the chirality-flipping property stated above.

Why this forces the Dirac equation. Suppose one postulates a single Weyl spinor $\bar{\psi}$ of type $(0,\frac{1}{2})$ and seeks a linear, Lorentz-covariant wave equation involving a mass. The only Lorentz-covariant linear operator built from $\bar{\psi}$ and one p is $\sigma^{\mu}p_{\mu}\bar{\psi}$, but this is $(\frac{1}{2},0)$ — the *wrong* chirality. An equation of the form "$\sigma^{\mu}p_{\mu}\bar{\psi} = m\bar{\psi}$" is therefore not Lorentz-covariant: its two sides live in inequivalent representations. The minimal remedy is to introduce a partner spinor χ of opposite chirality and write the coupled system

$$\boxed{\sigma^{\mu}p_{\mu}\bar{\psi} = m\chi, \qquad \bar{\sigma}^{\mu}p_{\mu}\chi = m\bar{\psi}}\,, \tag{9.80}$$

in which both sides of each equation transform with the same Weyl matrix. The price of Lorentz covariance plus mass is doubling the number of components from 2 to 4 — and that is the Dirac field.

Consistency check. Applying $\bar{\sigma}^{\nu} p_{\nu}$ to the first equation and using the second,

$$\bar{\sigma}^{\nu} p_{\nu} \sigma^{\mu} p_{\mu} \bar{\psi} = m \bar{\sigma}^{\nu} p_{\nu} \chi = m^2 \bar{\psi} .$$

The anticommutator identities (9.73)–(9.74) yield $\bar{\sigma}^{\nu} p_{\nu} \sigma^{\mu} p_{\mu} = p^2$, so $(p^2 - m^2)\bar{\psi} = 0$ and analogously for χ: both Weyl components satisfy the Klein–Gordon equation with the *same* mass.

Weyl-basis Dirac equation. Stacking the two Weyl fields as $\Psi = (\chi, \bar{\psi})^{\top}$, equations (9.80) take the matrix form $(\gamma^{\mu} p_{\mu} - m)\Psi = 0$ with

$$\gamma^{\mu}_{\text{Weyl}} = \begin{pmatrix} 0 & \sigma^{\mu} \\ \bar{\sigma}^{\mu} & 0 \end{pmatrix}, \qquad \gamma^5 = \begin{pmatrix} -\mathbb{1}_2 & 0 \\ 0 & +\mathbb{1}_2 \end{pmatrix}.$$

The off-diagonal block structure of γ^{μ} is *precisely* the statement that σ^{μ} and $\bar{\sigma}^{\mu}$ intertwine L- and R-type. The diagonal form of γ^5 reflects the fact that chirality intrinsically labels the two irreducible pieces of Ψ. The Clifford algebra $\{\gamma^{\mu}, \gamma^{\nu}\} = 2g^{\mu\nu}$ is a direct consequence of the $\sigma\bar{\sigma}$-anticommutators just displayed.

Physical content.

- *Chirality is Lorentz-invariant; helicity is not.* The labels L/R denote membership in inequivalent irreducible representations — a Lorentz-invariant statement. Helicity $\hat{p} \cdot \vec{S}$ is frame-dependent for a massive particle: a sufficiently large boost overtakes the particle and flips its helicity. Chirality and helicity coincide only in the massless limit.
- *Massless decoupling.* For $m = 0$, (9.80) decouples into two independent equations $\sigma^{\mu} p_{\mu} \bar{\psi} = 0$ and $\bar{\sigma}^{\mu} p_{\mu} \chi = 0$. A massless spin-$\frac{1}{2}$ particle may therefore live entirely in a single Weyl sector (the historical motivation for two-component neutrinos).
- *Parity exchanges L and R.* Under $\vec{p} \to -\vec{p}$, $\sigma^{\mu} p_{\mu} \leftrightarrow \bar{\sigma}^{\mu} p_{\mu}$, so parity *interchanges* the two equations in (9.80). A theory built from a single Weyl spinor is therefore necessarily parity-violating; the Dirac construction restores parity by including both halves on an equal footing.
- *Uniqueness of the mass term.* The mass term is the unique Lorentz-invariant, non-derivative coupling between χ and ψ that can be built from the two spinors alone. Hence m is the only free parameter of the free Dirac Lagrangian.

Connection to the homomorphism construction. Recalling the group-theoretic perspective developed in the SL(2,$\mathbb{C}$) homomorphism subsection above, the four-vector representation decomposes as $(\frac{1}{2},\frac{1}{2}) = (\frac{1}{2},0)\otimes(0,\frac{1}{2})$, and $\sigma^\mu, \bar{\sigma}^\mu$ are the Clebsch–Gordan-like objects realising this decomposition concretely. Contracting with p_μ produces the unique linear Lorentz-covariant map between the two Weyl sectors; coupling these maps with a mass yields (9.80), the Dirac equation in the Weyl basis.

The antisymmetric tensor generators $\sigma^{\mu\nu}$ and $\bar{\sigma}^{\mu\nu}$

We define the antisymmetric tensor generators:

$$(\sigma^{\mu\nu})_\alpha{}^\beta \equiv \tfrac{i}{4}\left(\sigma^\mu_{\alpha\dot{\alpha}}\bar{\sigma}^{\nu\dot{\alpha}\beta} - \sigma^\nu_{\alpha\dot{\alpha}}\bar{\sigma}^{\mu\dot{\alpha}\beta}\right) \tag{9.81}$$

$$(\bar{\sigma}^{\mu\nu})^{\dot{\alpha}}{}_{\dot{\beta}} \equiv \tfrac{i}{4}\left(\bar{\sigma}^{\mu\dot{\alpha}\alpha}\sigma^\nu_{\alpha\dot{\beta}} - \bar{\sigma}^{\nu\dot{\alpha}\alpha}\sigma^\mu_{\alpha\dot{\beta}}\right). \tag{9.82}$$

Equivalently, using eqs. (9.73) and (9.74):

$$(\sigma^\mu\bar{\sigma}^\nu)_\alpha{}^\beta = g^{\mu\nu}\delta_\alpha^\beta - 2i(\sigma^{\mu\nu})_\alpha{}^\beta \tag{9.83}$$

$$(\bar{\sigma}^\mu\sigma^\nu)^{\dot{\alpha}}{}_{\dot{\beta}} = g^{\mu\nu}\delta^{\dot{\alpha}}_{\dot{\beta}} - 2i(\bar{\sigma}^{\mu\nu})^{\dot{\alpha}}{}_{\dot{\beta}}. \tag{9.84}$$

The components of $\sigma^{\mu\nu}$ and $\bar{\sigma}^{\mu\nu}$ are explicitly:

$$\boxed{\begin{aligned} &\sigma^{ij} = \bar{\sigma}^{ij} = \frac{1}{2}\epsilon^{ijk}\sigma^k \\ &\sigma^{i0} = -\sigma^{0i} = -\bar{\sigma}^{i0} = \bar{\sigma}^{0i} = \frac{i}{2}\sigma^i \end{aligned}}. \tag{9.85}$$

The hermiticity of $\vec{\sigma}$ implies that $(\sigma^{\mu\nu})^\dagger = \bar{\sigma}^{\mu\nu}$, or equivalently:

$$((\sigma^{\mu\nu})_\alpha{}^\beta)^* = (\bar{\sigma}^{\mu\nu})^{\dot{\alpha}}{}_{\dot{\beta}}. \tag{9.86}$$

Proposition 9.4 (Self-duality relations). The matrices $\sigma^{\mu\nu}$ and $\bar{\sigma}^{\mu\nu}$ satisfy the self-duality relations:

$$\boxed{\sigma^{\mu\nu} = -\tfrac{i}{2}\epsilon^{\mu\nu\rho\kappa}\sigma_{\rho\kappa}, \quad \bar{\sigma}^{\mu\nu} = +\tfrac{i}{2}\epsilon^{\mu\nu\rho\kappa}\bar{\sigma}_{\rho\kappa}}, \tag{9.87}$$

where $\epsilon^{0123} = +1$.

The self-duality relations imply that $\sigma^{\mu\nu}$ and $\bar{\sigma}^{\mu\nu}$ each possess only three independent components. Hence, the total number of independent components of the bilinear covariants involving $\sigma^{\mu\nu}$ is indeed 3 (not 6).

These matrices satisfy:

$$S^{\mu\nu} = \begin{cases} \sigma^{\mu\nu} & \text{for the } (\frac{1}{2},0) \text{ representation} \\ \bar{\sigma}^{\mu\nu} & \text{for the } (0,\frac{1}{2}) \text{ representation} \end{cases}. \tag{9.88}$$

With the same active convention used in (9.44), the exponentiated Lorentz transformation matrices are:

$$M = \exp\left(-\frac{i}{2}\theta_{\mu\nu}\sigma^{\mu\nu}\right) \tag{9.89}$$

$$(M^{-1})^\dagger = \exp\left(-\frac{i}{2}\theta_{\mu\nu}\bar{\sigma}^{\mu\nu}\right). \tag{9.90}$$

From (9.13), $S^i = \frac{1}{2}\epsilon^{ijk}S^{jk} = \frac{1}{2}\sigma^i$ for both representations. Therefore $w^2 = -m^2\vec{S}^2$ has eigenvalues $-\frac{3}{4}m^2$ as expected for a spin-$\frac{1}{2}$ fermion.

Transformation laws for all spinor index positions

So far we have established the transformation laws for the two fundamental Weyl spinors in their *natural* index positions: lower undotted for L-type (eq. (9.45)) and upper dotted for R-type (eq. (9.47)):

$$\chi'_\alpha = M_\alpha{}^\beta \chi_\beta, \qquad \bar{\chi}'^{\dot{\alpha}} = \left[(M^{-1})^\dagger\right]^{\dot{\alpha}}{}_{\dot{\beta}}\bar{\chi}^{\dot{\beta}}. \tag{9.91}$$

But spinors also appear with *raised* undotted indices ($\chi^\alpha = \epsilon^{\alpha\beta}\chi_\beta$) and *lowered* dotted indices ($\bar{\chi}_{\dot{\alpha}} = (\chi_\alpha)^*$). How do *those* transform? The answer requires one group-theoretic ingredient: the invariance of the ϵ tensor under $SL(2,\mathbb{C})$.

The ϵ tensor is $SL(2,\mathbb{C})$-invariant

For any $M \in SL(2,\mathbb{C})$, the Levi-Civita symbol satisfies:

$$\epsilon^{\alpha\beta} M_\alpha{}^\gamma M_\beta{}^\delta = \epsilon^{\gamma\delta}. \tag{9.92}$$

Proof. The left-hand side is antisymmetric in (γ,δ), so it suffices to evaluate one component:

$$\epsilon^{\alpha\beta} M_\alpha{}^1 M_\beta{}^2 = M_1{}^1 M_2{}^2 - M_2{}^1 M_1{}^2 = \det M = 1 = \epsilon^{12}.$$

The general case follows by antisymmetry. □

This is the spinorial analogue of $g_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma = g_{\rho\sigma}$: just as the metric tensor is invariant under Lorentz transformations, the ϵ tensor is invariant under $SL(2,\mathbb{C})$. This invariance is what makes Lorentz-invariant spinor contractions possible. Since $\det M^* = (\det M)^* = 1$, the same identity holds for dotted indices:

$$\epsilon_{\dot{\alpha}\dot{\beta}}\,[(M^{-1})^\dagger]^{\dot{\alpha}}{}_{\dot{\gamma}}\,[(M^{-1})^\dagger]^{\dot{\beta}}{}_{\dot{\delta}} = \epsilon_{\dot{\gamma}\dot{\delta}}. \tag{9.93}$$

Spinors with raised undotted indices transform with (M^{-1})

Starting from $\chi'_\alpha = M_\alpha{}^\beta \chi_\beta$ and raising the free index with $\chi^\alpha = \epsilon^{\alpha\beta}\chi_\beta$:

$$\chi'^\alpha = \epsilon^{\alpha\gamma}\chi'_\gamma = \epsilon^{\alpha\gamma} M_\gamma{}^\beta \chi_\beta .$$

To express the right-hand side in terms of $\chi^\delta = \epsilon^{\delta\beta}\chi_\beta$, rewrite $\chi_\beta = \epsilon_{\beta\delta}\chi^\delta$:

$$\chi'^\alpha = \underbrace{\epsilon^{\alpha\gamma} M_\gamma{}^\beta \epsilon_{\beta\delta}}_{\equiv\ (M^{-1})^\alpha{}_\delta} \chi^\delta . \tag{9.94}$$

The notation $(M^{-1})^\alpha{}_\delta$ for the combination $\epsilon^{\alpha\gamma}M_\gamma{}^\beta\epsilon_{\beta\delta}$ is not merely a shorthand: contracting (9.92) with $\epsilon_{\gamma\sigma}$ one verifies that $(M^{-1})^\beta{}_\sigma M_\beta{}^\delta = \delta^\delta_\sigma$, so this matrix genuinely acts as the inverse of M in the appropriate index sense (in 2×2 matrix notation it equals $(M^\top)^{-1}$).

Spinors with lowered dotted indices: contragredient of $(M^{-1})^\dagger$

There are two equivalent routes to this law, and seeing both makes the structure transparent.

Route 1: lowering with ϵ from the $(M^{-1})^\dagger$ law. This is the exact dotted analogue of the undotted derivation above. Starting from $\bar\chi'^{\dot\alpha} = [(M^{-1})^\dagger]^{\dot\alpha}{}_{\dot\beta}\bar\chi^{\dot\beta}$ and lowering the free index with $\bar\chi_{\dot\alpha} = \epsilon_{\dot\alpha\dot\gamma}\bar\chi^{\dot\gamma}$:

$$\bar\chi'_{\dot\alpha} = \epsilon_{\dot\alpha\dot\gamma}\bar\chi'^{\dot\gamma} = \epsilon_{\dot\alpha\dot\gamma}[(M^{-1})^\dagger]^{\dot\gamma}{}_{\dot\delta}\bar\chi^{\dot\delta} .$$

Rewriting $\bar\chi^{\dot\delta} = \epsilon^{\dot\delta\dot\beta}\bar\chi_{\dot\beta}$:

$$\bar\chi'_{\dot\alpha} = \underbrace{\epsilon_{\dot\alpha\dot\gamma}[(M^{-1})^\dagger]^{\dot\gamma}{}_{\dot\delta}\epsilon^{\dot\delta\dot\beta}}_{\equiv\ [(M^{-1})^{\dagger -1}]_{\dot\alpha}{}^{\dot\beta}} \bar\chi_{\dot\beta} . \tag{9.95}$$

The structure is parallel to the undotted case: lowering a dotted index replaces $(M^{-1})^\dagger$ with its contragredient via ϵ, just as raising an undotted index replaces M with its contragredient (M^{-1}).

Route 2: complex conjugation. From $\bar\chi_{\dot\alpha} \equiv (\chi_\alpha)^*$ and $\chi'_\alpha = M_\alpha{}^\beta\chi_\beta$ one gets directly

$$\bar\chi'_{\dot\alpha} = (M_\alpha{}^\beta)^*\bar\chi_{\dot\beta} \equiv (M^*)_{\dot\alpha}{}^{\dot\beta}\bar\chi_{\dot\beta} ,$$

where $(M^*)_{\dot\alpha}{}^{\dot\beta} \equiv (M_\alpha{}^\beta)^*$. The two routes agree because $(M^{-1})^{\dagger -1} = M^\dagger = (M^*)^\top$, so the contragredient matrix $\epsilon_{\dot\alpha\dot\gamma}[(M^{-1})^\dagger]^{\dot\gamma}{}_{\dot\delta}\epsilon^{\dot\delta\dot\beta}$ has components $(M^*)_{\dot\alpha}{}^{\dot\beta}$.