

Contents

1	Special relativity	1
1.1	Introduction	1
1.2	Definitions and relativistic kinematics	1
1.3	Lorentz transformations	4
1.3.1	The Lorentz group	4
1.3.2	The Lorentz algebra	10
1.4	The Poincaré algebra and group	12
1.5	Summary	17
2	Quantum mechanics, relativity and fields	19
2.1	Introduction	19
2.2	The postulates of quantum mechanics	20
2.2.1	The state of a physical system	20
2.2.2	Measurements	22
2.2.3	Evolution of the system	28
2.3	The harmonic oscillator	32
2.4	Hamiltonians, Lagrangians and actions	35
2.4.1	Lagrangian formulation	36
2.4.2	Hamiltonian formulation	38
2.4.3	Noether's theorem	40
2.5	A first approach to quantum fields	42
2.5.1	A Hamiltonian for a multiparticle theory	42
2.5.2	The physical interpretation	46
2.6	Summary	49
3	Scattering and the S-matrix	51
3.1	Introduction	51
3.2	Cross sections	53
3.2.1	Generalities	53
3.2.2	Scattering of non-relativistic particles	54
3.2.3	Scattering in field theory	61
3.2.4	Spinless particle scattering in quantum electrodynamics	67
3.3	Decay rates	73
3.4	The LSZ reduction formula	74
3.5	The Feynman propagator	77
3.6	Feynman rules and Feynman diagrams	84
3.7	Summary	94

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),

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

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

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

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

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

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

$$\eta^{\mu\nu} x_\nu \equiv \sum_{\nu=0}^3 \eta^{\mu\nu} x_\nu. \quad (1.2.4)$$

Moreover, we can easily show that

$$\eta_{\mu\nu} \eta^{\nu\rho} = \delta_\mu^\rho \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \eta^{\mu\nu} \eta_{\nu\rho} = \delta^\mu_\rho \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (1.2.5)$$

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

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

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

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

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*),

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

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

$$\square = \partial_\mu \partial^\mu = \eta^{\mu\nu} \partial_\mu \partial_\nu = \frac{\partial^2}{\partial t^2} - \Delta. \quad (1.2.9)$$

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

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

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

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

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

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \quad (1.2.12)$$

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 ds^2 , we start from the most general transformation of coordinates,

$$x^\mu \rightarrow x'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} x^\nu. \quad (1.2.13)$$

As ds^2 must stay invariant under such a transformation (1.2.13), we have

$$ds^2 = dx^\mu dx^\nu \eta_{\mu\nu} = dx'^\alpha dx'^\beta \eta_{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} dx^\mu dx^\nu \eta_{\alpha\beta}. \quad (1.2.14)$$

This property must be satisfied regardless of dx so that we get

$$\frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} \eta_{\alpha\beta} = \eta_{\mu\nu}. \quad (1.2.15)$$

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

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad (1.2.16)$$

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,

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu, \quad (1.3.1)$$

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,

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

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

$$\eta_{\rho\sigma} = \Lambda^\mu{}_\rho \eta_{\mu\nu} \Lambda^\nu{}_\sigma \quad (1.3.3)$$

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

$$G \equiv \eta_{\mu\nu} \quad \text{and} \quad \Lambda \equiv \Lambda^\mu{}_\nu. \quad (1.3.4)$$

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

$$x \rightarrow x' = \Lambda x, \quad (1.3.5)$$

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

$$G = \Lambda^t G \Lambda. \quad (1.3.6)$$

Moreover, we have

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

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 ,

$$\Lambda_2^t \Lambda_1^t G \Lambda_1 \Lambda_2 = \Lambda_2^t G \Lambda_2 = G. \quad (1.3.8)$$

- *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

$$G = \Lambda^t G \Lambda \quad \Leftrightarrow \quad \eta_{\mu\nu} = \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu. \quad (1.3.9)$$

This leads to

$$\det \Lambda = \pm 1 \quad \text{and} \quad |\Lambda^0{}_0| \geq 1, \quad (1.3.10)$$

as well as to the fact that Λ has an inverse matrix Λ^{-1} defined by

$$(\Lambda^{-1})^\mu{}_\nu = \Lambda_\nu{}^\mu. \quad (1.3.11)$$

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 $x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu$.

1. Demonstrate that the matrix Λ satisfies

$$\eta_{\mu\nu} = \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu, \quad \det \Lambda = \pm 1 \quad \text{and} \quad |\Lambda^0{}_0| \geq 1.$$

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

$$P^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (1.3.12)$$

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

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \rightarrow x'^\mu = P^\mu{}_\nu x^\nu = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix}. \quad (1.3.13)$$

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 Λ^0_0 parameter allows for a classification of all Lorentz transformations as *orthochronous* Lorentz transformations ($\Lambda^0_0 \geq 1$ so that the direction of time is conserved) and *non-orthochronous* Lorentz transformations ($\Lambda^0_0 \leq -1$ so that the direction of time is reversed). Time reversal T consists of a particular non-orthochronous Lorentz transformation, with a transformation matrix T^μ_ν given by

$$T^\mu_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (1.3.14)$$

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

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \rightarrow x'^\mu = T^\mu_\nu x^\nu = \begin{pmatrix} -x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \quad (1.3.15)$$

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

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

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

$$\Lambda = \exp [i\vartheta\lambda]. \quad (1.3.16)$$

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 ($\vartheta \rightarrow \varepsilon$), we can expand (1.3.16) at $\mathcal{O}(\varepsilon^2)$. This leads to

$$\Lambda \simeq 1 + i\varepsilon\lambda. \quad (1.3.17)$$

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

$$(1 + i\varepsilon\lambda^t) G (1 + i\varepsilon\lambda) = G, \quad (1.3.18)$$

which implies, to first order in ε ,

$$\lambda^t G + G\lambda = 0 \quad \Leftrightarrow \quad \lambda^t = -G\lambda G. \quad (1.3.19)$$

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

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

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 λ^μ_ν 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

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

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

$$(J^{\alpha\beta})^\mu_\nu = i(\eta^{\alpha\mu}\delta^\beta_\nu - \eta^{\beta\mu}\delta^\alpha_\nu). \tag{1.3.22}$$

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* $\epsilon_{ij}{}^k$. The elements of this tensor are defined from $\epsilon_{12}{}^3 = 1$, 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 Λ^μ_ν 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 $\omega_{01}, \omega_{02}, \omega_{03}, \omega_{12}, \omega_{23}$ 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

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

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

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

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

$$R_3(\theta_z) = \exp \left[-\frac{i}{2} \theta_z (J^{12} - J^{21}) \right] = \exp \left[-i \theta_z J^{12} \right]. \quad (1.3.28)$$

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 $dx^2 + dy^2$ unchanged. Similarly, a Lorentz boost in the Oz direction would leave the quantity $dt^2 - dz^2$ invariant. This suggests that the corresponding Lorentz transformation matrix be written with hyperbolic functions rather than trigonometric functions,

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

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

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

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} ,

$$\tanh \varphi_z = \beta, \quad (1.3.31)$$

such that

$$\sinh \varphi_z = \beta \gamma \quad \text{and} \quad \cosh \varphi_z = \gamma, \quad (1.3.32)$$

with

$$\text{with} \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}. \quad (1.3.33)$$

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),

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

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

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

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,

$$f \rightarrow f, \quad (1.3.35)$$

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),

$$V^\mu \rightarrow V'^\mu = \Lambda^\mu{}_\nu V^\nu. \quad (1.3.36)$$

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

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

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 $T^{\mu_1 \mu_2 \dots \mu_n}$ 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

$$T^{\mu_1 \mu_2 \dots \mu_n} \rightarrow T'^{\mu_1 \mu_2 \dots \mu_n} = \Lambda^{\mu_1}{}_{\nu_1} \Lambda^{\mu_2}{}_{\nu_2} \dots \Lambda^{\mu_n}{}_{\nu_n} T^{\nu_1 \nu_2 \dots \nu_n}, \quad (1.3.38)$$

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 at most 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

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

and they satisfy the commutation relations

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

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

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

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),

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

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. Derive the symmetry properties of the tensor $\varepsilon_{\mu\nu}$, and demonstrate that $\varepsilon^\mu{}_\nu$ can consequently be written as

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

In these expressions, $\omega_{\alpha\beta}$ is real and antisymmetric under the exchange of its α and β indices.

2. With the help of the above definition for the generators of the Lorentz algebra $J^{\alpha\beta}$, show that they satisfy the commutation relations

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

3. We now consider a successive application of n infinitesimal Lorentz transformations such as those above, each defined from the parameters $\omega_{\alpha\beta}/n$. 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

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

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

$$[J^i, J^j] = iJ^k, \quad [J^i, K^j] = iK^k \quad \text{and} \quad [K^i, K^j] = -iJ^k, \quad (1.3.44)$$

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 $\varepsilon^{12}{}_3 = 1$,

$$[J^i, J^j] = i\varepsilon^{ij}{}_k J^k, \quad [J^i, K^j] = i\varepsilon^{ij}{}_k K^k \quad \text{and} \quad [K^i, K^j] = -i\varepsilon^{ij}{}_k J^k. \quad (1.3.45)$$

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

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

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

$$[N^i, N^j] = iN^k, \quad [\bar{N}^i, \bar{N}^j] = i\bar{N}^k \quad \text{and} \quad [N^i, \bar{N}^j] = 0, \quad (1.3.47)$$

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

$$[N^i, N^j] = i\varepsilon^{ij}{}_k N^k, \quad [\bar{N}^i, \bar{N}^j] = i\varepsilon^{ij}{}_k \bar{N}^k \quad \text{and} \quad [N^i, \bar{N}^j] = 0, \quad (1.3.48)$$

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,

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

can be rewritten as

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

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 $K^i \equiv J^{0i}$,

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

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 $\mathfrak{so}(1, 3) \sim \mathfrak{sl}(2, \mathbb{R}) \oplus \mathfrak{sl}(2, \mathbb{R})$.

The commutation relations

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

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 $|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2 - 1, j_1 + j_2$. 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

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad (1.4.1)$$

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

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

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

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

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.

We now consider an infinitesimal Poincaré transformation $(\Lambda, a) = (1 + \tilde{\omega}, \tilde{\varepsilon})$ such that $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \tilde{\omega}^\mu{}_\nu$ and $a^\mu = \tilde{\varepsilon}^\mu$, where $\omega^\mu{}_\nu$ and ε^μ are real and infinitesimal parameters.

2. From the exponential form of a Poincaré transformation involving the generators $J^{\mu\nu}$ of the Lorentz algebra in the vectorial representation and the generators P^μ of space-time translations,

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

relate the parameters $\omega_{\alpha\beta}$ and ε_α to the elements of the matrix $\tilde{\omega}$ and scalar $\tilde{\varepsilon}$.

3. Rewrite the combination of two finite and one infinitesimal three Poincaré transformations

$$(\Lambda, a) \left(1 + \tilde{\omega}, \tilde{\varepsilon} \right) (\Lambda, a)^{-1}$$

as a single Poincaré transformation.

4. Calculate in two ways $\Lambda \tilde{\omega} \Lambda^{-1}$ (in particular by using the fact that the single Poincaré transformation of the previous question is infinitesimal) to relate the quantities $\tilde{\omega}$, ω , Λ , a , $J^{\alpha\beta}$ and P^α . Show that this yields

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

5. 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,

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

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

$$\mathcal{M} = \langle\psi_1|\psi_2\rangle \quad \rightarrow \quad \mathcal{M}' = \langle\psi'_1|\psi'_2\rangle = \langle\psi_1|\mathcal{P}^\dagger\mathcal{P}|\psi_2\rangle = \langle\psi_1|\psi_2\rangle. \quad (1.4.5)$$

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,

$$\mathcal{C}_2 = P^\mu P_\mu. \quad (1.4.6)$$

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]

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

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

$$\mathcal{C}_4 = W^\mu W_\mu. \quad (1.4.8)$$

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,

$$|\psi\rangle = |m, j; \dots\rangle, \quad (1.4.9)$$

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

$$x^\mu \rightarrow x'^\mu = x^\mu + a^\mu. \quad (1.4.10)$$

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

$$f(x) \rightarrow f'(x') = f(x) \Leftrightarrow f'(x') = f(x' - a), \quad (1.4.11)$$

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 ,

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

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 2). The relation (1.4.12) can be compared to (1.4.2) which reads, once expanded to first order,

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

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,

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

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,

$$p^\mu = \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{with} \quad m > 0. \quad (1.4.15)$$

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

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

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

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

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 ($j_3 = -j, -j+1, \dots, j-1, j$),

$$|\psi\rangle \equiv |m, j; p^\mu, j_3\rangle. \quad (1.4.18)$$

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,

$$p^\mu = \begin{pmatrix} E \\ 0 \\ 0 \\ E \end{pmatrix} \quad \text{with} \quad p^\mu p_\mu = m^2 = 0. \quad (1.4.19)$$

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,

$$\mathcal{C}_2|\psi\rangle = \mathcal{C}_4|\psi\rangle = 0. \quad (1.4.20)$$

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,

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

that satisfy the algebra

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

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

$$|\psi\rangle \equiv |0, 0; p^\mu, \lambda\rangle, \quad (1.4.23)$$

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

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

A simple calculation leads to

$$W^\mu|0, 0; p^\mu, \lambda\rangle = \lambda p^\mu|0, 0; p^\mu, \lambda\rangle, \quad (1.4.25)$$

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,

$$W^\mu = \hat{h} P^\mu, \quad (1.4.26)$$

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

$$\hat{h} = \frac{\mathbf{p} \cdot \mathbf{J}}{|\mathbf{p}|}. \quad (1.4.27)$$

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 $j = 0, 1/2, 1, 3/2, \dots$ 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 $j_3 = -j, -j + 1, \dots, j - 1, j$.
- 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.

Chapter 2

Quantum mechanics, relativity and fields

2.1 Introduction

Quantum mechanics gradually appeared at the beginning of the 20th century, with the introduction of a finite minimal quantum of action to explain the observed spectral distribution of thermal radiation (the Planck's constant [9] that bears the name of the German physicist Max Planck (1858 – 1947)), the concept of a quantum of energy by Einstein [10], and finally the construction of the core equation of quantum mechanics by the German physicist Erwin Schrödinger (1887 – 1961) [11]. Though quantum mechanics is known today as an amazingly efficient theory to explain all phenomena at the microscopic scale, it took many years to find a suitable interpretation to the solution of the Schrödinger equation. This solution, called the *wave function*, corresponds to a probability amplitude that provides a probabilistic interpretation to the potential results of any experiment. This interpretation is known as the *Copenhagen interpretation*. It was developed by Niels Bohr (1885 – 1962), Max Born (1882 – 1970) and their group. Today the frontier between classical and quantum physics is much better explored and the naive separation between the two can be improved using a detailed theory of measurement. This is, however, not discussed in these notes, for which the old Copenhagen interpretation of quantum mechanics is sufficient.

The principles of quantum mechanics are set out through several *postulates* that make a connection between theory and experiment from a few hypotheses. They detail how to define a physical system, how to interpret the results of an experiment made on this system, and how it evolves with time. These postulates are collected in section 2.2, that additionally includes a reminder of all mathematical tools and definitions relevant to quantum mechanics (Hilbert spaces and operators acting on them). The *Dirac notation*, named after the physicist Paul Dirac (1902 – 1984) who introduced it in 1939 [12], is also reviewed as this elegant and abstract formulation of quantum mechanics is used consistently throughout this document. Finally, we highlight two special operators, the position and momentum operators \hat{x} and \hat{p} , as well as the Schrödinger and Heisenberg pictures addressing time evolution in quantum mechanics. The postulates of quantum mechanics are applied in section 2.3 to one of the most important systems in classical and quantum physics: the *simple harmonic oscillator*. We solve the associated problem in Dirac notation and derive the spectrum of the corresponding Hamiltonian in an algebraic manner. This leads to the introduction of *ladder operators*, the annihilation and creation operators a and a^\dagger , that are crucial in QFT for the description of particle and antiparticle creation and annihilation.

Section 2.4 highlights the classical case, setting the stage that will allow us to establish the relationship that exists between quantum fields and simple harmonic oscillators. We begin with a brief introduction to classical field theory, that we present as the continuous limit of a classical mechanical system of n particles localised in n positions. We define fields, their first-order derivatives and their conjugate momenta, and we next discuss existing possibilities to express the equations of motion dictating the evolution of the system. The latter can be written either by means of a system of second-order differential equations in the so-called *Lagrangian formalism* [13], or by means of a system of first-order differential equations in the so-called *Hamiltonian formalism* [14]. This discussion involves the definition of the *action*, the *Lagrangian* and the *Hamiltonian* of the system, several quantities that are central for the material presented in these notes. We next elaborate on the relationship between symmetries and conservation laws. This relationship is

manifest through *Noether's theorem* [15], a theorem that bears the name of the German physicist and mathematician Emmy Noether (1882 – 1935). Taking the example of the invariance of the laws of physics under space-time translations, we close the section with the introduction of the *energy-momentum tensor* and the reasons why energy and momentum are conserved quantities with respect to the laws of nature.

In section 2.5, we present a first approach to *quantum fields* and show how these objects allow for a *multiparticle theory* as required by special relativity. Special relativity indeed enables the conversion of energy into particle-antiparticle pairs, and conversely particles and antiparticles can either annihilate or lead to the creation of other particle species. QFT consists of the theoretical framework allowing us to handle such a situation, in which the number of particles is arbitrary and could change over time. We demonstrate this from the simplest relativistic equation of motion, the Klein-Gordon equation [16, 17] that bears the names of Oskar Klein (1894 – 1977) and Walter Gordon (1893 – 1939) who independently derived it. We consider the case of a real scalar field, which allows us to highlight how fields can be seen as infinite sets of harmonic oscillators and to introduce the procedure called *second quantisation* [18].

2.2 The postulates of quantum mechanics

Quantum mechanics aims to predict the evolution of a physical system as observed experimentally. It relies on a small number of postulates that allow not only for a precise definition of the *state* of the physical system, but also for an interpretation of the results of any *measurement* carried out on the system. Moreover, quantum mechanics provides methods to predict the *evolution* of the system over time. Clearly, the most surprising aspect of quantum mechanics is its *probabilistic nature*, that associates to all possible results of an experiment a probability of occurring. This contrasts with the classical case in which the evolution of the system is fully determined once the initial conditions are fixed.

2.2.1 The state of a physical system

In quantum mechanics, the state of a physical system is described by *wave functions*, *i.e.* complex-valued functions whose modulus squared can be interpreted as a probability density. In the following, we choose to rely on the abstract formalism developed by Dirac to describe quantum states. This involves *Hilbert spaces* and their generalisations, named after the German scientist David Hilbert (1862 – 1943), and a notation in terms of *kets* and *bras* to handle state vectors.

A Hilbert space \mathcal{H} is a specific type of vector space. Any element $|\psi\rangle$ of this space is called a ket, and it has all the usual properties of any element of any complex vector space. In addition, we associate to each state vector a dual vector represented by the bra $\langle\psi|$. This naturally leads to the introduction of the *dual space* of linear functionals acting on \mathcal{H} . It corresponds to a space of functions f that assign a complex scalar $f(\psi)$ to each state vector $|\psi\rangle$, with properties such as that for any two states $|\phi\rangle$ and $|\psi\rangle$, and for any two complex scalars a and b ,

$$f(a\phi + b\psi) = af(\phi) + bf(\psi). \quad (2.2.1)$$

The set of linear functionals forms a vector space once we define the sum of two functionals f and g by

$$(f + g)(\psi) = f(\psi) + g(\psi). \quad (2.2.2)$$

The bra vector $\langle\psi|$ is then a linear functional acting on the space of ket vectors,

$$\langle\psi| \equiv f_\psi(\cdot) = \langle\psi|\cdot\rangle, \quad (2.2.3)$$

where the dot represents the generic argument of the functional. There is a one-to-one correspondence between bras and kets, and we recall that the dual to $c|\psi\rangle$ is $c^*\langle\psi|$, \mathcal{H} being a complex vector space.

A Hilbert space is defined as a complex inner product space that is complete with respect to the distance function induced by the inner product. In other words, it fulfills the following properties:

- *Vector space* – The vector space \mathcal{H} is a set whose elements may be added together and multiplied by complex numbers (*i.e.* scalars). This equivalently means that for any two state vectors $|\psi\rangle$ and $|\phi\rangle$, the sum $|\psi + \phi\rangle \equiv |\psi\rangle + |\phi\rangle$ belongs to the space \mathcal{H} . Moreover, if c is a complex number, then the state vector $c|\psi\rangle$ also belongs to \mathcal{H} .
- *Inner product* – The inner product between a bra $\langle\phi|$ and a ket $|\psi\rangle$, $\langle\phi|\psi\rangle$, maps any pair of elements of \mathcal{H} to a complex number. This inner product allows for the introduction of the notion

of *orthogonality*. Two state vectors $|\phi\rangle$ and $|\psi\rangle$ are said to be orthogonal if $\langle\phi|\psi\rangle = 0$. Moreover, a vector is normalised if $\langle\phi|\phi\rangle = 1$.

★ *Conjugate symmetry* – The inner product is conjugate symmetric. For any two state vectors $|\psi\rangle$ and $|\phi\rangle$, we have

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^* . \quad (2.2.4)$$

The scalar product $\langle\psi|\psi\rangle$ is thus real, and it allows for the introduction of the norm of a state $\|\psi\|$ defined by

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle} . \quad (2.2.5)$$

★ *Linearity/anti-linearity* – The inner product is anti-linear in its first argument and linear in its second argument. For any two complex numbers λ_1 and λ_2 and any three vector states $|\psi_1\rangle$, $|\psi_2\rangle$ and $|\phi\rangle$, we have

$$\begin{aligned} \langle\lambda_1\psi_1 + \lambda_2\psi_2|\phi\rangle &= \lambda_1^*\langle\psi_1|\phi\rangle + \lambda_2^*\langle\psi_2|\phi\rangle , \\ \langle\phi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle &= \lambda_1\langle\phi|\psi_1\rangle + \lambda_2\langle\phi|\psi_2\rangle . \end{aligned} \quad (2.2.6)$$

★ *Positive definiteness* – The inner product is positive definite. This implies that

$$\langle\phi|\phi\rangle \geq 0 \quad \text{and} \quad \langle\phi|\phi\rangle = 0 \Leftrightarrow |\phi\rangle = 0 , \quad (2.2.7)$$

where the null vector is denoted by 0 (as an alternative to $|0\rangle$). The positive definiteness of the inner product is an essential ingredient to define probabilities, on which quantum mechanics as a whole relies.

□ *Metric space* – The inner product endowing the space allows for the definition of a distance between two state vectors $|\psi\rangle$ and $|\phi\rangle$. This distance is defined by $\|\psi - \phi\|$. It is symmetric under the exchange of the states $|\psi\rangle$ and $|\phi\rangle$, is always positive, and the distance between a state and itself is zero. Moreover, the triangle inequality holds: for any three state vectors $|\psi\rangle$, $|\phi\rangle$ and $|\chi\rangle$,

$$\|\phi - \psi\| \leq \|\phi - \chi\| + \|\chi - \psi\| . \quad (2.2.8)$$

This last property stems from the more general Cauchy-Schwarz inequality satisfied by the vector space inner product,

$$\langle\phi|\psi\rangle^2 \leq \langle\phi|\phi\rangle \langle\psi|\psi\rangle , \quad (2.2.9)$$

that carries the names of the mathematicians Augustin-Louis Cauchy (1789 – 1857) and Hermann Schwarz (1843 – 1921).

□ *Completeness* – If a series of state vectors $\sum_n |\psi_n\rangle$ converges absolutely, namely that $\sum_n \|\psi_n\|$ converges, then its limit belongs to \mathcal{H} .

Postulate on the state of a system – The physical state of a system is specified by a state vector of a Hilbert space \mathcal{H} represented by the ket $|\psi\rangle$. This ket contains all the information about the physical state.

Kets $|\psi\rangle$ represent a vectorial manner to handle the usual wave functions of quantum mechanics, and they provide a practical way to describe and understand the state of a system of particles and its evolution. When accounting for special relativity, the number of particles comprised in the system is, however, no longer fixed. This originates from the fact that mass is just one form of energy among others. Consequently energy can be converted into mass via the creation of a particle-antiparticle pair. Conversely, particles can annihilate or turn into other kinds of particles.

It actually turns out that relativistic corrections to any predictions for a microscopic-scale problem, that are of order v/c where v is a typical speed for the system considered, are generally not very relevant with respect to the possibility of producing new particles. This points to the necessity of constructing a *multiparticle theory* (i.e. quantum field theory) in which particles can be created and annihilated. This automatically leads the procedure called *second quantisation*, which consists of the canonical way to quantise relativistic objects in a manner that accounts for the possible creation and annihilation of an arbitrary number of particles. The use of the wording ‘second quantisation’ highlights that wave functions describing the state of a system are promoted to operators, allowing for the creation and annihilation of particles and antiparticles. This is discussed in detail in section 2.5.

2.2.2 Measurements

In order to relate measurements carried out on physics systems described by state vectors of a Hilbert space \mathcal{H} , we need to introduce *linear operators* acting on that space and detail several of their properties. Operators are mathematical objects that map vectors onto vectors. For instance, the relation

$$A|\phi\rangle = |\chi\rangle \quad (2.2.10)$$

indicates that the action of an operator A on the state $|\phi\rangle$ gives the state $|\chi\rangle$. All operators considered in these notes are linear operators so that for any two kets $|\psi_1\rangle$ and $|\psi_2\rangle$ and complex numbers λ_1 and λ_2 ,

$$A|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1 A|\psi_1\rangle + \lambda_2 A|\psi_2\rangle. \quad (2.2.11)$$

The action of a linear combination of two operators A and B on a ket $|\psi\rangle$ is also an operator, that is defined by

$$(\lambda A + \mu B)|\psi\rangle = \lambda A|\psi\rangle + \mu B|\psi\rangle, \quad (2.2.12)$$

for any two complex numbers λ and μ . In the following, the word ‘linear’ is always understood, and is therefore no longer specified. Likewise, the action of the product $A_1 A_2 \dots A_n$ of the n operators A_1, A_2, \dots, A_n on the ket $|\psi\rangle$ is an operator. It is defined by

$$A_1 A_2 \dots A_n |\psi\rangle = A_1 (A_2 \dots A_n |\psi\rangle) = A_1 (A_2 (\dots (A_n |\psi\rangle) \dots)). \quad (2.2.13)$$

This is equivalent to the successive actions of these operators starting from the right, that is of A_n , then of A_{n-1} , etc. The ordering of the operators is very important because for two operators A and B , the action of AB is generally not equivalent to that of BA . This property can be characterised in general through another operator called the *commutator* $[A, B]$ of the operators A and B ,

$$[A, B] = AB - BA. \quad (2.2.14)$$

We have so far defined the action of operators acting on the left, *i.e.* on ket vectors. We can additionally define the action of an operator on the right, *i.e.* on bra vectors. This is achieved by the introduction of the concept of a *matrix element* of the operator A between two state vectors $|\phi\rangle$ and $|\psi\rangle$,

$$\langle\phi|A|\psi\rangle = \langle\phi|(A|\psi\rangle) = (\langle\phi|A)|\psi\rangle. \quad (2.2.15)$$

The name ‘matrix element’ is justified below, when we describe a way to determine a representative matrix of an operator. Matrix elements further allow for the definition of the *adjoint operator* A^\dagger of an operator A through the relationship

$$\langle\phi|A^\dagger|\psi\rangle^* = \langle\psi|A|\phi\rangle. \quad (2.2.16)$$

The ket $A|\phi\rangle$ and the bra $\langle\phi|A^\dagger$ are thus dual to each other. Finally, the operator A is said to be *Hermitian* if it satisfies

$$A = A^\dagger, \quad (2.2.17)$$

which means that

$$\langle\phi|A|\psi\rangle = \langle\psi|A|\phi\rangle^*. \quad (2.2.18)$$

Any Hermitian operator whose eigenvectors (see below) form a complete set is called a *self-adjoint operator*, or equivalently an *observable operator* (also known as an *observable* for short). This should not be confused with the term ‘observable’ in common language, that refers instead to an ‘observable property’. Not all Hermitian operators are observables, but all usual Hermitian operators that are used in quantum mechanics are observables. The proof however lies beyond the scope of these notes.

Postulate on the definition of a measurable quantity – In quantum mechanics any measurable physical quantity \mathcal{A} is described by an observable A . The operator A is thus Hermitian, and its eigenvectors form a complete basis of the space \mathcal{H} of state vectors.

Two important examples of observables are the position operator $\hat{\mathbf{x}}$ and the momentum operator $\hat{\mathbf{p}}$. In the case of a system made of a single particle localised in a position \mathbf{x} , the action of these operators on the corresponding ket $|\psi\rangle$ is defined by

$$\hat{\mathbf{x}}|\psi\rangle = \mathbf{x}|\psi\rangle \quad \text{and} \quad \hat{\mathbf{p}}|\psi\rangle = -i\nabla|\psi\rangle. \quad (2.2.19)$$

The action of the position operator $\hat{\mathbf{x}}$ amounts to ‘multiply the state by the position of the state \mathbf{x} ’, and that of the momentum operator $\hat{\mathbf{p}}$ involves a first-order derivative. The circumflex accent put on the operators allows us to distinguish them from the position and the momentum themselves, and are thus introduced for clarity. In general, they are nevertheless omitted. The components of the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ satisfy the commutation relations

$$[\hat{x}^j, \hat{p}^k] = i\delta^{jk}, \quad (2.2.20)$$

in which we recall that the Latin indices j and k range from 1 to 3. Any two operators A and B satisfying a commutation relation of the form (2.2.20) are said to be *canonically conjugate*. This means that the two operators A and B satisfy the *canonical commutation relation*,

$$[A, B] = i. \quad (2.2.21)$$

Operators can also be constructed directly from a ket $|\psi\rangle$ and a bra $\langle\phi|$, using their *outer product* (also called their *dyadic product*)

$$|\psi\rangle\langle\phi|. \quad (2.2.22)$$

In order to understand that the previous expression corresponds to an operator, it is sufficient to investigate its action on a ket $|\chi\rangle$. By definition,

$$(|\psi\rangle\langle\phi|)|\chi\rangle = |\psi\rangle(\langle\phi|\chi\rangle). \quad (2.2.23)$$

The right hand side of this equality shows that the result of the application of the operator $|\psi\rangle\langle\phi|$ on the ket $|\chi\rangle$ gives as a result the ket $|\psi\rangle$ times the scalar $\langle\phi|\chi\rangle$. This precisely corresponds to the action of an operator.

If the application of the operator A on a non-zero ket $|\psi_a\rangle$ gives as a result the same ket times a complex constant a ,

$$A|\psi_a\rangle = a|\psi_a\rangle, \quad (2.2.24)$$

the ket $|\psi_a\rangle$ is called an *eigenket* or an *eigenvector* of the operator A . The proportionality coefficient a is the corresponding *eigenvalue* of A , and can be zero if $A|\psi_a\rangle = 0$. The eigenvalue is *non-degenerate* when all the eigenkets $|\psi_a\rangle$ associated with an eigenvalue a are proportional to each other. On the contrary, it is *degenerate* when there exist several linearly independent eigenvectors that correspond to a specific eigenvalue. Because of the linearity property (2.2.11), any linear combination of these eigenvectors is also an eigenvector of A , and the dimension of the vector subspace of \mathcal{H} generated by all kets $|\psi_a\rangle$ is the *degree of degeneracy* (or the *degeneracy*) of the eigenvalue a .

The eigenvalues and eigenvectors of a Hermitian operator have two important properties: the eigenvalues are real and the eigenkets corresponding to different eigenvalues are orthogonal to each other.

Exercise 2.1. Consider an arbitrary Hermitian operator A that is defined on a Hilbert space \mathcal{H} .

1. Prove that the eigenvalues of A are real.
2. Prove that eigenkets corresponding to different eigenvalues are orthogonal to each other.

It is often useful to normalise the eigenkets $|a_1\rangle, |a_2\rangle, \dots$ corresponding to different eigenvalues a_1, a_2, \dots of a Hermitian operator A to 1 so that they form an orthonormal set $\{|a_n\rangle\}$. Consequently, for any two vectors $|a_i\rangle$ and $|a_j\rangle$ of this set,

$$\langle a_i | a_j \rangle = \delta_{a_i a_j}. \quad (2.2.25)$$

In this expression, $\delta_{a_i a_j}$ is the Kronecker delta symbol that has values 1 for $a_i = a_j$ and 0 otherwise. If some or all the eigenvalues are degenerate, relation (2.2.25) must be generalised. A practical set of eigenvectors is then obtained by not only normalising the eigenvectors associated with different eigenvalues,

but also by orthonormalising those defining the eigensubspace associated with each degenerate eigenvalue. For each eigenvalue a_i that is degenerate we introduce a new label s_{ik} allowing for a distinction of the associated orthonormal eigenvectors. The values of the integer k range from 1 to the degeneracy of the eigenvalue a_i (i.e. the dimension of the corresponding eigensubspace). In this notation, relation (2.2.25) becomes

$$\langle a_i, s_{ik} | a_j, s_{j\ell} \rangle = \delta_{a_i a_j} \delta_{s_{ik} s_{j\ell}}. \quad (2.2.26)$$

The set of eigenkets $\{|a_n, s_{nm}\rangle\}$ is said to be *complete* if the full Hilbert space is spanned by the set of eigenkets $\{|a_n, s_{nm}\rangle\}$ of A . In order to keep the notation simple we omit the label s in the following. As the set $\{|a_n\rangle\}$ can be used as a basis of vectors for the ket space \mathcal{H} , any arbitrary ket $|\psi\rangle$ can be expanded on it,

$$|\psi\rangle = \sum_n c_n |a_n\rangle = \sum_n |a_n\rangle \langle a_n | \psi \rangle. \quad (2.2.27)$$

The *Fourier coefficients* $c_n = \langle a_n | \psi \rangle$ are defined as the projections of the ket $|\psi\rangle$ along the kets of the set $\{|a_n\rangle\}$. In addition, the relation (2.2.27) shows that

$$\sum_n |a_n\rangle \langle a_n| = 1, \quad (2.2.28)$$

where the 1 appearing on the right-hand side of this equation stands for the identity operator. The previous formula represents the *completeness relation* of the set $\{|a_n\rangle\}$. It involves a sum of projection operators $|a_1\rangle\langle a_1|, |a_2\rangle\langle a_2|, \dots$, in which each operator projects any ket of \mathcal{H} onto its ‘coordinate’ along the related ‘axis’ of the basis $\{|a_n\rangle\}$ of the ket vector space.

It may happen that two different observables A and B have a same complete set of eigenvectors. In this case, A and B commute. Conversely, two observables that commute must have the complete set of common eigenvectors.

Exercise 2.2. Consider an arbitrary observable A that is defined on a Hilbert space \mathcal{H} , and that possesses a complete set of eigenvectors $\{|a_n\rangle\}$.

1. Demonstrate that any vector $|\psi\rangle$ of \mathcal{H} can be written as

$$|\psi\rangle = \sum_n |a_n\rangle \langle a_n | \psi \rangle.$$

2. Now consider an observable B for which $\{|a_n\rangle\}$ is also a complete set of eigenvectors. Show that A and B commute.
3. If we consider two observables A and C , which commute, and assume that their spectrum of eigenvalues is non-degenerate, show that A and C have a complete set of common eigenkets.

The properties above must be generalised in the context of an observable possessing a continuous spectrum of eigenvalues. A practical (and non-trivial) example consists of the one-dimensional position operator \hat{x} . This example has the advantage that it additionally allows us to establish a connection with the wave-mechanical approach of quantum mechanics. We consider the eigenvalue x_0 of the operator \hat{x} , and write the corresponding eigenvector as $|x_0\rangle$. The one-dimensional version of the first relation in (2.2.19) reads

$$\hat{x}|x_0\rangle = x|x_0\rangle = x_0|x_0\rangle, \quad (2.2.29)$$

the second equality representing the eigenvalue equation. In this relation, x is a variable and can thus take any value, whereas x_0 is a constant. In order for such an equation to be solved, one must introduce the concept of distributions, that in particular includes the so-called *Dirac delta function*. This distribution is defined by

$$\int_a^b dx \delta(x) f(x) = \begin{cases} f(0) & \text{for } x \in [a, b] \\ 0 & \text{for } x \notin [a, b] \end{cases}, \quad (2.2.30)$$

for any function $f(x)$ for which $f(0)$ exists. In general the domain of integration is $]-\infty, +\infty[$, which leads to

$$f(0) = \int_{-\infty}^{+\infty} dx \delta(x) f(x), \quad (2.2.31)$$

where $f(x)$ is an arbitrary function defined on the set of real numbers and for which $f(0)$ exists. The definition (2.2.30) of the delta function yields

$$x \delta(x - a) = a \delta(x - a) \quad \forall a \in \mathbb{R}, \quad (2.2.32)$$

which implies that the delta function has the right properties to represent the eigenstates of the operator \hat{x} . Furthermore, it can be used to define the orthogonality between two eigenvectors $|x'\rangle$ and $|x''\rangle$ associated with the eigenvalues x' and x'' of the \hat{x} operator,

$$\langle x'' | x' \rangle = \delta(x'' - x'). \quad (2.2.33)$$

The states $\{|x\rangle\}$ do not consequently belong to the Hilbert space \mathcal{H} of state vectors. To deal with this problem mathematically, we need to introduce the concept of *rigged Hilbert spaces*, generalising Hilbert spaces to the case of distributions and continuous spectra (as for the \hat{x} operator for which all possible eigenvalues $x \in \mathbb{R}$ are allowed). These technicalities go beyond the scope of this document, and will consequently be ignored. We consider instead that the set $\{|x\rangle\}$ forms a basis of the Hilbert space \mathcal{H} since the operator \hat{x} is an observable, despite that strictly speaking these kets do not belong to \mathcal{H} . Relations (2.2.27) and (2.2.28) are generalised to

$$|\psi\rangle = \int_{-\infty}^{\infty} dx |x\rangle \langle x | \psi \rangle \quad \text{and} \quad \int_{-\infty}^{\infty} dx |x\rangle \langle x| = 1. \quad (2.2.34)$$

The inner product

$$\psi(x) \equiv \langle x | \psi \rangle \quad (2.2.35)$$

defines not only the wave function for the state $|\psi\rangle$, but also the action of the bra $\langle x|$. The quantity $\langle x | \psi \rangle$ can thus be interpreted as the probability amplitude that a position measurement on the system in a state $|\psi\rangle$ returns the value x , *i.e.* the precise meaning of the wave function. The representation (2.2.35) of the wave function corresponds to a particular choice of basis within Dirac's abstract Hilbert space approach, the basis of the position states $\{|x\rangle\}$. On the other hand, (2.2.35) shows that the bra $\langle x|$ is a linear functional that relates any ket $|\psi\rangle$ to a scalar quantity $\psi(x)$. This bra can therefore be considered as belonging to the dual space of \mathcal{H} , although strictly speaking this is not the case as $\langle x|$ is a distribution and not a functional. The second relation in (2.2.34) can also be used to define the projection of a state $|\psi\rangle$ on another state $|\phi\rangle$. The completeness relation of the states $\{|x\rangle\}$ indeed yields

$$\langle \phi | \psi \rangle = \langle \phi | 1 | \psi \rangle = \int_{-\infty}^{\infty} dx \langle \phi | x \rangle \langle x | \psi \rangle = \int_{-\infty}^{\infty} dx \phi^*(x) \psi(x). \quad (2.2.36)$$

The above discussions can straightforwardly be generalised to three dimensions. The position eigenbasis of the Hilbert state \mathcal{H} is now made of the kets $\{|\mathbf{x}\rangle\}$ that satisfy the properties

$$\hat{\mathbf{x}}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle, \quad \langle \mathbf{x} | \mathbf{x}' \rangle = \delta^{(3)}(\mathbf{x} - \mathbf{x}') \quad \text{and} \quad \int_{\mathbb{R}^3} d^3x |\mathbf{x}\rangle \langle \mathbf{x}| = 1, \quad (2.2.37)$$

for any states $|\mathbf{x}\rangle$ and $|\mathbf{x}'\rangle$ associated with the eigenvalues \mathbf{x} and \mathbf{x}' . The last relation consists of the completeness relation for the states $|\mathbf{x}\rangle$, and must be considered over the entire Euclidean space. In Cartesian coordinates $\mathbf{x} = (x, y, z)$, we recall that

$$d^3x = dx dy dz, \quad (2.2.38)$$

and that the tri-dimensional delta function is defined by

$$\delta^{(3)}(\mathbf{x} - \mathbf{x}') = \delta(x - x') \delta(y - y') \delta(z - z'). \quad (2.2.39)$$

The connection with the usual quantum mechanical wave functions can be retrieved similarly to (2.2.35) and (2.2.36),

$$\psi(\mathbf{x}) \equiv \langle \mathbf{x} | \psi \rangle \quad \text{and} \quad \langle \phi | \psi \rangle = \int_{\mathbb{R}^3} d^3x \langle \phi | \mathbf{x} \rangle \langle \mathbf{x} | \psi \rangle = \int_{\mathbb{R}^3} d^3x \phi^*(\mathbf{x}) \psi(\mathbf{x}). \quad (2.2.40)$$

Whilst this could be further generalised to a multiparticle situation, we only consider a one-particle case for now. The multiparticle case will be (properly) treated in the context of QFT in the last part of this chapter.

Instead of using the position basis, the treatment above can be repeated in momentum space. In this case, we employ the so-called momentum basis that is formed by the momentum eigenkets $\{|\mathbf{p}\rangle\}$. These kets satisfy the properties

$$\hat{\mathbf{p}}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle, \quad \langle\mathbf{p}|\mathbf{p}'\rangle = \delta^{(3)}(\mathbf{p} - \mathbf{p}') \quad \text{and} \quad \int_{\mathbb{R}^3} d^3p |\mathbf{p}\rangle\langle\mathbf{p}| = 1. \quad (2.2.41)$$

Making use of the second definition in (2.2.19), and projecting the first relation of (2.2.41) onto the bra $\langle\mathbf{x}|$, we obtain

$$\langle\mathbf{x}|\mathbf{p}\rangle = \frac{1}{(2\pi)^{3/2}} \exp[i\mathbf{p}\cdot\mathbf{x}], \quad (2.2.42)$$

with the normalisation factor stemming from the second relation of (2.2.41). The eigenvectors of the momentum operator are hence *plane waves*. As in the case of the position basis for the Hilbert space of state vectors, the eigenkets $\{|\mathbf{p}\rangle\}$ are not vectors of the Hilbert space, which originates here from the fact that the spectrum of eigenvalues of the operator $\hat{\mathbf{p}}$ is continuous. This technical problem is, however, ignored in these notes (similar to the case of the position basis $\{|\mathbf{x}\rangle\}$).

In our description of the position basis of \mathcal{H} , we have introduced the relation (2.2.40) that associates the (one-particle) state $|\psi\rangle$ to a wave function $\psi(\mathbf{x})$. Likewise, we can define a wave function in momentum space $\psi(\mathbf{p})$. The latter consists of the Fourier transform of $\psi(\mathbf{x})$,

$$\begin{aligned} \psi(\mathbf{p}) &\equiv \langle\mathbf{p}|\psi\rangle = \int_{\mathbb{R}^3} d^3x \langle\mathbf{p}|\mathbf{x}\rangle \langle\mathbf{x}|\psi\rangle \\ &= \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} d^3x e^{-i\mathbf{p}\cdot\mathbf{x}} \psi(\mathbf{x}), \end{aligned} \quad (2.2.43)$$

this relation being deduced from (2.2.37) and (2.2.42). The ‘coordinates’ of a state $|\psi\rangle$ in the momentum basis $\{|\mathbf{p}\rangle\}$ are thus given by the Fourier transform of the wave function. Likewise, the inverse Fourier transform gives

$$\begin{aligned} \psi(\mathbf{x}) &\equiv \langle\mathbf{x}|\psi\rangle = \int_{\mathbb{R}^3} d^3p \langle\mathbf{x}|\mathbf{p}\rangle \langle\mathbf{p}|\psi\rangle \\ &= \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} d^3p e^{i\mathbf{p}\cdot\mathbf{x}} \psi(\mathbf{p}). \end{aligned} \quad (2.2.44)$$

The last two expressions show that we can deduce the completeness relation for the states $\{|\mathbf{p}\rangle\}$ from the completeness relation of the states $\{|\mathbf{x}\rangle\}$ and *vice versa*. Moreover, the quantity $\langle\mathbf{x}|\mathbf{p}\rangle$ consists of the ingredients relating the two bases.

Exercise 2.3. Show that the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ are Hermitian, and that they are canonically conjugate. This should be addressed component by component.

Exercise 2.4. Consider a state vector $|\psi\rangle$ representing a single particle localised at a position \mathbf{x} , and show that while the momentum operator (2.2.19) is a derivative operator in position space, it consists instead of a multiplicative operator in momentum space. Equivalently, demonstrate that

$$\langle\mathbf{p}|\hat{\mathbf{p}}|\psi\rangle = \mathbf{p} \langle\mathbf{p}|\psi\rangle.$$

We now have all the technical ingredients to state the postulates of quantum mechanics on measurements. The number and form of these postulates may vary from author to author, but their content always describes the same physical hypotheses.

Postulates on measurements – Any measurement of the quantity \mathcal{A} on a physical system represented by the state vector $|\psi\rangle$ can only give, as a result, one of the eigenvalues of the observable A associated to \mathcal{A} . The measurement of \mathcal{A} is represented by the action of the operator A on $|\psi\rangle$.

We first suppose that the spectrum of the operator A is discrete. The probability $\mathcal{P}(a_n)$ of obtaining the eigenvalue a_n of the observable A for the result of the measurement of \mathcal{A} on the state $|\psi\rangle$ is given by

$$\mathcal{P}(a_n) = \frac{|\langle a_n | \psi \rangle|^2}{\langle \psi | \psi \rangle}, \quad (2.2.45)$$

where $|a_n\rangle$ stands for the normalised eigenvector corresponding to a_n . Immediately after the measurement, the state of the system collapses to the projection of $|\psi\rangle$ onto the eigenvector $|a_n\rangle$,

$$|\psi\rangle \rightarrow |a_n\rangle \langle a_n | \psi \rangle. \quad (2.2.46)$$

If the eigenvalue a_n is p times degenerate, the previous equations are generalised to

$$\mathcal{P}(a_n) = \sum_{j=1}^p \frac{|\langle a_{nj} | \psi \rangle|^2}{\langle \psi | \psi \rangle} \quad \text{and} \quad |\psi\rangle \rightarrow \sum_{j=1}^p |a_{nj}\rangle \langle a_{nj} | \psi \rangle, \quad (2.2.47)$$

where the kets $|a_{n1}\rangle, |a_{n2}\rangle, \dots, |a_{np}\rangle$ form a basis of the eigensubspace associated with the eigenvalue a_n of A . These relationships must be further generalised when A has continuous eigenvalues.

These postulates establish a connection between a physical notion, a quantity \mathcal{A} that can be measured by an apparatus during an experiment, and a particular mathematical operator A that is called an observable. The mathematical properties of A are such that any measurement always results in a real number. This is indeed ensured as the spectrum of an observable is always solely composed of real eigenvalues. Furthermore, the possibility for an observable to feature a *discrete spectrum* of eigenvalues (or at least a partially discrete spectrum of eigenvalues) leads to the notion of *quantisation* of the results of a measurement. This is a fundamental particularity of quantum mechanics.

On the other hand, these postulates emphasise the probabilistic interpretation of quantum mechanics through the relation (2.2.45). However, the sum of the probabilities to get all possible results for a measurement is always equal to 1,

$$\sum_n \mathcal{P}(a_n) = \sum_n \frac{|\langle a_n | \psi \rangle|^2}{\langle \psi | \psi \rangle} = \frac{1}{\langle \psi | \psi \rangle} \sum_n \langle \psi | a_n \rangle \langle a_n | \psi \rangle = \frac{\langle \psi | \psi \rangle}{\langle \psi | \psi \rangle} = 1, \quad (2.2.48)$$

after using the completeness relation (2.2.28). Finally, the collapse of the state (2.2.46) right after the measurement guarantees its *reproducibility*. The change (2.2.46) of the system, that originates from a measurement, is *irreversible*.

Before closing this section, we come back to operators and consider a basis of kets $\{|a_n\rangle\}$ that are associated with a Hermitian operator A . This basis can be used to determine a matrix representation of any operator B . This is achieved by utilising the completeness relation (2.2.28) twice, in the case of a discrete spectrum and its generalised integral form in the case of a continuous spectrum. For a discrete spectrum, we get

$$B = \left(\sum_j |\phi_j\rangle \langle \phi_j| \right) B \left(\sum_i |\phi_i\rangle \langle \phi_i| \right) = \sum_{i,j} |\phi_j\rangle \langle \phi_j | B | \phi_i \rangle \langle \phi_i|. \quad (2.2.49)$$

If the ket space is n -dimensional, the n^2 complex numbers $\langle \phi_j | B | \phi_i \rangle$ can be organised in a square matrix so that the index j refers to its row and the index i to its column,

$$B \doteq \begin{pmatrix} \langle \phi_1 | B | \phi_1 \rangle & \langle \phi_1 | B | \phi_2 \rangle & \dots \\ \langle \phi_2 | B | \phi_1 \rangle & \langle \phi_2 | B | \phi_2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}. \quad (2.2.50)$$

This justifies the name *matrix element* that is given to the quantity $\langle \phi_j | B | \phi_i \rangle$. The explicit form of the matrix representing the operator B obviously depends on the choice of the basis kets, several choices

being generally possible. For this reason, the symbol \doteq appearing in (2.2.50) stands for ‘represented by’. In addition, the knowledge of a matrix representation for the operator B automatically allows for the determination of a matrix representation for the adjoint operator B^\dagger , that is obtained by transposition and complex conjugation,

$$B^\dagger = (B^T)^* . \quad (2.2.51)$$

In a similar way, any arbitrary ket $|\psi\rangle$ can be represented by a column vector of components and its dual bra $\langle\psi|$ can be represented by a row vector of components,

$$|\psi\rangle \doteq \begin{pmatrix} \langle a_1|\psi\rangle \\ \langle a_2|\psi\rangle \\ \langle a_3|\psi\rangle \\ \vdots \end{pmatrix} \quad \text{and} \quad \langle\psi| \doteq \left(\langle\psi|a_1\rangle \quad \langle\psi|a_2\rangle \quad \langle\psi|a_3\rangle \quad \dots \right) . \quad (2.2.52)$$

The generalisation to the continuous case is immediate.

As an example, we opt to focus on the case of the position operator $\hat{\mathbf{x}}$ and its complete set of kets $\{|\mathbf{x}\rangle\}$. We consider a one-particle situation, and calculate the matrix element $\langle\phi|A|\psi\rangle$ involving two one-particle states $|\psi\rangle$ and $|\phi\rangle$. Using the completeness relation (2.2.37) twice and the definition (2.2.40), we obtain

$$\begin{aligned} \langle\phi|A|\psi\rangle &= \int_{\mathbf{R}^3} d^3x' \int_{\mathbf{R}^3} d^3x'' \langle\phi|\mathbf{x}'\rangle \langle\mathbf{x}'|A|\mathbf{x}''\rangle \langle\mathbf{x}''|\psi\rangle \\ &= \int_{\mathbf{R}^3} d^3x' \int_{\mathbf{R}^3} d^3x'' \phi^*(\mathbf{x}') \langle\mathbf{x}'|A|\mathbf{x}''\rangle \psi(\mathbf{x}'') . \end{aligned} \quad (2.2.53)$$

The set of matrix elements $\langle\mathbf{x}'|A|\mathbf{x}''\rangle$ provides a representation of the operator A in the position basis. This position representation of an operator A is particularly useful when the operator is a function of the position $A = f(\mathbf{x})$ such that

$$\langle\mathbf{x}'|f(\mathbf{x})|\mathbf{x}''\rangle = f(\mathbf{x}'') \delta^{(3)}(\mathbf{x}' - \mathbf{x}'') , \quad (2.2.54)$$

by virtue of the orthogonality properties (2.2.37) of the position states. The double integral in (2.2.53) consequently reduces to a single integral,

$$\langle\phi|f(\mathbf{x})|\psi\rangle = \int_{\mathbf{R}^3} d^3x' \phi^*(\mathbf{x}') f(\mathbf{x}') \psi(\mathbf{x}') . \quad (2.2.55)$$

2.2.3 Evolution of the system

The last postulate of quantum mechanics details the temporal evolution of the state of the physical system when no measurement is taken.

Postulate on evolution – The evolution over time of a system that is described by the time-dependent state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation

$$i \frac{\partial}{\partial t} |\psi(t)\rangle = H(t) |\psi(t)\rangle , \quad (2.2.56)$$

where the (possibly time-dependent) Hamiltonian $H(t)$ is the observable associated with the energy of the system. The observable $H(t)$ determines the dynamics of the system once the initial conditions at a time $t = t_0$ are fixed through the knowledge of the state vector $|\psi(t_0)\rangle$.

Equation (2.2.56) is a first-order differential equation with respect to time. We only need to choose the initial conditions to derive all information possible on the system at later times. In this sense, the evolution is deterministic, even though the information obtained is of a probabilistic nature at all times according to the other postulates of quantum mechanics.

This implies that if the state vector $|\psi(t)\rangle$ appearing in (2.2.56) is known at an initial time t_0 , its expression at time t can be obtained by means of a *time evolution operator* $U(t, t_0)$ to be determined,

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle . \quad (2.2.57)$$

The substitution of the above proposition for the state vector $|\psi(t)\rangle$ in the Schrödinger equation (2.2.56) shows that this operator satisfies the same evolution equation as $|\psi(t)\rangle$,

$$i\frac{\partial}{\partial t}U(t, t_0) = H(t)U(t, t_0) \quad (2.2.58)$$

with the initial condition $U(t_0, t_0) = 1$. Moreover, the operator $U(t, t_0)$ is unitary,

$$U^{-1}(t, t_0) = U^\dagger(t, t_0), \quad (2.2.59)$$

as required from probability conservation. This is easily shown by considering a basis of eigenkets $|a\rangle$ of an observable A , and the expansion of the states $|\psi(t_0)\rangle$ and $|\psi(t)\rangle$ in this basis as given by (2.2.27),

$$|\psi(t_0)\rangle = \sum_a c_a(t_0)|a\rangle \quad \text{and} \quad |\psi(t)\rangle = \sum_a c_a(t)|a\rangle. \quad (2.2.60)$$

In general the modulus of any individual coefficient c_a varies with time. Only in situations in which the Hamiltonian H commutes with the observable A , we have $|c_a(t_0)| = |c_a(t)|$. However, the total probability must be conserved such that the relation (2.2.48) holds at all times. In particular, if the initial state vector $|\psi(t_0)\rangle$ is normalised, then the state vector $|\psi(t)\rangle$ is normalised too, and for all times,

$$\langle\psi(t_0)|\psi(t_0)\rangle = 1 \quad \rightarrow \quad \langle\psi(t)|\psi(t)\rangle = 1. \quad (2.2.61)$$

This yields

$$\sum_a |c_a(t_0)|^2 = \sum_a |c_a(t)|^2 \quad \text{and} \quad U^\dagger(t, t_0)U(t, t_0) = 1. \quad (2.2.62)$$

Conversely, unitarity of the evolution operator $U^\dagger(t, t_0)U(t, t_0) = 1$ guarantees the conservation of total probability.

In addition, we impose that the evolution operator satisfies a *composition property* for successive time evolutions,

$$U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0) \quad \text{with} \quad t_2 > t_1 > t_0. \quad (2.2.63)$$

This means that the same result must be obtained from the initial state $|\psi(t_0)\rangle$ both when we evolve the system in time directly from t_0 to t_2 and when we utilise an intermediate step at time t_1 . Whereas in principle (2.2.63) holds for any times t_0, t_1 and t_2 (possibly satisfying another ordering), the unitarity properties of $U(t, t_0)$ guarantees that (2.2.63) can always be written in a time-ordered way, with $t_0 < t_1 < t_2$.

Exercise 2.5. In this exercise we study the properties of the time evolution operator $U(t, t_0)$.

1. Consider three times t_0, t_1 and t_2 such that $t_1 < t_0 < t_2$. Assuming that the evolution operator satisfies the relation

$$U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0),$$

show, with the help of the Hermiticity properties of the evolution operator, that this equivalently implies that

$$U(t_1, t_0)U(t_0, t_2) = U(t_1, t_2),$$

in which the times are ordered.

2. Consider an infinitesimal time evolution of parameter dt ,

$$U(t_0 + dt, t_0) = 1 + \Omega(t)dt.$$

Show that the operator $\Omega(t)$ is related to the Hamiltonian of the system through

$$\Omega(t) = -iH(t),$$

by making use of the composition property (2.2.63) and the evolution equation (2.2.58).

As $U(t_0, t_0) = 1$, an infinitesimal evolution from t_0 to $t_0 + dt$ can always be seen as a small deviation of order dt from the identity. In exercise 2.5, we have shown that this infinitesimal evolution can be written in terms of the Hamiltonian,

$$U(t_0 + dt, t_0) = 1 - iH(t) dt. \quad (2.2.64)$$

This allows for the extraction of the form of finite (non-infinitesimal) time evolutions, for which we distinguish three cases as a function of the time-dependence of $H(t)$. If the Hamiltonian $H(t) \equiv H$ is time-independent, the solution of (2.2.58) is a simple exponential,

$$U(t, t_0) = \exp[-iH(t - t_0)]. \quad (2.2.65)$$

We emphasise that $H(t - t_0)$ stands for the product of the time-independent Hamiltonian H and the time difference $t - t_0$. This relation can be proved by expanding the exponential as a Taylor series, computing its first-order time derivative, and showing that (2.2.58) is satisfied,

$$\begin{aligned} \exp[-iH(t - t_0)] &= 1 - iH(t - t_0) - \frac{1}{2}H^2(t - t_0)^2 + \dots, \\ \frac{\partial}{\partial t} \exp[-iH(t - t_0)] &= -iH - H^2(t - t_0) + \dots = -iH \exp[-iH(t - t_0)]. \end{aligned} \quad (2.2.66)$$

If the Hamiltonian $H(t)$ is instead time-dependent but the operators corresponding to different times t and t' commutes, *i.e.* $[H(t), H(t')] = 0$, the form of the solution is similar and can be determined from the composition of infinitesimal evolutions,

$$U(t, t_0) = \exp\left[-i \int_{t_0}^t dt' H(t')\right]. \quad (2.2.67)$$

The proof can be again obtained through the Taylor series expansion of the exponential. Finally if the Hamiltonian operator H taken at different times t and t' does not commute, *i.e.* $[H(t), H(t')] \neq 0$, the formal solution is given by the *Dyson series*

$$U(t, t_0) = 1 + \sum_{n=1}^{\infty} (-i)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2) \cdots H(t_n). \quad (2.2.68)$$

In the previous discussion we have implicitly assumed that time evolution affects states and that the form of the observables does not evolve with time, as shown for instance in the Schrödinger equation (2.2.56). This choice of attaching time evolution only to states is called the *Schrödinger picture*. A state $|a\rangle$ prepared at a time t_0 hence evolves to time t as

$$|a\rangle \rightarrow U(t, t_0)|a\rangle. \quad (2.2.69)$$

However, physical observations are associated with inner products. Taking an observable X as an example, we can calculate its matrix element over two specific states $|a\rangle$ and $|b\rangle$, and assess how this evolves with time from t_0 to t . This gives

$$\langle b|X|a\rangle \rightarrow \langle b|U^\dagger X U|a\rangle = \langle b|(U^\dagger X U)|a\rangle, \quad (2.2.70)$$

as only states evolve with time in the Schrödinger picture. In this expression, we have omitted the arguments of the time evolution operator for clarity, and the last equality stems from the associativity property of the multiplication of observables. There are, therefore, two equivalent approaches to unitary transformations: the Schrödinger picture in which operators stay unchanged and states vectors evolve as in (2.2.57) or (2.2.69); or the *Heisenberg picture* in which state vectors stay unchanged and time evolution is attached to operators as in

$$X \rightarrow U^\dagger X U. \quad (2.2.71)$$

Finally, there is also a third possibility, called the *interaction picture*, in which parts of time evolution are attached to operators and parts to state vectors. This case is discussed in more detail in the context of scattering theory in section 3.5.

In order to study further the relationship between the Schrödinger and Heisenberg pictures in a simple way, we assume that the Hamiltonian H is time-independent, and that the initial time is taken at the origin $t_0 = 0$. The evolution operator is thus given by (2.2.65),

$$U(t) \equiv U(t, 0) = \exp[-iHt] \quad \text{with} \quad U(0) = 1. \quad (2.2.72)$$

We consider an observable A that we represent by the operators $A_{\mathcal{H}}$ and $A_{\mathcal{S}}$ in the Heisenberg picture and Schrödinger picture respectively, and the time evolution operator $U(t)$ (the origin of time being $t_0 = 0$). These two versions of the same observable are related through relation (2.2.70), which gives

$$A_{\mathcal{H}}(t) = U(t)^\dagger A_{\mathcal{S}} U(t). \quad (2.2.73)$$

Within the same notation, we represent a specific state vector by $|\psi_{\mathcal{S}}\rangle$ and $|\psi_{\mathcal{H}}\rangle$ in the Schrödinger picture and Heisenberg picture respectively. These two versions of the same state are connected through the relation (2.2.69), which reads

$$|\psi_{\mathcal{S}}(t)\rangle = U(t)|\psi_{\mathcal{H}}\rangle. \quad (2.2.74)$$

At $t = 0$ the two operators and the two state vectors coincide,

$$A_{\mathcal{H}}(0) = A_{\mathcal{S}} \quad \text{and} \quad |\psi_{\mathcal{S}}(0)\rangle = |\psi_{\mathcal{H}}\rangle. \quad (2.2.75)$$

At any later time t , the Heisenberg state vector stays fixed to its initial value while the Heisenberg form of the operator evolves, whereas the Schrödinger state vector evolves while the Schrödinger form of the operator is frozen to its initial value.

Obviously measurable physical quantities are predicted to be the same in the two pictures, that are related through unitary transformations. The change from the Heisenberg picture to the Schrödinger picture can thus be seen as an analog to a change of coordinates. For example, the expectation value of the observable A always satisfies $\langle A \rangle \equiv \langle A \rangle_{\mathcal{S}} = \langle A \rangle_{\mathcal{H}}$,

$$\langle A \rangle_{\mathcal{S}} = \langle \psi_{\mathcal{S}}(t) | A_{\mathcal{S}} | \psi_{\mathcal{S}}(t) \rangle = \langle \psi_{\mathcal{S}}(0) | U^\dagger A_{\mathcal{S}} U | \psi_{\mathcal{S}}(0) \rangle = \langle \psi_{\mathcal{H}} | A_{\mathcal{H}}(t) | \psi_{\mathcal{H}} \rangle = \langle A \rangle_{\mathcal{H}}. \quad (2.2.76)$$

The main interest in the Heisenberg picture is that the state vectors are fixed and that time evolution only affects operators, whose studies can then be made through a plethora of standard techniques.

Differentiating (2.2.73) with respect to time and using the evolution equation (2.2.58) to get an expression of the first-order time-derivative of $U(t)$, we obtain

$$\begin{aligned} \frac{dA_{\mathcal{H}}(t)}{dt} &= i \left(U^\dagger(t) H A_{\mathcal{S}} U(t) - U^\dagger(t) A_{\mathcal{S}} H U(t) \right) + U^\dagger(t) \frac{\partial A_{\mathcal{S}}}{\partial t} U(t) \\ &= i \left[H_{\mathcal{H}}(t), A_{\mathcal{H}}(t) \right] + \left(\frac{\partial A(t)}{\partial t} \right)_{\mathcal{H}}, \end{aligned} \quad (2.2.77)$$

where we have introduced the operator $H_{\mathcal{H}}(t) = U^\dagger H U$ in the commutator. In the simple cases in which the Hamiltonian and the evolution operator commute, $H_{\mathcal{H}} = U^\dagger H_{\mathcal{S}} U = H_{\mathcal{S}} \equiv H$. In addition, the last term in (2.2.77) is only present if the operator A has an intrinsic time dependence, a case that lies beyond the scope of these notes.

In the Heisenberg picture, the time dependence of an operator $A_{\mathcal{H}}(t)$ satisfies the so-called Heisenberg equations of motion,

$$\frac{dA_{\mathcal{H}}}{dt} = i \left[H_{\mathcal{H}}, A_{\mathcal{H}} \right] + \left(\frac{\partial A}{\partial t} \right)_{\mathcal{H}}, \quad (2.2.78)$$

where $H_{\mathcal{H}}$ is the Hamiltonian of the system in the Heisenberg picture. The explicit time dependence of the operators has been omitted for simplicity.

There is a strong similarity between the Heisenberg equations of motion (2.2.78) and the classical equations of motion when written with Poisson brackets (see section 2.4.2). As an example we consider the position and the momentum of a particle evolving in a one-dimensional space, *i.e.* the operators x and p . These operators are time-independent in the Schrödinger picture so that we can ignore the last term in (2.2.78). On the contrary, they depend on time in the Heisenberg picture. Their equations of motion read, omitting the time dependence for clarity,

$$\frac{dx_{\mathcal{H}}}{dt} = i [H_{\mathcal{H}}, x_{\mathcal{H}}] \quad \text{and} \quad \frac{dp_{\mathcal{H}}}{dt} = i [H_{\mathcal{H}}, p_{\mathcal{H}}], \quad (2.2.79)$$

which can be rewritten in the form

$$\frac{dx_{\mathcal{H}}}{dt} = \frac{\partial H_{\mathcal{H}}}{\partial p_{\mathcal{H}}} \quad \text{and} \quad \frac{dp_{\mathcal{H}}}{dt} = -\frac{\partial H_{\mathcal{H}}}{\partial x_{\mathcal{H}}}. \quad (2.2.80)$$

The evolution of the position and momentum operators satisfy equations that are formally identical to Hamilton's canonical equations in classical mechanics.

Exercise 2.6. By making use of the canonical commutation relation (2.2.20) satisfied by the (one-dimensional) x and p operators, *i.e.* $[x, p] = i$, show that the Heisenberg equations of motion (2.2.79) can be written in the form (2.2.80).

The above results show that classical physics can be derived from quantum mechanics if there is a classical counterpart of the quantum operator. The opposite is, however, not true in general. For example, spin is a quantity that has no classical counterpart. The associated spin operator therefore satisfies the Heisenberg equations of motion (2.2.78), but there is no classical analogue as spin can not be written in terms of classical generalised coordinates.

In the opposite direction, *i.e.* from classical physics to quantum physics, there is a simple empirical rule allowing us to guess an observable A from its expression in classical mechanics. This rule is called the *correspondence principle*, and its validity is based on the fact that classical mechanics is a limit of quantum mechanics for macroscopic objects.

The observable $A(\mathbf{x}, \mathbf{p}, t)$, which describes a physical quantity \mathcal{A} defined in classical mechanics, is obtained through the replacement of the variables \mathbf{x} and \mathbf{p} by the associated operators in the classical expression. This amounts to using the relationship

$$\mathbf{x} \rightarrow \hat{\mathbf{x}} \quad \text{and} \quad \mathbf{p} \rightarrow \hat{\mathbf{p}} = -i \nabla. \quad (2.2.81)$$

Due to the non-commuting nature of operators, this replacement must, however, be done only after having appropriately symmetrised the classical expression.

For instance, we would obtain the observable associated to the classical expression $\mathbf{x} \cdot \mathbf{p}$ as

$$\mathbf{x} \cdot \mathbf{p} = \frac{1}{2}(\mathbf{x} \cdot \mathbf{p} + \mathbf{p} \cdot \mathbf{x}) \quad \rightarrow \quad \frac{1}{2}(\hat{\mathbf{x}} \cdot \hat{\mathbf{p}} + \hat{\mathbf{p}} \cdot \hat{\mathbf{x}}) = -\frac{i}{2}(\hat{\mathbf{x}} \cdot \nabla + \nabla \cdot \hat{\mathbf{x}}). \quad (2.2.82)$$

Though the above example is simple and leads to a unique quantum solution, this is not always the case. Recall that the fundamental theory is quantum mechanics, so that under certain conditions it tends to an approximation that we call classical mechanics at the macroscopic scale. The correspondence principle is only a means to deduce an observable from a limit valid in particular circumstances. It may not be applicable to complicated cases, and subtle effects that have no classical counterpart cannot be obtained from this rule.

2.3 The harmonic oscillator

Harmonic oscillators consist of one of the best examples to illustrate the efficiency of describing the dynamics of a quantum system with Dirac notation. In addition, harmonic oscillators play an important role both in quantum and classical physics, as they naturally appear when we consider physical systems described by small-amplitude movements around an equilibrium position. Whereas any system featuring a linear-restoring force (like a spring, a pendulum or a wave) is a harmonic oscillator, it turns out that any conservative force behaves that way when the system is considered in a state sufficiently close to its equilibrium position. In one dimension and in classical physics, the equation of motion driving the position $x(t)$ of the relevant object reads

$$\frac{d^2x(t)}{dt^2} + \omega^2x(t) = 0, \quad (2.3.1)$$

where the quantity ω stands for the *angular frequency* of the oscillator. This equation describes a *simple harmonic oscillator*, which consists of a special case of the more general situation of the *driven damped oscillator*, in which a term with a first-order time-derivative must be added to (2.3.1). Since only simple harmonic oscillators are relevant for quantum field theory, we ignore any potential damping terms in the study undertaken in this section.

In the present study, we focus on a particle of mass m that evolves in a one-dimensional space close to a local minimum of its potential energy $V(x)$. We choose this minimum to be at the origin $x = 0$. Moreover, we consider an approximate expression for the exact form of the potential, that is valid close to its minimum and that is given by the *harmonic potential*

$$V(x) = \frac{1}{2}m\omega^2x^2. \quad (2.3.2)$$

The frequency ω of the oscillator can in general be related to the second-order derivative of the exact potential through a Taylor series expansion. Consequently, measurements of the oscillation frequency of the system provide insights into the exact potential to which the particle is subjected. The dynamics of the system can be obtained from the classical Hamiltonian H derived from (2.3.1), that is as usual given by the sum of the particle's kinetic and potential energies,

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2, \quad (2.3.3)$$

where x and p represents the (one-dimensional) position and momentum. To quantise the problem, we promote the variables x and p to operators, in agreement with the correspondence principle (2.2.81), and we recall that these two operators satisfy the canonical commutation relation (2.2.20),

$$[x, p] = i. \quad (2.3.4)$$

The Hamiltonian (2.3.3) is further simplified through the choice of an appropriate normalisation for the position and momentum operators,

$$x \rightarrow \sqrt{\frac{x}{m\omega}} \quad \text{and} \quad p \rightarrow \sqrt{m\omega} p, \quad (2.3.5)$$

which keeps the commutation relation (2.3.4) unchanged and yields a rewriting of the Hamiltonian (2.3.3) as

$$H = \frac{\omega}{2}(p^2 + x^2). \quad (2.3.6)$$

The global factor of ω shows that the eigenvalues of the Hamiltonian, that correspond to the possible values for the energy of the system, will be expressed as some ω -independent factor times ω . The oscillator frequency is thus analogous to a unit of energy.

The expression (2.3.6) for the Hamiltonian of the system suggests the introduction of a pair non-Hermitian operators a and a^\dagger defined by

$$a = \frac{1}{\sqrt{2}}(x + ip) \quad \text{and} \quad a^\dagger = \frac{1}{\sqrt{2}}(x - ip). \quad (2.3.7)$$

These are traditionally known as the *annihilation operator* (or *lowering operator*) and the *creation operator* (or *raising operator*), and they satisfy the commutation relation

$$[a, a^\dagger] = 1. \quad (2.3.8)$$

Conversely, the x and p operators can be retrieved from the annihilation and creation operators as the relations (2.3.7) are invertible,

$$x = \frac{1}{\sqrt{2}}(a + a^\dagger) \quad \text{and} \quad p = -\frac{i}{\sqrt{2}}(a - a^\dagger). \quad (2.3.9)$$

Consequently, any operators that can be written in terms of the x and p operators can always be expressed in terms of annihilation and creation operators. This is applicable to the Hamiltonian (2.3.3), which becomes

$$H = \omega\left(a^\dagger a + \frac{1}{2}\right) \equiv \omega\left(N + \frac{1}{2}\right). \quad (2.3.10)$$

This form for the Hamiltonian hence introduces the so-called *number operator* $N = a^\dagger a$. Unlike the *ladder operators* a and a^\dagger , the operator N is Hermitian so that its eigenvalues are real. Moreover, the set of operators $\{1, a, a^\dagger, N\}$ forms an algebra. Expression (2.3.10) shows that it is sufficient to quantise N in order to quantise H , or in other words that energy eigenstates are eigenstates of N .

In order to derive the spectrum of the harmonic oscillator, we first define a normalised state $|0\rangle$ that is annihilated by the operator a ,

$$a|0\rangle = 0. \quad (2.3.11)$$

This state is an eigenvector of N associated with the eigenvalue 0, as well as an eigenvector of H with the eigenvalue $\omega/2$ that is called the *zero-point energy* or the *energy of the vacuum*. Defining a set of normalised states $\{|n\rangle\}$ such that

$$|n\rangle \propto (a^\dagger)^n |0\rangle, \quad (2.3.12)$$

the relation (2.3.8) implies that

$$N|n\rangle = n|n\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad \text{and} \quad a|n\rangle = \sqrt{n}|n-1\rangle. \quad (2.3.13)$$

These last relations can be proved easily by induction (see exercise 2.7), and they are not too complicated to interpret. The state $|n\rangle$ is an eigenstate of N corresponding to the eigenvalue n and an eigenstate of H with the eigenvalue $(n+1/2)\omega$. Relative to the zero-point energy, the state $|n\rangle$ corresponds to a situation in which n quanta of energy ω have been created, remembering that ω also represents, in our notation, the unit of energy. This property explains the name ‘number operator’ for the operator N . The action of the operator a lowers the eigenvalue of N by one unit, whereas that of the operator a^\dagger raises it by one unit. In other words, the action of a^\dagger on a state $|n\rangle$ consists of the creation of an additional quantum of energy, while conversely the action of a consists of the annihilation of one of the existing quanta of energy.

Exercise 2.7. Consider a one-dimensional simple harmonic oscillator.

1. Determine the commutation relation satisfied by the creation and annihilation operators a and a^\dagger from the canonical commutation relation to which the operators x and p obey.
2. Determine the algebra formed by the operators $\{1, a, a^\dagger, N\}$.
3. Consider a normalised state $|0\rangle$ such that $a|0\rangle = 0$. Demonstrate by induction that the states $|n\rangle \propto (a^\dagger)^n |0\rangle$ satisfy

$$N|n\rangle = n|n\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad \text{and} \quad a|n\rangle = \sqrt{n}|n-1\rangle,$$

with n being an integer.

In the above discussion, we have implicitly assumed that the spectrum of N is only made of positive integers. This can be demonstrated by *reductio ad absurdum* such that the set of states $\{|n\rangle\}$ exhausts the spectrum of the Hamiltonian H .

Before closing this section, we note that we can derive from the Heisenberg equation of motion (2.2.78) the way in which the creation and annihilation operators evolve with time. This yields

$$\frac{da}{dt} = i[H, a] = -i\omega a \quad \text{and} \quad \frac{da^\dagger}{dt} = i[H, a^\dagger] = i\omega a^\dagger. \quad (2.3.14)$$

These equations have a simple solution for the time-dependence of the ladder operators,

$$a(t) = a(0) e^{-i\omega t}. \quad (2.3.15)$$

The properties of the one-dimensional simple harmonic oscillator in quantum mechanics that we have reviewed in this section can be summarised as follows:

The dynamics of a one-dimensional harmonic oscillator is governed by the Hamiltonian

$$H = \omega \left(a^\dagger a + \frac{1}{2} \right), \quad (2.3.16)$$

where the angular frequency ω can be seen as a unit of energy, *i.e.* the energy eigenvalues of H are expressed as some factors of ω . This Hamiltonian involves the creation and annihilation operators a^\dagger and a , that obey to the commutation relation

$$[a, a^\dagger] = 1. \quad (2.3.17)$$

The spectrum of H is made of the eigenvalues $(n + 1/2)\omega$, where n stands for any positive integer, and the eigenvector associated with the eigenvalue $(k + 1/2)\omega$ is noted by $|k\rangle$. The set of eigenvectors $\{|n\rangle\}$ satisfies

$$N|n\rangle = n|n\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad \text{and} \quad a|n\rangle = \sqrt{n}|n-1\rangle, \quad (2.3.18)$$

with $N = a^\dagger a$. These properties justify the name number operator, creation operator and annihilation operator for the operators N , a^\dagger and a . Their action indeed corresponds to the counting of the number of quanta of energies associated with a specific state, to the creation of an additional quantum of energy, and to the annihilation of an existing quantum of energy respectively.

In the Heisenberg picture, the (conjugate) annihilation and creation operators evolve with time as

$$a(t) = a(0) e^{-i\omega t} \quad \text{and} \quad a^\dagger(t) = a^\dagger(0) e^{+i\omega t}. \quad (2.3.19)$$

Exercise 2.8. In this exercise, we generalise the results derived so far to the three-dimensional case. We consider an isotropic three-dimensional harmonic oscillator, so that the harmonic frequency and mass are the same in all three space directions.

1. Write the Hamiltonian governing the dynamics of the system.
2. Demonstrate the system can be separated into three independent harmonic oscillators acting each in a different direction of space.
3. Study the energy spectrum of this harmonic oscillator and the degeneracy of each eigenvalue.

2.4 Hamiltonians, Lagrangians and actions

The results derived and presented in the two previous sections rely on the Hamiltonian of the system. This an operator which represents the total energy of the system, and consists of a quantity that is conserved in physical processes. However, energy is not a Lorentz-invariant quantity as it is only the first component of the four-momentum (1.3.37) in special relativity. Consequently, whereas non-relativistic quantum mechanics is traditionally formulated by means of Hamiltonian operators, relativistic mechanics (and thus QFT) rather relies on Lagrangians. In contrast to Hamiltonians, Lagrangians are manifestly Lorentz invariant. However they do not represent a conserved quantity.

In this section, we briefly recap the connection between Hamiltonians and Lagrangians in the case of a non-relativistic system described by a finite set of generalised coordinates. We next generalise the discussion to the continuous case, and introduce in this way the concept of *classical field theory*. Consequently, this section slowly paves the way for QFT and the first simple approach presented in section 2.5.

Lagrangian mechanics is a formulation of classical mechanics founded on the *principle of least action*. It was introduced by the Italian-French mathematician Joseph-Louis Lagrange (1736 – 1813), and it relies on generalised coordinates $\{q_n\}$ and velocities $\{\dot{q}_n\}$. In contrast, Hamiltonian mechanics was a reformulation of Lagrangian mechanics introduced by William Rowan Hamilton (1805 – 1865), and in which the generalised velocities $\{\dot{q}_n\}$ are replaced with generalised momenta $\{p_n\}$.

2.4.1 Lagrangian formulation

We consider a system of n particles that are localised at positions $\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)$ and that evolve over time with velocities $\dot{\mathbf{x}}_1(t), \dots, \dot{\mathbf{x}}_n(t)$. After accounting for the constraints that are applicable on the system, only $s \leq 3n$ coordinates generally turn out to be independent. The number s is thus the minimal number of coordinates that are needed to characterise the positions of all particles in the system. We therefore define a set of generalised coordinates $q_1(t), \dots, q_s(t)$ such that the positions $\{\mathbf{x}_n(t)\}$ become functions of these generalised coordinates and time t ,

$$\mathbf{x}_i(t) \equiv \mathbf{x}_i(q_1(t), \dots, q_s(t), t) \quad \text{for } i = 1, \dots, n. \quad (2.4.1)$$

Accordingly, we define the generalised velocities $\dot{q}_1(t), \dots, \dot{q}_s(t)$.

The evolution of the system, or equivalently the manner in how the coordinates $\{q_s(t)\}$ change over time, can be derived from the principle of least action. This principle, usually credited to Pierre Louis Maupertuis (1698 – 1759) and Leonhard Euler (1707 – 1783), is a generalisation of the principle governing light propagation determined by Pierre de Fermat (1607 – 1665) to mechanics. The principle of least action states that there exists a scalar quantity called the *action*, a functional $S[q]$ of the generalised coordinates, that is stationary to first order when the system (smoothly) evolves from the configuration $\{q_s(t_1)\}$ at time t_1 to the configuration $\{q_s(t_2)\}$ at time t_2 . In other words, the variation of the action δS is zero on the path followed by the system in configuration space. The most general form of the action is given by

$$S[q] = \int_{t_1}^{t_2} dt L(q(t), \dot{q}(t), t) \equiv \int_{t_1}^{t_2} dt L(q_1(t), \dots, q_s(t), \dot{q}_1(t), \dots, \dot{q}_s(t), t), \quad (2.4.2)$$

where the function $L(q(t), \dot{q}(t), t)$ is the *Lagrangian* of the system. From now on, we adopt a notation in which $q(t)$ stands for the entire set of generalised coordinates $\{q_s(t)\}$, and $\dot{q}(t)$ stands for the entire set of generalised velocities $\{\dot{q}_s(t)\}$. In principle, the Lagrangian L could also include a dependence on higher-order time-derivatives of the coordinates q . Such a possibility can nevertheless be avoided by extending the number of generalised coordinates, the additional coordinates being mapped to higher-order time-derivatives of other coordinates. The functional form of the Lagrangian L as introduced in (2.4.2) (in which L only depends on the coordinates q and their first-order derivatives \dot{q}) is thus general.

The minimisation of the action (2.4.2) leads to a system of second-order ordinary differential equations,

$$\frac{\partial L(q, \dot{q}, t)}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_i} \right) = 0 \quad \text{for } i = 1, \dots, s, \quad (2.4.3)$$

in which the time-dependence of the coordinates and velocities has been omitted to simplify the notation. These equations, known as *Euler-Lagrange equations*, represent the equations of motion of the system, and their solution provides a description of the evolution of the system in coordinate space.

The form of Lagrangian can be more precisely determined from Newton's laws. It turns out that it is given by the difference between the kinetic energy of the system T (that depends in full generality on both the generalised coordinates $q(t)$ and velocities $\dot{q}(t)$, as well as on time), and on its potential energy V (that only depends on the generalised coordinates $q(t)$ and time),

$$L(q, \dot{q}, t) = T(q, \dot{q}, t) - V(q, t). \quad (2.4.4)$$

The transition to classical field theory is made by considering the limit in which the discrete set of generalised coordinates $\{q_s(t)\}$ becomes continuous. Instead of considering n positions in space as in (2.4.1), we associate to each point \mathbf{x} in space a dynamical variable, or a 'coordinate depending on time',

$$\phi_{\mathbf{x}}(t) \equiv \phi(t, \mathbf{x}). \quad (2.4.5)$$

The position \mathbf{x} hence labels the generalised coordinates $\phi_{\mathbf{x}}(t)$. This continuous set of dynamical variables is called a *classical field*, and it can be seen as a function $\phi(t, \mathbf{x})$ of the four space-time coordinates. This contrasts with the example above describing the motion of n particles. The position variable \mathbf{x} does not depend on time anymore, and it therefore does not describe any motion at all. The field represents instead some physical quantity (like temperature or any given density) that has a value for each point in space, and that changes over time.

Similarly, the path that the system follows in configuration space must now be seen as the manner in which the value of the field at each point in space evolves with time, the field being the dynamical variable. As such, the path followed by the system is determined by a function of the values of the field and of all its first-order derivatives. The time coordinate no longer plays any special role relative to the position coordinates \mathbf{x} .

By analogy with (2.4.2), we introduce the *Lagrangian density* \mathcal{L} , a function of the field and of its first-order derivatives,

$$\mathcal{L}(\phi, \partial_\mu \phi, x) \equiv \mathcal{L}(\phi(t, \mathbf{x}), \dot{\phi}(t, \mathbf{x}), \nabla \phi(t, \mathbf{x}), t, \mathbf{x}). \quad (2.4.6)$$

In the notation used in this expression, we omitted the arguments of the field and of its first-order derivatives, $\phi \equiv \phi(x)$ and $\partial_\mu \phi \equiv \partial_\mu \phi(x)$. Moreover, we have grouped the different relevant objects into four-vectors, *i.e.* $x \equiv x^\mu = (t, \mathbf{x})$ and $\partial_\mu \phi(x) \equiv (\dot{\phi}(x), -\nabla \phi(x))$. However, we still consider a non-relativistic situation so that four-vectors are only used to simplify the notation. The action is defined as a time-integral of the Lagrangian, as in (2.4.2). Introducing the Lagrangian density \mathcal{L} , it can be rewritten as a space-time integral,

$$S[\phi] = \int_{t_1}^{t_2} dt L(t) = \int_{\Omega} d^4x \mathcal{L}(\phi, \partial_\mu \phi, x), \quad (2.4.7)$$

where Ω stands for the space-time volume enclosing the system. In general, this volume is taken as the entire space, $\Omega \equiv \mathbb{R}^3$.

The principle of least action states that the field evolves from one configuration at $t = t_1$ to another configuration at $t = t_2$ according to a ‘path’ $\phi(x)$ along which the action is an extremum (*i.e.* the first-order variational derivative of the action $\delta S = 0$). In this case, a configuration corresponds to a setup in which the values of the field are known for all points in space. In practice, the classical equations of motion are derived from infinitesimal variations of the field and the condition that the corresponding variation of the action vanishes,

$$\phi \rightarrow \phi + \delta\phi \quad \text{and} \quad \delta S = 0. \quad (2.4.8)$$

This has, however, to be considered together with the assumption that the field falls off to zero when approaching the boundaries of the integration domain. Technically, such an assumption allows us to ignore the integration of any total derivative when handling the action. With these considerations in mind, we can deduce a generalised version of the Euler-Lagrange equations (2.4.3) valid in the continuous case,

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = 0. \quad (2.4.9)$$

We consider a scalar field $\phi(x) = \phi(t, \mathbf{x})$ describing the evolution over time of a physical quantity that has values at each point in space. In the Lagrangian formalism, the dynamics of the system is governed by a Lagrangian density $\mathcal{L} \equiv \mathcal{L}(\phi, \partial_\mu \phi, x)$ depending on the field ϕ and its first-order derivatives $\partial_\mu \phi$. The equations of motion associated with the system are Euler-Lagrange equations,

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = 0, \quad (2.4.10)$$

where the arguments of the Lagrangian density have been omitted for simplicity. In principle the Lagrangian density could explicitly depend on the space-time coordinates x , though such a situation will not be considered in these notes.

In general, most Lagrangian densities associated with theories useful for phenomenology depend on multiple fields, and those fields can be scalar fields (like in the example above) or fields of another nature (fermionic fields, vector fields, *etc.*). The generalisation to such a case is straightforward. The equations of motion dictating the dynamics of the system then correspond to a set of coupled Euler-Lagrange equations that relate the evolutions of the different fields.

From now on, the ‘Lagrangian density’ \mathcal{L} will simply be called the ‘Lagrangian’, as is traditionally done in high-energy physics in particular, and in field theory in general.

Exercise 2.9. This exercise is dedicated to the proof that Euler-Lagrange equations originate from the principle of least action.

1. Consider a physical system described by a set of generalised coordinates $\{q_s(t)\}$ and velocities $\{\dot{q}_s(t)\}$, and study its evolution from its configuration $\{q_s(t_1)\}$ at a time t_1 to its configuration $\{q_s(t_2)\}$ at a time t_2 . Assuming an infinitesimal variation of the coordinates $q_i \rightarrow q_i + \delta q_i$ (for $i = 1, \dots, s$) relative to the path chosen by the system to evolve, show that enforcing the action S to be an extremum (such that $\delta S = 0$) leads to Euler-Lagrange equations (2.4.3).
2. Demonstrate that the evolution of a scalar field $\phi(x)$ is driven by the Euler-Lagrange equations (2.4.10). In this case consider the variation of the action $\delta S = 0$ for an infinitesimal variation $\phi(x) \rightarrow \phi(x) + \delta\phi(x)$ of the field. This variation $\delta\phi$ has to be considered relative to the field values along the path chosen by the system when it evolves, and along which the action is an extremum.

2.4.2 Hamiltonian formulation

In the Lagrangian formalism described in section 2.4.1, a system with s degrees of freedom is described through s generalised coordinates $q \equiv \{q_s\}$ and s associated generalised velocities $\dot{q} \equiv \{\dot{q}_s\}$, the time-dependence of the different quantities being omitted for simplicity. In the Hamiltonian formulation of classical mechanics, the velocities are traded for generalised momenta $\{p_s\}$ defined by

$$p_i \equiv \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_i} \quad \text{for } i = 1, \dots, s, \quad (2.4.11)$$

where $L(q, \dot{q}, t)$ stands for the Lagrangian of the system. Inserting this definition in the Euler-Lagrange equations (2.4.3), we obtain

$$\dot{p}_i \equiv \frac{\partial L(q, \dot{q}, t)}{\partial q_i} \quad \text{for } i = 1, \dots, s. \quad (2.4.12)$$

The *Hamiltonian* H is defined by

$$H = \sum_{i=1}^s p_i \dot{q}_i - L(q, \dot{q}, t). \quad (2.4.13)$$

Its derivatives lead, after treating the coordinates $q \equiv \{q_s\}$ and the associated momenta $p \equiv \{p_s\}$ as independent variables, to the so-called *Hamilton's equations*,

$$\dot{q}_i = \frac{\partial H(q, p, t)}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H(q, p, t)}{\partial q_i} \quad \text{for } i = 1, \dots, s. \quad (2.4.14)$$

This pair of equations shows that instead of having to handle a system of second-order differential equations (*i.e.* Euler-Lagrange equations), as enforced by the Lagrangian formulation of classical mechanics, the Hamiltonian formalism involves a system of differential equations that is twice the number, but with equations that have the advantage of being only first order.

It can be shown that Lagrangian's and Hamiltonian's equations are related through a Legendre transformation, and that the Hamiltonian corresponds to the total energy of the system. It is thus given by the sum of the kinetic energy and the potential energy of the system,

$$H(q, p, t) = T(q, p, t) + V(q, t). \quad (2.4.15)$$

One particularity of the Hamiltonian formalism can be noted when we consider a general function $f(q, p, t)$ of the coordinates, the momenta and time. The total time-derivative of this function is given by

$$\frac{df}{dt} = \sum_{i=1}^s \left(\frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i \right) + \frac{\partial f}{\partial t} = \sum_{i=1}^s \left(\frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) + \frac{\partial f}{\partial t}, \quad (2.4.16)$$

where all dependence of the variables has again been omitted for simplicity. The second equality originates from (2.4.14), and it involves a sum that is called the *Poisson bracket* of the function f and the

Hamiltonian H . We can observe that Hamilton's equations can be rewritten in a more symmetric way by means of Poisson brackets,

$$\dot{q}_i = \{q_i, H\} \quad \text{and} \quad \dot{p}_i = \{p_i, H\} \quad \text{for } i = 1, \dots, s. \quad (2.4.17)$$

In general, Poisson brackets do not necessarily involve the Hamiltonian. The Poisson bracket $\{f, g\}$ of two arbitrary functions $f(q, p, t)$ and $g(q, p, t)$ can be defined as

$$\{f, g\} = \sum_{i=1}^s \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \quad (2.4.18)$$

Poisson brackets are antisymmetric, linear and they additionally satisfy the properties

$$\{fg, h\} = f\{g, h\} + \{f, g\}h \quad \text{and} \quad \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0, \quad (2.4.19)$$

for any three functions $f(q, p, t)$, $g(q, p, t)$ and $h(q, p, t)$.

Any function $f(q, p)$ of the dynamical variables q and p that does not explicitly depend on time and that has a zero Poisson bracket with the Hamiltonian H , *i.e.* any function that satisfies $\{f, H\} = 0$, is a constant of motion. Moreover, 'fundamental' Poisson brackets read

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \quad \{p_i, q_j\} = \delta_{ij} \quad \text{for } i, j = 1, \dots, s. \quad (2.4.20)$$

These last relations could be equivalently written by making use of standard commutators to which Poisson brackets are equivalent in this case,

$$[q_i, q_j] = [p_i, p_j] = 0, \quad [p_i, q_j] = \delta_{ij} \quad \text{for } i, j = 1, \dots, s. \quad (2.4.21)$$

As in the Lagrangian formalism, the above description can be generalised to the continuous case via the introduction of a field $\phi(x)$. In the discrete case, the velocities $\dot{q}(t)$ are traded for the conjugate momenta $p(t)$. The treatment of the continuous case is analogous: the first-order time-derivative of the field $\dot{\phi}(x)$ is traded for the so-called *momentum density* $\pi(x)$ conjugate to $\phi(x)$. We begin with the introduction of the definition of a *Hamiltonian density* \mathcal{H} , similar to what was done for the Lagrangian density introduced in (2.4.7). This density is related to the Hamiltonian H by integration over space,

$$H(t) = \int_{\Omega} d^3x \mathcal{H}(\phi, \partial_{\mu}\phi, x), \quad (2.4.22)$$

where Ω stands for the (possibly infinite) volume enclosing the system studied. The definition of the density \mathcal{H} is given by the generalisation to the continuous case of the relation (2.4.13),

$$\mathcal{H}(\phi, \partial_{\mu}\phi, x) = \pi(x) \dot{\phi}(x) - \mathcal{L}(\phi, \partial_{\mu}\phi, x). \quad (2.4.23)$$

As in the discrete case, the Lagrangian (density) \mathcal{L} and the Hamiltonian (density) \mathcal{H} are related through a Legendre transform. We still need to provide a definition of the momentum density $\pi(x)$. This is achieved analogously to (2.4.11),

$$\pi(x) \equiv \frac{\partial \mathcal{L}(\phi, \partial_{\mu}\phi, x)}{\partial \dot{\phi}}. \quad (2.4.24)$$

The commutation relations (2.4.21) associating a field with its conjugate momentum needs also to be generalised. This leads to the so-called *equal-time commutation relations*

$$[\phi(t, \mathbf{x}), \phi(t', \mathbf{y})] \Big|_{t=t'} = [\pi(t, \mathbf{x}), \pi(t', \mathbf{y})] \Big|_{t=t'} = 0, \quad [\phi(t, \mathbf{x}), \pi(t', \mathbf{y})] \Big|_{t=t'} = i\delta^{(3)}(\mathbf{x} - \mathbf{y}). \quad (2.4.25)$$

In these expressions, we have explicitly indicated the dependence on the space-time coordinates to highlight the fact that all fields and momenta appearing in the commutators must be evaluated at a given time $t' = t$.

In order to assess the dynamics of the system, we make use of Hamilton's equations for the field and its conjugate momentum, which generalises (2.4.14) to the continuous case,

$$\dot{\phi} = \frac{\partial \mathcal{H}(\phi, \partial_{\mu}\phi, x)}{\partial \pi} \quad \text{and} \quad \dot{\pi} = -\frac{\partial \mathcal{H}(\phi, \partial_{\mu}\phi, x)}{\partial \phi}. \quad (2.4.26)$$

We consider a scalar field $\phi(x) = \phi(t, \mathbf{x})$ describing the evolution over time of a physical quantity that has values at each point in space, and the Lagrangian density $\mathcal{L} \equiv \mathcal{L}(\phi, \partial_\mu \phi, x)$ embedding the dynamics of the system. In the Hamiltonian formalism, the associated equations of motion involve the field $\phi(x)$ and its conjugate momentum $\pi(x)$ defined by

$$\pi(x) \equiv \frac{\partial \mathcal{L}(\phi, \partial_\mu \phi, x)}{\partial \dot{\phi}}. \quad (2.4.27)$$

These two quantities satisfy equal-time commutation relations,

$$[\phi(t, \mathbf{x}), \phi(t', \mathbf{y})] \Big|_{t=t'} = [\pi(t, \mathbf{x}), \pi(t', \mathbf{y})] \Big|_{t=t'} = 0, \quad [\phi(t, \mathbf{x}), \pi(t', \mathbf{y})] \Big|_{t=t'} = i\delta^{(3)}(\mathbf{x} - \mathbf{y}), \quad (2.4.28)$$

and they are related to the Hamiltonian \mathcal{H} through Hamilton's equations,

$$\dot{\phi} = \frac{\partial \mathcal{H}(\phi, \partial_\mu \phi, x)}{\partial \pi}, \quad \dot{\pi} = -\frac{\partial \mathcal{H}(\phi, \partial_\mu \phi, x)}{\partial \phi}. \quad (2.4.29)$$

Once again, the previous discussion can easily be extended to a setup in which several fields are involved, the commutation relations (2.4.28) being imposed on each field in the theory separately.

2.4.3 Noether's theorem

Field transformations leaving the action $S[\phi]$ of the system invariant, up to a surface term (as such a term does not impact the derivation of the equations of motion, see exercise 2.9), are called *symmetries of the system*. Furthermore, when such a symmetry transformation depends continuously on some parameters, it is additionally said to be a *continuous symmetry*. Continuous symmetries play a special role in field theory as they can be associated with conservation laws, and in particular with conserved charges and currents.

To highlight this point, we consider a theory involving a set of n fields $\{\phi_n\}$, which also allows us to illustrate a more general setup than that considered in the previous subsections (in which we have systematically considered setups with a single field). We focus on field transformations associated to a continuous symmetry and depending on an infinitesimal parameter ε such that

$$\phi_i(x) \rightarrow \phi_i(x) + \delta_\varepsilon \phi_i(x) = \phi_i(x) + \varepsilon \Delta \phi_i(x) \quad \text{for } i = 1, \dots, n. \quad (2.4.30)$$

We emphasise that in this expression, $\Delta \phi_i$ represents a small variation of the field ϕ_i , and that the requirement that the transformation parameter ε is infinitesimal enables us to write the field transformation laws as above. The transformation being a symmetry, we can further impose that the action $S[\phi]$ is invariant (up to a surface term). Consequently, the variation of the Lagrangian \mathcal{L} must be equal to a total derivative such that at first order

$$\mathcal{L} \rightarrow \mathcal{L} + \delta_\varepsilon \mathcal{L} = \mathcal{L} + \varepsilon \partial_\mu \mathcal{J}^\mu, \quad (2.4.31)$$

where $\mathcal{J}_\varepsilon^\mu$ stands for some function of the fields and their derivatives, yet to be determined. Alternatively, the variation of the Lagrangian can be computed explicitly from the variation of the fields,

$$\delta_\varepsilon \mathcal{L} = \sum_{i=1}^n \left[\frac{\partial \mathcal{L}}{\partial \phi_i} \delta_\varepsilon \phi_i + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \delta_\varepsilon (\partial_\mu \phi_i) \right]. \quad (2.4.32)$$

Integrating this equality by parts, we can rewrite it as

$$\delta_\varepsilon \mathcal{L} = \sum_{i=1}^n \left[\left(\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right) \delta_\varepsilon \phi_i \right] + \partial_\mu \left(\sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \delta_\varepsilon \phi_i \right). \quad (2.4.33)$$

This equation holds both in the *off-shell* case, namely when the equations of motion of the system are not satisfied and for a field configuration in which the action is not extremal, and in the *on-shell* case

when the equations of motion hold. In this last case, the first term of (2.4.33) vanishes by virtue of Euler-Lagrange equations, and (2.4.33) simplifies to

$$\delta_\varepsilon \mathcal{L} = \partial_\mu \left(\sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \delta_\varepsilon \phi_i \right). \quad (2.4.34)$$

Comparing this results with (2.4.31) and recalling that we assumed that $\delta_\varepsilon \phi_i = \varepsilon \Delta \phi_i$, we can define a *conserved current* J^μ , independent of the transformation parameter ε . This constitutes *Noether's theorem*.

Noether's theorem – If a Lagrangian \mathcal{L} depending on a set of fields $\{\phi_n\}$ features a continuous symmetry, then there exists a conserved current J^μ when the equations of motion are satisfied. This current is defined by

$$J^\mu = \sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \Delta \phi_i - \mathcal{J}^\mu \quad \text{with} \quad \partial_\mu J^\mu = 0. \quad (2.4.35)$$

Here, $\varepsilon \Delta \phi_i$ represents the (infinitesimal) variation of the field ϕ_i under an infinitesimal symmetry transformation of parameter ε , and the quantity \mathcal{J}^μ is determined from the variation of the Lagrangian under this transformation,

$$\mathcal{L} \rightarrow \mathcal{L} + \varepsilon \partial_\mu \mathcal{J}^\mu. \quad (2.4.36)$$

The current $J^\mu = (J^0, \mathbf{J})$ is said to be ‘conserved’ because the associated total charge Q defined by the integral over space of its temporal component,

$$Q = \int_{\mathbb{R}^3} d^3x J^0, \quad (2.4.37)$$

does not vary with time. We indeed have

$$\frac{dQ}{dt} = \int_{\mathbb{R}^3} d^3x \frac{dJ^0}{dt} = \int_{\mathbb{R}^3} d^3x \nabla \cdot \mathbf{J} = 0, \quad (2.4.38)$$

the last equality stemming from $\partial_\mu J^\mu = 0$.

As an (important) example, consider the case of a space-time translation (1.4.10) of an infinitesimal parameter ε^μ . Under such a transformation, the coordinates transform as

$$x^\mu \rightarrow x'^\mu = x^\mu - \varepsilon^\mu, \quad (2.4.39)$$

the minus sign being conventional in light of the definitions to come. In addition, a scalar field $\phi(x)$ transforms in a way similar to (1.4.11),

$$\phi(x) \rightarrow \phi(x) + \varepsilon^\mu \partial_\mu \phi(x). \quad (2.4.40)$$

The Lagrangian \mathcal{L} of the system being a scalar quantity as well, we similarly get

$$\mathcal{L} \rightarrow \mathcal{L} + \varepsilon^\mu \partial_\mu \mathcal{L} = \mathcal{L} + \varepsilon^\nu \partial_\mu (\delta^\mu_\nu \mathcal{L}). \quad (2.4.41)$$

This expression explicitly involves the identity δ^μ_ν so that its form matches that of (2.4.36). There is, however, a notable difference originating from the fact that an infinitesimal space-time translation involves four infinitesimal parameters $\varepsilon^\nu = (\varepsilon^0, \varepsilon^1, \varepsilon^2, \varepsilon^3)$. Consequently, we can determine four quantities $\mathcal{J}^\mu_\nu = (\mathcal{J}^\mu_0, \mathcal{J}^\mu_1, \mathcal{J}^\mu_2, \mathcal{J}^\mu_3)$, one for each of the transformation parameters associated with the different values of the index ν . These four quantities can be compactly written in a tensorial form,

$$\mathcal{J}^\mu_\nu = \delta^\mu_\nu \mathcal{L}. \quad (2.4.42)$$

This provides an expression for the second term contributing to the four conserved currents in (2.4.35).

On the other hand, (2.4.40) shows that the infinitesimal variations of the field ϕ associated with each of the four parameters ε^ν are given by $\Delta_\nu \phi = \partial_\nu \phi$. This allows for the determination of the first contribution to the conserved currents in (2.4.35),

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta_\nu \phi = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi. \quad (2.4.43)$$

The two relations (2.4.42) and (2.4.43) allow us to write the four conserved currents in a tensorial form $\mathcal{T}^\mu{}_\nu$,

$$\mathcal{T}^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \delta^\mu{}_\nu \mathcal{L}. \quad (2.4.44)$$

The quantity $\mathcal{T}^\mu{}_\nu$ is called the (canonical) *energy-momentum tensor*, or alternatively the *stress-energy tensor*.

An important component of this tensor is its element T^{00} (with two upper indices), that allows for the derivation of the conserved charge associated with time translations once integrated over space. The quantity T^{00} indeed corresponds to the Hamiltonian density or energy density of the system,

$$\mathcal{T}^{00} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \dot{\phi} - \mathcal{L} = \mathcal{H}, \quad (2.4.45)$$

where the last equality stems from (2.4.23) and (2.4.24). Noether's theorem therefore tells us that the invariance of physics under time translations is the reason why energy is conserved. The time at which an experiment is conducted does not impact the laws of physics.

Three other important components of the energy-momentum tensor consist of its elements T^{01} , T^{02} and T^{03} (with upper indices). When integrated over space, they provide three conserved charges P^1 , P^2 and P^3 that are associated with space translations. We can collectively group these charges in a vector $\mathbf{P} = (P^1, P^2, P^3)$, and they are defined by

$$P^i = \int_{\mathbb{R}^3} d^3x T^{0i} \quad \Leftrightarrow \quad \mathbf{P} = \int_{\mathbb{R}^3} d^3x T^{0i} = - \int_{\mathbb{R}^3} d^3x \frac{\partial \mathcal{L}}{\partial(\dot{\phi})} \nabla \phi = - \int_{\mathbb{R}^3} d^3x \pi \nabla \phi. \quad (2.4.46)$$

We naturally associate these charges with the components of the total momentum carried by the field. Space translations are indeed related to momentum conservation. Noether's theorem thus provides with an explanation for momentum conservation: it arises from the fact that the laws of physics are invariant under space translations.

2.5 A first approach to quantum fields

2.5.1 A Hamiltonian for a multiparticle theory

In this section, we connect special relativity and quantum mechanics within quantum fields, which involves the simple harmonic oscillators presented in section 2.3. To this aim, we start with the simplest Lorentz-invariant equation of motion, the *Klein-Gordon equation*, which was the first relativistic quantum mechanical equation to have been proposed. It was developed following the observation that quantum mechanics, as briefly summarised in section 2.2, does not satisfy the underlying principles of special relativity. In particular, the Schrödinger equation (2.2.56) involves a first-order derivative in the time coordinate and second-order derivatives in the position coordinates. The Schrödinger equation is thus manifestly not Lorentz invariant. Lorentz transformations indeed mix coordinates of time and space, which enforces that space and time must be treated similarly.

The Klein-Gordon equation is instead invariant regardless of the choice of inertial frame of reference, and thus satisfies the underlying principle of special relativity. It can be obtained from a generalised version of the correspondence principle (2.2.81), in which energy and momentum are treated on equal footings. Such a joint treatment is imposed by special relativity, as energy E and momentum \mathbf{p} are the components of a single four-vector, the four-momentum $p^\mu = (E, \mathbf{p})$ defined by (1.3.37). Consequently, we extend the second relation in (2.2.81) to the four-momentum, and it becomes

$$p^\mu \equiv \begin{pmatrix} E \\ \mathbf{p} \end{pmatrix} \quad \rightarrow \quad \hat{p}^\mu = i\partial^\mu \equiv \begin{pmatrix} \hat{E} \\ \hat{\mathbf{p}} \end{pmatrix} = i \begin{pmatrix} \frac{\partial}{\partial t} \\ -\nabla \end{pmatrix}. \quad (2.5.1)$$

The Schrödinger equation (2.2.56) can be obtained from the correspondence principle, once we apply it to the classical symmetrised Hamiltonian function H that corresponds to the energy of the system. As mentioned above, in special relativity momentum and energy are grouped into a unique four-vector whose components satisfy Einstein's definition of the energy,

$$p^2 = p^\mu p_\mu = E^2 - \|\mathbf{p}\|^2 = m^2, \quad (2.5.2)$$

where m stands for the mass of a particle of energy E and momentum \mathbf{p} . This relation therefore provides a better starting point to derive a relativistic quantum equation than that provided by the Hamiltonian. The Klein-Gordon equation for a free particle is precisely obtained in this way. We begin with the above equation (2.5.2), next promote energy and momentum to operators by means of the correspondence principle (2.5.1), and finally apply these operators to a state $|\psi\rangle$. This yields

$$\hat{E}^2|\psi\rangle = (m^2 + \|\hat{\mathbf{p}}\|^2)|\psi\rangle \Leftrightarrow \left(\frac{\partial^2}{\partial t^2} - \Delta + m^2\right)|\psi\rangle = (\square + m^2)|\psi\rangle = 0. \quad (2.5.3)$$

As expected, the equation derived in the above way describes the dynamics of a single particle of energy E and momentum \mathbf{p} . As already mentioned, this however yields problems as in special relativity particle-antiparticle pairs can annihilate into energy or other particles and antiparticles, and conversely energy can be converted to create new particles. The number of particles therefore varies with time, and this needs to be accounted for. In addition, such a naive derivation starting from (2.5.2) and relying on (2.5.3) leads to violations of causality. This is best visible from the computation of the propagation amplitude $U(t)$ for the particle to travel from a position \mathbf{x} to a position \mathbf{y} in space. The latter can be derived from (2.2.72) once we match the Hamiltonian with the relativistic energy operator \hat{E} defined in (2.5.1). This gives

$$U(t) = \langle \mathbf{y} | e^{-i\hat{E}t} | \mathbf{x} \rangle = \langle \mathbf{y} | e^{-i\sqrt{m^2 + \|\hat{\mathbf{p}}\|^2} t} | \mathbf{x} \rangle, \quad (2.5.4)$$

which can be further simplified to

$$\begin{aligned} U(t) &= \int_{\mathbb{R}^3} d^3p \langle \mathbf{y} | e^{-i\sqrt{m^2 + \|\hat{\mathbf{p}}\|^2} t} | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{x} \rangle \\ &= \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} d^3p e^{-i\sqrt{m^2 + \|\mathbf{p}\|^2} t} e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \\ &= \frac{1}{2\pi^2} \int_0^\infty dp p \frac{\sin\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}-\mathbf{y}\|} e^{-i\sqrt{m^2+p^2} t}. \end{aligned} \quad (2.5.5)$$

In these relations, we have used for the first equality the completeness relation (2.2.41), whereas the second equality relies on the projection (2.2.42) of $|\mathbf{p}\rangle$ states onto $|\mathbf{x}\rangle$ states and on the definition of the momentum states $|\mathbf{p}\rangle$. The third equality is obtained after angular integration. This last integral can be written in terms of Bessel functions, and it can be shown that it is non-zero even for space-like separated points \mathbf{x} and \mathbf{y} . This consequently violates causality as the propagation speed must be bounded from above by the speed of light.

Quantum field theory solves this problem in the following way. We start again from (2.5.3), but we assume this time that it is applied to a field $\phi(x)$ which we consider to be real for simplicity,

$$\left(\frac{\partial^2}{\partial t^2} - \Delta + m^2\right)\phi(x) = (\square + m^2)\phi(x) = 0. \quad (2.5.6)$$

The more general case of a complex scalar field will be addressed in chapter ???. The equation that has emerged from the simplest relativistic equation, namely the total energy definition (2.5.2), corresponds to the equation of motion dictating the dynamics of any free scalar field. It thus describes particles belonging to the trivial representation of the Poincaré group (see section 1.4). Other possibilities are discussed in chapter ??. Equation (2.5.6) is easily solved once we recognise that it consists of the equation of a plane wave. Its solution for a specific value of the momentum \mathbf{p} reads

$$\phi(x) = a_{\mathbf{p}}(t) e^{i\mathbf{p}\cdot\mathbf{x}}, \quad (2.5.7)$$

where the coefficient $a_{\mathbf{p}}(t)$ is constrained such that

$$\left(\frac{\partial^2}{\partial t^2} + \mathbf{p}\cdot\mathbf{p} + m^2\right)a_{\mathbf{p}}(t) = 0. \quad (2.5.8)$$

The coefficient $a_{\mathbf{p}}(t)$ consequently satisfies the equation of motion (2.3.1) of a simple harmonic oscillator of frequency $\omega_{\mathbf{p}} = \sqrt{\mathbf{p} \cdot \mathbf{p} + m^2}$. We now have all the ingredients to write the general solution of (2.5.6). It is given by an integral over all possible values for the momentum \mathbf{p} ,

$$\phi(x) = \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \left(a_{\mathbf{p}}(t) e^{i\mathbf{p} \cdot \mathbf{x}} + a_{\mathbf{p}}^*(t) e^{-i\mathbf{p} \cdot \mathbf{x}} \right), \quad (2.5.9)$$

in which the time-dependent coefficients $a_{\mathbf{p}}(t)$ obey (2.5.8). These coefficients are complex-valued scalar functions, and the fact that we consider a real field $\phi(x)$ imposes that the second term makes use of coefficients $a_{\mathbf{p}}^*(t)$ conjugate to $a_{\mathbf{p}}(t)$. In the case of a complex field, the coefficients $a_{\mathbf{p}}^*(t)$ are simply replaced by independent quantities $b_{\mathbf{p}}^*(t)$, as described in chapter ??.

The above solution (2.5.9) can be interpreted as a Fourier decomposition of the field $\phi(x)$ into plane waves, in which each Fourier mode is an independent harmonic oscillator. Such an interpretation justifies the introduction of the conventional factor of $1/(2\pi)^3$ in the integral, that is included in our definition of Fourier transforms. Equation (2.5.8) dictates the time-dependence of the conjugate functions $a_{\mathbf{p}}(t)$ and $a_{\mathbf{p}}^*(t)$, which thus reads

$$a_{\mathbf{p}}(t) = a_{\mathbf{p}} e^{-i\omega_{\mathbf{p}} t}. \quad (2.5.10)$$

In this last expression, the coefficients $a_{\mathbf{p}}$ are constants, the time-dependence being entirely embedded in the exponential factor. The expression (2.5.9) for the field $\phi(x)$ can thus be rewritten in the form,

$$\phi(x) = \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \left(a_{\mathbf{p}} e^{-ip \cdot x} + a_{\mathbf{p}}^* e^{ip \cdot x} \right), \quad (2.5.11)$$

where the scalar product in the exponential is now the scalar product of two four vectors, the position four-vector x^μ and the four-momentum p^μ . Moreover, the coefficients $a_{\mathbf{p}}$ and their conjugate counterparts $a_{\mathbf{p}}^*$ are time-independent complex constants, and we have one of such constant coefficient for each value of the momentum \mathbf{p} .

We are now ready to quantise the field (2.5.11) in a similar fashion as what has been done for the harmonic oscillator in section 2.3. In practice, this is achieved through the introduction of an independent pair of annihilation and creation operators $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$ for each Fourier mode (of wave vector \mathbf{p}). Those operators being independent, they satisfy the commutation relations

$$[a_{\mathbf{p}}, a_{\mathbf{q}}] = [a_{\mathbf{p}}^\dagger, a_{\mathbf{q}}^\dagger] = 0 \quad \text{and} \quad [a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}). \quad (2.5.12)$$

The factor of $(2\pi)^3$ appearing in the last relation originates from the convention used for the Fourier transform in (2.5.9). The three-dimensional delta function $\delta^{(3)}(\mathbf{p} - \mathbf{q})$ is, however, not Lorentz invariant. Embedding the momenta \mathbf{p} and \mathbf{q} in the four-momenta $p = (E_p, \mathbf{p})$ and $q = (E_q, \mathbf{q})$, we can show that a good choice of a Lorentz-invariant quantity could be $E_p \delta^{(3)}(\mathbf{p} - \mathbf{q})$. This observation suggests that we modify the normalisation (2.5.11) of the classical field $\phi(x)$ so that its quantum version would read, also in analogy with the rescaling (2.3.5),

$$\phi(x) = \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{-ip \cdot x} + a_{\mathbf{p}}^\dagger e^{ip \cdot x} \right), \quad (2.5.13)$$

since the oscillator frequency $\omega_{\mathbf{p}} = \sqrt{\mathbf{p} \cdot \mathbf{p} + m^2}$ consists of the energy of the oscillator associated to each mode. In addition, we can note that we have included an extra explicit factor of $1/\sqrt{2}$ in the integral. This choice can be mapped to the factor of $1/\sqrt{2}$ included in the expression (2.3.9) of the position operator in terms of creation and annihilation operators. As shown below (in exercise 2.10), such a normalisation yields the standard equal-time commutation relations (2.4.28).

The Hamiltonian formulation requires us to associate a conjugate momentum $\pi(x)$ defined by (2.4.24) to the field $\phi(x)$. It can be shown that the Klein-Gordon equation (2.5.6) originates, by means of Euler-Lagrange equations (2.4.10), from the Klein-Gordon Lagrangian

$$\mathcal{L}_{\text{KG}} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{m^2}{2} \phi^2. \quad (2.5.14)$$

The conjugate momentum is thus given by

$$\pi(x) = \frac{\partial \mathcal{L}_{\text{KG}}(\phi, \partial_\mu \phi, x)}{\partial \dot{\phi}} \equiv \frac{\partial \mathcal{L}_{\text{KG}}(\phi, \partial_\mu \phi, x)}{\partial (\partial_0 \phi)} = \dot{\phi}. \quad (2.5.15)$$

The definition (2.5.13) of the field $\phi(x)$ thus yields

$$\pi(x) = -i \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \left(a_{\mathbf{p}} e^{-ip \cdot x} - a_{\mathbf{p}}^\dagger e^{ip \cdot x} \right). \quad (2.5.16)$$

This definition is analogous to the expression (2.3.9) of the operator p in terms of annihilation and creation operators in the case of the simple harmonic oscillator, once taken together with the rescaling (2.3.5). With the chosen normalisation for the field and conjugate momentum, we can show that the equal-time commutation relations that should be satisfied actually hold,

$$[\phi(x), \phi(y)] = [\pi(x), \pi(y)] = 0 \quad \text{and} \quad [\phi(x), \pi(y)] = i\delta^{(3)}(\mathbf{x} - \mathbf{y}), \quad (2.5.17)$$

for $x^\mu = (t, \mathbf{x})$ and $y^\mu = (t, \mathbf{y})$. We emphasise that the time appearing in the two four-vectors is the same.

Exercise 2.10. Consider a real scalar (quantum) field $\phi(x)$ whose dynamics are governed by the Klein-Gordon Lagrangian (2.5.14).

1. Show that the equation of motion that can be derived from this Lagrangian is the Klein-Gordon equation. The proof should rely on the Euler-Lagrange equations (2.4.10).
2. Verify that such a scalar field and its conjugate momentum $\pi(x)$, whose expressions in terms of creation and annihilation operators are given by (2.5.13) and (2.5.16) respectively, satisfy the canonical equal-time commutation relations (2.5.17).
3. Derive the Hamiltonian H_0 of the system from its definition (2.4.23) and relation (2.4.22). Compare this with the case of the simple harmonic oscillator.

From the respective expressions of the fields $\phi(x)$ and its conjugate momentum $\pi(x)$ in terms of creation and annihilation operators (2.5.13) and (2.5.16), we can determine an expression for the Hamiltonian H_0 of the free real scalar field (see also exercise 2.10). We obtain

$$H_0 = \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} \left(a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \frac{1}{2} [a_{\mathbf{p}}, a_{\mathbf{p}}^\dagger] \right). \quad (2.5.18)$$

This Hamiltonian consists of a straightforward generalisation of the Hamiltonian (2.3.16) to an infinite set of harmonic oscillators. The second term can be interpreted as a sum over all modes of the zero-point energies $\omega_{\mathbf{p}}/2$, and it is formally infinite by virtue of the non-zero nature of the commutator of an annihilation operator and a creation operator associated with the same momentum \mathbf{p} . This is not surprising as we consider an infinite set of oscillators per volume element together with an infinite volume in space. Physically, such an infinite expression cannot be measured experimentally since experiments only measure energy differences from the ground state, as for potential energies in classical mechanics. The second term of (2.5.18) can thus be ignored, although there are situations, not covered in these notes, in which this term matters. This omission can be seen as a renormalisation procedure in which a constant is added to the Hamiltonian to compensate for the zero-point energy, and that makes no difference in any physical process.

We next define the *ground state* $|0\rangle$ of the theory, that consists of a normalised state that is annihilated by all annihilation operators $a_{\mathbf{p}}$. The properties of these states are thus

$$\langle 0|0\rangle = 1 \quad \text{and} \quad a_{\mathbf{p}}|0\rangle = 0 \quad \forall \mathbf{p} \in \mathbb{R}^3. \quad (2.5.19)$$

Once the infinite constant appearing in the Hamiltonian \mathcal{H}_0 is dropped, it turns out that this state has a zero energy $E = 0$. As in section 2.3, the rest of the spectrum is built from the action of creation operators on the ground states. For instance, the state $a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger \dots |0\rangle$ would be an eigenstate of \mathcal{H}_0 corresponding to the eigenvalue $\omega_{\mathbf{p}} + \omega_{\mathbf{q}} + \dots$. As in the case of the harmonic oscillator, the full set of states built in this way exhausts the spectrum.

Second quantisation – We consider a set of particles of mass m that are of the same particle species. Whereas each particle has a different energy E and momentum \mathbf{p} , these two quantities always satisfy the definition of energy according to Einstein: $E^2 = \omega_{\mathbf{p}}^2 = \|\mathbf{p}\|^2 + m^2$.

By virtue of special relativity, we need to describe the dynamics of these particles through a multiparticle theory in which energy can be converted into particle-antiparticle pairs, and particles and antiparticles can annihilate. The Hamiltonian associated with such a theory can be obtained through the quantisation of a classical field.

In this subsection, we have taken the example of a real scalar field $\phi(x)$. The quantisation procedure, known as second quantisation, consists of starting with a classical theory for the field $\phi(x)$, solving the associated equations of motion (in our example, the Klein-Gordon equation), and writing the field $\phi(x)$ as the most general solution to these equations (in our example, a Fourier decomposition into plane waves). The field is next promoted to an operator such that the field and its conjugate momentum $\pi(x)$ satisfy equal-time commutation relations.

Through this procedure, we have obtained a field and an associated conjugate momentum given by

$$\begin{aligned}\phi(x) &= \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{-ip \cdot x} + a_{\mathbf{p}}^\dagger e^{ip \cdot x} \right), \\ \pi(x) &= -i \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \left(a_{\mathbf{p}} e^{-ip \cdot x} - a_{\mathbf{p}}^\dagger e^{ip \cdot x} \right).\end{aligned}\tag{2.5.20}$$

Both quantities involve an infinite set of creation and annihilation operators $a_{\mathbf{p}}^\dagger$ and $a_{\mathbf{p}}$, each such pair of operators being associated with a specific value of the momentum \mathbf{p} .

2.5.2 The physical interpretation

This subsection is dedicated to the physical interpretation of the eigenstates of the Hamiltonian (2.5.18). We have so far defined the ground state of the theory, $|0\rangle$, that consists of a normalised state whose associated energy is zero. The action of any of the annihilation operators of the theory on this state, that is also known as the *vacuum*, is given by (2.5.19): the vacuum is annihilated by any of the annihilation operators.

We now consider a *one-particle state* that is defined from the action of one of the creation operators of the theory, $a_{\mathbf{p}}^\dagger$, on the vacuum,

$$a_{\mathbf{p}}^\dagger |0\rangle.\tag{2.5.21}$$

As already mentioned, this consists of an eigenstate of the Hamiltonian \mathcal{H}_0 with energy $E = \omega_{\mathbf{p}}$. In order to further characterise this state, we estimate the action of the total momentum operator (2.4.46) on it,

$$\begin{aligned}\mathbf{P} a_{\mathbf{p}}^\dagger |0\rangle &= - \int_{\mathbb{R}^3} d^3x \pi(x) \nabla \phi(x) a_{\mathbf{p}}^\dagger |0\rangle \\ &= - \frac{1}{2} \int_{\mathbb{R}^3} d^3x \int_{\mathbb{R}^3} \frac{d^3q}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{d^3q'}{(2\pi)^3} \sqrt{\frac{\omega_{\mathbf{q}}}{\omega_{\mathbf{q}'}}} \mathbf{q}' \left(a_{\mathbf{q}} e^{-iq \cdot x} - a_{\mathbf{q}}^\dagger e^{iq \cdot x} \right) \\ &\quad \times \left(a_{\mathbf{q}'} e^{-iq' \cdot x} - a_{\mathbf{q}'}^\dagger e^{iq' \cdot x} \right) a_{\mathbf{p}}^\dagger |0\rangle \\ &= - \frac{1}{2} \int_{\mathbb{R}^3} d^3x \int_{\mathbb{R}^3} \frac{d^3q}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{d^3q'}{(2\pi)^3} \sqrt{\frac{\omega_{\mathbf{q}}}{\omega_{\mathbf{q}'}}} \mathbf{q}' e^{i(\mathbf{q}+\mathbf{q}') \cdot \mathbf{x}} \left(a_{\mathbf{q}} e^{-i\omega_{\mathbf{q}}t} - a_{-\mathbf{q}}^\dagger e^{i\omega_{\mathbf{q}}t} \right) \\ &\quad \times \left(a_{\mathbf{q}'} e^{-i\omega_{\mathbf{q}'}t} + a_{-\mathbf{q}'}^\dagger e^{i\omega_{\mathbf{q}'}t} \right) a_{\mathbf{p}}^\dagger |0\rangle.\end{aligned}\tag{2.5.22}$$

To derive the second equality, we have employed the definitions (2.5.20) of the field and its conjugate momentum. For the third equality we enforced the change of variables $\mathbf{q} \rightarrow -\mathbf{q}$ and $\mathbf{q}' \rightarrow -\mathbf{q}'$ in terms involving positive exponentials, and we recall that $\omega_{\mathbf{q}} = \omega_{-\mathbf{q}}$. Such a change of variables manifestly shows that the integral over the position space corresponds to the Fourier transform of the exponential. We recall that in our normalisation conventions for the Fourier transform, it is given by

$$\delta^{(3)}(\mathbf{p}) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} d^3x e^{-i\mathbf{p} \cdot \mathbf{x}}.\tag{2.5.23}$$

This further allows for the evaluation of one of the integrals over the momenta. We get

$$\mathbf{P} a_{\mathbf{p}}^\dagger |0\rangle = \frac{1}{2} \int_{\mathbb{R}^3} \frac{d^3q}{(2\pi)^3} \mathbf{q} \left(a_{\mathbf{q}} e^{-i\omega_{\mathbf{q}}t} - a_{-\mathbf{q}}^\dagger e^{i\omega_{\mathbf{q}}t} \right) \left(a_{-\mathbf{q}} e^{-i\omega_{\mathbf{q}}t} + a_{\mathbf{q}}^\dagger e^{i\omega_{\mathbf{q}}t} \right) a_{\mathbf{p}}^\dagger |0\rangle = \mathbf{p} a_{\mathbf{p}}^\dagger |0\rangle, \quad (2.5.24)$$

the last equality originating from the symmetry properties of the integrand and the commutation relations (2.5.12). This shows that the state $a_{\mathbf{p}}^\dagger |0\rangle$ corresponds to a state with momentum \mathbf{p} . Furthermore, its energy $\omega_{\mathbf{p}}$ is always positive and it satisfies

$$E = \omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}. \quad (2.5.25)$$

This precisely corresponds to the energy, according to Einstein, of a particle of mass m and momentum \mathbf{p} . We therefore say that the state $a_{\mathbf{p}}^\dagger |0\rangle$ ‘contains one particle of mass m and momentum \mathbf{p} ’.

Exercise 2.11. Consider a real scalar field $\phi(x)$ and its conjugate momentum $\pi(x)$, and a vacuum state denoted by $|0\rangle$. Demonstrate in detail that

$$\mathbf{P} a_{\mathbf{p}}^\dagger |0\rangle = \mathbf{p} a_{\mathbf{p}}^\dagger |0\rangle,$$

where \mathbf{P} stands for the total momentum operator and $a_{\mathbf{p}}^\dagger$ the creation operator associated with the specific value \mathbf{p} of the momentum.

With the derivation of (2.5.24), we have shown that the states $|\mathbf{p}\rangle$ which represents a single particle of momentum $|\mathbf{p}\rangle$ and the state $a_{\mathbf{p}}^\dagger |0\rangle$ are proportional to each other. Keeping Lorentz invariance in mind, we define

$$|\mathbf{p}\rangle = \sqrt{2\omega_{\mathbf{p}}} a_{\mathbf{p}}^\dagger |0\rangle. \quad (2.5.26)$$

The choice of such a prefactor $\sqrt{2\omega_{\mathbf{p}}}$ is motivated by the fact that it yields, for two states $|\mathbf{p}\rangle$ and $|\mathbf{q}\rangle$,

$$\langle \mathbf{p} | \mathbf{q} \rangle = 2 \sqrt{\omega_{\mathbf{p}} \omega_{\mathbf{q}}} \langle 0 | a_{\mathbf{p}} a_{\mathbf{q}}^\dagger | 0 \rangle = 2\omega_{\mathbf{p}} (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}). \quad (2.5.27)$$

The energy dependence in (2.5.26) guarantees that the scalar product of two states is Lorentz-invariant, whereas the factor of two is convenient to match that included in (2.5.20). Subsequently, the completeness relation shown in (2.2.41) has to be modified too. It now reads, for one-particle states $|\mathbf{p}\rangle$,

$$\int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} |\mathbf{p}\rangle \langle \mathbf{p}| = 1. \quad (2.5.28)$$

With the definition (2.5.26), we are ready to evaluate the action of the field operator (2.5.20) on the vacuum $|0\rangle$. This gives

$$\phi(x)|0\rangle = \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} e^{ip \cdot x} |\mathbf{p}\rangle. \quad (2.5.29)$$

Such a result can be seen as a linear combination of an infinite set of one-particle states, each having a definite momentum \mathbf{p} . Furthermore, projecting this expression on the bra $\langle \mathbf{q}|$ and using (2.5.27) leads to

$$\langle \mathbf{q} | \phi(x) | 0 \rangle = e^{iq \cdot x}. \quad (2.5.30)$$

This can be interpreted analogously to (2.2.42), after a generalisation to the case of space-time. Through the introduction of $|x\rangle \equiv \phi(x)|0\rangle$, we conclude that the quantity $\langle x | \mathbf{p} \rangle$ consists of the space-time representation of the one-particle state $|\mathbf{p}\rangle$.

The action of a scalar field $\phi(x)$ on the vacuum $|0\rangle$, *i.e.* $\phi(x)|0\rangle$, corresponds to the creation of a particle at position \mathbf{x} and time t .

We can perform a similar exercise on a state defined by the successive application of two creation operators,

$$a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger |0\rangle. \quad (2.5.31)$$

This results in the creation of a *two-particle state* with total energy $\omega_{\mathbf{p}} + \omega_{\mathbf{q}}$ and momentum $\mathbf{p} + \mathbf{q}$. As all creation operators commute (as shown in (2.5.12)), the two particles are freely interchangeable. In other words, we have

$$a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger |0\rangle = a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle. \quad (2.5.32)$$

Particles whose dynamics are described by the Klein-Gordon equation therefore obey *Bose-Einstein* statistics. Conversely, a single mode of total momentum \mathbf{p} can contain arbitrarily many particles. A given state of specific momentum \mathbf{p} and energy $E = \omega$ that is constructed by the successive application of creation operators to the vacuum therefore consists of a linear combination of n -particle states with n not fixed. In this context, the Hilbert space \mathcal{H} suitable to describe the eigenvectors of a single harmonic oscillator is promoted to a *Fock space* \mathcal{F} that is defined as a direct sum of copies of the single-particle Hilbert space \mathcal{H} ,

$$\mathcal{F} = \bigoplus_{n=0}^{\infty} \mathcal{V}_n = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n} = \mathbb{C} + \mathcal{H} + \mathcal{H} \otimes_s \mathcal{H} + \dots \quad (2.5.33)$$

In this notation, $\mathcal{V}_0 \equiv \mathbb{C}$ stands for the set of complex scalars and contains the vacuum $|0\rangle$. The first term in the sum, $\mathcal{V}_1 \equiv \mathcal{H}$, contains the full set of one-particle states $|\mathbf{p}\rangle$ of generic momentum \mathbf{p} and positive energy defined according to special relativity (so that the four components of the four-momentum are not independent of the particle mass). The following terms \mathcal{V}_k (with $k \geq 2$) are defined as the symmetric tensor product of k copies of the Hilbert space \mathcal{H} , and a specific term \mathcal{V}_k includes states containing k identical particles.

Second quantisation is a procedure in which an infinite set of quantum mechanical systems are treated simultaneously within a quantity known as a quantum field. Each of these systems is associated with a specific value of the momentum \mathbf{p} , and it is represented by a harmonic oscillator of frequency $\omega_{\mathbf{p}}$. We interpret the n^{th} excitation $|n\rangle$ of this oscillator as a state containing n particles of mass m and energy E related through

$$E = \omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}. \quad (2.5.34)$$

In QFT a state with total momentum \mathbf{p} and total energy E is a combination of multiparticle states containing an arbitrary number n of particles of momenta $\{\mathbf{p}_n\}$ and energies $\{\omega_{\mathbf{p}_n}\}$ such that

$$E = \sum_{i=1}^n \omega_{\mathbf{p}_i} \quad \text{and} \quad \mathbf{p} = \sum_{i=1}^n \mathbf{p}_i. \quad (2.5.35)$$

Moreover, (2.5.34) holds for any individual particle.

A QFT state defined as above lives in a Fock space \mathcal{F} that is defined as an infinite direct sum of symmetric tensorial products of the Hilbert spaces \mathcal{H} describing one-particle states. Each component of the QFT state within this sum is therefore associated with a specific number of particles, but this number is not bounded from above as the sum is infinite. QFT therefore provides the relativistic multiparticle theory that was needed from the beginning.

The generalisation to a theory describing particles of different types is immediate, the Fock space being then defined from a sum of copies of all Hilbert spaces associated with the different particle species.

From the above considerations, we can show that non-relativistic multiparticle quantum mechanics is retrieved as the low-energy limit of QFT, *i.e.* when the amount of available energy is both bounded from above and much smaller than the particle mass. It corresponds to a restricted region of the Fock space \mathcal{F} in which the number of states n is either fixed or bounded from above. Finally, we can also demonstrate that in QFT, the causality problems arising from particle propagation are solved. This is further elaborated on in chapter ??.

2.6 Summary

This chapter provides a first simplified approach to quantum field theory.

Our journey started with the postulates of quantum mechanics. We formulated them using the elegant and abstract formalism of Dirac, which relies on (rigged) Hilbert spaces. Our discussion additionally included all mathematical tools necessary for the understanding of these notes, and we particularly emphasised the momentum and position operators \mathbf{x} and \mathbf{p} whose eigenstates form a basis of the Hilbert space. As an example of the application of the postulates of quantum mechanics, we next considered the quantisation of the simple harmonic oscillator, a quantum system that is central in QFT. In Dirac notation, the quantisation of the associated classical and non-relativistic problem involves creation and annihilation operators, allowing to create and annihilate quanta of energy. These operators are crucial because they also appear in QFT. However, in QFT they are alternatively interpreted as operators yielding the creation and annihilation of (relativistic) particles of mass m , momentum \mathbf{p} and energy $E = \omega_{\mathbf{p}} = \sqrt{m^2 + \|\mathbf{p}\|^2}$.

The usual non-relativistic formulation of quantum mechanics is, however, not adapted to the relativistic case. Because of special relativity, mass indeed consists only of one form of energy among others, which means that particles can annihilate and be created over time. Consequently, the number of particles in a relativistic system varies with time, and it is clear that such a feature cannot be accounted for immediately from a non-relativistic quantum mechanical system comprising a fixed number of particles. In addition, non-relativistic quantum mechanics leads to problems with causality as due to the postulates of special relativity; propagation in space cannot be faster than light. Quantum fields are the objects that provide a solution to these problems.

Before moving deeply into the main subject of these notes in the last part of this chapter, we enjoyed an *intermezzo* to recapitulate the Lagrangian formulation and Hamiltonian formulation of classical mechanics. Here, we considered the description of a system containing a fixed number n of particles localised at specific positions. Taking the continuous limit with n tending to infinity, the concept of fields and conjugate momenta emerged from our discussion. We next related them to the Lagrangian and Hamiltonian densities that play an important role in classical field theory. This naturally leads us to introduce the quantity known as the action of the system, and how the associated principle of least action can be used to derive its evolution over time. Moreover, we provided some details on the consequences of the symmetries of the action, which leads us to the celebrated Noether's theorem. As an illustrative application, we focused on the derivation of one of the most golden rules of physics: energy and momentum conservation.

With this in mind, we had all the ingredients to begin a discussion on quantum fields. We began from the simplest relativistic equation, namely the Klein-Gordon equation. We then derived from it a consistent definition of a quantum field that solves all problems mentioned above, and that lies at the heart of a relativistic and quantum multiparticle theory. This field and associated conjugate momentum are operators given by the expressions

$$\begin{aligned}\phi(x) &= \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{-ip \cdot x} + a_{\mathbf{p}}^\dagger e^{ip \cdot x} \right), \\ \pi(x) &= -i \int_{\mathbb{R}^3} \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \left(a_{\mathbf{p}} e^{-ip \cdot x} - a_{\mathbf{p}}^\dagger e^{ip \cdot x} \right).\end{aligned}$$

This shows that these two objects consist of a combination of an infinite set of creation and annihilation operators $a_{\mathbf{p}}^\dagger$ and $a_{\mathbf{p}}$ associated with the creation and annihilation of a particle of mass m , momentum \mathbf{p} and energy $\omega_{\mathbf{p}}$ at a space-time point x . Through several applications of the field operator to the vacuum state $|0\rangle$, we can successively build a state containing a specific number of particles k , although the number of applications of the field operator can be arbitrary large. Alternatively, a QFT state of definite total momentum \mathbf{p}_{tot} and total energy E_{tot} can be seen as the direct sum of all one-particle state, two-particle state, *etc.* of the same total energy and momentum. Whereas for any component in this direct sum the number of particle is fixed, the number of terms included is not bounded from above. Consequently, the concept of Hilbert spaces inherent to quantum mechanics is promoted to that of a Fock space.

This chapter, together with the previous one, achieves our presentation of all the building blocks relevant for QFT. In the next chapter, we change topic and focus onto scattering theory, that we present

with the goal of using quantum fields to practically calculate quantities useful for high-energy physics phenomenology, which includes, in particular, cross sections and decay rates.

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.
- [9] M. Planck, *Ueber das gesetz der energieverteilung im normalspectrum*, *Annalen der Physik* **309** (1901) 553–563.
- [10] A. Einstein, *Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt*, *Annalen der Physik* **322** (1905) 132–148.
- [11] E. Schrödinger, *An undulatory theory of the mechanics of atoms and molecules*, *Phys. Rev.* **28** (1926) 1049–1070.
- [12] P. A. M. Dirac, *A new notation for quantum mechanics*, *Mathematical Proceedings of the Cambridge Philosophical Society* **35** (1939) 416–418.
- [13] J. L. Lagrange, *Analytical Mechanics*, vol. 191 of *Boston Studies in the Philosophy and History of Science*. Springer Dordrecht, 2010, [10.1007/978-94-015-8903-1](https://doi.org/10.1007/978-94-015-8903-1).
- [14] W. R. Hamilton, *On a General Method of Expressing the Paths of Light & of the Planets by the Coefficients of a Characteristic Function*. P.D. Hardy, Dublin, 1833.
- [15] E. Noether, *Invariante variationsprobleme*, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918** (1918) 235–257.
- [16] O. Klein, *Quantentheorie und fünfdimensionale relativitätstheorie*, *Zeitschrift für Physik* **37** (1926) 895–906.
- [17] W. Gordon, *Der comptoneffekt nach der schrödingerschen theorie*, *Zeitschrift für Physik* **40** (1926) 117–133.
- [18] P. A. M. Dirac, *The quantum theory of the emission and absorption of radiation*, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **114** (1927) 243–265.
- [19] R. P. Feynman, *The Theory of positrons*, *Phys. Rev.* **76** (1949) 749–759.

- [20] H. Lehmann, K. Symanzik and W. Zimmermann, *On the formulation of quantized field theories*, *Nuovo Cim.* **1** (1955) 205–225.
- [21] G. C. Wick, *The Evaluation of the Collision Matrix*, *Phys. Rev.* **80** (1950) 268–272.
- [22] H. Geiger and E. Marsden, *On a diffuse reflection of the α -particles*, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **82** (1909) 495–500.
- [23] H. Geiger, *The scattering of α -particles by matter*, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **83** (1910) 492–504.
- [24] E. Rutherford, *The scattering of alpha and beta particles by matter and the structure of the atom*, *Phil. Mag. Ser. 6* **21** (1911) 669–688.
- [25] S. Mandelstam, *Determination of the pion - nucleon scattering amplitude from dispersion relations and unitarity. General theory*, *Phys. Rev.* **112** (1958) 1344–1360.
- [26] G. Goldberg, *A Rule for the Combinatoric Factors of Feynman Diagrams*, *Phys. Rev. D* **32** (1985) 3331.
- [27] E. H. Wichmann and J. H. Crichton, *Cluster Decomposition Properties of the S Matrix*, *Phys. Rev.* **132** (1963) 2788–2799.
- [28] S. Weinberg, *What is quantum field theory, and what did we think it is?*, in *Conference on Historical Examination and Philosophical Reflections on the Foundations of Quantum Field Theory*, pp. 241–251, 3, 1996. [hep-th/9702027](https://arxiv.org/abs/hep-th/9702027).