THE MATHEMATICAL PHYSICS MONOGRAPH SERIES

A. S. Wightman, Editor
Princeton University

George W. Mackey (*Harvard*)	THE MATHEMATICAL FOUNDATIONS OF QUANTUM MECHANICS
R. F. Streater (*Imperial College of Science and Technology*) A. S. Wightman (*Princeton*)	PCT, SPIN AND STATISTICS, AND ALL THAT

PCT, SPIN AND STATISTICS, AND ALL THAT

R. F. STREATER
Imperial College of Science and Technology

A. S. WIGHTMAN
Princeton University

W. A. BENJAMIN, INC.
New York Amsterdam
1964

CONTENTS

1-3. THE LORENTZ AND POINCARÉ GROUPS

Among the most important symmetries of relativistic quantum theory are those which arise from the Lorentz transformations themselves. The next few paragraphs are devoted to establishing our notation and summarizing the main facts about these symmetries.

The Lorentz-invariant scalar product of two four-vectors $x = (x^0, x^1, x^2, x^3)$ and $y = (y^0, y^1, y^2, y^3)$ will be written

$$x \cdot y = x^0 y^0 - \mathbf{x} \cdot \mathbf{y} \equiv x^\mu g_{\mu\nu} x^\nu \equiv x^\mu y_\mu \tag{1-4}$$

with the usual summation convention on repeated indices.

Here

$$x_\mu = g_{\mu\nu} x^\nu$$

and $g^{\mu\nu} = g_{\mu\nu}$ is the μ,ν component of the matrix G,

$$G = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix}.$$

A Lorentz transformation Λ is a linear transformation mapping space-time onto space-time which preserves the scalar product (1-4): $(\Lambda x) \cdot (\Lambda y) = x \cdot y$. If $(\Lambda x)^\mu = \Lambda^\mu{}_\nu x^\nu$ the (real) matrix, $\Lambda^\mu{}_\nu$, of the transformation must satisfy

$$\Lambda^\kappa{}_\mu \Lambda_{\kappa\nu} = g_{\mu\nu} \quad \text{or} \quad \Lambda^T G \Lambda = G, \tag{1-5}$$

where the transpose Λ^T of Λ is defined by $(\Lambda^T)^\mu{}_\nu = \Lambda^\nu{}_\mu$ and indices on Λ are lowered according to

$$\begin{aligned} \Lambda_{\kappa\nu} &= g_{\kappa\sigma} \Lambda^\sigma{}_\nu \\ &= (G\Lambda)_{\kappa\nu}. \end{aligned}$$

If Λ and M satisfy (1-5), so do ΛM and Λ^{-1}. Here

$$\begin{aligned} (\Lambda M)^\mu{}_\nu &= \Lambda^\mu{}_\kappa M^\kappa{}_\nu \\ (\Lambda^{-1})^\mu{}_\kappa \Lambda^\kappa{}_\nu &= g^\mu{}_\nu = \begin{cases} 0 & \mu \neq \nu \\ 1 & \mu = \nu, \end{cases} \end{aligned} \tag{1-6}$$

so the Lorentz transformations form a group, the Lorentz group, L. Two Lorentz transformations Λ and M are defined to be close to one another if the numbers $\Lambda^\mu{}_\nu$ and $M^\mu{}_\nu$ are close for all $\mu,\nu = 0,1,2,3$. Clearly, with this definition, Λ^{-1} and ΛM are continuous functions of Λ and M, respectively. Furthermore, it makes sense to say that two Lorentz transformations can be connected to one another by a continuous curve of Lorentz transformations.

L has four components, each of which is connected in the sense that any one point can be connected to any other, but no Lorentz transformation in one component can be connected to another in another component. To see this, note that $\det \Lambda$ and $\operatorname{sgn} \Lambda^0{}_0$ are both continuous functions of the matrix elements $\Lambda^\mu{}_\nu$. Furthermore, $\det \Lambda = \pm 1$ and $\Lambda^0{}_0 \geqslant 1$ or $\leqslant -1$. [The first follows if one takes the determinant of (1–5); the second becomes evident if one looks at the 00 element of (1–5). It reads

$$(\Lambda^0{}_0)^2 - \sum_{j=1}^{3} (\Lambda^j{}_0)^2 = 1.$$

Therefore, $|\Lambda^0{}_0| \geqslant 1$.] Thus, $\det \Lambda$ and $\operatorname{sgn} \Lambda^0{}_0$ must be constant on any one component. The four possibilities are

$$\begin{aligned} L_+^\uparrow&: \quad \det \Lambda = +1, \operatorname{sgn} \Lambda^0{}_0 = +1 \text{ which contains } 1 \\ L_-^\uparrow&: \quad \det \Lambda = -1, \operatorname{sgn} \Lambda^0{}_0 = +1 \text{ which contains } I_s \\ L_+^\downarrow&: \quad \det \Lambda = +1, \operatorname{sgn} \Lambda^0{}_0 = -1 \text{ which contains } I_{st} \\ L_-^\downarrow&: \quad \det \Lambda = -1, \operatorname{sgn} \Lambda^0{}_0 = -1 \text{ which contains } I_t\,. \end{aligned} \tag{1–7}$$

Here, the Lorentz transformations I_s (*space inversion*), I_t (*time inversion*), and I_{st} (*space-time inversion*) are defined by

$$\begin{aligned} (I_s x)^0 &= x^0 & (I_s x)^j &= -x^j, j = 1,2,3 \\ (I_t x)^0 &= -x^0 & (I_t x)^j &= x^j, j = 1,2,3 \\ (I_{st} x) &= -x = (I_s I_t x). \end{aligned} \tag{1–8}$$

Clearly I_s maps $L_-^\uparrow$ one to one onto $L_+^\uparrow$, I_t maps $L_-^\downarrow$ one to one onto $L_+^\uparrow$, and I_{st} maps $L_+^\downarrow$ one to one onto $L_+^\uparrow$. All Λ for which $\Lambda^0{}_0 \geqslant +1$ are called *orthochronous*, Λ for which $\det \Lambda = +1$ *proper*, and Λ for which $\operatorname{sgn} \Lambda^0{}_0 \det \Lambda = +1$ *orthochorous*. To complete the proof of our assertion it has to be shown that $L_+^\uparrow$ is connected. This is customarily done by proving that any $\Lambda \in L_+^\uparrow$ has a decomposition,

$$\Lambda = \Lambda_1 \Lambda_2 \Lambda_3, \tag{1–9}$$

where Λ_1 and Λ_3 are rotations and Λ_2 is a pure Lorentz transformation along the three-axis, defined by

$$x \to \hat{x} = \Lambda_2 x,$$

where

$$\begin{aligned} \hat{x}^0 &= x^0 \cosh \chi + x^3 \sinh \chi \\ \hat{x}^3 &= x^0 \sinh \chi + x^3 \cosh \chi \qquad \tanh \chi = \frac{v}{c} \\ \hat{x}^1 &= x^1, \qquad \hat{x}^2 = x^2. \end{aligned} \tag{1–10}$$

We can then get from any one transformation of the form (1–9) to any other by varying the axes and angles of rotation of Λ_1 and Λ_3 and the parameter χ of Λ_2 continuously. We shall not prove (1–9) here but refer the reader to Ref. 7 of the bibliography of this chapter. This completes the proof that L has just four components. They are displayed in Figure 1–1.

In Figure 1–1 we have also indicated three important subgroups of L:

$$\begin{aligned} L^\uparrow &= L_+^\uparrow \cup L_-^\uparrow && \text{the } \textit{orthochronous Lorentz group} \\ L_+ &= L_+^\uparrow \cup L_+^\downarrow && \text{the } \textit{proper Lorentz group} \\ L_0 &= L_+^\uparrow \cup L_-^\downarrow && \text{the } \textit{orthochorous Lorentz group.} \end{aligned}$$

Associated with the restricted Lorentz group, $L_+^\uparrow$, is the group of 2×2 complex matrices of determinant one, which we shall denote $SL(2, C)$. (S stands for special meaning determinant one, L for linear, 2 is the dimension, and C stands for complex.) It is important in describing the transformation properties of spinors. The connection between $L_+^\uparrow$ and $SL(2, C)$ is obtained in the following standard fashion.

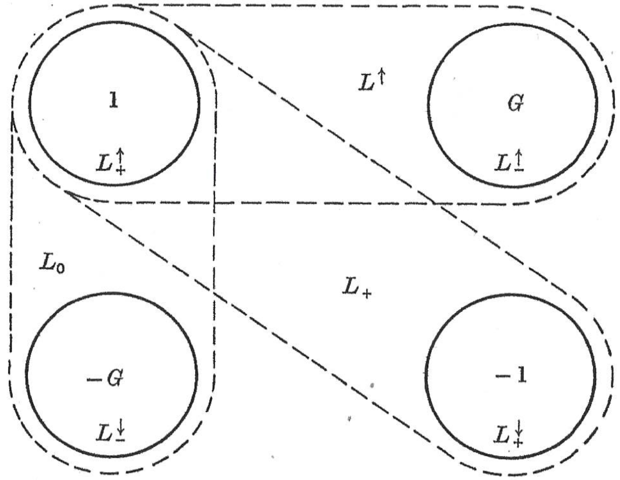

FIGURE 1-1. Connectivity properties of the Lorentz group, L, and its subgroups: the proper Lorentz group, L_+; the orthochronous Lorentz, $L^\uparrow$; the orthochorous Lorentz group, L_0; and the restricted Lorentz group, $L_+^\uparrow$.

If x is any four-vector there is an associated 2×2 matrix given by

$$\underset{\sim}{x} = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} = \sum_{\mu=0}^{3} x^\mu \tau^\mu = x^0 \mathbf{1} + \mathbf{x} \cdot \boldsymbol{\tau}, \tag{1–11}$$

where

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

Conversely, every 2×2 matrix X determines a four-vector via

$$x^\mu = \tfrac{1}{2} \operatorname{tr} (X\tau^\mu) \quad \text{and} \quad X = \underset{\sim}{x}. \tag{1–12}$$

When x^μ is real, $\underset{\sim}{x}$ is hermitian: $\underset{\sim}{x}^* = \underset{\sim}{x}$; if X is hermitian, (1–12) yields a real four-vector. Note that

$$\det \underset{\sim}{x} = x^\mu x_\mu, \qquad \tfrac{1}{2}[\det (\underset{\sim}{x} + \underset{\sim}{y}) - \det \underset{\sim}{x} - \det \underset{\sim}{y}] = x^\mu y_\mu, \tag{1–13}$$

so if A is any 2×2 matrix of determinant 1, then

$$\hat{\underset{\sim}{x}} = A\underset{\sim}{x}A^* \tag{1–14}$$

defines a real linear mapping of four-vectors x onto four-vectors $\hat{x}$ which satisfies $(\hat{x})^\mu(\hat{y})_\mu = x^\mu y_\mu$. It is therefore a Lorentz transformation, which will be denoted by $\Lambda(A)$. Actually, $\Lambda(A)$ is a restricted Lorentz transformation. This follows from a continuity argument: in (1–14) we can vary A continuously until it is the identity. The corresponding $\Lambda(A)$ varies continuously to reach the identity. Therefore, $\Lambda(A) \in L_+^\uparrow$. Clearly, $\Lambda(-A) = \Lambda(A)$. It is left to the reader to verify that if $\Lambda(A) = \Lambda(B)$ then $A = \pm B$. The correspondence $A \to \Lambda(A)$ has the properties

$$\Lambda(A)\Lambda(B) = \Lambda(AB), \qquad \Lambda(1) = 1. \tag{1–15}$$

In other words, $A \to \Lambda(A)$ is a homomorphism of $SL(2, C)$ onto $L_+^\uparrow$.

We shall have to make use later of another correspondence between vectors and 2×2 matrices

$$\tilde{x} = \sum_{\mu=0}^{3} x_\mu \tau^\mu = x^0 \mathbf{1} - \mathbf{x} \cdot \boldsymbol{\tau}. \tag{1–16}$$

Note that

$$A^* \tilde{x} A = \tilde{y} \qquad \text{where } y = \Lambda(A^{-1})x, \tag{1–17}$$

because

$$(\tau^2)(\tau^\mu)^T(\tau^2)^{-1} = g^{\mu\mu}\tau^\mu,$$

with τ^2 given by (1–11). This implies

$$\tau^2 (\underset{\sim}{x})^T (\tau^2)^{-1} = \tilde{x}$$

and

$$\tau^2 A^T (\tau^2)^{-1} = A^{-1} \tag{1–18}$$

for A of determinant 1.

Another group associated with the Lorentz group, L, is the *complex Lorentz group*, which we shall denote by $L(C)$. It is essential in the proof of the PCT theorem as we shall see. It is composed of all complex matrices satisfying (1–5). The argument given before to show that $\det \Lambda = \pm 1$ is also valid for the complex Lorentz group $L(C)$, which therefore has at least two components, $L_\pm(C)$, according to the sign of $\det \Lambda$. Actually, it has just two connected components. The transformations 1 and -1, which are disconnected in L are connected in $L(C)$. Consider for example the continuous curve of $\Lambda^\mu{}_\nu$:

$$\Lambda_t = \begin{Bmatrix} \cosh it & 0 & 0 & \sinh it \\ 0 & \cos t & -\sin t & 0 \\ 0 & \sin t & \cos t & 0 \\ \sinh it & 0 & 0 & \cosh it \end{Bmatrix} \qquad 0 \leqslant t \leqslant \pi.$$

As t varies from 0 to π, this curve connects 1 with -1. The connectivity properties of $L(C)$ are illustrated in Figure 1–2.

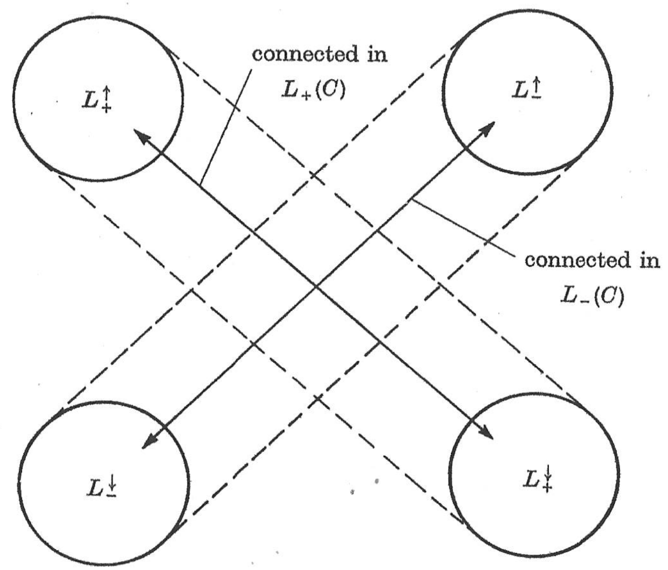

FIGURE 1–2. Connectivity properties of the complex Lorentz group, $L(C)$. There are two connected components: $L_+(C)$, the proper complex Lorentz group, which contains $L_+^\uparrow$ and $L_+^\downarrow$; and $L_-(C)$, which contains $L_-^\uparrow$ and $L_-^\downarrow$.

Just as the restricted Lorentz group, $L_+^\uparrow$, is associated with $SL(2,C)$, the proper complex Lorentz group is associated with $SL(2,C) \otimes SL(2,C)$. This latter group is the set of all pairs of 2×2 matrices of determinant one with the multiplication law

$$\{A_1, B_1\}\{A_2, B_2\} = \{A_1A_2, B_1B_2\}.$$

The connection between the matrix pair $\{A, B\}$ and the corresponding complex Lorentz transformation $\Lambda(A, B)$ is given by the analogue of (1–14).

$$\hat{\underset{\sim}{x}} = A\underset{\sim}{x}B^T. \tag{1–19}$$

Here $\underset{\sim}{x}$ is again defined by (1–11), but the vector x^μ is complex. Clearly

$$\Lambda(A_1, B_1)\Lambda(A_2, B_2) = \Lambda(A_1A_2, B_1B_2) \tag{1–20}$$

and $\Lambda(1, 1) = 1$. It is easy to see that the only matrix pairs which yield a given $\Lambda(A, B)$ are $(\pm A, \pm B)$. In particular,

$$\Lambda(-1, 1) = \Lambda(1, -1) = -1. \tag{1–21}$$

To each of the groups we have considered so far there is a corresponding inhomogeneous group, whose elements are pairs consisting of a translation and a homogeneous transformation. For example, the *Poincaré group*, $\mathscr{P}$, has elements $\{a, \Lambda\}$, where $\Lambda \in L$, with the multiplication law

$$\{a_1, \Lambda_1\}\{a_2, \Lambda_2\} = \{a_1 + \Lambda_1a_2, \Lambda_1\Lambda_2\}, \tag{1–22}$$

as follows from $\{a, \Lambda\}$'s interpretation as a linear transformation

$$x \to \Lambda x + a.$$

$\mathscr{P}$ has four components distinguished by det Λ and sgn $\Lambda^0{}_0$, namely, $\mathscr{P}_+^\uparrow$, $\mathscr{P}_-^\uparrow$, $\mathscr{P}_+^\downarrow$, and $\mathscr{P}_-^\downarrow$. The *complex Poincaré group*, $\mathscr{P}(C)$ admits complex translations but also the multiplication law (1–22). It has two components $\mathscr{P}_\pm(C)$ which are distinguished by det Λ. The inhomogeneous group corresponding to $SL(2,C)$ has no special name; we shall call it the *inhomogeneous* $SL(2,C)$. It consists of pairs $\{a, A\}$, where a is a translation and $A \in SL(2,C)$, and a multiplication law

$$\{a_1, A_1\}\{a_2, A_2\} = \{a_1 + \Lambda(A_1)a_2, A_1A_2\}. \tag{1–23}$$

Sometimes, we shall find it convenient to use the obvious shorthand Aa for $\Lambda(A)a$ and we shall do so in the following. Analogously, the inhomogeneous group of $SL(2,C) \otimes SL(2,C)$ is defined as the set of pairs $\{a, \{A, B\}\}$, where a is a complex translation and $A, B \in SL(2,C)$, and a multiplication law

$$\{a_1, \{A_1, B_1\}\}\{a_2, \{A_2, B_2\}\} = \{a_1 + \Lambda(A_1, B_1)a_2, \{A_1A_2, B_1B_2\}\}. \tag{1–24}$$

In all the following, all the theories considered will be invariant under $\mathscr{P}_+^\uparrow$ or some other subgroup of $\mathscr{P}$. The corresponding unitary or anti-unitary transformations of states will be denoted $U(a, \Lambda)$.

This completes our discussion of the groups themselves. We need in addition some information on the matrix representations of $SL(2,C)$. This is the basic raw material out of which the transformation law of fields will be constructed in Chapter 3.† We recall that an arbitrary matrix representation of $SL(2,C)$, i.e., a correspondence $A \to S(A)$ such that $S(\mathbf{1}) = 1$ and $S(A)S(B) = S(AB)$, can be written in the form

$$T\begin{Bmatrix} S_1(A) & 0 & 0 & \cdots \\ 0 & S_2(A) & 0 & \cdots \\ \vdots & \vdots & & \ddots \end{Bmatrix}T^{-1}, \tag{1–25}$$

where the $A \to S_1(A)$, $A \to S_2(A)$, etc., are irreducible and T is some fixed linear transformation. Thus, we need only describe the irreducible representations.

Consider a set of quantities $\xi_{\alpha_1\ldots\alpha_j\dot\beta_1\ldots\dot\beta_k}$, where the α's and $\dot\beta$'s take the values 1 and 2, and ξ is symmetric under permutations of the α's and also under permutations of the $\dot\beta$'s. For each $A \in SL(2,C)$ we define a linear transformation of the ξ's according to

$$\xi_{\alpha_1\ldots\alpha_j\dot\beta_1\ldots\dot\beta_k} \to \sum_{(\rho)(\dot\sigma)} A_{\alpha_1\rho_1}\ldots A_{\alpha_j\rho_j}\bar{A}_{\beta_1\sigma_1}\ldots\bar{A}_{\beta_k\sigma_k}\xi_{\rho_1\ldots\rho_j\dot\sigma_1\ldots\dot\sigma_k}. \tag{1–26}$$

[The dot over the index simply means that this index transforms according to $\bar{A}$ instead of A; the symbol (ρ) stands for $\rho_1\ldots\rho_j$; the symbol $(\dot\sigma)$ for $\dot\sigma_1\ldots\dot\sigma_k$.] This representation of $SL(2,C)$ is usually denoted $\mathscr{D}^{(j/2,k/2)}$. Every irreducible representation is equivalent to one of these.

If the A are restricted to be unitary as well as of determinant one we get a subgroup SU_2 of $SL(2,C)$. Clearly, SU_2 corresponds via (1–14) to the rotation group of three-space. [The trace of (1–14) reads $2\hat{x}^0 = 2x^0$ if A is unitary.] The irreducible representations of SU_2 are

† The reader who is unfamiliar with the finite-dimensional representations of $SL(2,C)$ might do well to skip the rest of this section on first reading. Readable accounts of these representations are to be found in Refs. 9 and 10. For the complex Lorentz group see also Ref. 7. We have included this material here because we want to treat an arbitrary spinor field in Chapters 3 and 4.