MASTER GENERAL PHYSICS

université
PARIS-SACLAY

## Major Particles, Nuclei and Universe

# Particles and Symmetries

## Samuel Wallon
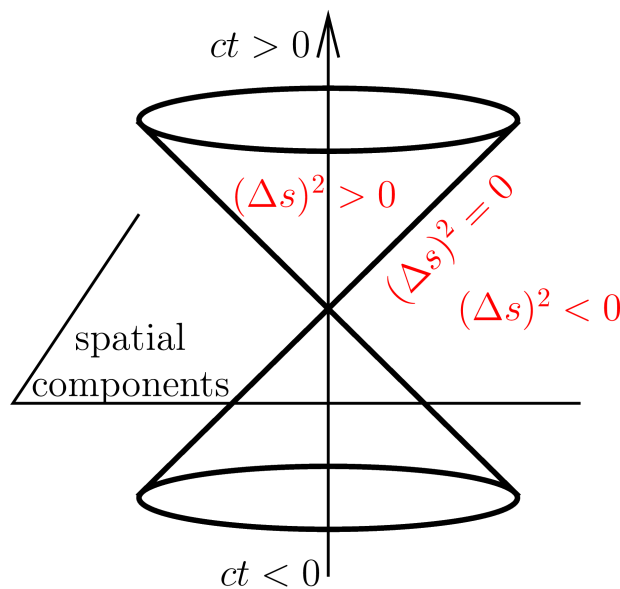
version: 28th November 2022

# Contents

# Contents

# II. Introduction to Quantum Electrodynamics     39

## 3. From nonrelativistic perturbation theory to Feynman diagrams    41

## 4. Electrodynamics     51

# Part I.

# Relativistic symmetries

# 1. Lorentz group

Maxwell equations, which describe electric and magnetic phenomena, are not invariant under the Galilean group. Thus, contrarily to what states the galilean principle, one should be able to reveal an absolute reference system, named ether. Several experiments led by Michelson and Morley (1881) invalidated such an hypothesis: speed of light is independent of the direction of propagation. Time cannot be considered as an absolute variable: simultaneity of two phenomena in reference frame does not imply simultaneity in another one, in contrast with classical mechanics.

## 1.1. Space-time interval

Let us briefly remind here the notion of space-time interval in special relativity. First consider an event described by an observer, assumed to be fixed in an inertial frame, by its space-time coordinate $(ct_1, \vec{x}_1)$.

*Example* : Emission of a photon in $\pi^0 \to \gamma\gamma$ disintegration:

The signal propagates at the speed of light (photon). It is detected at the space-time coordinates $(ct_2, \vec{x}_2)$.

By definition, the square of the space-time interval between two events is given by

$$\boxed{(\Delta s)^2 = c^2 \left(t_2 - t_1\right)^2 - \left(\vec{x}_2 - \vec{x}_1\right)^2 .} \qquad (1.1)$$

Of course in the present case, $\vec{x}_2 - \vec{x}_1 = c(t_2 - t_1)\vec{u}$ where $\vec{u}$ is a unitary vector pointing in the direction of the photon trajectory, and thus $(\Delta s)^2 = 0$.
A key assumption, made by Einstein, is the fact that $c$ does not depend on the inertial frame, therefore $(\Delta s)^2 = 0$ in **every** inertial frame in which one can wish to describe this experiment.

From the definition (1.1), the infinitesimal interval square reads

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 . \qquad (1.2)$$

## 1. Lorentz group

If $ds^2 = 0$ in an inertial frame $K$, then $ds'^2 = 0$ in any other inertial frame $K'$. In the infinitesimal limit, one should thus have $ds^2 = a\,ds'^2$.

Using the homogeneity of time and space, and isotropy of space, one can show that $a = \text{constant} = 1$.

*Exercise 1.1*

Show this (see Landau, Lifshitz, *The Classical Theory of Fields*, p. 12).

The space of points in space-time with the interval (1.2) is called Minkowski space[1]. This is called a pseudo-euclidean geometry, in contradistinction with usual euclidean geometry.

A point in Minkowski space $M$ is described by its **contravariant** coordinates $x^\mu = (ct, \vec{x})$. This is an example of a quadrivector.

**Metric:** it is given by the matrix

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{1.3}$$

and thus

$$(\Delta s)^2 = g_{\mu\nu}\left(x_2{}^\mu - x_1{}^\mu\right)\left(x_2{}^\nu - x_1{}^\nu\right),$$

which is identical to the square of interval defined previously.
One can now define 3 kinds of intervals depending on the sign of $(\Delta s)^2$:

⋄ $(\Delta s)^2 > 0$ time-like interval

⋄ $(\Delta s)^2 < 0$ space-like interval

⋄ $(\Delta s)^2 = 0$ light-like interval,

as illustrated in Fig. 1.1.

The infinitesimal square of the space-time interval thus reads

$$\boxed{ds^2 = g_{\mu\nu}\,dx^\mu dx^\nu.} \tag{1.4}$$

---

[1] Hermann Minkowski introduced this concept of physical space in 1907. Henri Poincaré introduced the same idea in 1905 from a mathematical point of view, in connexion with the invariance of Maxwell equations under change of inertial frame.
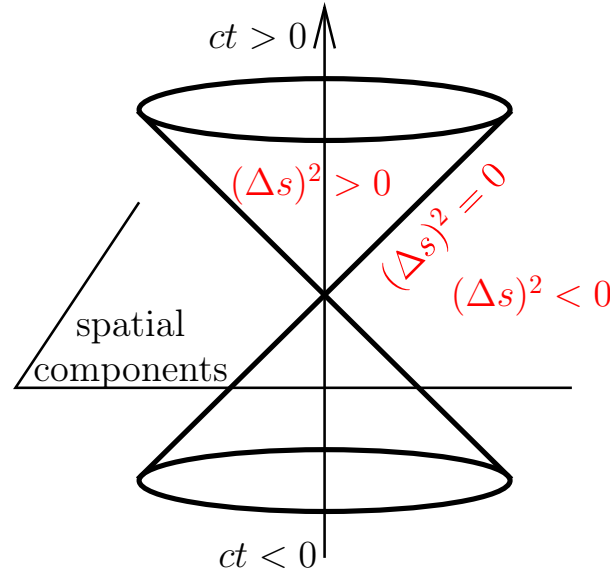
Figure 1.1.: The 3 different kinds of intervals.

*Exercise 1.2*

Show that the interval separating the emission and the detection of a particle in any process is either time-like or light-like.

A trajectory is given by a curve $X^\mu(s)$ where $s$ is an arbitrary parametrization of this trajectory. The infinitesimal proper time $d\tau$ is defined through

$$(cd\tau)^2 = (cdt)^2 - (d\vec{x})^2 \,. \tag{1.5}$$

Denoting the velocity as

$$\vec{v} = \frac{d\vec{x}}{dt} \tag{1.6}$$

we have

$$cd\tau = \sqrt{(cdt)^2 - (\vec{v})^2 dt^2} = cdt\sqrt{1 - \frac{\vec{v}^2}{c^2}} \,, \tag{1.7}$$

i.e.

$$dt = \frac{d\tau}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \tag{1.8}$$

or equivalently

$$\boxed{dt = \gamma d\tau \quad \text{with} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} \quad \text{and} \quad \vec{\beta} = \frac{\vec{v}}{c} \,.} \tag{1.9}$$

From the invariance of $ds^2$ under change of inertial frame, we deduce that the proper time $d\tau$ does not depend on the choice of inertial frame: it is a Poincaré scalar. We will characterize Poincaré transformations in the next section.

The relation (1.9) reveals a key feature. Indeed, suppose that a particle is at rest in a given inertial frame, named proper frame. This is always possible if its velocity differs from $c$, using a pure Lorentz boost along $\vec{v}$, see next section. Passing to another arbitrary inertial frame, of non zero velocity with respect to the proper frame, we will have $\gamma > 1$, which implies the famous phenomena of **time dilation**: $dt > d\tau$ : this means that time runs slower for a particle in any inertial frame with a non zero velocity with respect to a proper frame.

Throughout the rest of the text, unless otherwise stated, we will use a set of units in which $c = 1$, and thus $x^0 = t$, $x^1 = x$, $x^2 = y$, $x^3 = z$.

## 1.2. Poincaré and Lorentz groups

Having postulated that the speed of light is independent of the inertial frame, we have seen the important consequence that the infinitesimal interval $ds^2$ is identical in every inertial frame. We now want to determine the relativistic transformations relating inertial frames. We will then study the group structure of these transformations.

### 1.2.1. Characterization of $\mathcal{P}$

Consider a point $x$ of $M$, measured in frame $K$, and $x'$ the coordinates of the same point in frame $K'$.

The invariance of the infinitesimal interval

$$ds'^2 = ds^2 \tag{1.10}$$

implies that

$$\boxed{\begin{array}{c} \text{translations (4 real parameters)} \\ \swarrow \\ x'^\alpha = \Lambda^\alpha{}_\beta x^\beta + a^\alpha \ : \text{Poincaré group } \mathcal{P} \text{ (10 real parameters)} \\ \nwarrow \\ \text{Lorentz group } \mathcal{L} \text{ (6 real parameters)} \end{array}} \tag{1.11}$$

*Proof :*

The condition (1.10) reads

$$g_{\alpha\beta}\, dx'^{\alpha} dx'^{\beta} = g_{\mu\nu}\, dx^{\mu} dx^{\nu} \, ,$$

i.e.

$$g_{\alpha\beta} \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} dx^{\mu} dx^{\nu} = g_{\mu\nu}\, dx^{\mu} dx^{\nu}$$

thus

$$g_{\alpha\beta} \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} = g_{\mu\nu} \, . \tag{1.12}$$

Taking the determinant of each side of this equation, one thus gets

$$\det g \left( \det \frac{\partial x'}{\partial x} \right)^2 = \det g \, .$$

Since $g$ is regular, this implies that $\det \dfrac{\partial x'}{\partial x} \neq 0$ : Poincaré transformations are thus invertible.

Acting with $\dfrac{\partial}{\partial x^{\rho}}$ on Eq. (1.12), one gets

$$g_{\alpha\beta} \frac{\partial^2 x'^{\alpha}}{\partial x^{\mu}\partial x^{\rho}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} + g_{\alpha\beta} \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial^2 x'^{\beta}}{\partial x^{\nu}\partial x^{\rho}} = 0$$

which symbolically reads

$$A_{(\mu\rho)\nu} + A_{(\nu\rho)\mu} = 0 \, , \tag{1.13}$$

where $A$ is symmetric with respect to indexes inside brackets.

$$\mu \leftrightarrow \rho \quad \text{gives} \quad A_{(\mu\rho)\nu} + A_{(\mu\nu)\rho} = 0 \, , \tag{1.14}$$

$$\nu \leftrightarrow \rho \quad \text{gives} \quad A_{(\mu\nu)\rho} + A_{(\nu\rho)\mu} = 0 \, . \tag{1.15}$$

The combination (1.13) + (1.14) -(1.15) thus gives $2A_{(\mu\rho)\nu} = 0$, i.e.

$$g_{\alpha\beta} \frac{\partial^2 x'^{\alpha}}{\partial x^{\mu}\partial x^{\rho}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} = 0$$

and thus (since $\det \dfrac{\partial x'}{\partial x} \neq 0$) : $\dfrac{\partial^2 x'^{\alpha}}{\partial x^{\mu}\partial x^{\rho}} = 0$

which proves that $x'$ are linear functions of $x$ .

———

## 1. Lorentz group

We use matrix notations, namely

$$\Lambda^\alpha{}_\beta \quad \nearrow \text{ line index} \quad \searrow \text{ column index}$$

One deduces from Eqs. (1.12) and (1.11) that $\Lambda$ is a real matrix satisfying

$$\boxed{g_{\alpha\beta}\,\Lambda^\alpha{}_\mu\Lambda^\beta{}_\nu = g_{\mu\nu}\,.} \tag{1.16}$$

The set $\mathcal{L}$ of Lorentz transforms is a group (see Appendix A for a very short introduction to group theory).

*Proof :*

• this is obvious from the definition: the product of two such homogeneous transformations preserving $ds^2$ is an homogeneous transformation which preserves $ds^2$ ; on the hand, the inverse of an homogeneous transformation preserving $ds^2$ is an homogeneous transformation preserving $ds^2$ . And of course the identity preserves $ds^2$.

• algebraically, if $\Lambda$ and $\Lambda'$ belong to $\mathcal{L}$, then

$$\begin{cases} g_{\alpha\beta}\Lambda^\alpha{}_\mu\Lambda^\beta{}_\nu = g_{\mu\nu} \\ g_{\alpha\beta}\Lambda'^\alpha{}_\mu\Lambda'^\beta{}_\nu = g_{\mu\nu}\,. \end{cases}$$

This implies that

$$g_{\alpha\beta}\,\Lambda^\alpha{}_\mu\Lambda'^\mu{}_{\mu'}\Lambda^\beta{}_\nu\Lambda'^\nu{}_{\nu'} = g_{\mu\nu}\Lambda'^\mu{}_{\mu'}\Lambda'^\nu{}_{\nu'} = g_{\mu'\nu'}\,,$$

and thus that $\Lambda\,\Lambda' \in \mathcal{L}\,.$

The set $\mathcal{P}$ of Poincaré transformations, characterized by a pair $(a,\Lambda)$ where $a$ is a 4-vector and $\Lambda \in \mathcal{L}$, is a group.

*Proof :*

The first of the two above proofs immediately extends to the case of the inhomogeneous Lorentz group (or Poincaré group). The second proof relies on the group law on $\mathcal{P}$ :

given two Poincaré transformations $(a, \Lambda)$ and $(a', \Lambda')$, it is an easy exercise to show that their product reads

$$(a', \Lambda') \cdot (a, \Lambda) = (\Lambda' a + a', \Lambda' \Lambda) \tag{1.17}$$

of the required form $(a'', \Lambda'')$.

On should notice that the defining relation (1.16) of $\mathcal{L}$ reads, in matrix form,

$$\boxed{\Lambda^t \, g \, \Lambda = g \,.} \tag{1.18}$$

This is easily recovered in the following way: $\Lambda$ being a linear transformation, which leaves the scalar product invariant, one has, for arbitrary quadrivectors $v$ and $v'$ of $M$, denoted as column vectors,

$$v' \cdot v = (v')^t g v = (\Lambda v')^t g (\Lambda v) = (v')^t \Lambda^t g \Lambda v \tag{1.19}$$

which is equivalent to (1.18).

## 1.2.2. Structure of the Lorentz group

From Eq. (1.16) one deduces that $\det \Lambda = \pm 1$ :

$$\left|\begin{array}{lll} \det \Lambda = +1 : & \text{proper transformations} & \mathcal{L}_+ \\ \det \Lambda = -1 : & \text{improper transformations} & \mathcal{L}_- \,. \end{array}\right.$$

Again considering Eq. (1.16) with $\mu = \nu = 0$, one gets $\left(\Lambda^0{}_0\right)^2 - \sum\limits_{i=1}^{3} \left(\Lambda^i{}_0\right)^2 = 1$ :

$$\left|\begin{array}{lll} \Lambda^0{}_0 \geqslant 1 : & \text{no time reversal (orthochronous)} & \mathcal{L}^\uparrow \\ \Lambda^0{}_0 \leqslant -1 : & \text{with time reversal (antichronous)} & \mathcal{L}^\downarrow \,. \end{array}\right.$$

One can show (see Gelfand p. 165) that

> The Lorentz group $O(3,1)$ has 4 connected component. Each of them is doubly connected, but not simply connected.

Connected component: in a connected component, one can pass from one element to another one continuously.

Simply and multiply connected space:

A space is simply connected if any loop in this space can be contracted to a point. Equivalently, considering two paths, one can pass continuously from the

first one to the second one. If not, it is multiply connected. In our present case, doubly connected means that in each of these components, there are two classes of paths; considering two paths in a given class, one can pass from the first one to the second one by continuous transformation of the parameters.
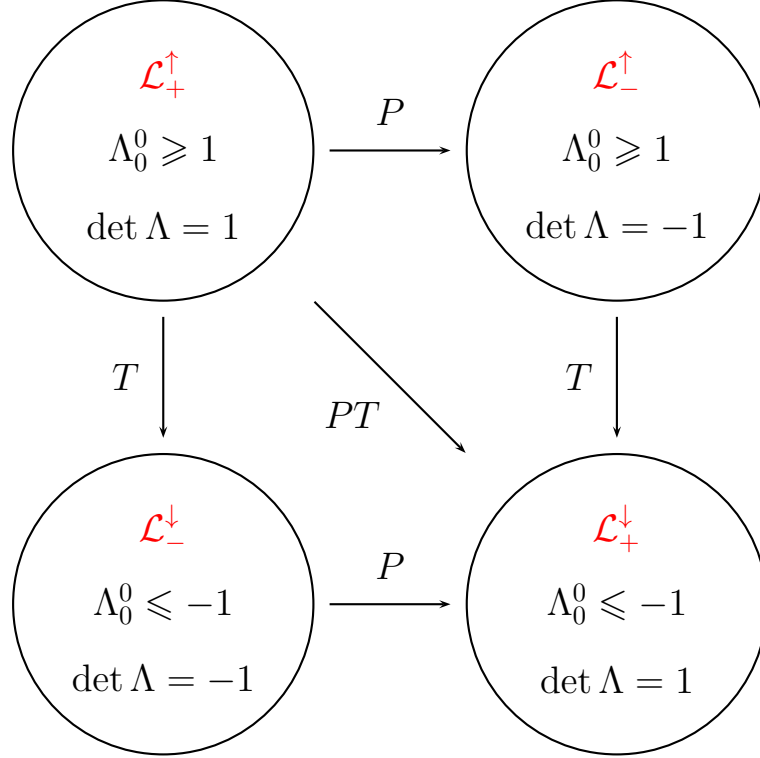


Figure 1.2.: The four connected components of $\mathcal{L}$.

These various connected components are related by discrete transformations:

- **space inversion (or parity):**

$$P = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

- **time reversal:**

$$T = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

This is illustrated in Fig. 1.2.

$\mathcal{L}$ leaves the 3 regions $x^2 > 0$, $x^2 = 0$ (light-cone) and $x^2 < 0$ invariant.

$\mathcal{L}^\uparrow_+$ leaves the 4 regions $\left| \begin{array}{l} x^2 > 0 \\ x_0 > 0 \end{array} \right.$, $\left| \begin{array}{l} x^2 > 0 \\ x_0 < 0 \end{array} \right.$, $x^2 = 0$ , $x^2 < 0$ invariant, as shown in Fig 1.1.

$\mathcal{L}^\uparrow_+$ is the **restricted** Lorentz group

$\mathcal{L}^\uparrow_+$ and $\mathcal{L}^\uparrow_-$ form together the **complete** Lorentz group

$\mathcal{L}^\uparrow_+$ and $L^\downarrow_+$ form together the **proper** Lorentz group

*Exercise 1.3*

Check that the restricted Lorentz group, the complete Lorentz group and the proper Lorentz group are subgroups of the Lorentz group $\mathcal{L}$.

**Compactness:**

> The restricted Lorentz group $\mathcal{L}^\uparrow_+$ is non compact.

This is due to the existence of boosts, also called special Lorentz transforms, which we will examine later, encoded by a rapidity which varies from $-\infty$ to $+\infty$.

## 1.2.3. The two types of Lorentz transformations

We restrict ourselves to the restricted Lorentz group.

**Rotations**

The first set of Lorentz transformations is well known: they are the usual rotations, elements of $SO(3)$, defined as 3-dimensional transformations which preserve the euclidean scalar product. They are characterized by the angle $\psi$ of the rotation, and the axis of the rotation, itself defined through a unitary vector $\vec{n}$ depending on two spherical angles $\theta, \phi$. Thus, a rotation is defined by three real numbers varying in compact sets

$$0 \leqslant \theta \leqslant \pi, \ 0 \leqslant \phi < 2\pi, \ 0 \leqslant \psi \leqslant \pi \,, \tag{1.20}$$

using the fact that a rotation around an oriented axis spanned by $\vec{n}$ and of angle $\psi$ is identical to a rotation of axis $-\vec{n}$ and angle $2\pi - \psi$, which allows to

restrict ourselves to $0 \leqslant \psi \leqslant \pi$.

*Exercise 1.4*

Show that a rotation of axis $x^1$ and angle $\Psi$ in the passive point of view (i.e. the axes rotates) reads

$$\begin{cases} x^{2\prime} = \cos\psi\, x^2 + \sin\psi\, x^3 \\ x^{3\prime} = -\sin\psi\, x^2 + \cos\psi\, x^3. \end{cases} \qquad (1.21)$$

*Exercise 1.5*

Show that a passive rotation of axis $\vec{n}$ ($\vec{n}^2 = 1$) and angle $\Psi$ reads

$$\vec{x}' = (\vec{x} - \vec{n}(\vec{n} \cdot \vec{x}))\cos\theta - \vec{n} \wedge \vec{x}\sin\theta + \vec{n}(\vec{n} \cdot \vec{x}) \qquad (1.22)$$

Indication: consider how each part of the decomposition of an arbitrary vector $\vec{x}$ as a sum of a component $\vec{x}_{\parallel}$ along the axis of the rotation and a component $\vec{x}_{\perp}$ perpendicular to it

$$\vec{x} = \vec{x}_{\parallel} + \vec{x}_{\perp} = [(\vec{x} \cdot \vec{n})\vec{n}] + [\vec{x} - (\vec{x} \cdot \vec{n})\vec{n}]$$

transforms under a rotation.

## Boosts

Besides, pure Lorentz transformations, also named boosts, affect a single space direction and time.

*Exercise 1.6*

Show that restricted Lorentz transforms which leaves invariant $x^2, x^3$ reads

$$\begin{cases} x^{0\prime} = \cosh\phi\, x^0 - \sinh\phi\, x^1 \\ x^{1\prime} = -\sinh\phi\, x^0 + \cosh\phi\, x^1 \end{cases} \qquad (1.23)$$

where $\phi \in ]-\infty, \infty[$.
Indication: consider a generic $2 \times 2$ linear transformation mixing $x^0$ and $x^1$, leaving the space-time interval invariant and preserving the sign of $x^0$, and solve for the 4 coefficients.

In the passive point of view, $\Lambda$ encodes the transformation from a frame $F$ to a frame $F'$, while in the active point of view, it encodes the change of a point

in a fixed frame. In the passive point of view, the movement of the origin $O'$ of the new frame $F'$ with respect to the origin $O$ of the initial frame $F$ is obtained by setting $x^{1'} = 0$, i.e. $x^1 = \tanh \phi\, x^0$ corresponding to a speed $\vec{v} = c \tanh \phi\, \vec{e_1}$ of the frame $F'$ with respect to the frame $F$, along the $x$ axis.

One should note that introducing $\beta(-1 \leqslant \beta \leqslant 1)$ as

$$\vec{v} = c\beta \vec{e_1}\,, \tag{1.24}$$

and $\gamma\ (0 \leqslant \gamma)$ as

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}}, \tag{1.25}$$

we have

$$\beta\ =\ \tanh \phi\,, \tag{1.26}$$

$$\gamma\ =\ \frac{1}{\sqrt{1 - \tanh^2 \phi}} = \cosh \phi\,, \tag{1.27}$$

$$\gamma\beta\ =\ \sinh \phi\,, \tag{1.28}$$

so that the pure boost (1.23) can be equivalently rewritten, $\beta$ being algebraic,

$$\begin{cases} x^{0'} &=\quad\ \gamma\, x^0\quad -\quad \gamma\beta\, x^1 \\ x^{1'} &=\ -\gamma\beta\, x^0\quad +\quad \gamma\, x^1 \end{cases} \tag{1.29}$$

*Exercise 1.7*

Show that a boost in the direction $\vec{n}\ (\vec{n}^2 = 1)$, of rapidity $\phi$, i.e. with a speed $\vec{v} = c\,\mathrm{th}\phi\,\vec{n}$ of the frame $F'$ with respect to the frame $F$, reads

$$\begin{cases} x^{0'} &=\ x^0 \cosh\ \phi - (\vec{n} \cdot \vec{x})\ \sinh\ \phi \\ \vec{x}' &=\ \vec{x} - \vec{n}\,(\vec{x} \cdot \vec{n}) + \left(-x^0 \sinh\ \phi + (\vec{n} \cdot \vec{x}) \cosh\ \phi\right) \vec{n} \end{cases} \tag{1.30}$$

Indication: consider how each part of the decomposition $\vec{x} = [(\vec{x} \cdot \vec{n})\vec{n}] + [\vec{x} - (\vec{x} \cdot \vec{n})\vec{n}]$ transforms under a boost.

Generally speaking, a boost is therefore characterized by three real numbers $\phi_1, \phi_2, \phi_3$.

## 1.2.4.  Number of parameters

One can demonstrate that any element $\Lambda$ of the restricted Lorentz group can be written in the form

$$\Lambda = R\,B \tag{1.31}$$

where $R$ is a rotation and $B$ is a boost. This can be understood geometricaly. Indeed, going from one 4-vector to another 4-vector can be done in two steps: first, a boost makes it possible to bring the norm of the spatial component of the first 4-vector to the norm of the second one, and at the same time to bring the temporal part of the first 4-vector to the temporal part of the second one; then, a rotation makes it possible to bring the spatial direction of the 4-vector obtained from the first step to the direction of the final 4-vector.

We thus deduce that

$$\boxed{\mathcal{L} \text{ depends on 6 real parameters.}} \tag{1.32}$$

and

$$\boxed{\mathcal{P} \text{ depends on 10 real parameters:}} \tag{1.33}$$

it depends on 6 real parameters describing $\mathcal{L}$ to which one should add the 4 space-time translations.

# 1.3. Covariance

## 1.3.1. Contravariant and covariant vectors*

In a formal way, a vector belongs to a vector space $E$ (over the field $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$), spanned by a basis. This naturally gives a set of contravariant coordinates for this vector on this basis. A covector belongs to the dual space $E^*$ of linear forms[2] $E \rightarrow \mathbb{K}$. In the special case of Hilbert spaces, there is a one-to-one correspondence between $E$ and $E^*$, which is extremely convenient in practice. Let us explain these different steps in detail.

**Contravariance**

Consider a vector space $E$ (over the field $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$), of dimension $n$. A change of basis from $B = \{\vec{e}_1, \cdots \vec{e}_n\}$ to $B' = \{\vec{e}_1', \cdots \vec{e}_n'\}$ (conveniently, basis are denoted as row vectors) is given by

$$B' = BA \tag{1.34}$$

where $A$ is an invertible $n \times n$ matrix, with entries $A^i{}_j$. This reads explicitly[3]

$$\vec{e}_j' = \vec{e}_i A^i{}_j \,. \tag{1.35}$$

---

[2] We recall that a linear form is a linear map from a vector space to its field of scalars .

[3] If not stated explicitly, we will now use Einstein convention, which states that repeated indexes are summed.

A vector $\vec{v}$ can be expanded in the basis $B$, with notations which emphasize the fact that the basis $B$ is used,

$$\boxed{\vec{v} = v^i[B]\vec{e}_i = B\,v[B]\,,} \tag{1.36}$$

where the *column* vector of components $v^i[B]$ is denoted as $v[B]$,

$$v[B] = \begin{pmatrix} v^1[B] \\ v^2[B] \\ \vdots \\ v^n[B] \end{pmatrix} \tag{1.37}$$

which allows the use of a matrix product in the second equality of Eq. (1.36).

In the new basis $B'$, the vector $\vec{v}$ being invariant, one has the two equivalent expansions

$$\vec{v} = B\,v[B] = B'\,v[B'] = BA\,v[BA] \tag{1.38}$$

and thus

$$v[BA] = A^{-1}v[B]\,, \tag{1.39}$$

or equivalently

$$v[B] = A\,v[BA] = A\,v[B']\,. \tag{1.40}$$

Such a transformation law is named **contravariant**, since one has under the transformation $A$

$$B \longrightarrow B' \tag{1.41}$$
$$v[B] \longleftarrow v[B']\,. \tag{1.42}$$

## Covariance

Consider now the dual space $E^*$, the space of linear forms from $E$ to $\mathbb{K}$.

A natural (thus named canonical) dual basis can be defined as a set of $n$ linear forms $e^i$, through their action on the chosen basis of $E$, as

$$e^i(\vec{e}_j) = \delta^i_j\,. \tag{1.43}$$

Any linear form $\ell \in E^*$ can then be expanded in this basis as

$$\ell = \ell_i[B]e^i = \ell[\vec{e}_i]e^i\,. \tag{1.44}$$

Indeed any coefficient $\ell_i[B]$ of this expansion can be easily obtained through the action of $\ell$ on a vector basis $\vec{e}_i$:

$$\ell(\vec{e}_i) = \ell_j[B]e^j(\vec{e}_i) = \ell_j[B]\delta^i_j = \ell_i[B]\,. \tag{1.45}$$

This expansion can be written in a more compact and useful form, as

$$\ell = \ell_i[B]e^i = \ell[\vec{e}_i]e^i = \ell[B]B^* \,, \tag{1.46}$$

in which the *row* vector of components $\ell_i[B]$ is denoted as

$$\ell[B] = (\ell_1[B], \cdots , \ell_n[B]) = (\ell(\vec{e}_1), \cdots , \ell(\vec{e}_n)) \tag{1.47}$$

and a *column* vector of components $e^i$ is denoted as $B^*$

$$B^* = \begin{pmatrix} e^1 \\ e^2 \\ \vdots \\ e^n \end{pmatrix} \tag{1.48}$$

which allows the use of a matrix product in the last equality of Eq. (1.46).

The action of $\ell \in E^*$ on an arbitrary vector $\vec{v} \in E$ thus reads

$$\ell(\vec{v}) = \ell(\vec{e}_i)e^i(\vec{v}) = \ell(\vec{e}_i)e^i(v^j\vec{e}_j) = \ell(\vec{e}_i)v^j e^i(\vec{e}_j) = \ell(\vec{e}_i)v^j\delta_j^i = \ell(\vec{e}_i)v^i \,, \tag{1.49}$$

were we have used the expansion of $\vec{v}$ on the basis $B$, see Eq. (1.36). Again using our compact notations (row vectors for forms, column vectors for vectors), we can write this as

$$\boxed{\ell(\vec{v}) = \ell(\vec{e}_i)v^i = \ell[B]v[B]\,.} \tag{1.50}$$

Let us now investigate how the change of basis (1.34) for $E$ translates in the dual space. We have

$$\boxed{\ell_i[B'] = \ell_j[B]A^j{}_i\,.} \tag{1.51}$$

Indeed, from Eqs. (1.45) and (1.35)

$$\ell_i[B'] = \ell_i[BA] = \ell(\vec{e}_i') = \ell(\vec{e}_j A^j{}_i) = \ell(\vec{e}_j)A^j{}_i = \ell_j[B]A^j{}_i\,. \tag{1.52}$$

Written in a more compact form, this means that

$$\ell[B'] = \ell[B]A \tag{1.53}$$

Such a transformation law is named **covariant**, since one has under the transformation $A$

$$B \longrightarrow B' \tag{1.54}$$

$$\ell[B] \longrightarrow \ell[B']\,. \tag{1.55}$$

Forms are therefore also named co-vectors.

**Correspondence between vectors and co-vectors in the case of a Hilbert space**

In the case of a Hilbert space (in finite dimension, a vector space equipped with a scalar product is automatically a Hilbert space; this is not always true for infinite dimension spaces), the Riesz representation theorem ensure that there is a one to one correspondence between linear forms and vectors:

Denoting the metric as $g$, and $(\,,\,)_g$ the associated scalar product,

$$\forall \ell \in E^*, \exists! \, \vec{\ell}_g \in E \, / \, \forall \vec{v} \in E, \ \ell(\vec{v}) = (\vec{\ell}_g, \vec{v})_g \,. \tag{1.56}$$

In practice, the scalar product is usually denoted without explicit reference to the underlying metric, so that the index $g$ is just removed.

The above discussion is familiar in quantum mechanics, where a form $\phi$ of the dual space $\mathcal{E}_H^*$ of the Hilbert space $\mathcal{E}_H$ is denoted as $\langle\phi|$ (bra) and means, by the correspondence theorem, $\langle\phi| = (\phi, \cdot)_g$ or explicitly, when acting on a vector $|\psi\rangle$ (ket), $\langle\phi|\psi\rangle = (\phi, \psi)_g$.

Coming back to Eq. (1.50), and removing any reference to $g$, we can now write

$$\boxed{\ell(\vec{v}) = \ell_i \, v^i = (\vec{\ell}, \vec{e}_i) \, v^i \,.} \tag{1.57}$$

## 1.3.2. Covariance and contravariance applied to Minkowski space

Let us consider a space of dimension $d$, with a basis $\underline{e}_{\mu\,(\mu=1,\cdots,d)}$, equipped with a metric $g$. As a special case, the Minkowski $M$ can be described through a basis $(\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4)$ denoted $\underline{e}_{\mu\,(\mu=0,\cdots,3)}$, the metric being given by Eq. (1.3). Note that we avoid here using arrows, since they are traditionally used for the spatial part of a 4-d vector.

**Contra-variant components**

An arbitrary quadrivector $\underline{A}$ can be decomposed in such a basis as

$$\underline{A} = A^\mu \underline{e}_\mu \,. \tag{1.58}$$

The $A^\mu$ are the *contra-variant components* of $\underline{A}$.

*Remark :* the $\underline{e}_\mu$ are not vectors since they are not invariant when changing the axis (recall that by definition a vector is an object with $d$ components which is invariant under changes of axis).

## Co-variant components

Using the metric $g$ which allows through (1.56) to define a co-vector associated to the vector $\underline{A}$, the *co-variant components* $A_\mu$ of this co-vector reads, according to Eq. (1.57),

$$A_\mu = \underline{A} \cdot \underline{e}_\mu \tag{1.59}$$

In general $A^\mu \neq A_\mu$. As an example, consider the euclidean plane, see Fig. 1.3, with a non orthonormal basis.
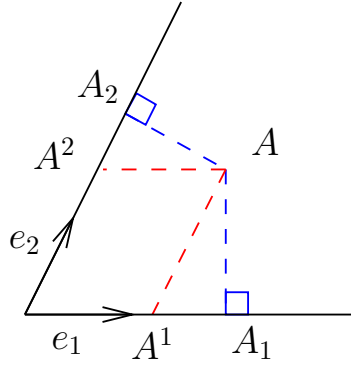


Figure 1.3.: Contravariant and covariant components.

## Relation between co-variant and contra-variant components

By definition, the metric tensor (or metric) is related to the vector basis as

$$\underline{e}_\mu \cdot \underline{e}_\nu = g_{\mu\nu}$$

with $g_{\mu\nu}$ given by (1.3) in the case of Minkowski space. We thus have

$$A_\mu = \underline{A} \cdot \underline{e}_\mu = A^\nu \underline{e}_\nu \cdot \underline{e}_\mu = g_{\mu\nu} A^\nu$$

We define $g^{\mu\nu}$ as the inverse of $g_{\mu\nu}$ : $g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho$.

In the case of Minkowski space, since $g_{\mu\nu}$ is its own inverse, $g_{\mu\nu} = g^{\mu\nu}$.
Thus

$$\boxed{\begin{aligned} A_\mu &= g_{\mu\nu} A^\nu \\ A^\mu &= g^{\mu\nu} A_\nu \,. \end{aligned}} \tag{1.60}$$

This implies that

$$\underline{A} \cdot \underline{B} = A^\mu B^\nu g_{\mu\nu} = A^\mu B_\mu = A_\mu B^\mu \,. \tag{1.61}$$

**Back to space-time interval**

The expression of the scalar-product (1.61) implies in particular that

$$ds^2 = d\underline{x} \cdot d\underline{x} = dx^\mu dx^\nu \underline{e}_\mu \cdot \underline{e}_\nu = dx^\mu dx^\nu g_{\mu\nu} \,,$$

in agreement with the above given definition (1.4).

One can also rely on co-variant space-time interval, so that we have the following identities

$$\boxed{ds^2 = g_{\mu\nu}dx^\mu dx^\nu = dx^\mu dx_\mu = g^{\mu\nu}dx_\mu dx_\nu \,.} \qquad (1.62)$$

**Inverse of a Lorentz transformation in covariant notations**

Let us consider a transformation $\Lambda$ of $\mathcal{L}$. From

$$g_{\alpha\beta}\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu = g_{\mu\nu} \,,$$

one gets

$$g_{\alpha\beta}\,\Lambda^\alpha{}_\mu = g_{\mu\nu}\left(\Lambda^{-1}\right)^\nu{}_\beta$$

thus

$$\left(\Lambda^{-1}\right)^\nu{}_\beta = g_{\beta\alpha}\Lambda^\alpha{}_\mu\,g^{\mu\nu} \,.$$

We denote

$$\Lambda_\beta{}^\nu = g_{\beta\alpha}\,\Lambda^\alpha{}_\mu\,g^{\mu\nu} \,. \qquad (1.63)$$

Thus

$$\boxed{\left(\Lambda^{-1}\right)^\nu{}_\beta = \Lambda_\beta{}^\nu \,.} \qquad (1.64)$$

Eq. (1.63) means that $\Lambda^\alpha{}_\mu$ is a tensor of order 2 which transforms as $A^\alpha B_\mu$, see section 1.3.3.

We thus have $\Lambda_\beta{}^\nu = \left(\Lambda^{-1}\right)^\nu{}_\beta = \left({}^t\Lambda^{-1}\right)^\beta{}_\nu$ and one can check explicitly that $\Lambda_\beta{}^\nu \Lambda^\rho{}_\nu = \left(\Lambda^{-1}\right)^\nu{}_\beta \Lambda^\rho{}_\nu = \Lambda^\rho{}_\nu \left(\Lambda^{-1}\right)^\nu{}_\beta = \delta^\rho_\beta$.

*Remark :*

With these notations, since $\mathbb{1} = \mathbb{1}^{-1} = \mathbb{1}^t$, one can write

$$\delta^\mu{}_\nu = \delta_\nu{}^\mu = \delta^\nu{}_\mu = \delta_\mu{}^\nu = \delta^\mu_\nu \,.$$

Conclusion:

The inverse of a matrix of $\mathcal{L}$ is easily obtained by rising and lowering indexes using the metric tensor, and then transpose.

## 1. Lorentz group

**Transformation law of co-variant components under $\mathcal{L}$**

Under a transformation $\Lambda$ of $\mathcal{L}$, the contra-variant components transform as

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}\, x^{\nu} \tag{1.65}$$

according to Eq. (1.11). This implies that

$$\boxed{x'_{\mu} = \Lambda_{\mu}{}^{\nu}\, x_{\nu}} \tag{1.66}$$

*Proof :*

This is just paste and copy of the relation (1.51) with here $A = \Lambda^{-1}$ :

$$x'_{\mu} = x_{\nu}(\Lambda^{-1})^{\nu}{}_{\mu} = \Lambda_{\mu}{}^{\nu}\, x_{\nu}\,. \tag{1.67}$$

More directly, one has indeed

$$\underline{x} = x^{\mu}\underline{e}_{\mu} = x'^{\nu}\underline{e}'_{\nu} = \Lambda^{\nu}{}_{\mu}x^{\mu}\underline{e}'_{\nu} \tag{1.68}$$

and thus

$$\underline{e}_{\mu} = \Lambda^{\nu}{}_{\mu}\underline{e}'_{\nu} \quad \text{and} \quad \underline{e}'_{\mu} = \Lambda_{\mu}{}^{\nu}\underline{e}_{\nu}$$

according to the tensor notation introduced for $\Lambda^{-1}$. Thus $x'_{\mu} = \underline{x} \cdot \underline{e}'_{\mu} = \underline{x} \cdot \Lambda_{\mu}{}^{\nu}\underline{e}_{\nu} = \Lambda_{\mu}{}^{\nu}x_{\nu}$.

*Remark :*

The fact that covariant components transform under ${}^{t}\Lambda^{-1}$, which is the inverse transpose of the matrix encoding the transformation of contra-variant components is due to the invariance of $ds^2 = dx^{\mu}dx_{\mu}$ which implies that $\Lambda^{t}g\Lambda = g$, see Eq. (1.18).

Indeed, this also reads ${}^{t}\Lambda^{-1} = g\,\Lambda\,g^{-1}$. Starting from $v' = \Lambda v$ for a contravariant $v$, one has

$$\underbrace{g\,v'}_{\text{covariant vector}} = \underbrace{g\,\Lambda\,g^{-1}}_{{}^{t}\Lambda^{-1}}\underbrace{g\,v}_{\text{covariant vector}}$$

## 1.3.3. Tensors

Based on the tools introduced above, we can introduce relativistic covariance in a more general context: physical quantities describing the state of a system

at generally speaking not the same for different observers. We can therefore distinguish:

• scalar quantities, which do not depend on a specific choice of reference frame. These are defined by scalar functions: given a function $\phi$ defined on $M$, in a given frame $F$, the corresponding scalar function in a new frame $F'$ is such that $\phi'(x') = \phi(x)$ in order to leave invariant the associated physical quantity.

• the contravariant components $v^\mu$ of a vector transform as coordinates (in the general case of $\mathcal{P}$, translations only affect coordinates) :

$$v'^\mu = \Lambda^\mu{}_\nu v^\nu \left( = \frac{\partial x'^\mu}{\partial x^\nu}\, v^\nu \right).$$

• the covariant components $v_\mu$ of a vector transform as $x_\mu$ :

$$v'_\mu = \Lambda_\mu{}^\nu v_\nu.$$

• tensor of rank $(\underbrace{m}_{\text{co-}}, \underbrace{n}_{\text{contra-}})$

$$F'_{\alpha_1 \cdots \alpha_m}{}^{\mu_1 \cdots \mu_n} = \Lambda_{\alpha_1}{}^{\beta_1} \cdots \Lambda_{\alpha_m}{}^{\beta_m} \Lambda^{\mu_1}{}_{\rho_1} \cdots \Lambda^{\mu_n}{}_{\rho_n} F_{\beta_1 \cdots \beta_m}{}^{\rho_1 \cdots \rho_n}.$$

• derivation :
$$\frac{\partial \phi}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial \phi}{\partial x^\nu} \quad \text{with} \quad \Lambda_\mu{}^\nu = \frac{\partial x^\nu}{\partial x'^\mu}.$$

Indeed,

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \quad \text{and} \quad x^\nu = (\Lambda^{-1})^\nu{}_\mu x'^\mu = \Lambda_\mu{}^\nu x'^\mu \quad \text{so that} \quad \Lambda_\mu{}^\nu = \frac{\partial x^\nu}{\partial x'^\mu}.$$

Thus $\dfrac{\partial \phi}{\partial x^\nu}$ transforms as a covariant vector, denoted as $\partial_\nu \phi$.

From the definition of contravariant coordinates $x^\mu = (x^0, \vec{x})$ one gets

$$\partial_\mu \;=\; \frac{\partial}{\partial x^\mu} = \left( \frac{\partial}{\partial x^0}, \vec{\nabla} \right), \tag{1.69}$$

$$\partial^\mu \;=\; \frac{\partial}{\partial x_\mu} = \left( \frac{\partial}{\partial x_0}, -\vec{\nabla} \right). \tag{1.70}$$

Beware of signs : $\vec{x}$ is contravariant but $\vec{\nabla}$ is covariant.

The quadri-divergence of a quadri-vector is the scalar

$$\partial^\mu A_\mu = \partial_\mu A^\mu = \frac{\partial A^0}{\partial x^0} + \vec{\nabla} \cdot \vec{A}$$

This leads naturally to the introduction of the scalar d'Alembertian operator:

$$\Box = \partial_\mu \partial^\mu = \frac{\partial^2}{\partial x_0^2} - \Delta \,. \tag{1.71}$$

Summary :

•  one can pass from one to another form of a given tensor by raising or lowering indexes through the metric tensor, as one does for quadri-vectors.

•  Lorentz transformation should act by letting its representative matrix acting on the left on a quadri-vector (either on its contravariant or covariant components), summing over repeated indexes (one up, one down). For contravariant components, this is just the matrix $\Lambda$ itself, and for covariant components, its transpose inverse.

## 1.3.4.  A few examples

Let us consider a few examples of tensors.

### Proper time

The first trivial example of Lorentz scalar is of course provided by the proper infinitesimal time $d\tau$ of any particle: it is a Lorentz invariant (which equals zero for a massless particle).

### Four-velocity

It is defined has

$$V^\mu = \frac{dX^\mu}{d\tau} = \left( c\frac{dt}{d\tau}, \frac{d\vec{x}}{d\tau} \right) = (\gamma c, \gamma \vec{v}) \tag{1.72}$$

which satisfies

$$V^2 = \gamma^2(c^2 - \vec{v}^2) = c^2 \,.$$

The above definition makes sense for a massive particle. For a massless particle, the above discussion fails since $d\tau = 0$ and $\gamma \to \infty$. Still, in any reference frame, we have $\frac{d\vec{x}}{dt} = c\vec{n}$ where $\vec{n}$ is a unit vector pointing in the direction of the motion of the particle, so that differentiating with respect to $t$ instead of $\tau$ leads to

$$V^\mu = (c, c\,\vec{n}) \,,$$

which now satisfies $V^2 = 0 \,.$

### Four-momentum (or four-momentum)

For a massive particle, it is defined as

$$P^\mu = mV^\mu = m\frac{dX^\mu}{d\tau} = (\gamma mc, \gamma m\vec{v}) = \left(\frac{E}{c}, \vec{p}\right) \qquad (1.73)$$

which satisfies

$$P^2 = m^2 c^2 \,.$$

The above definition makes sense for a massive particle. For a massless particle, the last equality is still valid, so that for arbitrary mass, we define

$$\boxed{P^\mu = \left(\frac{E}{c}, \vec{p}\right)} \,. \qquad (1.74)$$

Note that in any case,

$$\frac{\vec{v}}{c^2} = \frac{\vec{p}}{E}. \qquad (1.75)$$

### Four-current

From the usual charge density and the current density, one can define a four-current as

$$\boxed{J^\mu(x) = \left(c\rho(x), \vec{j}(x)\right)} \,. \qquad (1.76)$$

Introducing the rest charge density $\rho_0$ (i.e. the charge density in a rest frame), related to the charge density $\rho = \rho_v$ in a frame moving at velocity $\vec{v}$ with respect to the proper frame, one easily shows that

$$J^\mu(x) = (c\rho_v(x), \rho_v(x)\vec{v}) = \rho_0 V^\mu \,.$$

with[4] $\rho_v = \gamma \rho_0$, with as usual $\gamma = 1/\sqrt{1 - v^2/c^2}$, which shows that $J^\mu$ are the contravariant components of a four-vector ($\rho_0$ is a Lorentz scalar).

Let us consider the special case of the four-current associated to the motion of a charged particle, of charge $q$, moving along a trajectory $\vec{X}(t)$. We then have

$$\rho(ct, \vec{x}) = q\,\delta^{(3)}(\vec{x} - \vec{X}(t)) \qquad (1.77)$$

$$\vec{j}(ct, \vec{x}) = q\,\frac{d\vec{X}}{dt}\delta^{(3)}(\vec{x} - \vec{X}(t)) \,. \qquad (1.78)$$

---

[4] Indeed $\rho = \frac{d^3q}{d^3x}$ and $d^3x_v = \frac{1}{\gamma}d^3x_r$ when expressing the infinitesimal volume in the boosted frame ("v") in terms of the infinitesimal volume in the boosted frame ("r"), thus $\rho_v = \gamma\rho_0$.

*1. Lorentz group*

The trajectory reads

$$X^\mu(T) = (cT, \vec{X}(T)) \, . \tag{1.79}$$

It is therefore natural to write correspondingly, based on the idea that the four-current "feels" the charge at point $x$ if and only if it is there, covering any possibility by integrating over the whole trajectory,

$$J^\mu(x) = qc \int dT \, \frac{dX^\mu}{dT}(T) \, \delta^{(4)}(x - X(T)) \, . \tag{1.80}$$

First, we note that this expression is invariant under reparameterization since $dT$ cancels in this expression. Furthermore, this expression satisfies covariance. Indeed, $d^4x$ is a scalar (under a boost, $dx^0 \to \gamma dx^0$ and $d^3x \to \frac{d^3x}{\gamma}$) and thus $\delta^{(4)}(x)$ is also a scalar, since

$$\int d^4x \, \delta^{(4)}(x) = 1 \, .$$

Thus, $J^\mu(x)$ transforms as $dX^\mu$, i.e. as the contravariant components of a four-vector.
Inserting (1.79) inside the current (1.80), and since

$$\frac{dX^\mu}{dT}(T) = \left( c, \frac{d\vec{X}}{dT}(T) \right) , \tag{1.81}$$

we get for the four-current

$$J^\mu(ct, \vec{x}) = qc \int dT \, \delta^{(3)}(\vec{x} - \vec{X}(t)) \, \delta(ct - cT) \left( c, \frac{d\vec{X}}{dT}(T) \right) \tag{1.82}$$

$$= \left( cq\delta^{(3)}(\vec{x} - \vec{X}(t)), q\frac{d\vec{X}}{dt}\delta^{(3)}(\vec{x} - \vec{X}(t)) \right) \tag{1.83}$$

$$= \left( c\rho(x), \vec{j}(x) \right) \tag{1.84}$$

as expected.

# 2. The Klein-Gordon equation

In this chapter, our aim is to write a relativistic equation describing a massive spin 0 particle. Meanwhile, we will see that such a description naturally leads to states of negative energy, which will turn to be negative probability density solutions. This will be reinterpreted as antiparticles.

## 2.1. Reminder: nonrelativistic quantum mechanics

### 2.1.1. Correspondence principle and Schrödinger equation

We first start with the way Schrödinger equation can be obtained from the correspondence principle. Starting from the nonrelativistic dispersion equation

$$E = \frac{p^2}{2m} \tag{2.1}$$

and making the replacement, according to the correspondence principle,

$$E \rightarrow i\hbar\frac{\partial}{\partial t} \tag{2.2}$$

$$\vec{p} \rightarrow \frac{\hbar}{i}\vec{\nabla}\,, \tag{2.3}$$

we get

$$-\frac{\hbar^2}{2m}\Delta\psi(\vec{x},t) = i\hbar\frac{\partial\psi}{\partial t}\,, \tag{2.4}$$

which is the Schrödinger equation for a free particle.

### 2.1.2. Probability current

The probability density

$$\boxed{\rho = |\psi|^2} \tag{2.5}$$

has the physical meaning that

$$d^3P = |\psi|^2 d^3V \tag{2.6}$$

## 2. The Klein-Gordon equation

is the probability of finding a particle in the elementary volume $d^3V$.

On the one hand, starting from the Schrödinger equation (2.4), multiplied by $\psi^*$, and on the other hand writing its complex conjugate, multiplied by $\psi$, we get respectively

$$-\frac{\hbar^2}{2m}\psi^*\Delta\psi = i\hbar\,\psi^*\frac{\partial\psi}{\partial t} \tag{2.7}$$

$$-\frac{\hbar^2}{2m}\psi\Delta\psi^* = -i\hbar\,\psi\frac{\partial\psi^*}{\partial t} \tag{2.8}$$

and by subtraction

$$-\frac{\hbar^2}{2m}\left(\psi^*\Delta\psi - \psi\Delta\psi^*\right) = i\hbar\left[\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right] \tag{2.9}$$

so that

$$\frac{\partial\rho}{\partial t} = \psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t} = \frac{i\hbar}{2m}\left(\psi^*\Delta\psi - \psi\Delta\psi^*\right) \tag{2.10}$$

i.e., since $\Delta = \vec{\nabla}^2$,

$$\frac{\partial\rho}{\partial t} = \frac{i\hbar}{2m}\left(\psi^*\vec{\nabla}^2\psi - \psi\vec{\nabla}^2\psi^*\right) = \frac{i\hbar}{2m}\vec{\nabla}\cdot\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right) \tag{2.11}$$

which reads

$$\boxed{\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{j} = 0} \tag{2.12}$$

with the probability current, also named probability flux density, given by

$$\boxed{\vec{j} = -\frac{i\hbar}{2m}\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right).} \tag{2.13}$$

Eq. (2.12) expresses the local conservation of probability. The global point of view is obtained by integration over a given volume $V$: the increase of the probability for particles to be inside $V$, of boundary $S$, is given by

$$\frac{\partial}{\partial t}\iiint_V \rho\,d^3V = -\iiint_V \vec{\nabla}\cdot\vec{j}\,d^3V = -\iint_S \vec{j}\cdot d^2\vec{S} \tag{2.14}$$

where we have used the Green-Ostrogradsky theorem in the last step. Eq. (2.14) simply states that this increase is equal to the opposite of the flux of particles out of the volume $V$, i.e. through $S$.

A plane wave solution of the Schrödinger equation (2.4)

$$\psi = N \, e^{\frac{i}{\hbar}(\vec{p}\cdot\vec{x} - Et)}, \tag{2.15}$$

which describes a free particle of momentum $\vec{p}$ and energy $E$ has therefore a probability density and a current

$$\rho = |N|^2, \tag{2.16}$$

$$\vec{j} = \frac{\vec{p}}{m}|N|^2 = \vec{v}\,|N|^2\,. \tag{2.17}$$

## 2.2. The Klein-Gordon equation

### 2.2.1. Looking for a linear relativistic equation

The Schrödinger equation (2.4) explicitly violates Lorentz covariance. In special relativity, we have

$$p^\mu p_\mu = \frac{E^2}{c^2} - \vec{p}^2 = m^2 c^2 \tag{2.18}$$

and thus

$$E = \sqrt{\vec{p}^2 c^2 + m^2 c^4}\,. \tag{2.19}$$

Looking for a first order equation, in the spirit of the Schrödinger equation, the correspondence rule would thus leads to write

$$i\hbar\frac{\partial\psi}{\partial t} = \sqrt{-\hbar^2 c^2 \Delta + m^2 c^4}\,\psi = mc^2\left(1 - \frac{\hbar^2}{2m^2c^2}\Delta + \cdots\right)\psi\,. \tag{2.20}$$

This equation is very non-local, and the symmetry between $ct$ and $\vec{x}$ is completely hidden.

### 2.2.2. Second order differential equation

Giving up with such first order equations[1], and writing the four-momentum operator as[2]

$$p^\mu = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, -i\hbar\vec{\nabla}\right) = i\hbar\,\partial^\mu, \tag{2.21}$$

---

[1]Such an attempt makes sense when looking for a relativistic equation describing a spin 1/2 particle: this is the essence of the construction of the Dirac equation historically.

[2]Indeed $\vec{\nabla} = \partial_i = -\partial^i$.

the quadratic dispersion (2.18) leads, using the correspondence principle, to

$$\left[ -\hbar^2 \partial^\mu \partial_\mu - m^2 c^2 \right] \Phi = 0 \tag{2.22}$$

i.e.

$$\left[ \partial^\mu \partial_\mu + \frac{m^2 c^2}{\hbar^2} \right] \Phi(x) = 0 \tag{2.23}$$

or equivalently, setting $c = \hbar = 1$,

$$\boxed{\left[ \Box + m^2 \right] \Phi(x) = 0 \,,} \tag{2.24}$$

which is the Klein-Gordon equation.

This equation satisfies Lorentz covariance since under a transformation $x' = \Lambda x$, denoting $\phi'(x') = \phi(x)$, Eq. (2.24) becomes

$$\left[ \Box' + m^2 \right] \Phi'(x') = 0 \,, \tag{2.25}$$

since $\Box' = \partial'^\mu \partial'_\mu = \partial^\mu \partial_\mu = \Box$, as any scalar operator.

### 2.2.3. Non relativistic limit

In order to study the classical (i.e. non relativistic) limit, let us extract the rest-mass energy in the time-dependence of the wave function. We thus write

$$\Phi(t, \vec{x}) = \frac{1}{\sqrt{2m}} e^{-i \frac{mc^2}{\hbar} t} \Psi(t, \vec{x}) \,, \tag{2.26}$$

and thus

$$\left[ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - i \frac{2m}{\hbar} \frac{\partial}{\partial t} - \frac{\partial^2}{\partial \vec{x}^2} \right] \Psi(t, \vec{x}) = 0 \,. \tag{2.27}$$

In the limit $c \to \infty$, neglecting the first term, this equation simplifies into

$$i\hbar \frac{\partial}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi(t, \vec{x}) \,, \tag{2.28}$$

which is the Schrödinger equation, as expected.

## 2.3. Klein-Gordon equation and scalar field action

As a discursive remark, let us show that the Klein-Gordon equation is the equation of motion of a free scalar field.

## 2.3.1. Lagrangian formulation

We consider the generic situation of a field theory, i.e. a system with an infinite number of degrees of freedom, indexed by the space-time position[3].

In any local field theory, the action reads

$$S = \int d^4x \, \mathcal{L}(x) \tag{2.29}$$

where $\mathcal{L}(x)$ is the lagrangian density (usually called lagrangian), which is assumed to depend on a finite number of derivatives of fields[4]. In practice, in general $\mathcal{L}(x)$ only depends on fields and their first order derivatives, just like in mechanics, the Lagrange function only depends on generalized coordinates and their first order time derivative, so that the equations of motion would be second order equations.[5] Thus, $\mathcal{L}(x)$ has the generic form

$$\mathcal{L}(x) = \mathcal{L}(\phi_i(x), \, \partial_n \phi_i(x), \, x) \,. \tag{2.30}$$

The index $i$ labels the various types of fields or, if field are not scalars (this will not be the case in the present lectures), their components (example: $A_\mu(x)$ in the case of electromagnetism, or $\Psi(x)$, a bispinor field (a collection of 4 complex numbers) describing a spin $1/2$ relativistic particle). Generally speaking, we will always assume that fields vanish fast enough at infinity, so that boundary terms can be safely ignored in usual integration by parts. Note that at that stage, the space-time group of transformation could be classical (galilean) or relativistic (lorentzian). For simplicity, we will now use covariant relativistic notations.

## 2.3.2. Euler-Lagrange equations

We denote as $\Omega$ the space-time domain[6] on which we integrate the lagrangian $\mathcal{L}$ to get the action $S$. The method to get the equations of motion is the same as the one used in analytical mechanics: one should vary the action assuming

---

[3]The case of a finite number of degrees of freedom boils down to analytical mechanics. We refer to any lecture on analytical mechanics, as well as to the famous books of L. Landau, E. Lifchitz, *Mechanics*, or H. Goldstein, *Classical Mechanics*.

[4]This is assumed in order to satisfy locality: in order that $\mathcal{L}(x)$ would depend on the values of fields at point $y$, distinct from $x$, one should be able to reconstruct the field value $\phi(y)$ from $\phi(x)$ and its derivatives at position $x$, which requires generally speaking *an infinite number of derivatives* from Taylor exansion...

[5]This is due to the Ostrogradsky instability, for theories having equations of motion with more than two time derivatives (higher-derivative theories). A theorem of Mikhail Ostrogradsky in classical mechanics indeed states that a non-degenerate Lagrangian dependent on time derivatives higher than the first corresponds to a Hamiltonian unbounded from below, therefore non physical.

[6]In general, $\Omega$ is the whole space-time, in which case each $x$ component covers $]-\infty, +\infty[$, but it is sometime convenient to keep $\Omega$ arbitrary in full generality, in particular in view of the Noether theorem.

## 2. The Klein-Gordon equation

that it should be *extremal* around the actual values of the fields describing the physical system. Under the transformation

$$\phi(x) \to \phi(x) + \delta\phi(x), \tag{2.31}$$

with the constraint that fields on the boundary $\partial\Omega$ of $\Omega$ have fixed values, which means that their variation $\delta\phi(x)$ vanish on $\partial\Omega$, the action varies as

$$\delta S = \int_\Omega d^4x \left[ \frac{\delta\mathcal{L}(x)}{\delta\phi_i(x)} \delta\phi_i(x) + \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} \delta(\partial_\mu\phi_i(x)) \right], \tag{2.32}$$

where we have performed a first order Taylor expansion of $\mathcal{L}$, which is a functional of the fields and their first derivatives. Integrating by parts, we thus get

$$\delta S = \int_\Omega d^4x \left[ \frac{\delta\mathcal{L}(x)}{\delta\phi_i(x)} - \partial_\mu \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} \right] \delta\phi_i(x) + \int_{\partial\Omega} \delta\phi_i(x) \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} d^3\sigma_\mu \tag{2.33}$$

where $d^3\sigma_\mu$ is the infinitesimal 3-dimensional element of integration on the boundary $\partial\Omega$. The last term in Eq. (2.33) vanishes due to the vanishing of the variation $\delta\phi(x)$ of fields on the boundary[7]. The action being stationary by assumption, therefore for any field variation $\delta\phi_i(x)$, this implies from Eq. (2.33) that

$$\boxed{\frac{\delta\mathcal{L}(x)}{\delta\phi_i(x)} - \partial_\mu \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} = 0} \tag{2.34}$$

which are the Euler-Lagrange equations for the fields $\phi_i(x)$.

One should note that adding a 4-divergence term to the lagrangian does not modifies the action, and therefore not the equations of motion.

### 2.3.3. Application: field equations of a massive scalar field theory

Consider the lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{1}{2}m^2\Phi^2. \tag{2.35}$$

Its equation of motion can readily be obtained: rewriting

$$\mathcal{L} = \frac{1}{2}\partial_\mu\Phi g^{\mu\nu}\partial_\nu\Phi - \frac{1}{2}m^2\Phi^2, \tag{2.36}$$

---

[7]If $\Omega$ is just the entire space-time, this is consistent with the assumed vanishing values of fields at infinity.

we have

$$\frac{\delta \mathcal{L}}{\delta \Phi} = -m^2 \Phi \tag{2.37}$$

and

$$\frac{\delta \mathcal{L}}{\delta \partial_\mu \Phi} = \partial^\mu \Phi \,. \tag{2.38}$$

Indeed $\mu$ and $\nu$ are dummy indexes, and thus one should not forget to differentiate both terms in Eq. (2.36). This finally leads to the equation of motion

$$\left[ \partial^\mu \partial_\mu + \frac{m^2 c^2}{\hbar^2} \right] \Phi(x) = 0 \tag{2.39}$$

which is our celebrated Klein-Gordon equation.

## 2.4. Physical content

### 2.4.1. Probability density and current

Let us construct a conserved four-current which extends the probability density (2.16) and current density (2.17) to the relativistic case. Starting from the Klein-Gordon equation

$$\left[ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right] \Phi + \frac{m^2 c^2}{\hbar^2} \Phi = 0 \,, \tag{2.40}$$

and multiplying by $i\Phi^*$ , we get

$$\frac{1}{c^2} i\Phi^* \frac{\partial^2 \Phi}{\partial t^2} - i\Phi^* \vec{\nabla}^2 \Phi + \frac{m^2 c^2}{\hbar^2} i\Phi^* \Phi = 0 \,, \tag{2.41}$$

while considering the complex conjugate of the Klein-Gordon equation (2.40) multiplied by $i\Phi$ leads to

$$\frac{1}{c^2} i\Phi \frac{\partial^2 \Phi^*}{\partial t^2} - i\Phi \vec{\nabla}^2 \Phi^* + \frac{m^2 c^2}{\hbar^2} i\Phi^* \Phi = 0 \,, \tag{2.42}$$

so that the subtraction of Eqs. (2.41, 2.42) gives

$$\frac{1}{c^2} i \left( \Phi^* \frac{\partial^2 \Phi}{\partial t^2} - \Phi \frac{\partial^2 \Phi^*}{\partial t^2} \right) - i \left( \Phi^* \vec{\nabla}^2 \Phi - \Phi \vec{\nabla}^2 \Phi^* \right) = 0 \,, \tag{2.43}$$

# 2. The Klein-Gordon equation

or equivalently

$$\frac{\partial}{c\partial t}\left[i\left(\Phi^*\frac{\partial\Phi}{c\partial t}-\Phi\frac{\partial\Phi^*}{c\partial t}\right)\right]+\vec{\nabla}\cdot\left[-i\left(\Phi^*\vec{\nabla}\Phi-\Phi\vec{\nabla}\Phi^*\right)\right]=0\,,\qquad(2.44)$$

i.e.

$$\frac{\partial\rho}{\partial t}+\vec{\nabla}\cdot\vec{j}=0\,,\qquad(2.45)$$

where

$$\rho = \frac{i\hbar}{c^2}\left(\Phi^*\frac{\partial\Phi}{\partial t}-\Phi\frac{\partial\Phi^*}{\partial t}\right),\qquad(2.46)$$

$$\vec{j} = -i\hbar\left(\Phi^*\vec{\nabla}\Phi-\Phi\vec{\nabla}\Phi^*\right).\qquad(2.47)$$

Equation (2.45) just reflects the conservation of probability if (2.46) and (2.47) are to be interpreted as the relativistic probability density and the probability current respectively.

Introducing the four-current

$$j^\mu = \left(\rho c,\vec{j}\right),\qquad(2.48)$$

which has the covariant expression

$$j^\mu = i\hbar\left(\Phi^*\partial^\mu\Phi-\Phi\partial^\mu\Phi^*\right),\qquad(2.49)$$

the conservation of probability (2.45) which reads

$$\frac{\partial}{c\partial t}c\rho+\vec{\nabla}\cdot\vec{j}=0\,,\qquad(2.50)$$

can be presented in the covariant form

$$\partial_\mu j^\mu = 0\,.\qquad(2.51)$$

Note that the fact that $\rho$ is the time component of a four-vector is consistent with the fact that $\rho\,d^3V$ is the probability that a particle would be in the volume $d^3V$. Under a boost of Lorentz factor $\gamma$, $d^3V$ is Lorentz contracted by a factor of $1/\gamma$, which is compensated by the fact that $\rho$ is dilated by a factor of $\gamma$ as any time component of a four-vector, so that the probability is indeed invariant.

## 2.4.2. Non relativistic limit

Let us check that the relativistic probability density and probability current reduce to the non-relativistic expressions (2.5) and (2.13) respectively. Again presenting the wave function in the form (2.26), we get

$$\rho = \frac{i\hbar}{2mc^2}\left[-i\frac{mc^2}{\hbar}2\Psi^*\Psi + (\Psi^*\partial_t\Psi - \Psi\partial_t\Psi^*)\right], \quad (2.52)$$

$$j^i = -\frac{i\hbar}{2m}\left[\Psi^*\nabla^i\Psi - \Psi\nabla^i\Psi^*\right]. \quad (2.53)$$

The probability current (2.53) has already the appropriate form (2.13). As far as $\rho$ is concerned, since the term in parenthesis in the RHS of Eq. (2.52) is equal to $-2i\Psi\Psi^*(E-mc^2)/\hbar$, it is negligible with respect to the first term (rest energy) in the limit $c \to \infty$, so that $\rho$ reduces to the non-relativistic expressions (2.5), as expected.

## 2.4.3. Energy spectrum and probability density

Let us consider a plane wave solution[8]

$$\Phi = Ne^{-ip\cdot x}. \quad (2.54)$$

We thus get, from Eq. (2.49),

$$j^\mu = 2p^\mu|N|^2. \quad (2.55)$$

This looks very simple and elegant. However, substituting the plane wave solution (2.54) in the Klein-Gordon equation (2.24), we get the dispersion relation

$$p^2 = m^2 \quad i.e. \quad E = \pm\sqrt{\vec{p}^2 + m^2}, \quad (2.56)$$

so that we encounter two very serious problems:

  - there are negative energy solution, and the energy spectrum is not bounded from below, so that under an arbitrary perturbation, the system could furnish an arbitrary amount of energy to the outside.

  - the probability density being proportional to the energy, negative energy solutions have negative probability densities!

The problem is very serious since any consistent treatment of a physical system requires to deal with a complete set of states, as we are used to do in quantum mechanics.

---

[8]We use the natural units $c = 1$ and $\hbar = 1$.

## 2.4.4. Charge current and reinterpretation of the negative energy solution

It turns out that multiplying the current (2.49) by the elementary charge $-e$ is a way to escape the two previous problems. Let us indeed consider the current

$$j^\mu = -ie\left(\Phi^*\partial^\mu\Phi - \Phi\partial^\mu\Phi^*\right). \tag{2.57}$$

The corresponding $\rho = j^0$ is now the charge density, and not anymore the probability density. Leaving aside spin effects, let us use the Klein-Gordon equation as a relativistic equation describing "electrons". Spin effects require the introduction of the Dirac equation, which will not be discussed in these lectures.

Consider a plane wave (2.54). For an electron $e^-$, of charge $-e$, energy $E$ and momentum $\vec{p}$, we get, using the definition (2.57) of the charge current,

$$j^\mu(e^-) = -2e|N|^2(E,\vec{p}). \tag{2.58}$$

Besides, for a positron $e^+$, of charge $+e$, of energy $E$ and momentum $\vec{p}$, we get, using the definition (2.57) of the charge current,

$$j^\mu(e^+) = +2e|N|^2(E,\vec{p}) = -2e|N|^2(-E,-\vec{p}). \tag{2.59}$$

The last algebraic manipulation may look somehow awkward and trivial. However, it just means that there is no need to introduce positron degrees of freedom: they are already there in the Klein-Gordon equation, since according to the previous identity, a positron of energy $E$ and momentum $\vec{p}$ is the same as an electron of *negative* energy $-E$, of momentum $-\vec{p}$, i.e. propagating backward in time.

Pictorially, we can thus draw

$$\tag{2.60}$$



From the time evolution point of view of a wave function in quantum mechanics, this relies on the fact that the time depend part of a free particle moving forward in time, of energy $E$, is the same as the one of a free particle of energy $-E$ propagating backward in time, since

$$e^{-i(-E)(-t)} = e^{-iEt}. \tag{2.61}$$

# 2.5. In-depth study: global symmetries and Noether theorem

In this section, we want to show that whenever an action is invariant under a given global symmetry, there exists a conserved current, and a related conserved charge. As an application, we will recover the expression of the conserved charge current (2.57).

## 2.5.1. Variation of the action with respect to boundary configurations

As a first step, let us study how the action is modified when any field configuration, solution of the equations of motion, faces an infinitesimal change to another solution of the equations of motion, over the whole space-time domain $\Omega$, without fixing its value at the boundary $\delta\Omega$. This is in contradistinction with the situation studied previously in the section 2.3.2 when establishing the Euler-Lagrange equation, where the boundary field configuration was fixed. We will show that the variation of the action as a simple expression as a function of the variation of the field configuration at the boundary.

Consider an arbitrary transformation of the coordinates and fields[9]

$$
\begin{aligned}
x'^{\mu} &= x^{\mu} + \delta x^{\mu}, \\
\phi'(x') &= \phi(x) + \delta\phi(x).
\end{aligned}
\tag{2.62}
$$

The local variation of the field then reads

$$
\begin{aligned}
\bar{\delta}\phi(x) &= \phi'(x) - \phi(x) \\
&= \phi(x - \delta x) + \delta\phi(x - \delta x) - \phi(x) \\
&= \delta\phi(x) - \delta x^{\mu}\partial_{\mu}\phi(x) + ... \,(\text{higher order terms}).
\end{aligned}
\tag{2.63}
$$

The action for fields solutions of the equation of motion, which reads, before the transformation,

$$
S = \int_{\Omega} d^4x \, \mathcal{L}(\phi(x), \, \partial_{\mu}\phi(x), \, x)
\tag{2.64}
$$

thus becomes

$$
S' = \int_{\Omega'} d^4x \, \mathcal{L}(\phi'(x), \, \partial_{\mu}\phi'(x), \, x)
\tag{2.65}
$$

---

[9]Although we only consider scalar fields in the present chapter, the following discussion also applies to arbitrary fields.

where $\Omega'$ is obtained from $\Omega$ by the transformation (2.62). Besides, $S'$ also reads

$$S' = \int_{\Omega} d^4x\, \mathcal{L}(\phi',\, \partial_\mu\phi',\, x) + \int_{\delta\Omega} d^3\sigma_\mu\, \delta x^\mu \mathcal{L}(\phi',\, \partial_\mu\phi',\, x)\,, \qquad (2.66)$$

which gives, after a first order Taylor expansion, neglecting terms of order 2,

$$S' = S + \int_{\Omega} d^4x\, \left(\frac{\delta\mathcal{L}}{\delta\phi}\bar{\delta}\phi + \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\partial_\mu\bar{\delta}\phi\right) + \int_{\delta\Omega} d^3\sigma_\mu\, \delta x^\mu \mathcal{L}(\phi,\, \partial_\mu\phi,\, x) + \cdots\,.$$
$$(2.67)$$

Integrating by parts, one thus gets

$$\begin{aligned}
S' &= S + \int_{\Omega} d^4x\, \left(\frac{\delta\mathcal{L}}{\delta\phi} - \partial_\mu\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\right)\bar{\delta}\phi + \int_{\Omega} d^4x\, \partial_\mu\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\bar{\delta}\phi\right) \quad (2.68)\\
&\quad + \int_{\delta\Omega} d^3\sigma_\mu\, \delta x^\mu \mathcal{L}(\phi,\, \partial_\mu\phi,\, x) + \cdots\,.
\end{aligned}$$

The second term vanishes since fields satisfy the equations of motion. Integrating the third term, we thus get

$$\begin{aligned}
S' &= S + \int_{\delta\Omega} d^3\sigma_\mu\, \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\bar{\delta}\phi + \delta x^\mu \mathcal{L}\right) + \cdots \qquad (2.69)\\
&= S + \int_{\delta\Omega} d^3\sigma_\mu\, \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\delta\phi - \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\partial_\nu\phi\,\delta x^\nu + \delta x^\mu \mathcal{L}\right) + \cdots
\end{aligned}$$

where we have used Eq. (2.63). Finally, we obtain

$$\delta S = \int_{\delta\Omega} d^3\sigma_\mu\, \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\delta\phi - \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\partial_\nu\phi - g^\mu{}_\nu\mathcal{L}\right)\delta x^\nu\right) + \cdots\,, \qquad (2.70)$$

after using the fact that $g^\mu{}_\nu \equiv \delta^\mu{}_\nu$.

## 2.5.2. Noether theorem

In the special case where the action is invariant under the considered transformations, we get the Noether theorem, which states that *to any continuous group of symmetry of the action is associated a conserved quantity.*

The conserved current corresponding to this symmetry reads

$$\boxed{j^\mu = \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\delta\phi - \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\partial_\nu\phi - g^\mu{}_\nu\mathcal{L}\right)\delta x^\nu} \qquad (2.71)$$

since following Eq. (2.70),

$$\delta S = 0 = \int_{\delta\Omega} d^3\sigma_\mu j^\mu = \int_\Omega d^4x\, \partial_\mu j^\mu$$

whatever $\Omega$, which indeed shows that the current $j^\mu$ is conserved:

$$\partial_\mu j^\mu = 0\,.$$

The corresponding conserved charge reads

$$\boxed{Q = \int d^3x\, j^0(x)\,.}\tag{2.72}$$

Indeed,

$$\frac{\partial Q}{\partial t} = \int d^3x\, \partial_0\, j^0(x) = -\int d^3x\, \partial_i\, j^i(x) = 0$$

for rapidly decreasing field at infinity. More generally, the conserved charge can defined through the relation

$$Q = \int_{N_3} d^3\sigma_\mu\, j^\mu(x)\,,\tag{2.73}$$

where $N_3$ is a space-like 3-surface.

One should note that the above considered transformations are *global* (also named *rigid*). We do not assume any locality of these transformations in the proof. Imposing that the action would be invariant under *local* transformations would be an additional requirement, underlying gauge field theories. It turn out that this furnishes a dynamical principle allowing to *construct* various kinds of lagrangian. This is essence of Yang-Mills construction.

## 2.5.3. Charge current

Let us consider again the lagrangian (2.35) of a scalar field, assuming now that the field is complex:

$$\mathcal{L} = \partial_\mu\Phi^*\partial^\mu\Phi - m^2\Phi^*\Phi\,.\tag{2.74}$$

It means that $\Phi$ and $\Phi^*$ are to be considered as independent fields. Their two equations of motion thus read

$$\left[\partial^\mu\partial_\mu + \frac{m^2c^2}{\hbar^2}\right]\Phi(x) = 0\tag{2.75}$$

and

$$\left[\partial^\mu \partial_\mu + \frac{m^2 c^2}{\hbar^2}\right] \Phi^*(x) = 0 \,. \tag{2.76}$$

Obviously, the lagrangian (2.74), and therefore the corresponding action, are invariant under the so-called global gauge transforms

$$\Phi \;\; \to \;\; e^{ie\alpha}\Phi \tag{2.77}$$
$$\Phi^* \;\; \to \;\; e^{-ie\alpha}\Phi^* \tag{2.78}$$

where $e$ is an arbitrary parameter, which has the physical meaning of the elementary electric charge, due to its role in the conserved charge current which we will now construct.

Let us compute the Noether current associated to this transformation. We consider the infinitesimal versions of the transformations (2.77-2.78)

$$\delta\Phi \;\; = \;\; ie\delta\alpha\Phi \tag{2.79}$$
$$\delta\Phi^* \;\; = \;\; -ie\delta\alpha\Phi^* \tag{2.80}$$

with $\delta x = 0$. Therefore, according to Eq. (2.71), and using the fact that

$$\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Phi)} \;\; = \;\; \partial^\mu\Phi^*, \tag{2.81}$$

$$\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Phi^*)} \;\; = \;\; \partial^\mu\Phi \tag{2.82}$$

we get

$$j^\mu = \partial^\mu\Phi^*(ie\delta\alpha\Phi) - \partial^\mu\Phi(ie\delta\alpha\Phi^*) \tag{2.83}$$

valid for any $\delta\alpha$, which can thus be factorized out, so that the electromagnetic four-current

$$j^\mu = -ie\left(\Phi^*\partial^\mu\Phi - \Phi\,\partial^\mu\Phi^*\right) \tag{2.84}$$

is conserved. We have therefore recovered, now from a global symmetry principle, the current (2.57) which was obtained in a heuristic way.

# Part II.

# Introduction to Quantum Electrodynamics

# 3. From nonrelativistic perturbation theory to Feynman diagrams

## 3.1. Time dependent perturbation theory

### 3.1.1. General framework

We consider a particle which is described by the hamiltonian

$$H = H_0 + V(t)\,, \tag{3.1}$$

where $H_0$ is the free hamiltonian which has known spectrum and eigenstates, and $V(t)$ is an interaction potential, which generally speaking depends on time. The eigenstates of the free hamiltonian $H_0$ are denoted as $|\Psi_n^{(0)}\rangle$. They thus satisfy

$$H_0|\Psi_n^{(0)}\rangle = E_n|\Psi_n^{(0)}\rangle \tag{3.2}$$

with the orthonormalization properties

$$\langle\Psi_n^{(0)}|\Psi_m^{(0)}\rangle = \delta_{nm}\,. \tag{3.3}$$

The solutions of the free Schrödinger equation

$$H_0|\Psi_n^{(0)}(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi_n^{(0)}(t)\rangle \tag{3.4}$$

are thus given by

$$|\Psi_n^{(0)}(t)\rangle = e^{-i\frac{E_n t}{\hbar}}|\Psi_n^{(0)}\rangle\,. \tag{3.5}$$

We want to solve the full Schrödinger's equation governed by the hamiltonian $H$

$$H|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle\,. \tag{3.6}$$

### 3. From nonrelativistic perturbation theory to Feynman diagrams

For that, we aim at finding its solution through an expansion over the basis $|\Psi_n^{(0)}(t)\rangle$ of solutions of the free Schrödinger's equation (3.4) governed by the hamiltonian $H_0$.

We thus generically write

$$|\Psi(t)\rangle = \sum_n a_n(t)\,|\Psi_n^{(0)}(t)\rangle = \sum_n a_n(t)\,e^{-i\frac{E_n t}{\hbar}}|\Psi_n^{(0)}\rangle\,. \tag{3.7}$$

In terms of wave functions, this reads

$$\Psi(\vec{x},t) = \langle\vec{x}|\Psi(t)\rangle = \sum_n a_n(t)\,e^{-i\frac{E_n t}{\hbar}}\langle\vec{x}|\Psi_n^{(0)}\rangle = \sum_n a_n(t)\,e^{-i\frac{E_n t}{\hbar}}\Psi^{(0)}(\vec{x},t)\,. \tag{3.8}$$

The problem thus reduces to the determination of the coefficients $a_n(t)$.

Inserting the expansion (3.7) inside the Schrödinger's equation (3.6), we obtain

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle &= i\hbar\sum_n \frac{da_n(t)}{dt}e^{-i\frac{E_n t}{\hbar}}|\Psi_n^{(0)}\rangle + \sum_n a_n(t)E_n e^{-i\frac{E_n t}{\hbar}}|\Psi_n^{(0)}\rangle && (3.9)\\
&= \sum_n a_n(t)e^{-i\frac{E_n t}{\hbar}}V(t)|\Psi_n^{(0)}\rangle + \sum_n a_n(t)e^{-i\frac{E_n t}{\hbar}}H_0|\Psi_n^{(0)}\rangle && (3.10)
\end{aligned}
$$

and thus

$$i\hbar\sum_n \frac{da_n(t)}{dt}e^{-i\frac{E_n t}{\hbar}}|\Psi_n^{(0)}\rangle = \sum_n a_n(t)e^{-i\frac{E_n t}{\hbar}}V(t)|\Psi_n^{(0)}\rangle\,, \tag{3.11}$$

which is up to now an exact equation.

Before proceeding further, one should note that what we have done is nothing more than the method of variation of constant for a 1st order differential equation: indeed the Schrödinger's equation is just a (complicate) 1st order differential equation...!

Contracting Eq. (3.11) on the left with $\langle\Psi_f^{(0)}|$ leads to

$$\sum_n \frac{da_n(t)}{dt}e^{-i\frac{E_n t}{\hbar}}\langle\Psi_f^{(0)}|\Psi_n^{(0)}\rangle = \frac{-i}{\hbar}\sum_n a_n(t)e^{-i\frac{E_n t}{\hbar}}\langle\Psi_f^{(0)}|V(t)|\Psi_n^{(0)}\rangle\,, \tag{3.12}$$

and thus to the set of coupled linear differential equations

$$\boxed{\frac{da_f(t)}{dt} = \frac{-i}{\hbar}\sum_n a_n(t)\,\langle\Psi_f^{(0)}|V(t)|\Psi_n^{(0)}\rangle\,e^{i\frac{(E_f-E_n)t}{\hbar}}\,.} \tag{3.13}$$

To proceed, we now make an assumption: much before the potential $V$ starts to act, the system is in the state $|\Psi_i^{(0)}\rangle$, say at time $-T/2$, with $T$ very large. Thus,

$$a_n(-T/2) = \delta_{ni} . \tag{3.14}$$

The set (3.13) can be equivalently represented in integral form, namely

$$a_f(t) = a_f\left(-T/2\right) - \frac{i}{\hbar} \int_{-\frac{T}{2}}^{t} dt' \sum_n a_n(t') \, \langle\Psi_f^{(0)}|V(t)|\Psi_n^{(0)}\rangle \, e^{i\frac{(E_f - E_n)t'}{\hbar}} . \tag{3.15}$$

In general, this cannot be solved exactly.

## 3.1.2. Perturbation theory

We now assume that $V$ is small, so that a perturbative treatment can be performed. This is nothing more than a sophisticated series expansion in powers of $V$. We thus write

$$
\begin{array}{ccccccccc}
a_f(t) & = & a_f^{(0)} & + & a_f^{(1)} & + & a_f^{(2)} & + & \cdots \\
 & & \text{order } V^0 & & \text{order } V^1 & & \text{order } V^2 & &
\end{array} \tag{3.16}
$$

Let us examine the first terms in this expansion.

The generic term is obtained through the expansion

$$a_f(t) = a_f^{(0)} + a_f^{(1)} + a_f^{(2)} + \cdots \tag{3.17}$$

$$= a_f(-T/2) - \frac{i}{\hbar} \int_{-\frac{T}{2}}^{t} dt' \sum_n \langle\Psi_f^{(0)}|V(t')|\Psi_n^{(0)}\rangle \, e^{i\frac{(E_f - E_n)t'}{\hbar}} \tag{3.18}$$

$$\times \left[ a_n^{(0)}(t') + a_n^{(1)}(t') + a_n^{(2)}(t') + \cdots \right] \tag{3.19}$$

The order 0 term is trivial: it is solved by setting $V = 0$, i.e.

$$a_f^{(0)}(t) = a_f(-T/2) = \delta_{fi} . \tag{3.20}$$

The order 1 term reads:

$$a_f^{(1)}(t) = -\frac{i}{\hbar} \int_{-\frac{T}{2}}^{t} dt' \sum_n \langle\Psi_f^{(0)}|V(t')|\Psi_n^{(0)}\rangle \, e^{i\frac{(E_f - E_n)t'}{\hbar}} \delta_{ni} \tag{3.21}$$

$$= -\frac{i}{\hbar} \int_{-\frac{T}{2}}^{t} dt' \, \langle\Psi_f^{(0)}|V(t')|\Psi_i^{(0)}\rangle \, e^{i\frac{(E_f - E_i)t'}{\hbar}} \tag{3.22}$$

The order 2 term reads:

$$
\begin{aligned}
a_f^{(2)}(t) &= \left(-\frac{i}{\hbar}\right)^2 \int_{-\frac{T}{2}}^{t} dt' \sum_n \langle \Psi_f^{(0)}|V(t')|\Psi_n^{(0)}\rangle\, e^{i\frac{(E_f - E_n)t'}{\hbar}} \\
&\quad \times \int_{-\frac{T}{2}}^{t'} dt''\, \langle \Psi_n^{(0)}|V(t'')|\Psi_i^{(0)}\rangle\, e^{i\frac{(E_n - E_i)t''}{\hbar}}
\end{aligned}
\tag{3.23}
$$

etc.

## 3.2. Transition amplitude in perturbation theory

### 3.2.1. Transition amplitude

We are mostly interested in transitions between an initial state $(t = -T/2)$ and a final state $(t = +T/2)$, in which the interaction can be neglected both in the initial and in the final states. We therefore define the *transition amplitude*

$$
T_{fi} = a_f(T/2) - a_f(-T/2) = -\frac{i}{\hbar} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt' \sum_n a_n(t')\, \langle \Psi_f^{(0)}|V(t)|\Psi_n^{(0)}\rangle\, e^{i\frac{(E_f - E_n)t'}{\hbar}},
\tag{3.24}
$$

where we have subtracted the contribution $a_f(-T/2)$, since we are not interested in the situation in which nothing occurred.

The potential $V(t)$ is assumed to be a local space-time operator, i.e.

$$
\langle \vec{x}|V(t)|\vec{x}'\rangle = V(x)\delta^{(4)}(x - x').
\tag{3.25}
$$

Inserting twice the identity operator

$$
\mathbb{1} = \int d^3\vec{x}\, |\vec{x}\rangle\langle\vec{x}|
\tag{3.26}
$$

inside (3.24), we get

$$
T_{fi}
\tag{3.27}
$$
$$
= -\frac{i}{\hbar} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \int d^3\vec{x} \int d^3\vec{x}' \sum_n a_n(t)\, \langle \Psi_f^{(0)}|\vec{x}\rangle\langle\vec{x}|V(t)|\vec{x}'\rangle\langle\vec{x}'|\Psi_n^{(0)}\rangle\, e^{i\frac{(E_f - E_n)t}{\hbar}}
$$
$$
= -\frac{i}{\hbar} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \int d^3\vec{x}\, \Psi_f^{*(0)}(x)\, V(x) \sum_n \Psi_n^{(0)}(x) a_n(t),
$$

where we have used the fact, see Eq.(3.5), that

$$\Psi_n^{(0)}(x) \equiv \Psi_n^{(0)}(t, \vec{x}) = \Psi_n^{(0)}(0, \vec{x}) \, e^{-i\frac{E_n t}{\hbar}} \,. \tag{3.28}$$

In the $T \to \infty$ limit, we thus have the covariant expression

$$\boxed{T_{fi} = -\frac{i}{\hbar} \int d^4x \; \Psi_f^{*(0)}(x) \, V(x) \sum_n \Psi_n^{(0)}(x) \, a_n(t) \,.} \tag{3.29}$$

## 3.2.2. First order transition amplitude for scattering on a static potential

Let us focus now on the first order expansion in $V$. We thus have

$$
\begin{aligned}
T_{fi}^{(1)} &= -\frac{i}{\hbar} \int_{-\frac{T}{2} \to -\infty}^{\frac{T}{2} \to +\infty} dt \; \langle \Psi_f^{(0)} | V(t) | \Psi_i^{(0)} \rangle \, e^{i\frac{E_f - E_i}{\hbar}t} \\
&= -\frac{i}{\hbar} \int d^4x \; \Psi_f^{*(0)}(x) \, V(x) \Psi_i^{(0)}(x) \,.
\end{aligned} \tag{3.30}
$$

We now consider the special situation in which $V$ is time independent:

$$V(t, \vec{x}) = V(\vec{x}) \,. \tag{3.31}$$

Thus,

$$T_{fi}^{(1)} = -\frac{i}{\hbar} V_{fi} \int_{-\infty}^{+\infty} dt \; e^{i\frac{E_f - E_i}{\hbar}t} = -2\pi i \, V_{fi} \, \delta(E_f - E_i) \,. \tag{3.32}$$

where

$$V_{fi} = \langle \Psi_f^{(0)} | V | \Psi_i^{(0)} \rangle = \int d^3\vec{x} \; \Psi_f^{*(0)}(\vec{x}) \, V(\vec{x}) \, \Psi_i^{(0)}(\vec{x}) \,. \tag{3.33}$$

One should note that assuming $V$ to be time-independent is equivalent to have no energy transfer from the potential to the particle, since integration over time leads to a $\delta(E_f - E_i)$.

Thus, we recover the fact for a static potential, on which the scattering takes an infinite amount of time,

$$\Delta t = \infty \quad \Leftrightarrow \quad \Delta E = 0 \tag{3.34}$$

in accordance with Heisenberg's uncertainty $\Delta t \Delta E \gtrsim \hbar/2$.

### 3.2.3. Fermi's Golden rule

We are interested in the probability that the particle is scattered on $V$ from an initial state $i$ to a final state $f$. Quantum mechanics tells us that one should square $T_{fi}$, but we immediately face a serious technical problem, since this implies to compute the square $[\delta(E_f - E_i)]^2$ which looks like meaningless.

The solution is to compute the transition probability per unit of time

$$W_{fi} = \lim_{T \to \infty} \frac{|T_{fi}|^2}{T} \, . \tag{3.35}$$

The trick is to write $[\delta(E_f - E_i)]^2$ in a non symmetric way:

$$
\begin{aligned}
|T_{fi}^{(1)}|^2 &= \frac{2\pi}{\hbar} |V_{fi}^{(1)}|^2 \delta(E_f - E_i) \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \, e^{i \frac{E_f - E_i}{\hbar} t} \\
&= \frac{2\pi}{\hbar} |V_{fi}^{(1)}|^2 \delta(E_f - E_i) \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \\
&= \frac{2\pi}{\hbar} |V_{fi}^{(1)}|^2 \delta(E_f - E_i) \, T
\end{aligned}
\tag{3.36}
$$

and thus

$$W_{fi} = \frac{2\pi}{\hbar} |V_{fi}^{(1)}|^2 \delta(E_f - E_i) \, . \tag{3.37}$$

Note: one can easily check that $W$ has the correct dimension. Since $[\hbar] = [E] \, T$,

with $[V_{fi}^2 \, \delta(E_f - E_i)] = \frac{[E]^2}{[E]} = [E]$ we thus have $[W] = \frac{[E]}{[E] \, T} = T^{-1}$.

In practice, one usually deals with final states in a given set. We therefore introduce the density $\rho(E_f)$ of these states, so that $\rho(E_f) \, dE_f$ is the number of states in the energy range $E_f$ to $E_f + dE_f$. Thus, the transition probability per unit time $W_{fi}$ reads

$$W_{fi} = \frac{2\pi}{\hbar} \int dE_f \, \rho(E_f) \, |V_{fi}^{(1)}|^2 \delta(E_f - E_i) \, , \tag{3.38}$$

so that

$$\boxed{W_{fi} = \frac{2\pi}{\hbar} |V_{fi}^{(1)}|^2 \, \rho(E_i) \, ,} \tag{3.39}$$

known under the name of Fermi's Golden rule.

### 3.2.4. Order 2

From the expression (3.23) obtained for $a_f^{(2)}(t)$, one gets

$$
T_{fi}^{(2)} = \left(-\frac{i}{\hbar}\right)^2 \int_{-\frac{T}{2}}^{\frac{T}{2}} dt' \sum_n \langle \Psi_f^{(0)} | V(t') | \Psi_n^{(0)} \rangle \, e^{i\frac{E_f - E_n}{\hbar}t'}
$$

$$
\times \int_{-\frac{T}{2}}^{t'} dt'' \, \langle \Psi_n^{(0)} | V(t'') | \Psi_i^{(0)} \rangle \, e^{i\frac{E_n - E_i}{\hbar}t''} \,. \tag{3.40}
$$

Again working under the assumption of a time-independent potential, we have

$$
T_{fi}^{(2)} = \left(-\frac{i}{\hbar}\right)^2 \sum_n V_{fn} V_{ni} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt' \, e^{i\frac{E_f - E_n}{\hbar}t'} \int_{-\frac{T}{2}}^{t'} dt'' \, e^{i\frac{E_n - E_i}{\hbar}t''} \,. \tag{3.41}
$$

We should now take the limit $T \to \infty$. Meanwhile, the interaction should cease at $t = -\infty$ and $t = +\infty$. For that, we introduce an adiabatic parameter $\epsilon > 0$ to make the $t''$ integration meaningful. Physically, it is introduced in order to circumvent the fact that choosing a time-independent potential is at odd with the fact that it is supposed to vanish at $t = -\infty$ and $t = +\infty$.

One now has

$$
\int_{-\frac{T}{2}}^{t'} dt'' \, e^{i\frac{E_n - E_i - i\varepsilon}{\hbar}t''}
$$

$$
= i\hbar \frac{1}{E_i - E_n + i\varepsilon} \left[ e^{i\frac{E_n - E_i - i\varepsilon}{\hbar}t'} - e^{i\frac{E_n - E_i - i\varepsilon}{\hbar}\left(-\frac{T}{2}\right)} \right]
$$

$$
\xrightarrow[T \to +\infty]{} i\hbar \frac{1}{E_i - E_n + i\varepsilon} e^{i\frac{E_n - E_i - i\varepsilon}{\hbar}t'} \tag{3.42}
$$

where we have used the fact that second term in the second line vanishes in the limit $T \to +\infty$, thanks to the damping factor $\exp(-\varepsilon T/(2\hbar))$.

Thus,

$$
T_{fi}^{(2)} = \left(-\frac{i}{\hbar}\right)^2 i\hbar \sum_n \frac{V_{fn} V_{ni}}{E_i - E_n + i\varepsilon} \int_{-\infty}^{+\infty} dt' \, e^{i\frac{E_f - E_n + E_n - E_i}{\hbar}t'}
$$

$$
= -2\pi i \sum_n V_{fn} \frac{1}{E_i - E_n + i\varepsilon} V_{ni} \, \delta(E_i - E_f) \,, \tag{3.43}
$$

which exhibit the structure of the result: the transition amplitude is the product of two matrix elements of the potential, between which an energy denominator is to be inserted, describing the propagation of the particle from state $i$ to state $f$. The concept of propagation can be made more precise based on the formalism of Green's functions, see any advanced quantum mechanics lectures, e.g. Messiah Chap. XIX.

## 3.2.5. Higher orders

The same line of thought can be pursued at order 3. In this case, 3 matrix elements of the potential and two energy denominators will be involved in $T_{fi}^{(3)}$.

Finally, the general structure is the following

$$
\begin{aligned}
T_{fi} &= T_{fi}^{(1)} + T_{fi}^{(2)} + T_{fi}^{(3)} + \cdots \\[1mm]
&= -2\pi i \left[ V_{fi} + \sum_n V_{fn} \frac{1}{E_i - E_n + i\varepsilon} V_{ni} \right. \\[1mm]
&\left. + \sum_{n,m} V_{fn} \frac{1}{E_m - E_n + i\varepsilon} V_{nm} \frac{1}{E_i - E_m + i\varepsilon} V_{mi} + \cdots \right] \delta(E_i - E_f),
\end{aligned}
\tag{3.44}
$$

This is illustrated at order 1 in Fig. 3.1.
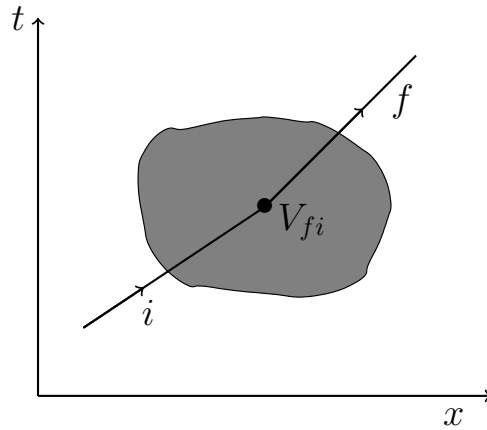


Figure 3.1.: Perturbative expansion of $T_{fi}$ at 1st order, drawn in space-time.

Fig. 3.2 illustrates the scattering at order 2. A possibility of producing a pair of virtual particles appears. Indeed one should pay attention to time ordering: reading Eq. (3.43) in normal order, i.e. from left to right corresponds to going backward with respect to the direction of the arrows. The physical interpretation of the configurations encountered at each step depends on the ordering of time:

⋄ in the first diagram, following this backward flow of arrows, one first faces the last scattering (at time $t_2$) and then the first scattering (at time $t_1$). Therefore, this first diagram involves a single particle at every stage.

⋄ in the second diagram, following this backward flow of arrows, one first
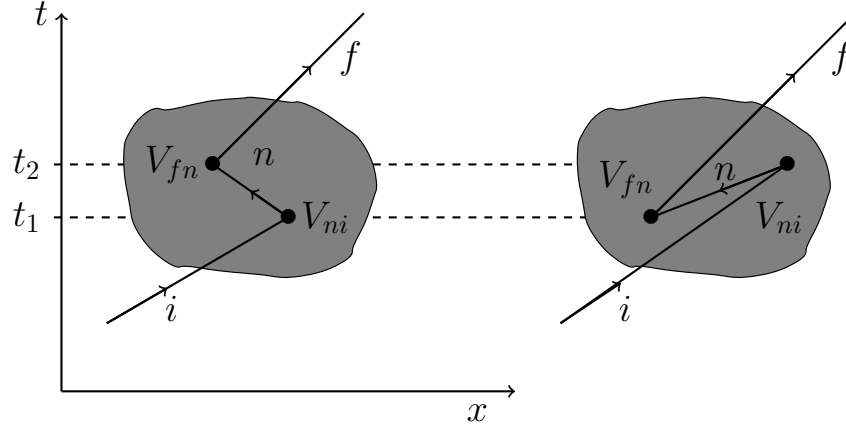
Figure 3.2.: Perturbative expansion of $T_{fi}$ at 2nd order, drawn in space-time, with two diagrams corresponding to two different time orderings of the double scattering of the electron on the external potential. In the second diagram, a virtual pair is produced in the first scattering (at time $t_1$), which is annihilated in the second scattering (at time $t_2$).

faces the first scattering (at time $t_1$) and then the second scattering (at time $t_2$).

As we have seen in Chap. 2, a particle propagating back in time should be interpreted as an antiparticle propagating forward in time. At time $t_1$, a particle-antiparticle pair is produced, so that between the two scatterings at $t_1$ and then $t_2$, the state is made of two particles and one antiparticle. At time $t_2$, the virtual pair annihilates, and the state is made of a single particle.

# 4. Electrodynamics

We use the Heaviside-Lorentz's units: Coulomb force is given by

$$F = \frac{QQ'}{4\pi r^2} \tag{4.1}$$

which means that $\varepsilon_0 = 1$ and $\mu_0 = 1$. Furthermore, we use a system of units in which $c = 1$. We refer to Jackson's appendix[1] for a detailed discussion on the various systems of units.

## 4.1. Maxwell equations

### 4.1.1. Local form

We recall that in local form, the 4 Maxwell equations read

$$\text{div}\vec{E} = \rho, \tag{4.2}$$

$$\overrightarrow{\text{rot}}\vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j}, \tag{4.3}$$

$$\text{div}\vec{B} = 0, \tag{4.4}$$

$$\overrightarrow{\text{rot}}\vec{E} + \frac{\partial \vec{B}}{\partial t} = 0. \tag{4.5}$$

The local conservation of charges states that

$$\frac{\partial \rho}{\partial t} + \text{div}\vec{j} = 0. \tag{4.6}$$

### 4.1.2. Integral form

Their corresponding integral form are:

◇ Gauss law: for any closed surface $S$ enclosing a volume $V$, one has

$$\iint_S \vec{E} \cdot d^2\vec{S} = \iiint_V \rho \, d^3x. \tag{4.7}$$

---

[1] J. D. Jackson, *Classical Electrodynamics.*

◇ Generalized Ampère's law: for any closed curve $C$ bordering a surface $S$,

$$\oint_C \vec{B} \cdot d\vec{\ell} = \iint_S \left( \vec{j} + \frac{\partial \vec{E}}{\partial t} \right) \cdot d^2\vec{S} \,, \tag{4.8}$$

where $\frac{\partial \vec{E}}{\partial t}$ is the displacement current, which was absent in the original Ampère's law.

◇ Absence of magnetic monopole:

for any closed surface $S$,

$$\iint_C \vec{B} \cdot d^2\vec{S} = 0 \,, \tag{4.9}$$

which states that there is no magnetic monopole, in contradistinction with Gauss's law which states that electric charges exists.

◇ Faraday's induction law:

for any closed curve $C$, bordering a surface $S$,

$$\oint_C \vec{E} \cdot d\vec{\ell} = - \iint_S \frac{\partial \vec{B}}{\partial t} \cdot d^2\vec{S} \,. \tag{4.10}$$

## 4.2. Covariant formulation of electrodynamics

As usual, we denote a space-time point in Minkowski space as

$$x^\mu = (t, \vec{x}) \tag{4.11}$$

and the 4-electric current as

$$j^\mu = (\rho, \vec{j}) \,. \tag{4.12}$$

### 4.2.1. Field strength tensor and its dual

In contra-variant form, the field strength tensor is an antisymmetric tensor, defined as

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^2 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix} \tag{4.13}$$

i.e.

$$F^{i0} = -F^{0i} = -F_{i0} = F_{0i} = E^i = -E_i \,, \tag{4.14}$$
$$F^{ij} = F_{ij} = \varepsilon_{ijk} B_k = -\varepsilon_{ijk} B^k \,. \tag{4.15}$$

Conversely, the magnetic field can be expressed from

$$\varepsilon_{ijk} F_{ij} = \varepsilon_{ijk} \varepsilon_{ijk'} B_{k'} = 2 B_k \tag{4.16}$$

since $\varepsilon_{ijk}\varepsilon_{ijk'} = 2\delta_{kk'}$. Thus,

$$B_k = \frac{1}{2}\varepsilon_{ijk} F_{ij} = -B^k \,. \tag{4.17}$$

Another useful tensor is obtained from $F^{\mu\nu}$, its dual. To define it, one should introduce the 4-dimensional Levi-Civita tensor $\varepsilon$ (see Appendix B for details). Its is defined as a fully antisymmetric tensor such that its contra-variant components are given by[2]

$$\varepsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 \text{ if } \{\mu\nu\rho\sigma\} \text{ is an even permutation of } \{0,1,2,3\} \\ -1 \text{ if } \{\mu\nu\rho\sigma\} \text{ is an odd permutation of } \{0,1,2,3\} \\ \phantom{+}0 \text{ otherwise.} \end{cases} \tag{4.18}$$

Since $\det g = -1$, we thus have

$$\varepsilon_{\mu\nu\rho\sigma} = -\varepsilon^{\mu\nu\rho\sigma} \,. \tag{4.19}$$

The dual tensor is then defined as

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} = -\tilde{F}^{\nu\mu} \tag{4.20}$$

since $\varepsilon^{\mu\nu\rho\sigma}$ is completely antisymmetric.

Conversely,

$$\begin{aligned} \varepsilon_{\mu\nu\rho\sigma}\tilde{F}^{\rho\sigma} &= \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\varepsilon^{\rho\sigma\lambda\tau} F_{\lambda\tau} \\ &= \frac{1}{2}\varepsilon_{\rho\sigma\mu\nu}\varepsilon^{\rho\sigma\lambda\tau} F_{\lambda\tau} \\ &= \frac{1}{2}(-2)(g_\mu^\lambda g_\nu^\tau - g_\mu^\tau g_\nu^\lambda)F_{\lambda\tau} = -2F_{\mu\nu} \end{aligned} \tag{4.21}$$

---

[2]Note that this definition is not universal. Some authors use this definition for the covariant components, which leads to an opposite tensor with respect to the present definition.

## 4. Electrodynamics

and thus

$$F_{\mu\nu} = -\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\tilde{F}^{\rho\sigma}\,. \tag{4.22}$$

In conclusion, we have

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \tag{4.23}$$

$$F_{\mu\nu} = -\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\tilde{F}^{\rho\sigma}\,. \tag{4.24}$$

The components of $\tilde{F}^{\mu\nu}$ can be obtained easily:

$$\tilde{F}^{0i} = \frac{1}{2}\varepsilon^{0i\rho\sigma}F_{\rho\sigma} = \frac{1}{2}\varepsilon^{0ijk}F_{jk} = \frac{1}{2}\varepsilon_{ijk}F_{jk} = B_i = -B^i\,, \tag{4.25}$$

and

$$\tilde{F}^{ij} = \frac{1}{2}\varepsilon^{ij\rho\sigma}F_{\rho\sigma} = \frac{1}{2}\varepsilon^{ij0k}F_{0k} + \frac{1}{2}\varepsilon^{ijk0}F_{k0} = \varepsilon^{0ijk}F_{0k} = \varepsilon_{ijk}F_{0k} = \varepsilon_{ijk}E^k\,. \tag{4.26}$$

Thus,

$$\tilde{F}^{\mu\nu} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & E^3 & -E^2 \\ B^2 & -E^3 & 0 & E^1 \\ B^3 & E^2 & -E^1 & 0 \end{pmatrix} \tag{4.27}$$

i.e.

$$\tilde{F}^{i0} = -\tilde{F}^{0i} = -\tilde{F}_{i0} = \tilde{F}_{0i} = B^i = -B_i\,, \tag{4.28}$$

$$\tilde{F}^{ij} = \tilde{F}_{ij} = -\varepsilon_{ijk}E_k = \varepsilon_{ijk}E^k\,. \tag{4.29}$$

Besides, the two tensors are related through the following transformation:

$$F^{\mu\nu} \xrightarrow[\vec{E}\to\vec{B}]{\vec{B}\to-\vec{E}} \tilde{F}^{\mu\nu}\,. \tag{4.30}$$

## 4.2.2. Lorentz transformations

As for any contra-variant 2-tensor, $F^{\mu\nu}$ transforms under a Lorentz transformation $\Lambda$ as

$$F'^{\mu\nu} = \Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\nu'}F^{\mu'\nu'}\,. \tag{4.31}$$

**Orthogonal transformations**

Le us first consider orthogonal transformations $O(3)$. They correspond to particular transformations $\Lambda$ of the form

$$\Lambda^0{}_0 = 1 \quad \Lambda^0{}_i = \Lambda^i{}_0 = 0 \quad \Lambda^i{}_j = O^i{}_j \tag{4.32}$$

i.e.

$$\Lambda = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & O & \\ 0 & & & \end{array}\right). \tag{4.33}$$

The $3 \times 3$ matrix $O$ satisfies the constraints

$$O^t O = O O^t = \mathrm{Id} \tag{4.34}$$

by definition of $O(3)$, which is the set of matrices of real orthogonal matrices, which equivalently leave the scalar product $\vec{x} \cdot \vec{y}$ invariant.

Taking the determinant of Eq. (4.34), one gets

$$(\det O)^2 = 1 \tag{4.35}$$

so that:

⋄ either $\det O = +1$: $SO(3)$ which are rotations

⋄ either $\det O = -1$: $O = R \cdot P$ where $R \in SO(3)$ and $P$ is the parity: $P = -\mathrm{Id}_{3\times3}$.

From Eq. (4.34), we get the following relations among components of $O$ :

$$(O^t)^i{}_{j'} O^{j'}{}_j = O^{j'}{}_i O^{j'}{}_j = \delta_{ij} \qquad \text{and} \qquad (O)^i{}_{j'}(O^t)^{j'}{}_j = O^i{}_{j'} O^j{}_{j'} = \delta_{ij} . \tag{4.36}$$

Let us now consider the way various components of $F$ transform under the Lorentz transform $\Lambda$.

⋄ Electric field:

$$F'^{i0} = \Lambda^i{}_j \Lambda^0{}_0 F^{j0}, \tag{4.37}$$

so that

$$E'^i = O^i{}_j E^j \tag{4.38}$$

## 4. Electrodynamics

as expected for a vector: this equation tells us that $\vec{E}$ transforms as a vector under a rotation $O = R \in SO(3)$, and gets reversed, just as $\vec{x}$, under parity.

Therefore

> $\vec{E}$ is a vector (also named polar vector).

◇ Magnetic field:

$$F'^{ij} = \Lambda^i{}_{i'}\Lambda^j{}_{j'}F^{i'j'} \,, \tag{4.39}$$

with

$$F'^{ij} = -\varepsilon_{ijk'}B'^{k'} \,. \tag{4.40}$$

We thus have

$$B'^k = -\frac{1}{2}\varepsilon_{ijk}F'^{ij} = \frac{1}{2}\varepsilon_{ijk}O^i{}_{i'}O^j{}_{j'}\varepsilon_{i'j'k'}B^{k'} \,. \tag{4.41}$$

Using the fact that

$$\varepsilon_{ijk} = \varepsilon_{ijn}\delta_{nk} \tag{4.42}$$

and

$$O^n{}_p O^k{}_p = \delta_{nk} \,, \tag{4.43}$$

see Eq. (4.36), we thus get

$$\varepsilon_{ijk}O^i{}_{i'}O^j{}_{j'} = \varepsilon_{ijn}O^i{}_{i'}O^j{}_{j'}O^n{}_p O^k{}_p = \varepsilon_{i'j'p}O^k{}_p \det O \,, \tag{4.44}$$

since

$$\varepsilon_{ijn}O^i{}_{i'}O^j{}_{j'}O^n{}_p = \varepsilon_{i'j'p}\det O \,, \tag{4.45}$$

see Appendix B. We finally get

$$B'^k = -\frac{1}{2}\varepsilon_{ijk}F'^{ij} = \frac{1}{2}\varepsilon_{i'j'p}\varepsilon_{i'j'k'}O^k{}_p B^{k'}\det O = \det O\, O^k{}_{k'}B^{k'}, \tag{4.46}$$

since $\varepsilon_{i'j'p}\varepsilon_{i'j'k'} = 2\delta_{pk'}$.

Due to the presence of the prefactor $\det O$, this equation tells us that $\vec{B}$ transforms as a vector under a rotation $O = R \in SO(3)$, and remains invariant under parity. Therefore,

> $\vec{B}$ is a pseudo-vector (also named axial vector).

**Pure boosts**

One can show, see tutorial, that under an arbitrary boost along the direction $\vec{n}$ ($\vec{n}^2 = 1$), i.e. with a velocity $\vec{v} = \beta\vec{n}$,

$$\vec{E}' = (\vec{E}\cdot\vec{n})\vec{n} + \gamma\left[\vec{E} - (\vec{E}\cdot\vec{n})\vec{n}\right] + \gamma\,\vec{v}\wedge\vec{B}\,, \qquad (4.47)$$

$$\vec{B}' = (\vec{B}\cdot\vec{n})\vec{n} + \gamma\left[\vec{B} - (\vec{B}\cdot\vec{n})\vec{n}\right] - \gamma\,\vec{v}\wedge\vec{E}\,. \qquad (4.48)$$

In the non-relativistic limit, the transformations (4.47) and (4.48) simplifies into

$$\vec{E}' = \vec{E} + \vec{v}\wedge\vec{B}\,, \qquad (4.49)$$

$$\vec{B}' = \vec{B} - \vec{v}\wedge\vec{E}\,, \qquad (4.50)$$

which are particularly important for induction.

## 4.2.3. Relativistic invariants

From the two tensors $F$ and $\tilde{F}$, on can easily build two Lorentz invariants.

⬦ First, consider the contraction of $F$ with itself:

$$\begin{aligned}
F_{\mu\nu}F^{\mu\nu} &= F_{0i}F^{0i} + F_{i0}F^{i0} + F_{ij}F^{ij} \\
&= 2F_{0i}F^{0i} + \varepsilon_{ijk}B_k\varepsilon_{ijk'}B_{k'} \\
&= -2E^iE^i + 2B_kB_k \\
&= -2\vec{E}^2 + 2\vec{B}^2 = -2(\vec{E}^2 - \vec{B}^2)\,. \qquad (4.51)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu} &= \frac{1}{4}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}\varepsilon^{\mu\nu\rho'\sigma'}F_{\rho'\sigma'} \\
&= \frac{1}{4}(-2)\left(g_\rho^{\rho'}g_\sigma^{\sigma'} - g_\rho^{\sigma'}g_\sigma^{\rho'}\right)F^{\rho\sigma}F_{\rho'\sigma'} = -F^{\rho\sigma}F_{\rho\sigma}\,, \quad (4.52)
\end{aligned}$$

therefore identical (up to the sign) with the contraction of $F$ with itself, see the expression (4.51), a fact which is obvious by just applying the symmetry (4.30) to the expression (4.51).

◇ Second, consider the contraction of $F$ with $\tilde{F}$ :

$$
\begin{aligned}
F_{\mu\nu}\tilde{F}^{\mu\nu} &= \frac{1}{2}F_{\mu\nu}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \\
&= \frac{1}{2}F_{ij}\varepsilon^{ijk0}F_{k0} + \frac{1}{2}F_{ij}\varepsilon^{ij0k}F_{0k} + \frac{1}{2}F_{0i}\varepsilon^{0ijk}F_{jk} + \frac{1}{2}F_{i0}\varepsilon^{i0jk}F_{jk} \\
&= -F_{ij}\varepsilon^{0ijk}F_{k0} + F_{0i}\varepsilon^{0ijk}F_{jk} \\
&= -F_{ij}\varepsilon_{ijk}F_{k0} + F_{0i}\varepsilon_{ijk}F_{jk} \\
&= 2F_{0i}\varepsilon_{ijk}F_{jk} = -4E^iB^i = -4\vec{E}\cdot\vec{B}\,. \qquad (4.53)
\end{aligned}
$$

In conclusion, we have built two relativistic invariants from the two tensors $F$ and $\tilde{F}$ :

$$
\begin{aligned}
F_{\mu\nu}F^{\mu\nu} &= -\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu} = -2(\vec{E}^2 - \vec{B}^2)\,, \qquad (4.54) \\
F_{\mu\nu}\tilde{F}^{\mu\nu} &= -4\vec{E}\cdot\vec{B}\,. \qquad (4.55)
\end{aligned}
$$

Note that under the transformation $(\vec{E}, \vec{B}) \to (\vec{B}, -\vec{E})$, the first invariant (4.54) is unchanged, while the second one (4.55) gets a minus sign, in accordance with (4.30). Indeed, the invariant (4.54) is a Lorentz scalar, while the invariant (4.55) is a Lorentz pseudo-scalar.

# 4.3. Covariant form of Maxwell's equations

## 4.3.1. Maxwell's equations

We will now show that the four Maxwell's equations (4.2) can be summarized in the following very elegant set of two equations, which are explicitly covariant

$$
\begin{aligned}
\partial_\mu F^{\mu\nu} &= j^\nu\,, \qquad (4.56) \\
\partial_\mu \tilde{F}^{\mu\nu} &= 0\,. \qquad (4.57)
\end{aligned}
$$

Proof:

◇ Gauss's law:

$$
\partial_i F^{i0} = j^0 \Leftrightarrow \vec{\nabla}\cdot\vec{E} = \rho \qquad (4.58)
$$

$\diamond$ Generalization of Maxwell-Ampère's law:

$$\partial_0 F^{0k} + \partial_i F^{ik} = j^k \tag{4.59}$$

with

$$F^{ik} = -\varepsilon_{ikn} B^n \text{ and } F^{0k} = -E^k \tag{4.60}$$

so that

$$-\partial_t \vec{E}^k - \varepsilon_{ikn} \overrightarrow{\nabla}^i \vec{B}^n = \left( -\partial_t \vec{E} + \overrightarrow{\nabla} \wedge \vec{B} \right)^k = j^k \tag{4.61}$$

i.e.

$$\overrightarrow{\mathrm{rot}} \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j} \,. \tag{4.62}$$

$\diamond$ Absence of magnetic monopole:

$$\partial_i \tilde{F}^{i0} = 0 \quad \Leftrightarrow \quad \partial_i B^i = 0 \quad \Leftrightarrow \quad \mathrm{div} \vec{B} = 0 \,. \tag{4.63}$$

$\diamond$ Maxwell-Faraday's equation:

$$\partial_0 \tilde{F}^{0k} + \partial_i \tilde{F}^{ik} = 0 \,. \tag{4.64}$$

with

$$\tilde{F}^{ik} = \varepsilon_{ikn} E^n \text{ and } \tilde{F}^{0k} = -B^k \tag{4.65}$$

so that

$$\overrightarrow{\mathrm{rot}} \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \,. \tag{4.66}$$

Compatibility condition:

The field strength tensor $F^{\mu\nu}$ is antisymmetric, therefore

$$\partial_\nu \partial_\mu F^{\mu\nu} = 0 = \partial_\nu j^\nu \,, \tag{4.67}$$

which means that the current should be conserved:

$$\boxed{\partial_\nu j^\nu = 0 \,.} \tag{4.68}$$

## 4.3.2. Four-potential

**From $F^{\mu\nu}$ to $A^\mu$**

Several remarks are in order:

⋄ Maxwell's equations, expressed in terms of $\vec{E}$ and $\vec{B}$, are not covariant.

⋄ They only involve first order time derivative of the dynamical variables $\vec{E}$ and $\vec{B}$. Thus, their conjugated momenta are not independent of the dynamical variable[3].

One should rather look for second order equations.

Let us introduce the *four-potential* $A^\mu$ such that

$$\boxed{F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu\,.} \tag{4.69}$$

Theorem (Poincaré):

The equation $\partial_\mu \tilde{F}^{\mu\nu} = 0$ is a necessary condition for $A^\mu$ to exist. This is a sufficient condition if the space is contractible,[4] which means that it can be smoothly contracted to a point.

Eq. (4.69) implies that

$$F^{i0} = \partial^i A^0 - \partial^0 A^i \quad \Leftrightarrow \quad \vec{E} = -\vec{\nabla} A^0 - \frac{\partial \vec{A}}{\partial t} \tag{4.70}$$

and

$$F^{ij} = \partial^i A^j - \partial^j A^i \tag{4.71}$$

i.e.

$$-\frac{1}{2}\varepsilon_{ijk}F^{ij} = -\frac{1}{2}\varepsilon_{ijk}\left(\partial^i A^j - \partial^j A^i\right) = -\varepsilon_{ijk}\partial^i A^j = \left(\vec{\nabla} \wedge \vec{A}\right)^k \tag{4.72}$$

and thus

$$\vec{E} = -\vec{\nabla} A^0 - \frac{\partial \vec{A}}{\partial t} \tag{4.73}$$

$$\vec{B} = \vec{\nabla} \wedge \vec{A}\,. \tag{4.74}$$

---

[3]Consider for example the trivial case $\mathcal{L} = \dot{\phi}\phi$, an action in which $\dot{\phi}$ does not appear quadratically.

[4]$S^1$ is neither simply connected nor contractible. Obviously, a domain which is contractible is simply connected, but the opposite might not be true. Indeed, $S^n$ for $n > 1$ is simply connected but not contractible.

Consistency check:

One should have $\partial_\mu \tilde{F}^{\mu\nu} = 0$ which is indeed satisfied since

$$\partial_\mu \tilde{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\partial_\mu\left(\partial_\rho A_\sigma - \partial_\sigma A_\rho\right) = \varepsilon^{\mu\nu\rho\sigma}\partial_\mu\partial_\rho A_\sigma = 0\,, \qquad (4.75)$$

since $\varepsilon^{\mu\nu\rho\sigma}$ is symmetric under $\mu \longleftrightarrow \rho$ exchange, while $\partial_\mu\partial_\rho$ is symmetric.

## Gauge invariance

The relation (4.69) does not provide a unique 4-potential $A^\mu$.

Indeed, the *gauge* transformation

$$A^\mu(x) \to A^\mu(x) + \partial^\mu\phi\,, \qquad (4.76)$$

where $\phi$ is an arbitrary function, leaves $F^{\mu\nu}$ invariant.

Assuming the space-time domain to be contractible, Maxwell's equations are equivalent to

$$\partial_\mu F^{\mu\nu} = \Box A^\nu - \partial^\nu\left(\partial_\mu A^\mu\right) = j^\nu\,. \qquad (4.77)$$

The equation

$$\Box A^\nu - \partial^\nu\left(\partial_\mu A^\mu\right) = j^\nu \qquad (4.78)$$

is gauge invariant, since under the gauge transformation (4.76), we have

$$\Box A^\nu - \partial^\nu\left(\partial_\mu A^\mu\right) \to \Box A^\nu + \partial^\nu\Box\phi - \partial^\nu\left(\partial_\mu A^\mu\right) - \partial^\nu\Box\phi = \Box A^\nu - \partial^\nu\left(\partial_\mu A^\mu\right)\,. \qquad (4.79)$$

Let us write the various components of Maxwell's equations (4.78) explicitly:

◇ Time component:

$$\Box A^0 - \frac{\partial}{\partial t}\left[\frac{\partial A^0}{\partial t} + \mathrm{div}\vec{A}\right] = \rho\,, \qquad (4.80)$$

i.e.

$$\boxed{-\Delta A^0 - \frac{\partial}{\partial t}\mathrm{div}\vec{A} = \rho\,.} \qquad (4.81)$$

◇ Space components:

$$\boxed{\Box\vec{A} + \vec{\nabla}\left(\frac{\partial A^0}{\partial t} + \mathrm{div}\vec{A}\right) = \vec{j}\,.} \qquad (4.82)$$

## Usual gauges

There is an infinite set of possible gauge choices. The most popular one are:

⬦ Landau gauge

It relies on the covariant gauge-fixing condition

$$\partial_\mu A^\mu = 0 \tag{4.83}$$

which thus implies that

$$\Box A^\nu = j^\nu \,. \tag{4.84}$$

⬦ Coulomb gauge

It relies on the non-covariant gauge-fixing condition

$$\mathrm{div}\vec{A} = 0 \tag{4.85}$$

which implies that the scalar potential $A^0$ satisfies

$$-\Delta A^0 = \rho \,. \tag{4.86}$$

The solution of this second-order differential equation is given by

$$A^0(t, \vec{x}) = \frac{1}{4\pi} \int d^3 y \frac{\rho(t, \vec{y})}{|\vec{y} - \vec{x}|} \tag{4.87}$$

since

$$-\Delta \frac{1}{|\vec{r} - \vec{r}'|} = 4\pi \delta(\vec{r} - \vec{r}') \,. \tag{4.88}$$

In this gauge, $A^0$ is therefore an instantaneous potential. The vector potential $\vec{A}$ is then, from Eq. (4.82), solution of the equation

$$\Box \vec{A} = \vec{j} - \vec{\nabla} \int \frac{d^3 x'}{4\pi} \frac{\partial \rho(t, \vec{x}')}{\partial t} \frac{1}{|\vec{x} - \vec{x}'|} \,. \tag{4.89}$$

Note the consistency with the fact that the current should be conserved:

$$\mathrm{div}\,\Box \vec{A} \;=\; \mathrm{div}\vec{j} - \Delta \int \frac{d^3 x'}{4\pi} \frac{\partial \rho(t, \vec{x}')}{\partial t} \frac{1}{|\vec{x} - \vec{x}'|} \tag{4.90}$$

$$\;=\; \mathrm{div}\vec{j} + \int d^3 x' \,\delta(\vec{x} - \vec{x}') \frac{\partial \rho(t, \vec{x}')}{\partial t} \tag{4.91}$$

$$\;=\; \mathrm{div}\vec{j} + \frac{\partial \rho(t, \vec{x})}{\partial t} = 0 \,, \tag{4.92}$$

which is therefore consistent with our gauge choice $\mathrm{div}\vec{A} = 0$.

◇ Temporal gauge

In this gauge, the scalar potential is chosen to vanish

$$A^0 = 0 \,. \tag{4.93}$$

◇ Axial gauge

In this gauge, specified by a given space-like direction, e.g. $z$,

$$A^3 = 0 \,. \tag{4.94}$$

◇ Light-cone gauge

This gauge, of particular use in high-energy particle physics, is specified by choosing a light-cone vector $n$ with $n^2 = 0$, such that

$$A \cdot n = 0 \,. \tag{4.95}$$

◇ General axial gauge: it includes the three previous cases, by introducing a vector $n$, imposing

$$A \cdot n = 0 \,. \tag{4.96}$$

Clearly, $n^2 > 0$ is the temporal gauge, $n^2 < 0$ is the pure axial gauge, and $n^2 = 0$ is the light-cone gauge.

## 4.4. Lagrangian for photons

We are looking for a lagrangian describing the photon degrees of freedom. It should be local, quadratic in the field $A^\mu$ and its derivatives, Lorentz invariant, and its equations of motion should be (4.78). We thus generically write

$$\mathcal{L}(x) = a\, A^\mu A_\mu + b\,(\partial_\mu A^\nu)(\partial_\nu A^\mu) + c\,(\partial_\mu A^\nu)(\partial^\mu A_\nu) + d\,(\partial_\mu A^\mu)^2 + e\, A_\mu j^\mu. \tag{4.97}$$

The Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial A^\rho} - \partial_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\sigma A^\rho)} = 0 \tag{4.98}$$

then leads to

$$2aA_\rho + ej_\rho = \partial_\sigma \left[ 2b\,\partial_\rho A^\sigma + 2c\,\partial^\sigma A_\rho + 2d\,\delta^\sigma_\rho \partial_\mu A^\mu \right] \tag{4.99}$$

*4. Electrodynamics*

Let us study the gauge invariance of the equation of motion for $j = 0$ : under the gauge transformation (4.76), we get

$$a\,\partial_\rho\phi = b\,\partial_\rho\partial_\sigma\partial^\sigma\phi + c\,\partial_\rho\partial_\sigma\partial^\sigma\phi + d\,\partial_\rho\partial_\sigma\partial^\sigma\phi = (b + c + d)\partial_\rho\Box\phi, \quad (4.100)$$

therefore one should have $a = 0$ and $b + c + d = 0$.

Besides, rewriting the term $(\partial_\mu A^\mu)^2$ in the Lagrangian density as

$$(\partial_\mu A^\mu)^2 = (\partial_\mu A^\nu)(\partial_\nu A^\mu) + \partial_\mu\left[A_\nu\left(g^{\mu\nu}\partial_\rho A^\rho - \partial^\nu A^\mu\right)\right], \quad (4.101)$$

and using the fact that the second term in the r.h.s. is a total derivative, which thus does not contribute to the action, it can be simply omitted. From the fact that $b + d = -c$, we get

$$\mathcal{L} = -c\left(\partial_\mu A^\nu\partial_\nu A^\mu - \partial_\mu A^\nu\partial^\mu A_\nu\right) + e\,A_\mu j^\mu. \quad (4.102)$$

Using the fact that

$$\partial_\mu A^\nu\partial_\nu A^\mu - \partial_\mu A^\nu\partial^\mu A_\nu = -\frac{1}{2}F^{\mu\nu}F_{\mu\nu}, \quad (4.103)$$

the Lagrangian density thus reads

$$\mathcal{L} = \frac{c}{2}F^{\mu\nu}F_{\mu\nu} + e\,A_\mu j^\mu. \quad (4.104)$$

It remains to adjust the values of $c$ and $e$ properly.

Let us come back to the equation of motion (4.98). Using the fact that

$$\frac{\partial}{\partial(\partial_\sigma A^\rho)}\left(F_{\mu\nu}F^{\mu\nu}\right) = 4F^\sigma{}_\rho, \quad (4.105)$$

we get

$$4\frac{c}{2}\partial_\sigma F^\sigma{}_\rho = e\,j_\rho \quad (4.106)$$

so that in order for this equation to be equivalent to the Maxwell equation (4.56), one should have $2c = e$.

Actually, the overall normalization, as well as the relative value of $c$ and $e$ can be obtained as follows:

◇ $\vec{A}$ is a dynamical variable. Thus, the 'kinetic' term should be

$$\frac{1}{2}\left(\frac{\partial\vec{A}}{\partial t}\right)^2 \quad (4.107)$$

in $\mathcal{L}$. Now, by inspection

$$\frac{c}{2}F_{\mu\nu}F^{\mu\nu} \tag{4.108}$$

provides

$$\frac{c}{2} \times 2 \times F^{0i}F_{0i} = c\,\partial^0 A^i \partial_0 A_i = -c\left(\partial_0 A^i\right)^2 \tag{4.109}$$

so that $c = -\frac{1}{2}$.

$\diamond$ In the Lagrangian density $\mathcal{L} = \mathcal{T} - \mathcal{U}$ where $\mathcal{T}$ is the kinetic energy and $\mathcal{U}$ is the potential energy, we know that $\mathcal{U}$ should contain the potential energy $\rho A^0$ in exactly this form. Besides, $j^\mu A_\mu$ provides exactly $\rho A^0$, so that $e = -1$.

In conclusion,

$$\boxed{\mathcal{L}(x) = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu\,,} \tag{4.110}$$

and the corresponding action reads

$$\boxed{I = \int \mathcal{L}(x)\,d^4x = \int d^4x \left[\frac{1}{2}\left(\vec{E}^2 - \vec{B}^2\right) - \rho A^0 + \vec{j}\cdot\vec{A}\right].} \tag{4.111}$$

One should note that the Lagrangian density (4.110) is not gauge invariant if $j(x)$ is an external current. Indeed under a gauge transformation,

$$\mathcal{L}(x) \xrightarrow{\;A^\mu \to A^\mu + \partial^\mu \phi\;} \mathcal{L}(x) - j_\mu \partial^\mu \phi\,. \tag{4.112}$$

Still, since

$$j_\mu \partial^\mu \phi = \partial^\mu\left(\phi j_\mu\right) - \phi\,\partial^\mu j_\mu \tag{4.113}$$

in which, in the r.h.s, the first term is a total derivative and the second one vanishes because $j_\mu$ is a conserved current. Therefore, both terms can be omitted from the action $I$.

## 4.5. Coupling between matter and electromagnetic field

### 4.5.1. Matter part

As a reminder, consider two real scalar fields $\phi_1$ and $\phi_2$, of identical masses. The Lagrangian densities for these two fields are

$$\mathcal{L}_1 = \frac{1}{2}\left(\partial_\mu\phi_1\right)\left(\partial^\mu\phi_1\right) - \frac{1}{2}m^2\phi_1^2, \tag{4.114}$$

$$\mathcal{L}_2 = \frac{1}{2}\left(\partial_\mu\phi_2\right)\left(\partial^\mu\phi_2\right) - \frac{1}{2}m^2\phi_2^2. \tag{4.115}$$

Introducing the fields

$$\Phi = \frac{1}{\sqrt{2}}\left[\Phi_1 + i\,\Phi_2\right], \tag{4.116}$$

$$\Phi^* = \frac{1}{\sqrt{2}}\left[\Phi_1 - i\,\Phi_2\right] \tag{4.117}$$

which are treated as independent fields, we have

$$\left|\Phi\right|^2 = \Phi\,\Phi^* = \frac{1}{2}\left[\Phi_1^2 + \Phi_2^2\right] \tag{4.118}$$

and

$$\left|\partial_\mu\Phi\right|^2 = \left(\partial_\mu\Phi\right)\left(\partial^\mu\Phi\right)^* \tag{4.119}$$

$$= \frac{1}{2}\left(\partial_\mu\Phi_1 + i\,\partial_\mu\Phi_2\right)\left(\partial^\mu\Phi_1 - i\,\partial^\mu\Phi_2\right) \tag{4.120}$$

$$= \frac{1}{2}\left(\partial_\mu\Phi_1\right)\left(\partial^\mu\Phi_1\right) + \frac{1}{2}\left(\partial_\mu\Phi_2\right)\left(\partial^\mu\Phi_2\right). \tag{4.121}$$

Thus, $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$ can be rewritten as

$$\mathcal{L} = \left(\partial_\mu\Phi\right)\left(\partial^\mu\Phi\right)^* - m^2\Phi\,\Phi^*. \tag{4.122}$$

The Euler-Lagrange equations read

$$\frac{\partial\mathcal{L}}{\partial\Phi} - \partial_\sigma\frac{\partial\mathcal{L}}{\partial(\partial_\sigma\Phi)} = 0 \quad\Rightarrow\quad \Box\Phi^* + m^2\,\Phi^* = 0 \tag{4.123}$$

$$\frac{\partial\mathcal{L}}{\partial\Phi^*} - \partial_\sigma\frac{\partial\mathcal{L}}{\partial(\partial_\sigma\Phi^*)} = 0 \quad\Rightarrow\quad \Box\Phi + m^2\,\Phi = 0 \tag{4.124}$$

As we have seen in Chap. 2, under global $U(1)$ transformations

$$
\begin{aligned}
\Phi &\rightarrow e^{ie\alpha}\Phi &&i.e. \ \delta\Phi = ie\delta\alpha\Phi\,, &&(4.125)\\
\Phi^* &\rightarrow e^{-ie\alpha}\Phi^* &&i.e. \ \delta\Phi^* = -ie\delta\alpha\Phi^*, &&(4.126)
\end{aligned}
$$

the Lagrangian $\mathcal{L}$ is invariant, which implies that there is a conserved Noether current, namely

$$
\begin{aligned}
j^\mu &= \frac{1}{\delta\alpha}\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\delta\Phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^*)}\delta\Phi^*\right]\\
&= ie\left[(\partial^\mu\Phi^*)\Phi - (\partial^\mu\Phi)\Phi^*\right]\\
&= -ie\left[\Phi^*(\partial^\mu\Phi) - (\partial^\mu\Phi^*)\Phi\right]\\
&= -ie\,\Phi^*\overleftrightarrow{\partial}^\mu\Phi\,. &&(4.127)
\end{aligned}
$$

One can add a potential term to $\mathcal{L}$, of the form $V(\Phi^*\Phi)$, without spoiling the gauge invariance, so that we denote

$$
\mathcal{L}_{\text{matter}} = \mathcal{L}_{\text{free}} - V(\Phi^*\Phi)\,. \tag{4.128}
$$

A minimal example is $V = \lambda(\Phi^*\Phi)^2$. With such a modification, the equations of motion then read

$$
\begin{aligned}
\Box\,\Phi^* + m^2\Phi^* &= -\frac{\partial V}{\partial\Phi}\,, &&(4.129)\\
\Box\,\Phi + m^2\Phi &= -\frac{\partial V}{\partial\Phi^*}\,. &&(4.130)
\end{aligned}
$$

Obviously, since $\mathcal{L}_{\text{matter}}$ and $\mathcal{L}_{\text{free}}$ differ only through powers of $\Phi^*\Phi$, the Noether current (4.127) remains identical.

## 4.5.2. Pure photon part

As we have shown above, the Lagrangian which describes the dynamics of photons reads

$$
\mathcal{L}_{\text{em}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,. \tag{4.131}
$$

## 4.5.3. Interaction between matter and photons

### Minimal Lagrangian through a dynamical current

We know that one can add a term $-j^\mu A_\mu$, so that the equation of motion for $A_\mu$ lead to the Maxwell equations. Doing, so, we immediately encounter a technical

## 4. Electrodynamics

problem: indeed, this is valid for $j_\mu$ being external, non dynamical, but what if $j^\mu$ is promoted to be dynamical, being itself constructed from dynamical fields?

We already have a candidate to start with:

$$-j^\mu_{\text{Noether}} A_\mu = ie\, \Phi^* \overleftrightarrow{\partial}{}^\mu \Phi A_\mu \,. \tag{4.132}$$

But this new term in the Lagrangian, which now involves derivatives of the fields $\Phi$ and $\Phi^*$, will modify the Noether current itself, so that this additional term should be itself modified, changing the current, etc. My gosh! Do we enter an endless loop?

Let us show that one can find a minimal and consistent solution to this difficulty. For that purpose, we write

$$\mathcal{L} = \mathcal{L}_{\text{em}} + \mathcal{L}_{\text{matter}} + \mathcal{L}_{\text{int}} \,. \tag{4.133}$$

Let us formulate the problem more precisely.

◇ The Euler-Lagrange equations applied to $\mathcal{L}$ for $A^\mu$

$$\partial_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\sigma A^\rho)} = \frac{\partial \mathcal{L}}{\partial A^\rho} \tag{4.134}$$

should lead to the Maxwell equations. Two remarks are then in order:

- Since the dynamics of $A^\mu$ (its time derivative, and by covariance any of its derivatives) is inside $\mathcal{L}_{\text{em}}$, only $\mathcal{L}_{\text{em}}$ contributes to the l.h.s. of Eq. (4.134).

- The r.h.s. of Eq. (4.134) gets contributions from $\mathcal{L}_{\text{int}}$ alone.

Thus, Eq. (4.134) actually reads

$$\partial_\sigma \frac{\partial \mathcal{L}_{\text{em}}}{\partial(\partial_\sigma A^\rho)} = \frac{\partial \mathcal{L}_{\text{int}}}{\partial A^\rho} \,, \tag{4.135}$$

or equivalently

$$-\partial_\sigma F^\sigma{}_\rho = \frac{\partial \mathcal{L}_{\text{int}}}{\partial A^\rho} = -J_\rho \tag{4.136}$$

where $J_\rho$ is the *full* Noether current, since Maxwell's equations with <u>this</u> current should hold.

◇ The full Noether current is given by

$$J^\mu = \frac{1}{\delta\alpha}\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\delta\Phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^*)}\delta\Phi^*\right] \tag{4.137}$$

so that

$$J^\mu = ie\left[\frac{\partial\mathcal{L}_{\text{matter}}}{\partial(\partial_\mu\Phi)}\Phi - \frac{\partial\mathcal{L}_{\text{matter}}}{\partial(\partial_\mu\Phi^*)}\Phi^*\right]$$
$$+ie\left[\frac{\partial\mathcal{L}_{\text{int}}}{\partial(\partial_\mu\Phi)}\Phi - \frac{\partial\mathcal{L}_{\text{int}}}{\partial(\partial_\mu\Phi^*)}\Phi^*\right], \tag{4.138}$$

In this expression, the first part, coming from $\mathcal{L}_{\text{matter}}$, is the current

$$j^\mu = -ie\,\Phi^*\overleftrightarrow{\partial}^\mu\Phi\,, \tag{4.139}$$

see Eq. (4.127), which is of course independent of $A^\mu$.

Our problem is thus to look for $\mathcal{L}_{\text{int}}$, solution of the coupled equations (4.136) and (4.138). We are looking for a minimal solution (the problem has per se no unique solution).

First, let us integrate Eq. (4.136) with respect to $A^\rho$ : separating the term $-j^\mu A_\mu$ out of $\mathcal{L}_{\text{int}}$, we write

$$\mathcal{L}_{\text{int}} = -j^\mu A_\mu + \mathcal{L}'_{\text{int}} = ie\,\Phi^*\overleftrightarrow{\partial}^\mu\Phi A_\mu + \mathcal{L}'_{\text{int}}\,. \tag{4.140}$$

We now make a minimal assumption: we look for a solution with $\mathcal{L}'_{\text{int}}$ independent of $\partial_\mu\Phi$ and $\partial_\mu\Phi^*$.
Thus,

$$\frac{\partial\mathcal{L}_{\text{int}}}{\partial(\partial^\rho\Phi)} = ieA_\rho\Phi^*, \tag{4.141}$$

$$\frac{\partial\mathcal{L}_{\text{int}}}{\partial(\partial^\rho\Phi^*)} = -ieA_\rho\Phi \tag{4.142}$$

and

$$J_\rho = j_\rho + ie\left[ieA_\rho\Phi^*\Phi - (-ie)A_\rho\Phi\Phi^*\right]$$
$$= j_\rho - 2e^2A_\rho\Phi^*\Phi\,, \tag{4.143}$$

i.e.

$$\boxed{J_\rho = -ie\,\Phi^*\overleftrightarrow{\partial}_\rho\Phi - 2e^2A_\rho\Phi^*\Phi\,.} \tag{4.144}$$

69

## 4. Electrodynamics

The last step is to use Eq. (4.136), which thus reads

$$\frac{\partial \mathcal{L}_{\text{int}}}{\partial A^\rho} = -J_\rho = -j_\rho + 2e^2 A_\rho \Phi^* \Phi \tag{4.145}$$

and to solve it for $\mathcal{L}_{\text{int}}$. We thus finally get

$$\mathcal{L}_{\text{int}} = -j^\mu A_\mu + e^2 A^2 \Phi^* \Phi \,. \tag{4.146}$$

In conclusion, we have constructed the minimal Lagrangian

$$
\begin{aligned}
\mathcal{L} &= \mathcal{L}_{\text{em}} + \mathcal{L}_{\text{matter}} + \mathcal{L}_{\text{int}} \\
&= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (\partial_\mu \Phi)(\partial^\mu \Phi)^* - m^2 \Phi \Phi^* - V(\Phi^* \Phi) \\
&\quad + ieA_\mu (\Phi^* \overleftrightarrow{\partial}^\mu \Phi) + e^2 A^2 \Phi^* \Phi \,.
\end{aligned}
\tag{4.147}
$$

Let us summarize the equations of motion satisfied by the fields $A^\mu$, $\Phi$ and $\Phi^*$.

⋄ First, the equation of motion for $A^\mu$ reads

$$\partial_\mu F^{\mu\nu} = J^\mu \tag{4.148}$$

with $J^\mu$ given by Eq. (4.144).

⋄ Second, the equations of motion for $\Phi$ and $\Phi^*$, using respectively

$$\partial_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\sigma \Phi)} = \frac{\partial \mathcal{L}}{\partial \Phi} \tag{4.149}$$

and

$$\partial_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\sigma \Phi^*)} = \frac{\partial \mathcal{L}}{\partial \Phi^*} \,, \tag{4.150}$$

read[5]

$$\left[ (\partial_\mu + ieA_\mu)(\partial^\mu + ieA^\mu) + m^2 \right] \Phi^* = -\frac{\partial V}{\partial \Phi} \,. \tag{4.151}$$

and

$$\left[ (\partial_\mu - ieA_\mu)(\partial^\mu - ieA^\mu) + m^2 \right] \Phi = -\frac{\partial V}{\partial \Phi^*} \,. \tag{4.152}$$

---

[5]Beware on the fact that in these two equations, the differential operators $\partial_\mu$ and $\partial^\mu$ act on any structure which is on their right, namely $A_\mu$, $\Phi$, $\Phi^*$.

## Conservation of the current $J^\mu$

Let us take a short breath and verify the consistency of what we have done so far. As we know, the current $J^\mu$ should be conserved. Let us check that our constructed current indeed satisfies this constraint, a fact which is not completely obvious from its explicit expression (4.144).

A direct computation gives

$$
\begin{aligned}
\partial_\mu J^\mu &= -ie(\partial_\mu \Phi^*)(\partial^\mu \Phi) + ie(\partial^\mu \Phi)(\partial_\mu \Phi^*) - ie\Phi^* \Box \Phi + ie(\Box \Phi^*)\Phi \\
&\quad -2e^2(\partial_\mu A^\mu)\Phi^*\Phi - 2e^2 A^\mu(\partial_\mu \Phi^*)\Phi - 2e^2 A^\mu \Phi^*(\partial_\mu \Phi)\,.
\end{aligned} \tag{4.153}
$$

The first two terms cancel each other. For the other ones, things are a little bit more tricky. The two equations of motion (4.151) and (4.152) can be respectively expanded as

$$
(\Box + m^2)\Phi^* + ie(\partial_\mu A^\mu)\Phi^* + 2ieA^\mu(\partial_\mu \Phi^*) - e^2 A^2 \Phi^* = -\frac{\partial V}{\partial \Phi} \tag{4.154}
$$

and

$$
(\Box + m^2)\Phi - ie(\partial_\mu A^\mu)\Phi - 2ieA^\mu(\partial_\mu \Phi) - e^2 A^2 \Phi = -\frac{\partial V}{\partial \Phi^*}\,. \tag{4.155}
$$

Computing $ie\Phi \times (4.154) - ie\Phi^* \times (4.155)$ thus gives

$$
ie\Phi(\Box \Phi^*) - ie\Phi^*(\Box \Phi) - 2e^2 A^\mu(\partial_\mu \Phi^*)\Phi - 2e^2 A^\mu \Phi^*(\partial_\mu \Phi) - 2e^2(\partial_\mu A^\mu)\Phi^*\Phi
$$
$$
= -ie\left[\Phi\frac{\partial V}{\partial \Phi} - \Phi^*\frac{\partial V}{\partial \Phi^*}\right] = 0 \tag{4.156}
$$

which explicitly shows, when inserted in Eq. (4.153), that the current $J^\mu$ is indeed conserved.

## Minimal Lagrangian from gauge invariance

One may have noticed that the structure of the Lagrangian (4.147) is very peculiar, as one may guess from the equations of motion (4.151) and (4.152). Indeed, $\mathcal{L}$ can be rewritten has

$$
\boxed{\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu \Phi)^*(D_\mu \Phi) - m^2\Phi^*\Phi - V(\Phi^*\Phi)} \tag{4.157}
$$

with the *covariant derivative* defined as[6]

$$
\boxed{D_\mu = \partial_\mu - ieA_\mu\,,} \tag{4.158}
$$

---

[6]The present convention is the same as the one used in Halzen-Martin. It is also consistent with Itzykson-Zuber and Peskin-Schroeder. In these two last references, one should pay attention to the fact that $e$ is the electric charge of the electron, i.e. $e = -|e|$, while we use here the convention that $e = |e|$.

## 4. Electrodynamics

so that the equations of motion (4.151) and (4.152) read

$$\left[ D^2 + m^2 \right]^* \Phi^* = -\frac{\partial V}{\partial \Phi}, \qquad (4.159)$$

$$\left[ D^2 + m^2 \right] \Phi = -\frac{\partial V}{\partial \Phi^*}. \qquad (4.160)$$

Let us make a step back, and re-examine the problem of coupling matter to photon, just starting from the Lagrangian $\mathcal{L}_{\text{matter}}$ on which we *impose* gauge invariance, as a general principle.

We thus consider $U(1)$ *local* transformation, named *gauge* transformation:

$$\Phi(x) \rightarrow \Phi(x) \, e^{ie\alpha(x)} \qquad (4.161)$$

$$\Phi^*(x) \rightarrow \Phi^*(x) \, e^{-ie\alpha(x)}. \qquad (4.162)$$

Thus,

$$\partial_\mu \Phi(x) \rightarrow \partial_\mu \Phi(x) \, e^{ie\alpha(x)} + ie \, \partial_\mu \alpha(x) \Phi(x) \, e^{ie\alpha(x)} \qquad (4.163)$$

$$\partial_\mu \Phi^*(x) \rightarrow \partial_\mu \Phi^*(x) \, e^{ie\alpha(x)} - ie \, \partial_\mu \alpha(x) \Phi^*(x) \, e^{-ie\alpha(x)}. \qquad (4.164)$$

Clearly, while $-m^2 \Phi^* \Phi - V(\Phi^* \Phi)$ is gauge invariant, this invariance is broken by the term $(\partial_\mu \Phi(x))^* \partial^\mu \Phi(x)$.

The way to restore this gauge invariance is to introduce a new field $A^\mu$, which carries a $\mu$ index like $\partial_\mu$ (it is thus a spin one field, as shown by the study of the representations of the Lorentz group). These two are combined to built up the covariant derivative

$$D_\mu = \partial_\mu - ieA_\mu.$$

Suppose now that under a gauge transformation, $A^\mu$ transforms as

$$A^\mu(x) \rightarrow A^\mu(x) + \partial^\mu \alpha(x), \qquad (4.165)$$

which we already encountered in Eq. (4.76) when discussing gauge invariance of Maxwell's equations when expressed in terms of $A^\mu$. One then immediately sees that

$$\begin{aligned} D_\mu \Phi(x) &\rightarrow [\partial_\mu \Phi(x) + ie\partial_\mu \alpha(x) \, \Phi(x) - ieA_\mu(x) \, \Phi(x) - ie\partial_\mu \alpha(x) \, \Phi(x)] \, e^{ie\alpha(x)} \\ &= [\partial_\mu \Phi(x) - ieA_\mu(x) \, \Phi(x)] \, e^{ie\alpha(x)} \\ &= [D_\mu \Phi(x)] \, e^{ie\alpha(x)} \end{aligned} \qquad (4.166)$$

and similarly

$$[D_\mu \Phi(x)]^* \rightarrow [D_\mu \Phi(x)]^* \, e^{-ie\alpha(x)}. \qquad (4.167)$$

Therefore, making the minimal replacement

$$\left[\partial_\mu \Phi(x)\right]^* \left[\partial_\mu \Phi(x)\right] \rightarrow \left[D_\mu \Phi(x)\right]^* \left[D_\mu \Phi(x)\right], \qquad (4.168)$$

the kinetic part of the Lagrangian becomes gauge invariant. Adding this term to the mass and potential terms for the fields $\Phi$ and $\Phi^*$, and to the pure QED Lagrangian, one therefore gets the full QED gauge invariant Lagrangian which we obtained earlier, see Eq. (4.157).

This construction can be extended to other groups. This is the essence of the Yang-Mills construction. This is THE way to construct a dynamical field theory which couples matter and gauge fields.

For example, passing from $U(1)$ to $U(1) \times SU(2)$ gauge invariance led to the construction of the electroweak theory, with quarks and leptons (electron, muon, tau and their associated neutrinos) as matter fields, carrying charges under this group, and $\gamma, W^\pm, Z^0$ as gauge fields.

Similarly, having a gauge theory based on the (color) group $SU(3)$ leads to the (quantum) chromodynamics (QCD), the modern theory of strong interaction, with quarks as matter fields, and gluons as gauge fields.

This leads to the Standard Model, based on the gauge group $U(1) \times SU(2) \times SU(3)$.

One should note that the gauge field dynamics itself is governed by a term of the type (4.131). In general, as it is the case for the Standard Model, the gauge group is non-abelian. The consequence is that the gauge fields themselves carry a charge: by construction, gauge fields live in the *adjoint representation* of the group. For an abelian group, this representation is trivial, and indeed we know that the photon as no charge, and therefore does not couple to itself. But for a non-abelian group, the gauge fields acquire a charge, so that they can couple to each other. This is the case of $W^\pm$ which carry an electric charge, of $Z^0, W^\pm$ which carry a weak isospin, of the gluons which carry a color charge. This is technically hidden in the field strength $F^{\mu\nu}$ which is not anymore linear in the gauge field.

We refer to lectures on group theory for more details.

# 5. From Lagrangians to cross-sections

In this chapter, we will show how to compute a physical observable for a given scattering process, relying on the Lagrangian built in the previous chapter. This will be limited to the case of scalar matter, in order to avoid any complication related to spin.

## 5.1. Perturbation theory

From the interaction part of the Lagrangian, since

$$\mathcal{L} = \mathcal{T} - \mathcal{U}, \tag{5.1}$$

we can identify

$$\mathcal{U} = -\mathcal{L}_{\text{int}} . \tag{5.2}$$

In the scattering of a "spinless" electron off an electromagnetic potential $A_\mu$, facing a transition between an initial state $i$ to a final state $f$, using the covariant notation of Chap. 3, the amplitude $T_{fi}$ thus reads, at lowest order, according to Eq. (3.30),

$$T_{fi} = i \int \langle f | \mathcal{L}_{\text{int}} | i \rangle \, d^4x \,, \tag{5.3}$$

where $\langle f | \mathcal{L}_{\text{int}} | i \rangle$, see Eq. (4.146), is given by

$$\langle f | \mathcal{L}_{\text{int}} | i \rangle = ie A^\mu \Phi_f^* \overleftrightarrow{\partial}_\mu \Phi_i + e^2 A^2 \Phi_f^* \Phi_i \,. \tag{5.4}$$

From now on, we only consider the dominant $e^1$ term, neglecting the $e^2$ term, so that the $\mathcal{L}_{\text{int}}$ as the simple structure of a current $j_\mu$ interacting with the field $A^\mu$ :

$$\mathcal{L}_{\text{int}} = -j_\mu A^\mu \,, \tag{5.5}$$

## 5. From Lagrangians to cross-sections

where $j_\mu$ is just the matter current

$$j_\mu = -ie\Phi^* \overleftrightarrow{\partial}_\mu \Phi. \tag{5.6}$$

Let us consider now the matrix element $j_\mu^{fi} = \langle f|j_\mu|i\rangle$ of this current:

◇ an incoming (spinless) electron of moment $p_i$ has a wave function

$$\Phi_i(x) = N_i\, e^{-ip_i \cdot x} \tag{5.7}$$

◇ an outgoing (spinless) electron of moment $p_f$ has a wave function

$$\Phi_f^*(x) = N_f\, e^{+ip_f \cdot x}. \tag{5.8}$$

The normalization constants $N_i$ and $N_f$ in Eqs. (5.7) and (5.8) will be specified later. We thus have

$$j_\mu^{fi} = -eN_i\, N_f(p_i + p_f)_\mu\, e^{i(p_f - p_i)\cdot x}. \tag{5.9}$$

The perturbation theory developed in Chap. 3 was time-dependent, but we later on restricted ourselves to a time-independent potential. However, nothing prevent us to take into account the time dependence of the field $A^\mu$.

To understand how things are going, let us consider as an example the spinless $e^-\mu^-$ scattering[1], in the kinematics

$$e^-(p_A)\mu^-(p_B) \rightarrow e^-(p_C)\mu^-(p_D) \tag{5.10}$$

as illustrated in Fig. 5.1. The current $j_\mu^{(2)}$ is a source term for the $A_\mu$, which itself couples to the current $j_\mu^{(1)}$. To find this field $A^\mu$, one should just solve the Maxwell's equation

$$\Box^{(2)} A_\mu = j_\mu^{(2)} \tag{5.11}$$

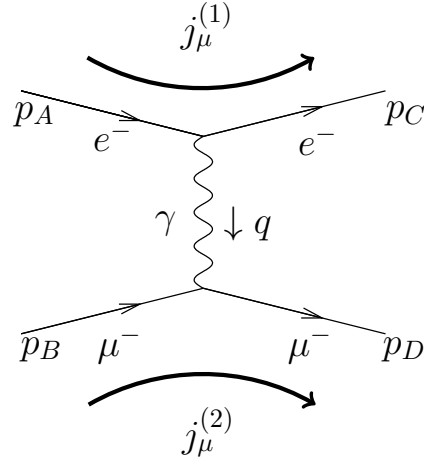for $A^\mu$, with, using Eq. (5.9),

$$j_\mu^{(2)} = -eN_B\, N_D(p_B + p_D)_\mu\, e^{i(p_D - p_B)\cdot x}. \tag{5.12}$$

Thus, since

$$\Box e^{i(p_D - p_B)\cdot x} = -(p_D - p_B)^2 e^{i(p_D - p_B)\cdot x}, \tag{5.13}$$

---

[1]We do not consider here the little bit more involved situation of $e^-e^-$ scattering, in which two contributions would occur, because the two particle are indistinguishable).

Figure 5.1.: $e^- \mu^-$ scattering.

we get

$$A^\mu = -\frac{1}{q^2} j_\mu^{(2)} \tag{5.14}$$

where $q = p_D - p_B$.

At lowest order, we get, after integration over $x$,

$$
\begin{aligned}
T_{fi} &= -i \int j_\mu^{(1)} \left(-\frac{1}{q^2}\right) j^{(2)\,\mu}(x)\, d^4x \tag{5.15} \\
&= i N_A N_B N_C N_D (2\pi)^4 \delta^{(4)}(p_A + p_B - p_C - p_D)\, \mathcal{M} \tag{5.16}
\end{aligned}
$$

where

$$
i\mathcal{M} = [ie(p_A + p_C)^\mu]\left[-i\frac{g^{\mu\nu}}{q^2}\right][ie(p_B + p_D)^\nu] \tag{5.17}
$$



$$\tag{5.18}$$

$$\text{vertex} \quad \bigg| \quad \text{photon} \quad \text{vertex} \tag{5.19}$$

$$\text{propagator} \tag{5.20}$$

Note: $q \neq 0$, therefore the photon which propagates is virtual, or off-mass shell.

## 5.2. Cross-sections

### 5.2.1. Normalization of a free particle wave-function

The wave-function of a particle of momentum $p$ reads

$$\Phi(x) = N\, e^{-ip\cdot x}. \tag{5.21}$$

We have shown in Chap. 2 that for such a plane wave, $\rho = 2E|N|^2$, see Eq. (2.55).

One should note the consistency and the Lorentz invariance under boosts. Indeed, $\rho\, d^3x$ is a Lorentz invariant, since

$$d^3x \ \overset{Boost}{\longrightarrow}\ \frac{d^3x}{\gamma} \tag{5.22}$$

$$2E \ \overset{Boost}{\longrightarrow}\ \gamma 2E\,. \tag{5.23}$$

In a given volume $V$, we normalize $\Phi$ such that there are $2E$ particles:

$$\int_V \rho\, d^3V = \int 2E|N|^2\, d^3V = 2E \tag{5.24}$$

i.e.

$$N = \frac{1}{\sqrt{V}}\,. \tag{5.25}$$

### 5.2.2. Transition rate per unit volume

Following Chap. 3, we introduce the transition rate per unit volume, for a process $A + B \rightarrow C + D$, as

$$W_{fi} = \frac{|T_{fi}|^2}{TV}\,. \tag{5.26}$$

Using the same trick which we used in Chap. 3 to pass from $T_{fi}$ to $|T_{fi}|^2$ in order to give a meaning to $[\delta^{(4)}(p_A + p_B - p_C - p_D)]^2$, one $\delta^{(4)}$ remains, and the other $(2\pi)^4\delta^{(4)}$ should be replaced by $TV$. Finally, one gets

$$W_{fi} = (2\pi)^4\frac{\delta^{(4)}(p_A + p_B - p_C - p_D)|\mathcal{M}|^2}{V^4}\,. \tag{5.27}$$

We define the concept of *cross-section* as

$$\text{cross-section} = \frac{W_{fi}}{\text{initial flux}} \times (\text{number of final states})\,. \tag{5.28}$$

The reason why one divides $W_{fi}$ by the initial flux, which is the density of incoming states is that one wants to normalize the result independently of any particular experimental setup (like the density of the target, of the beam).

### 5.2.3. Number of final states

In a volume $V$, with momenta inside $d^3p$, there are

$$\rho_V(p) = \frac{V d^3 p}{(2\pi)^3} \quad \text{states.} \tag{5.29}$$

This is a pure quantum mechanical effect. Consider first a particle in a one-dimensional box of length $L$. The boundary conditions implies, since the wave-function is of the form

$$\Phi \sim e^{ip_x x}, \tag{5.30}$$

that $p_x L = n2\pi$ with $n \in \mathbb{Z}$, and thus, passing from a sum over $n$ to an integration over $p_x$, the number of states in the range $p_x$ to $p_x + dp_x$ is

$$dn = \frac{dp_x \, L}{2\pi}, \tag{5.31}$$

from which the result (5.29) comes immediately after passing from the one-dimensional to the three-dimensional case, denoting $V = L^3$.

The number of final states/particle is thus

$$\frac{\rho_V(p)}{2E} = \frac{V d^3 p}{(2\pi)^3 2E}, \tag{5.32}$$

where we have normalized the wave-function such that there are $2E$ particles in the volume $V$.

For outgoing particles $C$ and $D$ scattered into $d^3 p_C$ and $d^3 p_D$, we thus have:

$$\text{number of final states} = \frac{V d^3 p_C}{(2\pi)^3 2E_C} \frac{V d^3 p_D}{(2\pi)^3 2E_D}. \tag{5.33}$$

### 5.2.4. Initial flux

In the lab frame, see Fig. 5.2, the number of $A$ particles which go through the surface of area $L^2$ during $t$ is
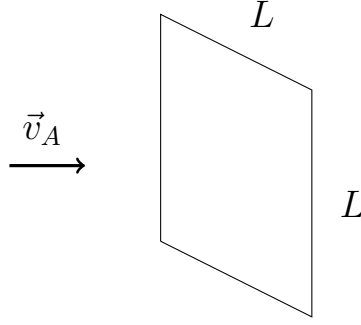
$$\frac{2E_A}{V} L^2 |\vec{v}_A| t. \tag{5.34}$$

Figure 5.2.: Evaluation of the initial flux.

Thus, the number of beam particles passing through a unit area per unit of time is

$$\frac{2E_A}{V}|\vec{v}_A| \, . \tag{5.35}$$

The number of target particles per unit volume is

$$\frac{2E_B}{V} \, . \tag{5.36}$$

We thus finally define the initial flux as

$$|\vec{v}_A|\frac{2E_A}{V}\frac{2E_B}{V} \, . \tag{5.37}$$

## 5.2.5. Cross-section

Putting all the factors together, we obtain

$$d\sigma = \frac{V^2}{|\vec{v}_A|2E_A\,2E_B}\frac{1}{V^4}|\mathcal{M}|^2(2\pi)^4\delta^{(4)}(p_A + p_B - p_C - p_D)\frac{V\,d^3p_C}{(2\pi)^32E_C}\frac{V\,d^3p_D}{(2\pi)^32E_D} \tag{5.38}$$

so that the $V$ dependence cancels, as expected. One can thus write the differential cross-section as

$$\boxed{d\sigma = \frac{|\mathcal{M}|^2}{F}\,d(PS)} \tag{5.39}$$

where the *phase-space* is given by

$$d(PS) = (2\pi)^4\delta^{(4)}(p_A + p_B - p_C - p_D)\frac{V\,d^3p_C}{(2\pi)^32E_C}\frac{V\,d^3p_D}{(2\pi)^32E_D} \tag{5.40}$$

which is Lorentz invariant, see tutorial, and $F$ is the incident flux,

$$F = |\vec{v}_A| 2E_A \, 2E_B \,. \tag{5.41}$$

*Exercise 5.1*

For an arbitrary collision along a given axis, show that

$$
\begin{aligned}
F &= |\vec{v}_A - \vec{v}_B| 2E_A \, 2E_B & (5.42) \\
&= 4 \left[ (p_A \cdot p_B)^2 - m_A^2 m_B^2 \right]^{1/2} = 2K, & (5.43)
\end{aligned}
$$

where we have introduced the standard notation $2K$ frequently used in the literature.

## 5.2.6. Cross-section in the center-of-mass system

A few technical steps are still needed before obtaining the expression of the differential cross-section in a compact form.

*Exercise 5.2*

In the center-of-mass system, introducing

$$
\begin{aligned}
|\vec{p}_A| &= |\vec{p}_A| = p_i^*, & (5.44) \\
|\vec{p}_C| &= |\vec{p}_D| = p_f^*, & (5.45)
\end{aligned}
$$

show that

$$F = 2K = 4p_i^* \, W^* \tag{5.46}$$

where

$$W^* = E_A^* + E_B^* \tag{5.47}$$

is the total center-of-mass energy.

*Exercise 5.3*

Introducing the solid angle $\Omega$ for the particle $C$, such that[2]

$$d^3 p_C = p_f^2 \, dp_f \, d^2\Omega \,, \tag{5.48}$$

---

[2]As usual, $d^2\Omega = d\varphi \sin\theta d\theta$ when expressed in terms of the usual spherical coordinates $(\theta, \varphi)$.

## 5. From Lagrangians to cross-sections

show that the phase-space can be expressed as

$$d(PS) = \frac{1}{4\pi^2} \frac{p_f^*}{4W^*} d^2\Omega \,. \tag{5.49}$$

Finally, the differential cross-section can be written as

$$\boxed{\left. \frac{d\sigma}{d^2\Omega} \right|_{\text{c.m.s}} = \frac{1}{64\pi^2 s} \frac{p_f^*}{p_i^*} |\mathcal{M}|^2 \,,} \tag{5.50}$$

where we have introduced the Mandelstam variable

$$s = (p_A + p_B)^2 = W^{*2} \,. \tag{5.51}$$

We refer to the tutorial for the explicit study of the $e^- e^- \to e^- e^-$ scattering.

# Appendices

# A. Elements of group theory

## A.1. Group

### A.1.1. Definitions

**Définition 1.1 :** *Group*

A group is a pair $(G, \cdot)$ made of a set $G$ and an operation acting on this set, which associates two a pair of elements $a$ and $b$ of $G$ an element $a \cdot b$.
This law should satisfy four axioms:

**Internal composition law:** $\forall a, b \in G,\ a \cdot b \in G$

**Associativity:** $\forall a, b, c \in G,\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$

**Existence of a neutral element:** $\exists\, e \in G / \forall a \in G, e \cdot a = a \cdot e = a.$

   The element $e$ is called the neutral element of the group.

**Existence of a symmetric:** $\forall a \in G,\ \exists\, \mathrm{sym}(a)$ such that
   $a \cdot \mathrm{sym}(a) = \mathrm{sym}(a) \cdot a = e.$

Depending on the context, the group law $\cdot$ can be denoted using various symbols:

◇ whenever the law is additive, it will be denoted as $+$, and the symmetric element of $a$ is named as its opposite, denoted as $-a$. The neutral element is then denoted as $0$.

Example: the $(\mathbb{Z}, +)$ group.

In practice, this is used only for abelian groups (see below).

◇ whenever the law is multiplicative, it will be denoted as $\times$, and the symmetric element of $a$ is named as its inverse, denoted as $a^{-1}$. The neutral element is then denoted as $1$.

Example: for $n \in \mathbb{N}^*$, the group made of the $n$-roots of unity, equipped with $\times$, is a multiplicative group.

⬦ one will also encounter the notations ∘ and ∗.

**Définition 1.2 :** *Abelian and non-abelian groups*

If the group law · is commutative, the group is named *abelian.*
In the opposite case, the group is named *non-abelian.*

**Définition 1.3 :** *Group order*

The *order* of a group $G$ is the cardinal of $G$.

⬦ If the cardinal is finite, $G$ is said to be a finite group (!) and its order is denoted as $|G|$.

⬦ When the group has an infinite order, the group can be *discrete* (topologically), or continuous.

## A.1.2. Examples of groups

⬦ The trivial group $G = (\{0\}, +)$, also denoted as 0, for an additive group (or $G = (\{1\}, \times)$, also denoted as 1, for a multiplicative group).

⬦ $(\mathbb{Z}, +)$ is a discrete abelian group.

⬦ $U(1) = \{e^{i\theta}/\theta \in \mathbb{R}\}$, the group of phases, is abelian.

⬦ If $K$ is a field (examples: $K = \mathbb{R}, \mathbb{C}$), $(K, +)$ and $(K^\star, \times)$ are abelian groups.

⬦ For $n \in \mathbb{N}^\star$, the set of integers modulo $n$ equipped with addition $G = (\mathbb{Z}/n\mathbb{Z}, +)$ is an abelian group, of order $n$.

⬦ $G = (\mathcal{S}(E), \circ)$ :

On a set $E$, the set $\mathcal{S}(E)$ of bijections from $E$ to $E$, equipped with the composition law ∘ for maps is a group.

⬦ *Symmetric group $\mathcal{S}_n$ :*

In the special case where $E = \{1, \cdots, n\}$, one denotes $\mathcal{S}_n$ as the set $\mathcal{S}(E)$, which is called the symmetric group of $n$ elements. Its order is $n!$, and it is non abelian for $n \geqslant 3$, made of permutations of the various $n$ elements $\{1, \cdots, n\}$ .

$\diamond$ *Linear group* $\mathrm{GL}_n(K)$ :

The set of invertible $n \times n$ matrices with coefficient belonging to a field $K$, equipped with matrix multiplication, is a group (non-abelian for $n \geqslant 2$).

$\diamond$ *Special Linear group* $\mathrm{SL}_n(K)$ :

The $n \times n$ matrices with coefficient belonging to a field $K$, equipped with matrix multiplication, of determinant 1 (therefore invertible), is a group (non-abelian for $n \geqslant 2$).

$\diamond$ *Orthogonal group* $O(n)$:

The set $O(n)$ of $n \times n$ orthogonal matrices is the set of real matrices satisfying $O \cdot O^t = O^t \cdot O = \mathbb{1}$, or equivalently, which leaves the bilinear form (the scalar product) $\vec{x} \cdot \vec{y} = \sum_{i=1}^{n} x_i \, y_i$ invariant, i.e. $(O\vec{x}) \cdot (O\vec{y}) = \vec{x} \cdot \vec{y}$.

Equipped with matrix multiplication, it is non-abelian group for $n \geqslant 2$. The defining relation of $O(n)$ obviously imposes that $\det O = \pm 1$.

$\diamond$ *Special orthogonal group* $SO(n)$:

Subset of matrices of $O(n)$ of determinant $+1$.

$\diamond$ *Unitary group* $U(n)$:

The set $U(n)$ of unitary matrices $n \times n$ is made of complex matrices satisfying $U \, U^\dagger = U^\dagger U = \mathbb{1}$, or equivalently, which leaves the sesquilinear form $(x, y) = \sum_{i=1}^{n} x_i^* \, y_i$ invariant (this form is left-sesquilinear: linear with respect to $y$, antilinear with respect to $x$, see bra-ket in quantum physics), i.e. $(Ux, Uy) = (x, y)$.

Equipped with matrix multiplication, it is a non-abelian group for $n \geqslant 2$. The defining relation for $U(n)$ imposes that $|\det U| = 1$.

$\diamond$ *Special unitary group* $(SU(n), \cdot)$:

Subset of $U(n)$ matrices of determinant $+1$.

$\diamond$ *Symplectic group* $Sp(2n, K)$ :

Consider a field $K$ (typically $\mathbb{R}$ or $\mathbb{C}$).

The set $Sp(2n, K)$ of symplectic matrices $2n \times 2n$ is made of matrices with coefficients belonging to $K$, which leaves the antisymmetric form

$x^t g\, y$ invariant, where $g$ is the antisymmetric matrix

$$g = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}$$

*i.e.* satisfying the condition

$$S^t g S = g\,.$$

# A.2. Morphisms and subgroups

## A.2.1. Maps, injections, surjections

**Définition 1.4 :** *Function (or map)*

A function $f : E \to F$ maps elements from $E$ to elements of $F$: any element of $E$ as a single image in $F$

**Définition 1.5 :** *Image of a function*

The image of a function $f : E \to F$ is the set of images of elements of $E$, denoted as $f(E)$, or Im $f$:

$$f(E) = \{f(x)/x \in E\}\,.$$

**Définition 1.6 :** *Direct image of a subset*

More generally, given a function $f : E \to F$, the image (or direct image) $f(A)$ of a subset $A \subset E$ is the set of images of elements of $A$:

$$f(A) = \{f(x)/x \in A\}\,.$$

**Définition 1.7 :** *Inverse image of a subset*

Given a function $f : E \to F$, the inverse image $f^{-1}(B)$ of a subset $B \subset F$ is the set of inverse image of elements of $B$:

$$f^{-1}(B) = \{x \in E/f(x) \in B\}\,.$$

**Définition 1.8 :** *Injective function (or injection)*

The function $f : E \to F$ is injective, or one-to-one, if each element of $F$ is mapped to by *at most* one element of the domain:

$$\forall x, x' \in E, f(x) = f(x') \Longrightarrow x = x',$$

or equivalently, if distinct elements of $E$ map to distinct elements in $F$:

$$\forall x, x' \in E, x \neq x' \Longrightarrow f(x) \neq f(x').$$

**Définition 1.9 :** *Surjective function (or surjection)*

The function $f : E \to F$ is surjective, or onto, if each element of $F$ is mapped to by at least one element of $E$. That is, the image $F(E)$ is equal to $F$:

$$\forall y \in F, \exists x \in E \,/ y = f(x).$$

**Définition 1.10 :** *Bijective function (or bijection)*

The function $f : E \to F$ is bijective if each element of $F$ is mapped to by exactly one element of $E$. In other words, the function is both injective and surjective:

$$\forall y \in F, \exists! \, x \in E \,/ y = f(x).$$

**Définition 1.11 :** *Inverse of a bijection*

To any bijection $f : E \to F$, one can associate the inverse function (which is also a bijection) denoted as $f^{-1}$, which maps any element $y \in F$ to its unique inverse image $x \in E$, i.e.

$$x = f^{-1}(y) \Leftrightarrow y = f(x).$$

Be aware that conventionally, the same notation $f^{-1}$ is used for both the inverse bijection and for the inverse image of a set.

## A.2.2. Morphism

**Définition 1.12 :** *Group morphism*

Let $(G, \star)$ et $(G', \cdot)$ be two groups. A morphism (or homomorphism) from $G$ to $G'$ is a map $f : G \to G'$ which satisfies

$$\forall g_1, \, g_2 \in G, \, f(g_1 \star g_2) = f(g_1) \cdot f(g_2). \tag{A.1}$$

In other words, it means that the map $f$ preserves the group law on $G$ and $G'$.

**Définition 1.13 :** *Group isomorphism*

If furthermore $f$ is bijective, it is easy to show that $f^{-1}$ is also a group morphism. Then, $f$ is called an isomorphism. In other words, the image of the symmetric of $g \in G$ is the symmetric of $f(g)$ in $G'$.

The two groups $G$ et $G'$ are said to be *isomorph*, a fact denoted as $G \simeq G'$ or $G \cong G'$.

**Définition 1.14 :** *Group automorphism*

If $G = G'$, then the group isomorphism $f$ is called an automorphism.

**Proposition 1.15 :** $\mathrm{Aut}(G)$

The set of automorphisms of a group $G$, equipped with the composition of maps $\circ$, is a group denoted as $\mathrm{Aut}(G)$.

**Définition 1.16 :** *Kernel of a morphism*

The kernel of a morphism $f$ is

$$\mathrm{Ker}\,(f) = \{g \in G / f(g) = e\}. \tag{A.2}$$

**Définition 1.17 :** *Injective, surjective morphism*

$\diamond$ A morphism $f$ is injective (or on-to-one) if and only if $\mathrm{Ker}\, f = e$.

$\diamond$ Besides, as we have seen above for the general case of a function, a morphism $f$ is surjective if and only if $\mathrm{Im}\, f = G'$.

$\diamond$ A morphism is an isomorphism if and only if it is both injective and surjective.

## A.2.3. Subgroup

**Définition 1.18 :** *Subgroup*

Let $H$ be a subset of $G$. A *subgroup* $(H, \cdot)$ of a group $(G, \cdot)$ is a group whose law $\cdot$ is obtained by restricting the group law $\cdot$ on $H \times H$.

**Proposition 1.19 :** For a subset $H$ of $G$ to be a subgroup of $G$, it is necessary and sufficient that $\forall a, b \in H$, $a \cdot b^{-1} \in H$.

*Remark :*
Note that the criterion 1.19 makes it possible to dispense with checking the associativity, automatically satisfied by inclusion of $H$ in $G$.

*Examples :*

⬦ Evidently, $G$ and $\{e\}$ are subgroups of $G$.

⬦ The intersection of an arbitrary family (finite or infinite) of subgroups of a group $G$ is a subgroup of $G$.

⬦ The union of two subgroups of $G$ is a subgroup if an only if one of the two is included in the other.

⬦ The subgroups of $\mathbb{Z}$ are the $n\,\mathbb{Z}$ for $n \in \mathbb{N}$.

## Propositions 1.20 :

⬦ The image and the kernel of a morphism $f : G \to G'$ are subgroups of $G'$ and $G$ respectively.

More generally:

⬦ The inverse image by $f$ of any subgroup of $G'$ is a subgroup of $G$.

⬦ The image by $f$ of any subgroup of $G$ is a subgroup of $G'$.

*Examples :*

⬦ Let $\mathbb{K}$ be a field. Then $\det : \mathrm{GL}_n(\mathbb{K}) \to \mathbb{K}^*$ is a morphism.

If $E$ is a $\mathbb{K}-$vector space of dimension $n$, then $\mathrm{GL}_n(\mathbb{K}, \cdot) \cong (\mathrm{GL}(E), \circ)$, the group of linear bijections from $E$ to $E$.

The kernel of det is thus a subgroup of $\mathrm{GL}_n(\mathbb{K})$, named the special linear group, denoted as $\mathrm{SL}_n(\mathbb{K})$ (it is thus the set of matrices of determinant 1, with coefficients in $\mathbb{K}$).

# B. Levi-Civita symbols and tensors

## B.1. Levi-Civita symbol in an euclidean space

### B.1.1. 2d

In two dimensions, one denotes

$$
\begin{aligned}
\varepsilon_{ij} &= +1 \quad \text{if} \quad (i,j) = (1,2) \\
&= -1 \quad \text{if} \quad (i,j) = (2,1) \\
&= \phantom{-}0 \quad \text{if} \quad i = j\,.
\end{aligned}
\tag{B.1}
$$

One readily checks that

$$
\boxed{\varepsilon_{ij}\,\varepsilon_{mn} = \delta_{im}\,\delta_{jn} - \delta_{in}\,\delta_{jm}\,.}
\tag{B.2}
$$

After one contraction over a pair of indexes, one gets

$$
\boxed{\varepsilon_{ij}\,\varepsilon_{in} = \delta_{jn}\,,}
\tag{B.3}
$$

since

$$
\delta_{ii}\delta_{jn} - \delta_{in}\delta_{ji} = 2\delta_{jn} - \delta_{jn} = \delta_{jn}\,.
\tag{B.4}
$$

Starting from Eq. (B.3), one more contraction gives

$$
\boxed{\epsilon_{ij}\,\epsilon_{ij} = 2\,.}
\tag{B.5}
$$

### B.1.2. 3d

In three dimensions, one denotes

$$
\begin{aligned}
\varepsilon_{ijk} &= +1 \quad \text{if} \quad (i,j,k) = \text{ even permutation of } (1,2,3) \\
&= -1 \quad \text{if} \quad (i,j,k) = \text{ odd permutation of } (1,2,3) \\
&= \phantom{-}0 \quad \text{ otherwise}\,.
\end{aligned}
\tag{B.6}
$$

## B. Levi-Civita symbols and tensors

One can show easily that

$$\varepsilon_{ijk}\varepsilon_{\ell mn} = \begin{vmatrix} \delta_{i\ell} & \delta_{im} & \delta_{in} \\ \delta_{j\ell} & \delta_{jm} & \delta_{jn} \\ \delta_{k\ell} & \delta_{km} & \delta_{kn} \end{vmatrix}. \tag{B.7}$$

Indeed, there are 6 terms, as can be seen from the fact in 3d, $i$, $j$ and $k$, as well as $\ell$, $m$ and $n$ can take the 3 different values $1, 2, 3$: thus $i$ should be equal to $\ell$, $m$ or $n$, which forces $j$ to be one of the two remaining indexes among $\ell$, $m$ or $n$, so that $k$ is the last one, i.e. $3 \times 2 \times 1 = 6$ choices. The first choice $i = \ell$, $j = m$, $k = n$, i.e $\delta_{i\ell}\delta_{jm}\delta_{kn}$, gives, without sum, $\varepsilon_{ijk}^2 = +1$, while the 5 other terms are simply obtained by permutation of indexes, accounting for the sign provided by the product of the two $\varepsilon$.

One contraction gives

$$\varepsilon_{ijk}\,\varepsilon_{imn} = \delta_{jm}\,\delta_{kn} - \delta_{jn}\,\delta_{km}\,. \tag{B.8}$$

One more contraction leads to

$$\varepsilon_{ijk}\,\varepsilon_{ijn} = 2\delta_{kn}\,, \tag{B.9}$$

since

$$\delta_{jj}\delta_{kn} - \delta_{jn}\delta_{kj} = 3\delta_{kn} - \delta_{kn} = 2\delta_{kn}\,. \tag{B.10}$$

Finally,

$$\varepsilon_{ijk}\,\varepsilon_{ijk} = 3! = 6\,. \tag{B.11}$$

## B.1.3. General case

In dimension $n$, one denotes

$$\begin{aligned} \varepsilon_{i_1 i_2 \cdots i_n} &= +1 \quad \text{if} \quad (i_1, i_2, \cdots, i_n) = \text{ even permutation of } (1, 2, \cdots, n) \\ &= -1 \quad \text{if} \quad (i_1, i_2, \cdots, i_n) = \text{ odd permutation of } (1, 2, \cdots, n) \\ &= \quad 0 \quad \text{otherwise}\,. \end{aligned} \tag{B.12}$$

Consider a $n \times n$ matrix $A$, which has matrix elements $a_{ij}$. Its determinant is given by

$$\det A = \varepsilon_{i_1 \cdots i_n}\, a_{1i_1} \cdots a_{ni_n} = \frac{1}{n!}\, \varepsilon_{i_1 \cdots i_n}\, \varepsilon_{j_1 \cdots j_n}\, a_{i_1 j_1} \cdots a_{i_n j_n} \tag{B.13}$$

and

$$a_{i_1 j_1} \cdots a_{i_n j_n} \, \varepsilon_{i_1 \cdots i_n} = (\det A) \, \varepsilon_{j_1 \cdots j_n} \,, \tag{B.14}$$

which can be understood as the way the Levi-Civita symbol transforms under an arbitrary linear transformation.

## B.1.4. Transformation under $O(n)$ and $SO(n)$

We recall that orthogonal transformations, elements of the group $O(n)$,

$$\mathbb{R}^n \;\rightarrow\; \mathbb{R}^n \tag{B.15}$$
$$x \;\mapsto\; O\,x \text{ with } x^i = O^i{}_j x^j \tag{B.16}$$

are characterized by the constraint

$$O\,O^t = O^t\,O = \mathbb{1} \,. \tag{B.17}$$

Taking the determinant of both sides, one sees that $\det O = \pm 1$. The case $\det O = 1$ defines the subgroup $SO(n)$, named special orthogonal group, also named rotation group, by extension of the group of rotations in $n = 2$ dimension or $n = 3$ dimension.

Inserting this inside the general transformation (B.14) of the Levi-Civita symbol, we see that[1]

$$O_{i_1 j_1} \cdots O_{i_n j_n} \, \varepsilon_{i_1 \cdots i_n} = (\det O) \, \varepsilon_{j_1 \cdots j_n} \,, \;\; \text{with } \det O = \pm 1. \tag{B.18}$$

The Levi-Civita symbol is thus a *pseudo-tensor*, since it naturally extends what appends in 3d:

⋄ under a rotation, i.e. a special orthogonal transformation, of determinant $+1$, the Levi-Civita is unchanged.

⋄ under an orthogonal transformation of determinant $-1$, for example a reflection in an odd number of dimensions (like the familiar 3d space of euclidean geometry), named parity $P$, it should acquire a minus sign if it were a tensor, made for example as the direct product of the component of a vector $V^i$, as $V^{i_1} V^{i_2} \cdots V^{i_n}$. In the present case, it remains invariant, thus the name of pseudo-tensor.

---

[1]We are working here in the euclidean space $\mathbb{R}^n$, with a trivial metric given by the identity matrix, therefore the up or down position of indexes is arbitrary.

# B.2. Levi-Civita tensor in 4d Minkowski space

We now consider the Minkowski space, with the usual metric $g_{\mu\nu}$.

The 4-contravariant Levi-Civita tensor is defined as

$$
\begin{aligned}
\varepsilon^{\mu\nu\rho\sigma} &= +1 \quad \text{if} \quad (\mu,\nu,\rho,\sigma) = \text{ even permutation of } (0,1,2,3) \\
&= -1 \quad \text{if} \quad (\mu\nu\rho\sigma) = \text{ odd permutation of } (0,1,2,3) \\
&= \ \ 0 \quad \ \text{otherwise}.
\end{aligned}
\tag{B.19}
$$

Using as usual the tensor metric $g_{\mu\nu}$ and its inverse, one can descend or raise any index.

Since $\det g = -1$, one has

$$
\varepsilon_{\mu\nu\rho\sigma} = -\varepsilon^{\mu\nu\rho\sigma}.
\tag{B.20}
$$

One should be careful with the fact that the normalization of $\varepsilon^{\mu\nu\rho\sigma}$ is conventional. In some textbook/articles, it is defined with an opposite sign. We use here the convention of Jackson, Itzykson-Zuber and Peskin-Schroeder.

## B.2.1. Contraction identities

The following successive contraction identities are satisfied.

$$
\varepsilon^{\mu\nu\rho\sigma}\,\varepsilon^{\mu'\nu'\rho'\sigma'} = -\det g^{\alpha\alpha'} \quad
\begin{aligned}
\alpha &= \mu,\nu,\rho,\sigma &\quad \text{row} \\
\alpha' &= \mu',\nu',\rho',\sigma' &\quad \text{column}
\end{aligned}
\tag{B.21}
$$

$$
\varepsilon^{\mu\nu\rho\sigma}\,\varepsilon_\mu{}^{\nu'\rho'\sigma'} = -\det g^{\alpha\alpha'} \quad
\begin{aligned}
\alpha &= \nu,\rho,\sigma &\quad \text{row} \\
\alpha' &= \nu',\rho',\sigma' &\quad \text{column}
\end{aligned}
\tag{B.22}
$$

$$
\varepsilon^{\mu\nu\rho\sigma}\,\varepsilon_{\mu\nu}{}^{\rho'\sigma'} = -2\left(g^{\rho\rho'}g^{\sigma\sigma'} - g^{\rho\sigma'}g^{\rho'\sigma}\right).
\tag{B.23}
$$

$$
\varepsilon^{\mu\nu\rho\sigma}\,\varepsilon_{\mu\nu\rho}{}^{\sigma'} = -6g^{\sigma\sigma'}.
\tag{B.24}
$$

$$
\varepsilon^{\mu\nu\rho\sigma}\,\varepsilon_{\mu\nu\rho\sigma} = -4! = -24.
\tag{B.25}
$$

## B.2.2. Relationship with the Levi-Civita symbol

$$\boxed{\varepsilon^{0ijk} = \varepsilon_{ijk} = -\varepsilon_{0ijk}\,.} \tag{B.26}$$

## B.2.3. Behavior under Lorentz transformations

Under Lorentz transformations, the Levi-Civita tensor transforms as

$$\varepsilon^{\mu\nu\rho\sigma} \mapsto \Lambda^{\mu}{}_{\mu'}\,\Lambda^{\nu}{}_{\nu'}\,\Lambda^{\rho}{}_{\rho'}\,\Lambda^{\sigma}{}_{\sigma'}\varepsilon^{\mu'\nu'\rho'\sigma'} = \det\Lambda\,\varepsilon^{\mu\nu\rho\sigma}\,. \tag{B.27}$$

Thus:

⋄ under proper transformations ($\mathcal{L}_+ : \det\Lambda = +1$), $\varepsilon^{\mu\nu\rho\sigma}$ is invariant. This includes boosts and rotations.

⋄ under improper transformations ($\mathcal{L}_- : \det\Lambda = -1$), $\varepsilon^{\mu\nu\rho\sigma}$ is reversed. This includes $T$ and $P$ separately.

In conclusion,

$$\boxed{\varepsilon^{\mu\nu\rho\sigma} \text{ is a pseudo-tensor under Lorentz transformations.}} \tag{B.28}$$

# C. Dimensional analysis in high-energy physics

## C.1. Natural units in high-energy physics

In high-energy physics, it is customary to choose a system of units such that

$$c = 1, \tag{C.1}$$
$$\hbar = 1. \tag{C.2}$$

Since $[c] = L/T$, taking $c = 1$ means that $[L] = [T]$.

Besides, we know that $[\hbar] = ML^2T^{-1}$ (e.g. starting from $E = h\nu$). In the system of units in which $c = 1$, we thus have $[\hbar] = ML$ and taking $\hbar = 1$ leads to $M = L^{-1}$. Finally, since $[P] = M[c] = M$, we have the following rules for dimensional analysis:

$$\boxed{[P] = [E] = M = \frac{1}{L} = \frac{1}{T}.} \tag{C.3}$$

Since

$$\hbar \times c = 1.055 \times 10^{-34}\,\text{J} \cdot \text{s} \times 2.998 \times 10^8\,\text{m/s} = 3.161 \times 10^{-26}\text{J} \cdot \text{s} \cdot \text{m}$$
$$= 1.973 \times 10^{-7}\text{eV} \cdot \text{m}$$

it means that in natural units,

$$\boxed{1\,\text{fm} \sim 1/(200\,\text{MeV}).} \tag{C.4}$$

Similarly, $c = 1$ implies that

$$\boxed{1\,\text{fm} \sim 3.3 \times 10^{-24}\,\text{s},} \tag{C.5}$$

and from $1\,\text{J} = 1\,\text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2}$ one gets

$$1\,\text{GeV} \sim 1.6 \times 10^{-10}\text{J} = 1.6 \times 10^{-10}(3 \times 10^8)^{-2}$$

i.e.

$$\boxed{1\,\text{GeV} \sim 1.8 \times 10^{-27}\text{s}.} \tag{C.6}$$

## C.2. Dimensional analysis in field theory

Let us see how this applies to field theory. Since action as dimension of $\hbar$, it means that it is dimensionless in natural units. Now, from the fact that $S = \int d^4x \, \mathcal{L}$, it means that

$$[\mathcal{L}] = M^4. \tag{C.7}$$

The structure of the kinetic term, imposed by the fact that it should be quadratic in $\partial_\mu \phi$ (presence of a time derivative and Lorentz covariance, with thus, for a massless field,

$$\mathcal{L}_{\text{kin}} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) \tag{C.8}$$

thus leads to

$$\boxed{[\phi] = M \, .} \tag{C.9}$$

Similarly, from the Lagrangian

$$\mathcal{L}_{\text{em}} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \tag{C.10}$$

we deduce that a photon field has a dimension

$$\boxed{[A^\mu] = M \, .} \tag{C.11}$$