Sorbonne Université
Master NPAC

Quantum Field Theory[1]

Benjamin FUKS

Academic year 2023–2024

[1]Based on material and lectures given by A. Cornell at the University of Johannesburg, A. Deandrea at Master Science de la matière ENS Lyon, and B. Fuks at Master NPAC Sorbonne Université.

Contents

Chapter 1

Special relativity

1.1 Introduction

Quantum field theory combines two of the major pillars of modern physics, special relativity and quantum mechanics. The current chapter is dedicated to the former, and additionally includes some basics about group theory.

Chapter 1 begins in section 1.2 with a few reminders of special relativity, assuming that readers already have a solid understanding of it. We first introduce the notion of four-vectors and Minkowski space-time, that is named as such after the physicist Hermann Minkowski (1864 – 1909) and the work presented in his lecture from 1908 [1], and we then move on with the postulates of special relativity as introduced by Albert Einstein (1879 – 1955) in 1905 [2]. This naturally leads us to the notion of Lorentz and Poincaré transformations that leave the structure of space-time invariant, these two sets of transformations being named after the physicists Hendrik Lorentz (1853 – 1928) and Henri Poincaré (1854 – 1912).

In section 1.3, we focus on Lorentz transformations and their properties [3, 4], and we define objects as Lorentz scalars, vectors and tensors according to how they are modified by Lorentz transformations. After providing some basic and brief knowledge about group theory, we demonstrate that Lorentz transformations form a group, the Lorentz group, that we study in details together with the associated algebra. The representations of this algebra naturally yield, in the context of QFT, information about the spin of the particles.

In section 1.4, we consider coordinate transformations that do not only include a Lorentz transformation component, but also a space-time translation one. This leads us to the Poincaré group and algebra [5], that lie at the cornerstone of modern high-energy physics. We discuss its representations, that allow for a definition of the concept of a particle. We next determine the associated Casimir operators, that are named after the physicist Hendrik Casimir (1909 – 2000), as the eigenvalues of such operators provide a universal way to characterise any representation and therefore label any specific state. We further move on with a study of the Poincaré little group transformations, that form a special set of Poincaré transformations that preserve the four-momentum [6, 7]. Introduced by the physicist Eugene Wigner (1902 – 1995), little groups provide a powerful tool allowing for the classification of particles and fields, which we apply both to the massless and massive case. We demonstrate that any representation of the Poincaré algebra (that we link to particles) is characterised by its mass, and its spin or helicity in the massive and massless case respectively.

1.2 Definitions and relativistic kinematics

In special relativity, the description of space and time is unified into space-time coordinates so that vectors have four components. Moreover, differently from Euclidean space, the scalar product defined in Minkowski space is not positive definite. The standard notation therefore introduces upper and lower indices for vectors and tensors and a metric tensor giving the prescription to contract them. An event E

is represented by a *contravariant four-vector* in space-time (*i.e.* a vector with an upper Lorentz index),

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} , \tag{1.2.1}$$

where $x^0 = ct$ stands for the time of the event and $\mathbf{x} = (x^1, x^2, x^3)$ for its position in a given reference frame $\mathcal{R}$. In the following we adopt the system of units typical of high-energy physics in which the speed of light $c = 1$. Moreover, as Planck's constant is also set to unity ($\hbar = 1$), all quantities get dimensions of mass to some power. Notation-wise, we make use of Greek letters (μ, ν, ρ, *etc.*) ranging from 0 to 3 for space-time indices, and Latin letters (i, j, k, *etc.*) ranging from 1 to 3 for position indices.

Four-vector indices can be raised and lowered by means of the *Minkowski metric* or Minkowski tensor, that is given in our convention by

$$\eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \text{and} \qquad \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{1.2.2}$$

This sign convention $(+,-,-,-)$ is the typical convention used in particle physics, the time component (η^{00}) being associated with a plus sign and the space components (η^{11}, η^{22} and η^{33}) being associated with minus signs. *Covariant four-vectors* (*i.e.* vectors with a lower Lorentz index) and contravariant four-vectors are thus related through

$$x_\mu = \eta_{\mu\nu} x^\nu \equiv \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} \qquad \text{and} \qquad x^\mu = \eta^{\mu\nu} x_\nu \, . \tag{1.2.3}$$

These expressions make use of Einstein summation convention in which any pair of repeated (or contracted) indices is summed. For instance,

$$\eta^{\mu\nu} x_\nu \equiv \sum_{\nu=0}^{3} \eta^{\mu\nu} x_\nu \, . \tag{1.2.4}$$

Moreover, we can easily show that

$$\eta_{\mu\nu}\eta^{\nu\rho} = \delta_\mu{}^\rho \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu{}_\rho \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.2.5}$$

This demonstrates that the Minkowski tensor is its own inverse. The Minkowski tensor is also used to define the scalar product of two four-vectors $x^\mu = (x^0, \mathbf{x})$ and $y^\mu = (y^0, \mathbf{y})$,

$$x \cdot y = x^\mu y_\mu = x_\mu y^\mu = \eta_{\mu\nu}\, x^\mu y^\nu = x^0 y^0 - \mathbf{x} \cdot \mathbf{y} \, , \tag{1.2.6}$$

as well as the (squared) norm of a four-vector,

$$x^2 = x \cdot x = (x^0)^2 - ||\mathbf{x}||^2 \, . \tag{1.2.7}$$

The position of the repeated indices is not important, but it is important that in any pair of repeated indices, one index is an upper index and the other is a lower index.

We can further define derivative operators with respect to space-time coordinates (the upper index in the derivative being related to the lower index of position and *vice versa*),

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \begin{pmatrix} \frac{\partial}{\partial t} \\ \nabla \end{pmatrix} \qquad \text{and} \qquad \partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \begin{pmatrix} \frac{\partial}{\partial t} \\ -\nabla \end{pmatrix} , \tag{1.2.8}$$

as well as the d'Alembert operator $\Box$, sometimes also called the quabla operator (as a reference to the tri-dimensional nabla operator ∇),

$$\Box = \partial_\mu \partial^\mu = \eta^{\mu\nu} \partial_\mu \partial_\nu = \frac{\partial^2}{\partial t^2} - \Delta \, . \tag{1.2.9}$$

The special theory of relativity, or *special relativity*, relies on two fundamental principles postulated by Albert Einstein in his seminal article from 1905. These postulates have consequences that have been verified in countless experiments.

- The laws of nature and the results of experiments are identical in all inertial frames of reference.
- The speed of light in vacuum c is universal.

We recall that an inertial frame of reference is a frame in which Newton's first law of motion is valid. As a consequence of the first postulate of special relativity, physical laws have the same form in any two reference frames that are in relative motion at a constant speed. The second postulate tells us that the value of c is independent of the motion of the luminous source, and that it is identical in all inertial frames of reference. This last property determines the structure of space-time, that can be shown to be a pseudo-Euclidean space (a vector space in which a vector with zero norm can be non-zero, unlike in a Euclidean space) of dimension $D = 4$ equipped with a degenerate scalar product defined by (1.2.6). This space is called *Minkowski space-time*.

The most general transformations that preserve the Minkowskian inner product (1.2.6) are called Poincaré transformations, and they play a special role in high-energy physics. We consider two frames of reference $\mathcal{R}$ and $\mathcal{R}'$ that share a common spatial-temporal origin, and an event E that takes place at a space-time point $x^\mu = (x^0, \mathbf{x})$ in $\mathcal{R}$ and $x'^\mu = (x'^0, \mathbf{x}')$ in $\mathcal{R}'$. If the information between the event E and the spatial-temporal origin of $\mathcal{R}$ and $\mathcal{R}'$ is transmitted by a ray of light, we have

$$x^2 = (x^0)^2 - ||\mathbf{x}||^2 = 0 \qquad \text{and} \qquad x'^2 = (x'^0)^2 - ||\mathbf{x}'||^2 = 0\,, \tag{1.2.10}$$

as a consequence of the speed of light being the same in the two frames of reference. This suggests a definition of a space-time interval Δs between any two events E_1 and E_2, represented by the four-vectors $x^\mu = (x^0, \mathbf{x})$ and $y^\mu = (y^0, \mathbf{y})$, as

$$(\Delta s)^2 = (y^0 - x^0)^2 - ||\mathbf{x} - \mathbf{y}||^2\,. \tag{1.2.11}$$

If $(\Delta s)^2$ is respectively zero, positive or negative in a given inertial frame of reference, it is zero, positive or negative in any inertial reference frame. As this space-time interval is independent of the frame of reference, it can be used to classify events:

- If $(\Delta s)^2 > 0$ the interval is said to be *time-like*. If the norm of an event is positive, it is similarly said to be time-like.
- If $(\Delta s)^2 < 0$ the interval is *space-like*. If the norm of an event is negative, it is similarly said to be space-like.
- If $(\Delta s)^2 = 0$ the interval is *light-like*. If the norm of an event vanishes, the event is similarly said to be light-like.

For a given event E localised in $x^\mu = (x^0, \mathbf{x})$, the set of space-time points $\{E_i \equiv x_i^\mu = (x_i^0, \mathbf{x}_i)\}$ for which $(\Delta s)^2$ is zero forms a cone called a light cone. Its name originates from the fact that it consists of the path that a flash of light emanating from E and traveling in all directions would take through space-time. The events inside the cone are all time-like, with $(\Delta s)^2 > 0$. They form the past $(x^0 > x_i^0)$ and the future $(x^0 < x_i^0)$ of E. On the other hand, any event E_i lying outside the light cone is inaccessible from E, and is causally disconnected from E. The 'distance' between these events and E is too large so that they cannot be connected by a ray of light. Each event thus has its own past and future, and is associated with a set of space-time points for which there is no causal link. This shows that *time is not absolute*.

For two infinitesimally-spaced events, the expression (1.2.11) can be rewritten, using a Cartesian system of coordinates, as

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu\,. \tag{1.2.12}$$

The structure of space-time stems from enforcing that the space-time interval (1.2.12) stays invariant under a change of inertial reference frames from $\mathcal{R}$ to $\mathcal{R}'$. It is equivalent to enforcing that the Minkowskian scalar product (1.2.6) is invariant under such a change of frame of reference. In order to determine the most general set of transformations that preserve $\mathrm{d}s^2$, we start from the most general transformation of coordinates,

$$x^\mu \to x'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} x^\nu\,. \tag{1.2.13}$$

As $\mathrm{d}s^2$ must stay invariant under such a transformation (1.2.13), we have

$$\mathrm{d}s^2 = \mathrm{d}x^\mu \mathrm{d}x^\nu \eta_{\mu\nu} = \mathrm{d}x'^\alpha \mathrm{d}x'^\beta \eta_{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \eta_{\alpha\beta}\,. \tag{1.2.14}$$

This property must be satisfied regardless of $\mathrm{d}x$ so that we get

$$\frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu}\eta_{\alpha\beta} = \eta_{\mu\nu}\,. \tag{1.2.15}$$

We have demonstrated that the most general transformations preserving the Minkowskian metric and the scalar product (1.2.6) are linear in the coordinates. Any such transformation can thus generically be written as

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu\,, \tag{1.2.16}$$

where the 4×4 matrix Λ and four-vector a are the transformation parameters.

In section 1.3, we focus on the first term of (1.2.16) and show that the set of acceptable matrices Λ forms a group known as the *Lorentz group* $O(1,3)$. In section 1.4, we include the second term relevant for space-time translations of a four-vector a^μ, and discuss the resulting *Poincaré group* $ISO(1,3)$. The representations of these two groups are heavily used in high-energy physics and allow in particular for a definition of the concept of particles.

1.3 Lorentz transformations

1.3.1 The Lorentz group

We call *Lorentz transformations* the set of linear change of coordinates,

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu\,, \tag{1.3.1}$$

such that the scalar product (1.2.6) is preserved, with Λ being a real 4×4 matrix. As already mentioned, this equivalently means that the transformation matrix Λ is such that $x\cdot y = x'\cdot y'$ for any two four-vectors x^μ and y^μ. This property in particular implies that the squared norm x^2 of any four-vector is invariant under Lorentz transformations,

$$x'^2 = x'^\mu\,\eta_{\mu\nu}\,x'^\nu = \left(\Lambda^\mu{}_\rho\,x^\rho\right)\eta_{\mu\nu}\left(\Lambda^\nu{}_\sigma\,x^\sigma\right) = \eta_{\rho\sigma}\,x^\rho x^\sigma\,, \tag{1.3.2}$$

after applying (1.3.1) twice. Therefore, any coordinate transformation which satisfies

$$\eta_{\rho\sigma} = \Lambda^\mu{}_\rho\,\eta_{\mu\nu}\,\Lambda^\nu{}_\sigma \tag{1.3.3}$$

is a Lorentz transformation. In order to have compact formulas we can alternatively use a matrix notation defined by

$$G \equiv \eta_{\mu\nu} \qquad \text{and} \qquad \Lambda \equiv \Lambda^\mu{}_\nu\,. \tag{1.3.4}$$

Within this notation $\Lambda^\nu{}_\mu \eta_{\nu\alpha} \equiv \Lambda^t G$ and $G = G^{-1}$, using the explicit form of the metric tensor (1.2.2) and the fact that it is its own inverse as shown in (1.2.5). A Lorentz transformation is then a coordinate transformation

$$x \to x' = \Lambda x\,, \tag{1.3.5}$$

in which the matrix Λ satisfies the condition (1.3.3) that now reads

$$G = \Lambda^t G \Lambda\,. \tag{1.3.6}$$

Moreover, we have

$$Gx' = G\Lambda x = \left(\Lambda^t\right)^{-1} Gx \qquad \text{and} \qquad x^2 = x^T G x\,, \tag{1.3.7}$$

after making use of (1.3.6) and (1.3.2) respectively.

The set of Λ matrices forms *a group*, once we equip it with the usual matrix product as a binary operation (*i.e.* the operation that defines how to determine the ‘product’ of two elements of the set).

- *Closure* – For any two elements of the set Λ_1 and Λ_2, the matrix $\Lambda_1\Lambda_2$ belongs to the set. The condition (1.3.6) is indeed realised for the matrix $\Lambda_1\Lambda_2$ if it is individually realised for the matrices Λ_1 and Λ_2,
$$\Lambda_2^t\Lambda_1^t G\Lambda_1\Lambda_2 = \Lambda_2^t G\Lambda_2 = G\,. \tag{1.3.8}$$
- *Associativity* – For any three elements of the set Λ_1, Λ_2 and Λ_3, we have $(\Lambda_1\Lambda_2)\Lambda_3 = \Lambda_1(\Lambda_2\Lambda_3)$ as matrix multiplication is associative.
- *Identity* – There exists an element I in the set that satisfies $I\Lambda = \Lambda I = \Lambda$ for any element Λ of the set. It consists of the identity matrix $I^\mu{}_\nu = \delta^\mu{}_\nu$.
- *Inverse* – For any element Λ of the set, there exists an element Λ^{-1} such that $\Lambda\Lambda^{-1} = \Lambda^{-1}\Lambda = I$ (see exercise 1.1 for a proof).

The set of matrices Λ preserving the Minkowskian scalar product forms the *indefinite orthogonal group* $O(1,3)$, that is also known as the Lorentz group. These matrices satisfy
$$\eta = \Lambda^t\ \eta\ \Lambda \qquad \Leftrightarrow \qquad \eta_{\mu\nu} = \Lambda^\rho{}_\mu\ \eta_{\rho\sigma}\ \Lambda^\sigma{}_\nu\,. \tag{1.3.9}$$
This leads to
$$\det\Lambda = \pm 1 \qquad \text{and} \qquad \left|\Lambda^0{}_0\right| \geq 1\,, \tag{1.3.10}$$
as well as to the fact that Λ has an inverse matrix Λ^{-1} defined by
$$(\Lambda^{-1})^\mu{}_\nu = \Lambda_\nu{}^\mu\,. \tag{1.3.11}$$

By definition, an indefinite orthogonal group $O(p,q)$ is a group formed by linear transformations of a D-dimensional real vector space. These transformations additionally leave a non-degenerate, symmetric bilinear form of signature (p,q) with $D = p+q$ invariant. The metric (1.2.2) having a signature $(+,-,-,-)$, the Lorentz group corresponds to $O(1,3)$.

Exercise 1.1. We consider a Lorentz transformation of parameter Λ defined by $x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$.

1. Demonstrate that the matrix Λ satisfies
$$\eta_{\mu\nu} = \Lambda^\rho{}_\mu\ \eta_{\rho\sigma}\ \Lambda^\sigma{}_\nu\,, \qquad \det\Lambda = \pm 1 \qquad \text{and} \qquad \left|\Lambda^0{}_0\right| \geq 1\,.$$
2. Demonstrate that any matrix Λ has an inverse Λ^{-1} defined by $(\Lambda^{-1})^\mu{}_\nu = \Lambda_\nu{}^\mu$.
3. Conclude and show that the set of matrices Λ forms a group.

The sign of the determinant of the transformation matrix Λ allows for a classification of all Lorentz transformations as *proper* Lorentz transformations (with $\det\Lambda = 1$) and *improper* Lorentz transformations (with $\det\Lambda = -1$). In particular, space inversion P (or parity) is a special class of improper Lorentz transformations, with a transformation matrix $P^\mu{}_\nu$ given by
$$P^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\,. \tag{1.3.12}$$
Under the action of parity, any specific four-vector x^μ is transformed as
$$x^\mu = \begin{pmatrix} x^0\\ x^1\\ x^2\\ x^3\end{pmatrix} \to x'^\mu = P^\mu{}_\nu\ x^\nu = \begin{pmatrix} x^0\\ -x^1\\ -x^2\\ -x^3\end{pmatrix}\,. \tag{1.3.13}$$
Space inversion further allows any improper Lorentz transformation to be made proper as the product of two improper transformations is proper. If a given transformation Λ is improper, then $P\cdot\Lambda$ is indeed proper.

On the other hand, the sign of the $\Lambda^0{}_0$ parameter allows for a classification of all Lorentz transformations as *orthochronous* Lorentz transformations ($\Lambda^0{}_0 > 1$ so that the direction of time is conserved) and *non-orthochronous* Lorentz transformations ($\Lambda^0{}_0 < -1$ so that the direction of time is reversed). Time reversal T consists of a particular non-orthochronous Lorentz transformation, with a transformation matrix $T^\mu{}_\nu$ given by

$$T^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.3.14}$$

Under the action of time reversal, any specific four-vector x^μ is transformed as

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \to x'^\mu = T^\mu{}_\nu \, x^\nu = \begin{pmatrix} -x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} . \tag{1.3.15}$$

It further allows any non-orthochronous Lorentz transformation to be transformed to an orthochronous transformation, as if Λ is non-orthochronous, then $T \cdot \Lambda$ is orthochronous.

The set of proper and orthochronous Lorentz transformations forms a subgroup of $O(1,3)$ denoted by $SO_0(1,3)$, often equivalently written as $SO^+(1,3)$ and called the *restricted Lorentz group*. It is the identity component of the Lorentz group, and therefore includes all Lorentz transformations that can be connected to the identity by a continuous curve lying in the group. Any product of T, P and $SO_0(1,3)$ transformations is therefore equal to a transformation that is part of the whole Lorentz group. Conversely, it turns out that the set of these products actually saturates the Lorentz group, which can be symbolically written as

$$\text{Lorentz group} \quad = \quad SO_0(1,3) \text{ transformations} \quad + \quad T \quad + \quad P \, .$$

However, whereas any $SO_0(1,3)$ transformation can be continuously connected to the identity I, space inversion P and time-reversal T cannot (their determinant is different from 1). Those two discrete symmetries nevertheless play a special role in QFT, even though they cannot be written in terms of proper and orthochronous Lorentz transformations.

The restricted Lorentz group $SO_0(1,3)$ exhibits the structure of a *real Lie group*, named after the Norwegian mathematician Sophus Lie (1842 – 1899). Lie groups consist of a special class of groups related to continuous symmetries (like for example rotations or Lorentz transformations) that contain an infinite number of elements that can be derived from a *finite-dimensional set of generators*. Lie groups are an incontrovertible part of high-energy physics, as they are critical to the understanding of the fundamental interactions. As the restricted Lorentz group $SO_0(1,3)$ transformations are continuously connected to the identity, any $SO_0(1,3)$ transformation Λ can be generically written, using matrix notation, as

$$\Lambda = \exp\big[i\vartheta\lambda\big] \, . \tag{1.3.16}$$

The quantity ϑ is a real constant, the matrix λ is the generator of the $SO_0(1,3)$ transformation considered and the factor of i is there by convention. In fact, a rewriting, such as that in (1.3.16), is a general property of any continuous group.

If the ϑ parameter is infinitesimally small ($\vartheta \to \varepsilon$), we can expand (1.3.16) at $\mathcal{O}(\varepsilon^2)$. This leads to

$$\Lambda \simeq 1 + i\varepsilon\lambda \, . \tag{1.3.17}$$

With this in mind, the relation (1.3.6) defining a Lorentz transformation can be written as

$$\big(1 + i\varepsilon\lambda^t\big) \, G \, (1 + i\varepsilon\lambda) = G \, , \tag{1.3.18}$$

which implies, to first order in ε,

$$\lambda^t G + G\lambda = 0 \quad \Leftrightarrow \quad \lambda^t = -G\lambda G \, . \tag{1.3.19}$$

This last equation can be written with all indices made explicit as

$$\begin{pmatrix} \lambda^0{}_0 & \lambda^1{}_0 & \lambda^2{}_0 & \lambda^3{}_0 \\ \lambda^0{}_1 & \lambda^1{}_1 & \lambda^2{}_1 & \lambda^3{}_1 \\ \lambda^0{}_2 & \lambda^1{}_2 & \lambda^2{}_2 & \lambda^3{}_2 \\ \lambda^0{}_3 & \lambda^1{}_3 & \lambda^2{}_3 & \lambda^3{}_3 \end{pmatrix} = \begin{pmatrix} -\lambda^0{}_0 & \lambda^0{}_1 & \lambda^0{}_2 & \lambda^0{}_3 \\ \lambda^1{}_0 & -\lambda^1{}_1 & -\lambda^1{}_2 & -\lambda^1{}_3 \\ \lambda^2{}_0 & -\lambda^2{}_1 & -\lambda^2{}_2 & -\lambda^2{}_3 \\ \lambda^3{}_0 & -\lambda^3{}_1 & -\lambda^3{}_2 & -\lambda^3{}_3 \end{pmatrix} . \tag{1.3.20}$$

This relation has important consequence on the form of the generators of restricted Lorentz transformations. Recalling that Latin indices i and j range from 1 to 3, this gives the following three properties:

- ☐ All diagonal elements are zero, $\lambda^0{}_0 = \lambda^1{}_1 = \lambda^2{}_2 = \lambda^3{}_3 = 0$.
- ☐ $\lambda_{i0} = \lambda_{0i}$ for all values of $i = 1,\, 2,\, 3$. The λ matrix is thus symmetric in three of its elements.
- ☐ $\lambda_{ij} = -\lambda_{ji}$ for $i \neq j$. The λ matrix is thus antisymmetric in three of its elements.

A basis for all the Lorentz transformation generators thus have six independent elements. The standard choice for these six generators is to take all $\lambda^\mu{}_\nu$ parameters equal to zero, with the exception of one of them, that is conventionally fixed to $-i$. The reality condition of the transformation parameters ε and that of the transformation matrix Λ, taken together with the explicit factor of i included in (1.3.16), justify this choice. Using a standard notation $(J^{\alpha\beta})^\mu{}_\nu$ for the generator symbols (this choice of index structure will become clear below) gives

$$
(J^{01})^\mu{}_\nu = -i\begin{pmatrix} 0&1&0&0\\ 1&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix}, \quad
(J^{02})^\mu{}_\nu = -i\begin{pmatrix} 0&0&1&0\\ 0&0&0&0\\ 1&0&0&0\\ 0&0&0&0 \end{pmatrix}, \quad
(J^{03})^\mu{}_\nu = -i\begin{pmatrix} 0&0&0&1\\ 0&0&0&0\\ 0&0&0&0\\ 1&0&0&0 \end{pmatrix},
$$
$$
(J^{12})^\mu{}_\nu = -i\begin{pmatrix} 0&0&0&0\\ 0&0&-1&0\\ 0&1&0&0\\ 0&0&0&0 \end{pmatrix}, \quad
(J^{23})^\mu{}_\nu = -i\begin{pmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&-1\\ 0&0&1&0 \end{pmatrix}, \quad
(J^{31})^\mu{}_\nu = -i\begin{pmatrix} 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0\\ 0&-1&0&0 \end{pmatrix}. \tag{1.3.21}
$$

Alternatively, these definitions can be written in a compact form as

$$
\left(J^{\alpha\beta}\right)^\mu{}_\nu = -i\left(\eta^{\alpha\mu}\delta^\beta{}_\nu - \eta^{\beta\mu}\delta^\alpha{}_\nu\right). \tag{1.3.22}
$$

Whilst such a definition naively leads to 16 possible matrices $J^{\alpha\beta}$, only six of them are independent by virtue of the antisymmetric property $J^{\alpha\beta} = -J^{\beta\alpha}$. These six independent matrices are those given in (1.3.21). The generators shown in the first line of that equation are the so-called boost generators, while those in the second line are the rotation generators. This will be clarified in exercise 1.2.

Starting from (1.3.17), we can rewrite the most general infinitesimal $SO_0(1,3)$ transformation as

$$
\Lambda = 1 + i\xi_i J^{0i} + \frac{i}{2}\epsilon_{ij}{}^k\vartheta_k J^{ij}, \tag{1.3.23}
$$

with ξ_i and ϑ_i being the infinitesimal parameters of the transformation. This relation involves the *Levi-Civita tensor* or *totally anti-symmetric tensor* $\varepsilon_{ij}{}^k$. The elements of this tensor are defined from $\varepsilon_{12}{}^3 = 1$, all other elements being deduced from the rule indicating that the sign changes under the swap of any two indices (any element with two identical indices is thus equal to zero). By exponentiation we obtain expressions for finite Lorentz transformations,

$$
\Lambda = \exp\left[i\xi_i J^{0i} + \frac{i}{2}\epsilon_{ij}{}^k\vartheta_k J^{ij}\right]. \tag{1.3.24}
$$

Any element Λ of the Lorentz group can hence be cast in the compact form

$$
\Lambda = \exp\left[\frac{i}{2}\omega_{\alpha\beta}J^{\alpha\beta}\right]. \tag{1.3.25}
$$

This compact notation makes use of the fact that the transformation parameters ω are antisymmetric under the exchange of their indices, $\omega_{\alpha\beta} = -\omega_{\beta\alpha}$, whose proof is part of exercise 1.4. This property originates from the constraint (1.3.3) (or (1.3.6) in matrix notation). Moreover, the ω parameters are real, as we have shown.

As a consequence there are only six ω parameters to be fixed, in agreement with (1.3.23), and there are accordingly only six relevant matrices $J^{\alpha\beta}$, that thus represent the six independent generators of the group. This is not surprising, as any given real matrix $\Lambda^\mu{}_\nu$ has 16 entries that are constrained by the 10 independent relations included in (1.3.9). We are thus left with six degrees of freedom, which is the greatest strength of the expression (1.3.25). It is indeed completely general, and any element of the restricted Lorentz group can be written uniquely in this form. To this aim, it is sufficient to provide the six real numbers ω_{01}, ω_{02}, ω_{03}, ω_{12}, ω_{23} and ω_{31}.

The simplest examples of Lorentz transformations consist of *rotations*. For example, assuming a frame of reference expressed in terms of Cartesian coordinates, we consider a rotation of angle θ_z around the Oz axis. The associated Lorentz transformation matrix $R_3(\theta_z)$ reads

$$R_3(\theta_z) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_z & -\sin\theta_z & 0 \\ 0 & \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{1.3.26}$$

which obeys the property (1.3.9). This shows that the coordinates transform as

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \to x'^\mu = \big(R_3(\theta_z)\big)^\mu{}_\nu \, x^\nu = \begin{pmatrix} x^0 \\ x^1\cos\theta_z - x^2\sin\theta_z \\ x^1\sin\theta_z + x^2\cos\theta_z \\ x^3 \end{pmatrix} . \tag{1.3.27}$$

The expression (1.3.26) can be retrieved from (1.3.25) once we fix $\omega_{12} = -\omega_{21} = \theta_z$, and take all other parameters $\omega_{\alpha\beta}$ as vanishing. This indeed gives

$$R_3(\theta_z) = \exp\Big[\frac{i}{2}\theta_z\big(J^{12} - J^{21}\big)\Big] = \exp\Big[i\theta_z J^{12}\Big] . \tag{1.3.28}$$

Similarly, we can show that the two other basic rotations of angles θ_x and θ_y, around the axes Ox and Oy respectively, are related to the generators J^{23} and J^{31}.

Instead of mixing two of the spatial coordinates, we can define transformations mixing the temporal coordinate x^0 with one of the three spatial coordinates x^i (with $i = 1,2,3$) of a four-vector. This defines what we call the three *Lorentz boosts* in the Ox, Oy and Oz directions. Starting from the spatial interval (1.2.12), we observe that a rotation around the Oz axis leaves the quantity $\mathrm{d}x^2 + \mathrm{d}y^2$ unchanged. Similarly, a Lorentz boost in the Ox direction would leave the quantity $\mathrm{d}t^2 - \mathrm{d}z^2$ invariant. This suggests that the corresponding Lorentz transformation matrix be written with hyperbolic functions rather than trigonometric functions,

$$B_3(\varphi_z) = \begin{pmatrix} \cosh\varphi_z & 0 & 0 & \sinh\varphi_z \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\varphi_z & 0 & 0 & \cosh\varphi_z \end{pmatrix} , \tag{1.3.29}$$

where the 'angle' φ_z is called the *rapidity* and can take any real value. Acting on a four-vector x^μ, this yields

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \to x'^\mu = \big(B_3(\varphi_z)\big)^\mu{}_\nu \, x^\nu = \begin{pmatrix} x^0\cosh\varphi_z + x^3\sinh\varphi_z \\ x^1 \\ x^2 \\ x^0\sinh\varphi_z + x^3\cosh\varphi_z \end{pmatrix} . \tag{1.3.30}$$

Rapidities can be connected to the speed $\beta = v/c$ ($= v$ in our conventions with $c = 1$) of an inertial frame of reference $\mathcal{R}'$ in translation (at a constant speed in the Oz direction) with respect to another frame $\mathcal{R}$,

$$\tanh\varphi_z = \beta , \tag{1.3.31}$$

such that

$$\sinh\varphi_z = \beta\gamma \qquad \text{and} \qquad \cosh\varphi_z = \gamma , \tag{1.3.32}$$

with

$$\text{with} \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}} . \tag{1.3.33}$$

Those relations show that if an observer O observes an event of coordinates x in the frame $\mathcal{R}$, then an observer O' would observe the same event with coordinates x' in $\mathcal{R}'$. The coordinates x' in $\mathcal{R}'$ can be calculated from the coordinates x in $\mathcal{R}$ with the above formulas. We can relate the boost $B_3(\varphi_z)$ to the generators of the Lorentz algebra J^{03} and J^{30}, as in (1.3.28),

$$B_3(\varphi_z) = \exp\left[\frac{i}{2}\varphi_z\left(J^{03} - J^{30}\right)\right] = \exp\left[i\varphi_z J^{03}\right] . \tag{1.3.34}$$

This time, the only non-vanishing transformation parameters appearing in (1.3.25) are $\omega_{03} = -\omega_{30} = \varphi_z$, as the boost considered relates the temporal and the third spatial component of a four-vector.

Exercise 1.2. In this exercise, we study the relations between the generators $J^{\alpha\beta}$ of the Lorentz group and finite Lorentz transformations Λ.

1. Demonstrate that the Lorentz transformation $\Lambda_1 = \exp\left[i\varphi J^{03}\right]$ is a boost of rapidity φ along the Oz axis.
2. Demonstrate that the Lorentz transformation $\Lambda_2 = \exp\left[i\theta J^{12}\right]$ is a rotation of angle θ around the Oz axis.

To summarise, any Lorentz transformation can be written in terms of the three elementary rotations around the Ox, Oy and Oz axes, the three elementary boosts in the Ox, Oy and Oz direction, and the discrete transformations P and T. These eight basic transformations saturate the Lorentz group $O(1,3)$, whereas boosts and rotations saturate its subgroup $SO_0(1,3)$. Symbolically, we have

$$\text{Lorentz group} \quad = \quad \text{rotations} \quad + \quad \text{boosts} \quad + \quad T \quad + \quad P \, .$$

Up to now, we have only investigated the action of Lorentz transformations on four-vectors. Not all quantities, however, transform in this way. We call *scalar* or *Lorentz-invariant* quantities expressions f that are invariant under Lorentz transformations,

$$f \to f \, , \tag{1.3.35}$$

and that do not depend on the choice of the frame of reference. For instance, the scalar product of two four-vectors $x \cdot y = x_\mu y^\mu$ is invariant under Lorentz transformations, as for any expression in which all Lorentz indices are contracted. In contrast, quantities with free Lorentz indices are said to be *Lorentz covariant* and change with the choice of the frame of reference. Objects V^μ in which one Lorentz index is free are *four-vectors* or simply *vectors* under Lorentz transformation, and they transform as in (1.3.1),

$$V^\mu \to V'^\mu = \Lambda^\mu{}_\nu V^\nu \, . \tag{1.3.36}$$

Important examples include the space-time position x^μ, the derivative operator ∂^μ and the four-momentum p^μ defined, in a system of Cartesian coordinates, by

$$p^\mu = \begin{pmatrix} E \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} E \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} , \tag{1.3.37}$$

with E standing for the energy and $\mathbf{p} = (p^1, p^2, p^3)$ for the usual tri-dimensional momentum (actually $c\,\mathbf{p}$ after reinstating the appropriate factor of c). Lorentz-covariant expressions $T^{\mu_1\mu_2\ldots\mu_n}$ can carry more than a single free Lorentz index, and they are in this case generically called *tensors*. The number of free indices denotes the *rank* of the tensor. Lorentz tensors transform as

$$T^{\mu_1\mu_2\ldots\mu_n} \to T'^{\mu_1\mu_2\ldots\mu_n} = \Lambda^{\mu_1}{}_{\nu_1}\Lambda^{\mu_2}{}_{\nu_2}\ldots\Lambda^{\mu_n}{}_{\nu_n} T^{\nu_1\nu_2\ldots\nu_n} \, , \tag{1.3.38}$$

hence generalising the transformation law (1.3.36).

Exercise 1.3. In 1987 the supernova SN1987A exploded in a nearby galaxy, the Large Magellanic Cloud. Two neutrino detectors, one at Brookhaven in the US and one at Kamiokande in Japan, detected neutrino bursts that could be used to set an upper bound on the neutrino mass.

We consider the Brookhaven events in which the earliest neutrinos detected had an energy $E_1 \simeq 38$ MeV, while the latest ones had an energy $E_2 \simeq 22$ MeV with a difference in arrival times of about 5 seconds. The distance of the Large Magellanic Cloud is $L = 50$ kiloparsec (1.543×10^{21} m). As these neutrinos were likely produced at the same time, the most energetic ones should have travelled faster to us. Use the observed time delay to establish an upper limit on the neutrino mass.

1.3.2 The Lorentz algebra

In our study of the (restricted) Lorentz group $SO_0(1,3)$ and its elements, we have introduced its six generators given, in the *vectorial representation*, by (1.3.21). The name 'vectorial representation' comes from the fact that the transformation matrices that we studied act on four-vectors. Very importantly, we have shown that any element Λ of the group can be uniquely determined by providing six real numbers, as shown in (1.3.25).

By virtue of the properties of the elements of the Lorentz group, the matrices $J^{\alpha\beta}$ form a *Lie algebra*, that we denote $\mathfrak{so}(1,3)$, that we will investigate in the current subsection.

We consider the six generators $J^{\alpha\beta}$ of the Lorentz algebra $\mathfrak{so}(1,3)$ that act on four-vectors. We thus focus on what we call the *vectorial representation* of the algebra. These generators are defined by

$$\left(J^{\alpha\beta}\right)^{\mu}{}_{\nu} = -i\Big(\eta^{\alpha\mu}\delta^{\beta}{}_{\nu} - \eta^{\beta\mu}\delta^{\alpha}{}_{\nu}\Big)\,, \tag{1.3.39}$$

and they satisfy the commutation relations

$$\Big[J^{\alpha\beta}, J^{\gamma\delta}\Big] = i\Big(\eta^{\beta\gamma}J^{\alpha\delta} - \eta^{\alpha\gamma}J^{\beta\delta} + \eta^{\delta\beta}J^{\gamma\alpha} - \eta^{\delta\alpha}J^{\gamma\beta}\Big)\,. \tag{1.3.40}$$

In the context of Lie algebras, the rule defining the multiplication between two elements is called a *Lie bracket*. With the matrix representation considered so far, the Lie bracket is equivalent to a commutator, as for example given in (1.3.40). As any element of a Lie algebra can be written as a linear combination of the generators, the commutation relations between the generators are sufficient to uniquely define the algebra.

We have so far focused on the vectorial representation of the group, since the matrices Λ have been introduced as acting on four-vectors. By definition, generators in any other representation must satisfy the same relations (1.3.40), and may act on objects different from four-vectors. For instance, the generalisation of the quantum mechanical orbital momentum operator to the relativistic case leads to the operators $L^{\mu\nu}$ defined by

$$L^{\mu\nu} = i\Big(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu}\Big)\,. \tag{1.3.41}$$

They consist of an infinite-dimensional representation of the Lorentz group that acts on functions of the space-time coordinates. The operators $L^{\mu\nu}$ indeed satisfy commutation relations similar to those of (1.3.40),

$$\Big[L^{\alpha\beta}, L^{\gamma\delta}\Big] = i\Big(\eta^{\beta\gamma}L^{\alpha\delta} - \eta^{\alpha\gamma}L^{\beta\delta} + \eta^{\delta\beta}L^{\gamma\alpha} - \eta^{\delta\alpha}L^{\gamma\beta}\Big)\,. \tag{1.3.42}$$

In the next chapters, we additionally consider the spinorial representations of the Lorentz algebra, that are used to describe fermions. The generators will be different, and act on the elements of a different vector space.

Exercise 1.4. We consider an infinitesimal Lorentz transformation connected to the identity,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \varepsilon^\mu{}_\nu\,,$$

where the matrix ε is an infinitesimal quantity.

1. Demonstrate that $\varepsilon^\mu{}_\nu$ can be written as

$$\varepsilon^\mu{}_\nu = \frac{i}{2}\omega_{\alpha\beta}(J^{\alpha\beta})^\mu{}_\nu \qquad \text{with} \qquad (J^{\alpha\beta})^\mu{}_\nu = -i\Big(\eta^{\alpha\mu}\delta^\beta{}_\nu - \eta^{\beta\mu}\delta^\alpha{}_\nu\Big)\,.$$

In these expressions, $\omega_{\alpha\beta}$ is real and antisymmetric under the exchange of its α and β indices, and the generators of the Lorentz algebra $J^{\alpha\beta}$ satisfy the commutation relation

$$\Big[J^{\alpha\beta}, J^{\gamma\delta}\Big] = i\Big(\eta^{\beta\gamma}J^{\alpha\delta} - \eta^{\alpha\gamma}J^{\beta\delta} + \eta^{\delta\beta}J^{\gamma\alpha} - \eta^{\delta\alpha}J^{\gamma\beta}\Big)\,.$$

2. We now consider n infinitesimal Lorentz transformations of parameter $\omega_{\alpha\beta}/n$. Show that the matrix Λ defined by

$$\Lambda = \exp\left[\frac{i}{2}\omega_{\alpha\beta}J^{\alpha\beta}\right]$$

belongs to the restricted Lorentz group $SO_0(1,3)$, and that it hence represents a proper and orthochronous Lorentz transformation.

It is now time to investigate further the Lorentz algebra (1.3.40) in order to characterise the associated representations. To achieve this, we define $J^i \equiv J^{jk}$ with (i,j,k) being a circular permutation of $(1,2,3)$, and $K^i \equiv J^{0i}$. In this notation, we can write the $J^{\alpha\beta}$ and $J_{\alpha\beta}$ tensors as

$$J^{\alpha\beta} = \begin{pmatrix} 0 & K^1 & K^2 & K^3 \\ -K^1 & 0 & J^3 & -J^2 \\ -K^2 & -J^3 & 0 & J^1 \\ -K^3 & J^2 & -J^1 & 0 \end{pmatrix} \qquad \text{and} \qquad J_{\alpha\beta} = \begin{pmatrix} 0 & -K^1 & -K^2 & -K^3 \\ K^1 & 0 & J^3 & -J^2 \\ K^2 & -J^3 & 0 & J^1 \\ K^3 & J^2 & -J^1 & 0 \end{pmatrix}\,. \tag{1.3.43}$$

Each element of these tensors is itself a 4×4 matrix. It is useful to simplify the Lorentz algebra (1.3.40) to

$$\big[J^i, J^j\big] = iJ^k\,, \qquad \big[J^i, K^j\big] = iK^k \qquad \text{and} \qquad \big[K^i, K^j\big] = -iJ^k\,, \tag{1.3.44}$$

with any triplet (i,j,k) being a circular permutation of $(1,2,3)$. Alternatively, these relations can be written by means of the totally anti-symmetric tensor $\varepsilon^{ij}{}_k$ defined from $\varepsilon^{12}{}_3 = 1$,

$$\big[J^i, J^j\big] = i\varepsilon^{ij}{}_kJ^k\,, \qquad \big[J^i, K^j\big] = i\varepsilon^{ij}{}_kK^k \qquad \text{and} \qquad \big[K^i, K^j\big] = -i\varepsilon^{ij}{}_kJ^k\,. \tag{1.3.45}$$

We next introduce the pair of conjugate generators N^i and $\bar N^i$ defined by

$$N^i = \frac{1}{2}\Big(J^i + iK^i\Big) \qquad \text{and} \qquad \bar N^i = \frac{1}{2}\Big(J^i - iK^i\Big)\,. \tag{1.3.46}$$

By doing so, we have made the Lorentz algebra $\mathfrak{so}(1,3)$, that consists of a real vector space, to become $\mathfrak{so}(1,3)_\mathbb{C} = \mathfrak{so}(1,3)\times\mathbb{C}$, that is now a complex vector space. The definitions (1.3.46) allow us to rewrite (1.3.44) as

$$\big[N^i, N^j\big] = iN^k\,, \qquad \big[\bar N^i, \bar N^j\big] = i\bar N^k \qquad \text{and} \qquad \big[N^i, \bar N^j\big] = 0\,, \tag{1.3.47}$$

with (i,j,k) being again a circular permutation of $(1,2,3)$. These relations can be alternatively rewritten by means of the totally anti-symmetric tensor $\varepsilon^{ij}{}_k$ defined from $\varepsilon^{12}{}_3 = 1$. This yields

$$\big[N^i, N^j\big] = i\varepsilon^{ij}{}_kN^k\,, \qquad \big[\bar N^i, \bar N^j\big] = i\varepsilon^{ij}{}_k\bar N^k \qquad \text{and} \qquad \big[N^i, \bar N^j\big] = 0\,, \tag{1.3.48}$$

which indicates that the Lorentz algebra has two commuting sub-algebras.

Exercise 1.5. Demonstrate that the commutation relations satisfied by the generator $J^{\alpha\beta}$ of the Lorentz algebra,

$$\Big[J^{\alpha\beta}, J^{\gamma\delta}\Big] = i\Big(\eta^{\beta\gamma}J^{\alpha\delta} - \eta^{\alpha\gamma}J^{\beta\delta} + \eta^{\delta\beta}J^{\gamma\alpha} - \eta^{\delta\alpha}J^{\gamma\beta}\Big)\,,$$

can be rewritten as

$$\big[N^i, N^j\big] = iN^k\,, \qquad \big[\bar{N}^i, \bar{N}^j\big] = i\bar{N}^k \qquad \text{and} \qquad \big[N^i, \bar{N}^j\big] = 0\,.$$

In these last expressions, (i, j, k) stands for a circular permutation of $(1, 2, 3)$ and the generators N and $\bar{N}$ are linear combinations of the rotation generators $J^i \equiv J^{jk}$ and the boost generators $K^i \equiv J^{0i}$,

$$N^i = \frac{1}{2}\Big(J^i + iK^i\Big) \qquad \text{and} \qquad \bar{N}^i = \frac{1}{2}\Big(J^i - iK^i\Big)\,.$$

From the results in (1.3.47) or (1.3.48), we can deduce that the generators N and $\bar{N}$ independently satisfy the same well-known Lie algebra, that of tri-dimensional rotations. However, strictly speaking this does not consist of the algebra $\mathfrak{so}(3)$ as we had to make the vector space complex through the definitions (1.3.46). The algebra $\mathfrak{sl}(2,\mathbb{C})$ is nevertheless in a one-to-one correspondence with $\mathfrak{so}(3)$.

We thus demonstrated that $\mathfrak{so}(1,3)_\mathbb{C} \sim \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$, which is hence in one-to-one correspondence with $\mathfrak{so}(3) \oplus \mathfrak{so}(3)$. The complex Lorentz algebra $\mathfrak{so}(1,3)_\mathbb{C}$ is thus equivalent to two independent Lie algebras, $\mathfrak{sl}(2,\mathbb{C})$. Recalling that the generators N and $\bar{N}$ are complex conjugates of each other, we deduce that the (real) Lorentz algebra $\mathfrak{so}(1,3) \sim \mathfrak{sl}(2,\mathbb{R}) \oplus \overline{\mathfrak{sl}(2,\mathbb{R})}$.

The commutation relations

$$\big[N^i, N^j\big] = iN^k\,, \qquad \big[\bar{N}^i, \bar{N}^j\big] = i\bar{N}^k \qquad \text{and} \qquad \big[N^i, \bar{N}^j\big] = 0\,,$$

show that the study of the representations of the Lorentz algebra is similar to the (simpler) study of the finite-dimensional representations of $\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$ (or equivalently of $\mathfrak{so}(3) \oplus \mathfrak{so}(3)$).

We can therefore characterise any given representation of the Lorentz algebra by a couple of numbers (j_1, j_2), with j_1 and j_2 being either integer or half-integer. These two numbers refer to the representation under each of the $\mathfrak{sl}(2,\mathbb{C})$ algebras. For a given representation, the associated generators N^i and $\bar{N}^i$ then act on matrices of dimension $2j_1 + 1$ and $2j_2 + 1$ respectively. Such a representation has thus $(2j_1 + 1)(2j_2 + 1)$ degrees of freedom.

Since $\mathfrak{so}(3)$ is a sub-algebra of the Lorentz algebra, any finite-dimensional representation of the Lorentz algebra is also a representation of $\mathfrak{so}(3)$, which provides a handle on the spin of the particles (this will be further discussed in the next section). However, spins must be combined vectorially in quantum mechanics so that a given representation of the Lorentz algebra (j_1, j_2) generates many representations of $\mathfrak{so}(3)$ with spins $|j_1 - j_2|, |j_1 - j_2| + 1, \ldots, j_1 + j_2 - 1, j_1 + j_2$. For example, the representation $(1/2, 1/2)$ of the Lorentz algebra has four degrees of freedom, and it is the one that acts on real four-vectors. From what is mentioned above, such a representation can describe both spin-0 and spin-1 representations of $\mathfrak{so}(3)$, with one and three degrees of freedom respectively.

1.4 The Poincaré algebra and group

As mentioned in section 1.2, the most general transformations that preserve the space-time interval (1.2.12) has the structure (1.2.16), and is thus linear. The set of all such transformations form the so-called Poincaré group $ISO(1, 3)$, which includes the Lorentz group discussed in section 1.3 and space-time translations. We recall that under a Poincaré transformation, a four-vector transforms as

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu\,, \tag{1.4.1}$$

where a^μ corresponds to a space-time translation and $\Lambda^\mu{}_\nu$ to a Lorentz transformation preserving the metric $\eta^{\mu\nu}$. Such a transformation is denoted by (Λ, a). By introducing the generators $J^{\mu\nu}$ of Lorentz transformations in the vectorial representation defined in (1.3.39) and the generators P^μ of space-time translations (that will be more precisely defined below), any element of the Poincaré group can be written as

$$(\Lambda, a) = \exp\left[\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} + i\varepsilon_\mu P^\mu\right], \tag{1.4.2}$$

where $\omega_{\mu\nu}$ and ε_μ represent the parameters of the transformation. This expression generalises (1.3.25) when space-time translations are included.

We consider the six generators $J^{\mu\nu}$ of the Lorentz algebra $\mathfrak{so}(1,3)$ and the four generators of space-time translations P^μ. They satisfy the commutation relations

$$\begin{aligned}
\Big[J^{\mu\nu}, J^{\rho\sigma}\Big] &= i\Big(\eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} + \eta^{\sigma\nu}J^{\rho\mu} - \eta^{\sigma\mu}J^{\rho\nu}\Big),\\
\Big[J^{\mu\nu}, P^{\rho}\Big] &= i\Big(\eta^{\nu\rho}P^{\mu} - \eta^{\mu\rho}P^{\nu}\Big),\\
\Big[P^{\mu}, P^{\nu}\Big] &= 0.
\end{aligned} \tag{1.4.3}$$

These relations define the Poincaré algebra $\mathfrak{iso}(1,3)$.

Whilst we have focused on the vectorial representation of the Poincaré algebra thus far, the transformation (1.4.2) acting on four-vectors, the commutation relations above can be applied to any representation of the Poincaré group.

Exercise 1.6. In this exercise, we derive the Poincaré algebra from the properties of Poincaré transformations.

1. Show that the set of Poincaré transformations (Λ, a) forms a group.
2. Rewrite the combination of three Poincaré transformations

$$(\Lambda, a)\left(1 + \frac{1}{2}\omega_{\mu\nu}J^{\mu\nu}, \varepsilon_\mu P^\mu\right)(\Lambda, a)^{-1}$$

   as a single Poincaré transformation. In the above expression, $\omega^\mu{}_\nu$ and ε^μ are real and infinitesimal parameters, $J^{\mu\nu}$ are the generators of the Lorentz algebra in the vectorial representation, and P^μ are the generators of space-time translations.
3. Simplify the quantities

$$(\Lambda, a)J^{\mu\nu}(\Lambda, a)^{-1} \qquad \text{and} \qquad (\Lambda, a)P^{\mu}(\Lambda, a)^{-1},$$

   and show that this leads to

$$\begin{aligned}
(\Lambda, a)J^{\mu\nu}(\Lambda, a)^{-1} &= (\Lambda^{-1})^\mu{}_\alpha(\Lambda^{-1})^\nu{}_\beta\,\left(J^{\alpha\beta} + a^\alpha P^\beta - a^\beta P^\alpha\right),\\
(\Lambda, a)P^{\alpha}(\Lambda, a)^{-1} &= (\Lambda^{-1})^\alpha{}_\mu P^\mu.
\end{aligned}$$

4. We now consider that the finite Poincaré transformation (Λ, a) is infinitesimal too (with the same parameters $\omega^\mu{}_\nu$ and ε^μ). Deduce from the above relations the algebra $\mathfrak{iso}(1,3)$ spanned by the generators $J^{\mu\nu}$ and P^μ.

The *isometries of Minkowski space*, hence the name $ISO(1,3)$, are crucial to the laws of physics. As indicated by the postulates of special relativity, there is no place in space-time that is different from any other place, so that physics is translation-invariant. Moreover, physics additionally satisfies Lorentz invariance (*i.e.* the laws of nature are invariant under rotations and boosts). On the other hand, our universe is made of particles of different kinds, and a given particle has a mass, a spin (together with the value of the projection of this spin onto some axis of reference), other quantum numbers (like an electric charge), as well as a four-momentum. When we move from a specific inertial frame of reference to another one, the particle's four-momentum and projection of the spin change as determined by the Poincaré group. However, all other quantum numbers are invariant under such a transformation. A *particle* is defined as a set of states that mix only among themselves under Poincaré transformations.

This precisely defines what is called a representation of a group: a set of objects that mix under a transformation of the group. In general, we determine a basis of states $\{|\psi^i\rangle\}$ that allows us to express any state $|\psi\rangle$, and in particular any transformed state $|\psi'\rangle$, as a linear combination of the elements of the basis,

$$|\psi\rangle = c_i|\psi^i\rangle \quad \to \quad |\psi'\rangle = (\Lambda, a)|\psi\rangle = c'_i|\psi^i\rangle \,. \tag{1.4.4}$$

If there is no subset of states that transform only among themselves, then the representation is *irreducible*. The irreducible representations of the Poincaré algebra are known to be the elementary building blocks yielding a correct description of nature. They imply that through experiments allowing for the manipulation of momenta and spins, some states will mix (as embedded in a specific representation) and some will not (as lying in different representations).

Finally, whereas there are numerous representations of the Poincaré group (we have so far only discussed the vectorial one), only *unitary* representations are relevant to describe particles. This originates from the fact that matrix elements (that lie at the heart of any computation in QFT) must be invariant under Poincaré transformations. In other words, if $|\psi_1\rangle$ and $|\psi_2\rangle$ denote two different states, then the matrix element $\mathcal{M} = \langle\psi_1|\psi_2\rangle$ must be invariant under any Poincaré transformation $\mathcal{P} \equiv (\Lambda, a)$. This gives

$$\mathcal{M} = \langle\psi_1|\psi_2\rangle \quad \to \quad \mathcal{M}' = \langle\psi'_1|\psi'_2\rangle = \langle\psi_1|\mathcal{P}^\dagger\mathcal{P}|\psi_2\rangle = \langle\psi_1|\psi_2\rangle \,. \tag{1.4.5}$$

The transformation $\mathcal{P}$ must therefore be either a unitary transformation or an anti-unitary transformation. The latter are, however, not continuously connected to the identity that is unitary, and we consequently focus on the former. The task left to be achieved is thus to determine the set of irreducible unitary representations of the Poincaré group.

Particles are *defined* as objects that transform under irreducible unitary representations of the Poincaré group.

We show in the next part of this section, following the work done by Wigner, that irreducible unitary representations of the Poincaré algebra can be classified from the knowledge of only two numbers, the eigenvalues of the two *Casimir operators* associated with $\mathfrak{iso}(1,3)$. According to Schur's lemma, that is named after the mathematician Issai Schur (1875 – 1941), such Casimir operators must be proportional to the identity. They therefore consist of Lorentz-scalar quantities that automatically commute with the generators of the Lorentz algebra $J^{\mu\nu}$. Therefore, the determination of the Casimir operators of the Poincaré algebra is reduced to the determination of scalar operators that commute with the generators of space-time translations P^μ.

The first Casimir operator $\mathcal{C}_2$ is quadratic in the generators. It consists of the norm of the generators of space-time translations,

$$\mathcal{C}_2 = P^\mu P_\mu \,. \tag{1.4.6}$$

The second Casimir operator $\mathcal{C}_4$ is instead quartic in the generators, and it is built from the norm of the Pauli-Lubanski operator W^μ. This last operator is named from the work of Wolfgang Pauli (1900 -1958) and Józef Lubański (1914 – 1946). It is defined by [8]

$$W_\mu = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}P^\nu J^{\rho\sigma} \,, \tag{1.4.7}$$

with $\varepsilon_{\mu\nu\rho\sigma}$ being a fully antisymmetric tensor derived from $\varepsilon_{0123} = 1$. As usual in this case, the only other non-vanishing elements of the tensor are obtained by including an extra sign flip per permutation of the indices. The quartic Casimir operator then reads

$$\mathcal{C}_4 = W^\mu W_\mu \,. \tag{1.4.8}$$

Exercise 1.7. Calculate the commutator $[P^\mu, W^\nu]$ and deduce that $W^\mu W_\mu$ is a Casimir operator.

Any representation of the Poincaré algebra is thus characterised by two numbers, the eigenvalues of the Casimir operators $\mathcal{C}_2$ and $\mathcal{C}_4$. These numbers consist of a real non-negative number m representing the mass of the representation (*i.e.* the eigenvalue of $\mathcal{C}_2$), and a non-negative integer or half-integer number j representing its spin (*i.e.* the eigenvalue of $\mathcal{C}_4$). Any state can subsequently be labelled with at least two quantum numbers,

$$|\psi\rangle = |m, j; \dots\rangle\,, \tag{1.4.9}$$

in which m^2 and $m^2 j(j+1)$ are the eigenvalues of the operators $\mathcal{C}_2$ and $\mathcal{C}_4$ (as shown below). In these notations, the dots stand for extra quantum numbers such as the eigenvalue p^μ of the operator P^μ or the eigenvalues of the generators of the associated little algebra.

We begin by showing that the four-momentum p^μ is the eigenvalue of the space-time translation operator P^μ. To this end, we consider a scalar object $f(x)$ (that will be called a field later) depending on the space-time coordinates, and a translation of parameters a^μ. This choice of a scalar quantity is only a proxy for any object depending on space-time coordinates, that could thus be any kind of Lorentz tensor. As a result of a translation of parameters a^μ, the coordinates transform as

$$x^\mu \quad \to \quad x'^\mu = x^\mu + a^\mu\,. \tag{1.4.10}$$

If we assume that $f(x)$ is translation-invariant, then

$$f(x) \quad \to \quad f'(x') = f(x) \qquad \Leftrightarrow \qquad f'(x') = f(x'-a)\,, \tag{1.4.11}$$

where we use a primed notation for the transformed quantities. The second equality is deduced from (1.4.10), as $x^\mu = x'^\mu - a^\mu$. Considering an infinitesimal translation $a^\mu = \varepsilon^\mu$, the right-hand side of the last relation can be expanded to first order in ε. This gives, after replacing x' by x,

$$f'(x) = f(x) - \varepsilon^\mu \partial_\mu f(x) = f(x) + i\varepsilon^\mu p_\mu f(x)\,, \tag{1.4.12}$$

where in the last equality, we have made use of the relativistic version of the correspondence principle of quantum mechanics. The latter relates the four-momentum and the space-time derivative operator through $p^\mu = i\partial^\mu$ (see chapter **??**). The relation (1.4.12) can be compared to (1.4.2) which reads, once expanded to first order,

$$(1, \varepsilon) = 1 + i\varepsilon^\mu P_\mu + \mathcal{O}(\varepsilon^2)\,. \tag{1.4.13}$$

The four generators of the translations are identified with the four components of the momentum operator.

We now characterise the representations of the Poincaré algebra by considering a state $|\psi\rangle$ of mass m and four-momentum p^μ, which implies that $p^2 = p^\mu p_\mu = m^2$. The eigenvalue of the first Casimir operator $\mathcal{C}_2$ can be immediately derived,

$$\mathcal{C}_2|\psi\rangle = P^\mu P_\mu|\psi\rangle = p_\mu p^\mu|\psi\rangle = m^2|\psi\rangle\,. \tag{1.4.14}$$

The eigenvalue associated with the quadratic Casimir operator $\mathcal{C}_2$ is thus the squared mass of the state. We need to distinguish three situations according to the sign of m^2. We first ignore the case of *tachyonic representations* for which $p^2 < 0$. They correspond to particles moving with a speed larger than the speed of light, and there is currently no experimental indication that such a representation is realised in nature. We are thus left with the case of *massless particles* (with $m = 0$) and that of *massive particles* (with $m > 0$).

In order to assess the eigenvalue of the quartic Casimir operator $\mathcal{C}_4$, we consider the *standard frame* for the four-momentum. The eigenvalue of $\mathcal{C}_4$ being a Lorentz scalar, we are indeed free to choose the frame in which it will be calculated. In the massive case, the standard frame is the frame in which the particle is at rest,

$$p^\mu = \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{with} \qquad m > 0\,. \tag{1.4.15}$$

We note that such a frame can always be reached from any other frame of reference by applying a Lorentz boost. The Pauli-Lubanski operators reads in this case

$$W^\mu = -m \begin{pmatrix} 0 \\ J^{23} \\ J^{31} \\ J^{12} \end{pmatrix} \equiv -m \begin{pmatrix} 0 \\ J^1 \\ J^2 \\ J^3 \end{pmatrix}\,, \tag{1.4.16}$$

as $P^0 = m$ and $P^1 = P^2 = P^3 = 0$. We immediately deduce that $\mathcal{C}_4$ is related to the angular momentum operator $\mathbf{J} = (J^1, J^2, J^3)$, that is associated with the three generators of the rotations. We have

$$\mathcal{C}_4\big|\psi\big\rangle = -m^2\mathbf{J}^2\big|\psi\big\rangle = -m^2(J_1^2 + J_2^2 + J_3^2)\big|\psi\big\rangle = -m^2 j(j+1)\big|\psi\big\rangle\,. \tag{1.4.17}$$

The state $|\psi\rangle$ is therefore labelled by two quantum numbers, its mass m and its spin j, which arises as the quantum number associated with the squared norm of the angular momentum operator $\mathbf{J}^2$. Massive elementary particles are hence identified with irreducible representations of the Poincaré group with definite spin j. Their polarisation states are arranged in multiplets of size $2j+1$, each element differing in the projection j_3 of their spin that can take $2j+1$ different eigenvalues ($j_3 = -j, -j+1, \ldots, j-1, j$),

$$|\psi\rangle \equiv |m, j; p^\mu, j_3\rangle\,. \tag{1.4.18}$$

This last property can be alternatively recovered by working out the little algebra associated with the four-momentum (1.4.15). The little algebra is defined as the sub-algebra of the Lorentz algebra that leaves the momentum (1.4.15) invariant. It consists of the tri-dimensional rotation algebra $\mathfrak{so}(3)$, whose Casimir operator is $\mathbf{J}^2 = J_{12}^2 + J_{23}^2 + J_{31}^2$. We get to the same conclusion as in (1.4.18).

Massive representations $|m, j; p^\mu, j_3\rangle$ of the Poincaré algebra are classified according to their mass m (and their four-momentum p^μ given by (1.4.15) in the standard frame), as well as their spin quantum number j (related to the eigenvalue of $\mathbf{J}^2$, *i.e.* $j(j+1)$) and its projection j_3, that allows for the categorisation of all the components within a given multiplet.

In the massless case, a frame such as that provided in (1.4.15) does not exist, as it would lead to the unphysical consequence that the particle's energy vanishes. A different treatment is thus in order. We opt to choose as a standard frame the frame in which the particle's momentum is aligned with the Oz direction,

$$p^\mu = \begin{pmatrix} E \\ 0 \\ 0 \\ E \end{pmatrix} \qquad \text{with} \qquad p^\mu p_\mu = m^2 = 0\,. \tag{1.4.19}$$

In this expression, the energy E is an arbitrary positive real number (cases featuring negative energy are ignored). In this frame of reference, the eigenvalues of the two Casimir operators are zero,

$$\mathcal{C}_2\big|\psi\big\rangle = \mathcal{C}_4\big|\psi\big\rangle = 0\,. \tag{1.4.20}$$

In order to further characterise this representation, we opt to work out the little algebra associated with the four-momentum (1.4.19). It contains three operators,

$$J^{12} = J^3\,, \qquad T^1 \equiv J^{23} + J^{02} = J^1 + K^2 \qquad \text{and} \qquad T^2 \equiv J^{31} - J^{01} = J^2 - K^1\,, \tag{1.4.21}$$

that satisfy the algebra

$$[J^3, T^1] = iT^2\,, \qquad [J^3, T^2] = -iT^1 \qquad \text{and} \qquad [T^1, T^2] = 0\,. \tag{1.4.22}$$

This algebra is $\mathfrak{iso}(2)$, *i.e.* the algebra of the isometries of a two-dimensional Euclidean plane that is also known as the algebra of the translations and rotations in two dimensions. To avoid handling the continuous degrees of freedom related to the translation operators T^1 and T^2, that do not seem to be realised in nature, we set their eigenvalues to zero.

Exercise 1.8. Show that the little algebra for massless particles is that of $\mathfrak{iso}(2)$. To this aim, we propose to consider the standard frame of reference for massless particles in which the particle's momentum is aligned with the Oz axis, and to determine the constraints that are satisfied by a Lorentz transformation leaving the corresponding momentum operator invariant.

A massless representation of the Poincaré algebra is therefore labeled as

$$|\psi\rangle \equiv |0, 0; p^\mu, \lambda\rangle \,, \tag{1.4.23}$$

in which we denote the eigenvalue of the J^3 operator, that is either an integer or a half-integer, by λ. We hence have

$$J^3|0, 0; p^\mu, \lambda\rangle = \lambda \; |0, 0; p^\mu, \lambda\rangle \qquad \text{and} \qquad T^1|0, 0; p^\mu, \lambda\rangle = T^2|0, 0; p^\mu, \lambda\rangle = 0 \,. \tag{1.4.24}$$

A simple calculation leads to

$$W^\mu|0, 0; p^\mu, \lambda\rangle = \lambda \; p^\mu \; |0, 0; p^\mu, \lambda\rangle \,, \tag{1.4.25}$$

which shows that the Pauli-Lubanski operator and the momentum operator are linearly dependent. The proportionality factor λ, that is also the eigenvalue of the J^3 operator, is called the *helicity*. Promoting this relation to a relation between operators,

$$W^\mu = \hat{h} \; P^\mu \,, \tag{1.4.26}$$

we can derive a definition for the helicity operator $\hat{h}$ from (1.4.7),

$$\hat{h} = \frac{\mathbf{p} \cdot \mathbf{J}}{||\mathbf{p}||} \,. \tag{1.4.27}$$

The helicity is therefore intuitively seen as the projection of the particle's spin onto the particle's direction of motion. It can thus only take two values, a positive one and a negative one.

It can be shown that in the massless case, the *helicity operator* $\hat{h}$ commutes with all the generators of the Poincaré algebra and therefore consists of an additional Casimir operator. As a consequence, Lorentz transformations cannot mix states of different helicities, and each helicity eigenstate is a multiplet by itself. This contrasts with the massive case, in which all spin projection states form a multiplet of dimension $2j + 1$ (for a spin j) and mix under Lorentz transformations. In addition, the helicity eigenvalue is independent of the reference frame.

In general, it turns out that parity invariance is applicable when massless particles are considered. Consequently, a state of negative helicity $|0, 0; p^\mu, -\lambda\rangle$ must always be matched with a state of positive helicity $|0, 0; p^\mu, \lambda\rangle$, as a parity transformation flips the sign of the helicity (the direction of the three-momentum is flipped under a parity transformation). This applies in particular to the case of electromagnetism and the two states of polarisation of light, as well as that of quantum chromodynamics (the theory of the strong interaction).

Massless particles are organised in singlet representations $|0, 0; p^\mu, \lambda\rangle$ of the Poincaré algebra, with a definite helicity λ that corresponds to the projection of the particle's angular momentum onto the direction of the four-momentum p^μ. In the case of theories that respect parity invariance (like electromagnetism), we must always consider pairs of states that differ by the sign of their helicity (with the exception of the spin zero case for which there is only one state).

1.5 Summary

This chapter has been built on the postulates of special relativity stated by Einstein more than 100 years ago: the laws of physics must satisfy Poincaré invariance, and the speed of light is universal. Starting from these two principles, we have derived the structure of space-time and recovered the Lorentz and Poincaré groups that include all transformations which leave physics unchanged.

The study of the representations of the Lorentz and Poincaré groups consists of one of the cornerstones of modern particle physics and QFT. In particular, it enables the definition of the concept of a particle: a particle is an object that transforms under irreducible unitary transformations of the Poincaré group. In order to further characterise particles, we distinguished two situations, the massive and the massless ones.

- Massive particles of momentum p^μ and mass m are represented by states $|m, j; p^\mu, j_3\rangle$, where $j = 0$, 1/2, 1, 3/2, ... stands for the particle's spin. These states are arranged in multiplets of size $2j+1$, each component differing by the projection of the spin $j_3 = -j$, $-j+1$, ..., $j-1$, j.
- Massless particles of momentum p^μ are organised in multiplets $|0, 0; p^\mu, \lambda\rangle$ of definite helicity λ, that corresponds to the projection of the particle's spin onto its direction of motion.

Bibliography

[1] H. Minkowski, *Raum und Zeit*, *Physikalische Zeitschrift* **10** (1909) 75–88.

[2] A. Einstein, *Zur Elektrodynamik bewegter Körper*, *Annalen der Physik* **322** (1905) 891–921.

[3] H. A. Lorentz, *Simplified Theory of Electrical and Optical Phenomena in Moving Systems*, *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences* **1** (1898) 427–442.

[4] H. A. Lorentz, *Electromagnetic phenomena in a system moving with any velocity smaller than that of light*, *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences* **6** (1903) 809–831.

[5] H. M. Poincaré, *Sur la dynamique de l'électron*, *Rendiconti del Circolo Matematico di Palermo* **21** (1906) 129–176.

[6] E. P. Wigner, *On Unitary Representations of the Inhomogeneous Lorentz Group*, *Annals Math.* **40** (1939) 149–204.

[7] V. Bargmann and E. P. Wigner, *Group Theoretical Discussion of Relativistic Wave Equations*, *Proc. Nat. Acad. Sci.* **34** (1948) 211.

[8] J. Lubański, *Sur la theorie des particules élémentaires de spin quelconque I*, *Physica* **9** (1942) 310–324.