

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

Classical field theory and relativistic quantum mechanics

Relatore:
Prof. Fiorenzo Bastianelli

Presentata da:
Giancarlo Oancia

Anno Accademico 2022/2023

«Zwei Dinge erfüllen das Gemüt mit immer neuer und zunehmender Bewunderung und Ehrfurcht, je öfter und anhaltender sich das Nachdenken damit beschäftigt: der bestirnte Himmel über mir und das moralische Gesetz in mir.»

–Kant, Kritik der praktischen Vernunft.

«Due cose riempiono l’animo di ammirazione e venerazione sempre nuova e crescente, quanto più spesso e più a lungo la riflessione si occupa di esse: il cielo stellato sopra di me, e la legge morale in me.»

–Kant, Critica della ragion pratica.

ABSTRACT

Nowadays, the best theoretical framework we have to describe elementary particles' physics is the Standard Model, whose main language consists of quantum field theory. Historically, before its formulation, it has been attempted to make Schrödinger's quantum theory relativistic, in the framework of the so-called first quantization. The aim of this thesis is to make the reader aware of the problems of this procedure, necessary condition to understand the need to change paradigm and develop a new theory, known as second quantization.

Since, in this context, the new fundamental physical entity is the quantum field, we shall introduce field theory, starting from the classical description of electromagnetism, making use of a formalism to make Maxwell's equations manifestly covariant.

To obtain the latter directly from an action and to use the Lagrangian mechanics' tools, it's necessary to generalize the latter to a system with an infinite number of degrees of freedom. This will be achieved initially by discretizing the space and applying the known formalism into any elementary cell and, later, through a variational principle.

Furthermore, we'll try to apply quantum mechanics to a relativistic particle, obtaining the Klein-Gordon equation, which will be interpreted as representing a field whose quantum is a massive particle without spin. We'll notice how, forcing a particular global symmetry of this equation to be locally valid, it'll be necessary to add some terms on the Lagrangian which can be interpreted as an interaction with the electromagnetic field. This allows us to introduce Gauge's principle, which is a fundamental tool to describe interactions in the Standard Model. Finally, this principle will be critically analyzed, leading to the conclusion that it's not correct to distinguish between the object and the mediator of an interaction.

SOMMARIO

Il migliore quadro teorico che abbiamo attualmente a disposizione per descrivere la fisica delle particelle elementari è il Modello Standard, il cui linguaggio principale consiste nella teoria di campo quantizzato. Storicamente, prima della sua formulazione, si è provato a rendere relativistica la teoria quantistica di Schrödinger, nel contesto della così detta prima quantizzazione. Questa tesi si propone di far comprendere al lettore le problematiche di tale procedimento, condizione necessaria per capire l'esigenza di cambiare paradigma e di sviluppare una nuova teoria, nota come seconda quantizzazione.

Poiché, in questo contesto, il nuovo ente fisico fondamentale è il campo quantizzato, dovremo introdurre la teoria dei campi, partendo dalla descrizione classica dell'elettromagnetismo, servendoci di un formalismo che renda le equazioni di Maxwell manifestamente covarianti.

Per ricavare queste ultime direttamente da un'azione e usare gli strumenti della meccanica Lagrangiana, è necessario generalizzare quest'ultima al caso di un sistema con un numero infinito di gradi di libertà. Questo verrà realizzato dapprima discretizzando lo spazio e applicando il formalismo noto all'interno di ogni cella elementare e, successivamente, attraverso un principio variazionale.

Inoltre, si proverà ad applicare la meccanica quantistica a una particella relativistica, ottenendo l'equazione di Klein-Gordon, che verrà interpretata come descrivente un campo il cui quanto sia una particella massiva priva di spin. Noteremo come, imponendo che una particolare simmetria globale di questa equazione valga localmente, sarà necessario introdurre dei termini nella Lagrangiana che possono essere interpretati come un'interazione con il campo elettromagnetico. Questo permetterà d'introdurre il principio di Gauge, il quale è uno strumento fondamentale per descrivere le interazioni nel Modello Standard. Questo principio, infine, verrà analizzato criticamente, portando alla conclusione che non sia corretto distinguere tra oggetto e mediatore dell'interazione.

CONTENTS

1	Calculus of Variations	1
1.1	Introduction to variational calculus	1
1.2	Euler-Lagrange equation	3
1.3	Functional derivative	5
2	Classical Particle Theory	11
2.1	Lagrangian formalism	11
2.2	Special relativity and tensor notation	12
2.3	Charge and current densities	19
2.4	Electromagnetism	22
2.5	Gauge invariance	25
2.6	Solution for Maxwell's equations	26
3	Classical Field Theory	35
3.1	A first model: the rod	35
3.2	From discrete to continuous systems	38
3.3	Least action principle	41
3.4	Noether's theorem	44
3.5	Scalar fields and Klein-Gordon Lagrangian	46
3.6	Lagrangian for electromagnetic field	48
4	Relativistic quantum mechanics	53
4.1	Quantum mechanics in a nutshell	54
4.2	From Schrödinger to Klein-Gordon equation	57
4.3	Scalar electrodynamics	64
4.4	Gauge invariance as a principle	67

Appendix A Mathematical Appendix	71
A.1 Fourier transform	71
A.2 Dirac's delta	72
A.3 Residue theorem	73
References	77

CALCULUS OF VARIATIONS

Since throughout this whole thesis we'll refer to Euler-Lagrange equation, it is appropriate to begin with some basic concepts about variational calculus. Its purpose is to maximize or minimize functionals, namely, operators that map functions to real numbers. We'll use this mathematical tool to extract a particle's equations of motion from its Lagrangian, and then, using the functional derivative, we'll seek to generalize the formalism to continuous fields. Therefore, an introduction to variational derivatives is necessary.

1.1 Introduction to variational calculus

Variational calculus concerns with functionals, so we should start from there.

Definition 1 (Functional). A *functional* is an operator that maps functions (or curves) to real numbers.

In particular, we are interested in functionals which can be cast in the form

$$J[y] = \int_a^b dx f(x, y(x), y'(x)), \quad (1.1)$$

subject to the boundary conditions

$$y(a) = y_a, \quad y(b) = y_b. \quad (1.2)$$

In order to chase a rigorous mathematical approach, we should define the space in which y and f are defined, but since it is beyond our purposes, we should assume that $y(x) \in C^2(a, b)$, i.e., the space of functions which have continuous first and second derivative with respect to x , and $f(x, y(x), y'(x))$ is a function with continuous

first and second partial derivatives with respect to all its arguments. These conditions will ensure that all the quantities we'll deal with have right behaviour.

From now on, we'll deal with differentiable functionals, that is

Definition 2 (Differentiable). Let $J[y]$ be a functional defined on some normed space with norm $\|\cdot\|$. It's said to be *differentiable* if

$$\Delta J[y; h] \equiv J[y + h] - J[y] = \phi[h] + \varepsilon\|h\|, \quad (1.3)$$

where $\phi[h]$ depends linearly on h and $\|h\| \rightarrow 0$ as $\varepsilon \rightarrow 0$. $h = h(x)$ is the increment of the variable $y = y(x)$. The linear part of the increment ΔJ , $\phi[h]$, is called the *variation* of the functional and is denoted by $\delta J[y; h]$.

Leaving aside further technicalities about the normed space in which to define the functional and the uniqueness of the variation, let's move on defining what we mean for extremum, or rather, what is a maximum or a minimum for the functional considered.

Definition 3 (Extremum). Recalling the definition 2, a curve $y = y(x)$ is called an *extremum* of a differentiable functional $J[y]$ if $\delta J[h] = 0$ for every admissible h .

Furthermore, we shall make use of the following lemma.

Lemma L.1 (Fundamental lemma of the calculus of variations). If $M(x) \in C(a, b)$ and if

$$\int_a^b dx M(x)\eta(x) = 0$$

for every $\eta(x) \in C^1(a, b)$ such that

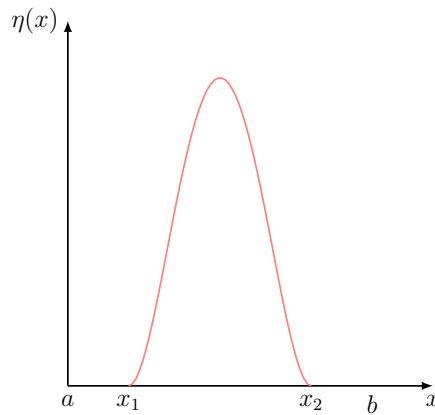
$$\eta(a) = \eta(b) = 0,$$

then

$$M(x) = 0$$

for all $x \in [a, b]$.

Proof. Let's tackle the problem by contradiction. Suppose that $M(x)$ is nonzero at some point in (a, b) . Without any loss of generality, let's suppose it is strictly positive. Then, by continuity, it is also positive in some interval $[x_1, x_2] \subset [a, b]$.

Figure 1.1 Definition of $\eta(x)$.

If we set, as depicted in fig. 1.1,

$$\eta(x) = \begin{cases} (x - x_1)^2(x - x_2)^2 & \text{if } x \in [x_1, x_2], \\ 0 & \text{otherwise,} \end{cases}$$

then $\eta(x)$ obviously satisfies the hypothesis of the lemma. However,

$$\int_a^b dx M(x)\eta(x) = \int_{x_1}^{x_2} dx M(x)(x - x_1)^2(x - x_2)^2 > 0,$$

since the integrand is positive. This contradicts our starting hypothesis, therefore it must be

$$M(x) = 0, \quad x \in (a, b),$$

and the continuity of the function guarantees it vanishes at a and b too. ■

1.2 Euler-Lagrange equation

Let's now consider the variational problem of our interest, that is, we want to find an extremum of the functional (1.1) with the conditions (1.2). We are able to prove that the extremum curve must satisfy the *Euler-Lagrange* equation.

Theorem T.1. *The curve $y = y(x) \in C^2(a, b)$ is an extremum of the functional*

$$J[y] = \int_a^b dx f(x, y(x), y'(x))$$

on the space of curves passing through the points $y(a) = y_a$ and $y(b) = y_b$, precisely when

$$\frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) - \frac{\partial f}{\partial y} = 0 \quad \text{along the curve } y(x). \quad (1.4)$$

Proof. Suppose we give $y(x)$ an increment $h(x)$, where, in order for the function

$$y(x) + h(x) \quad (1.5)$$

to continue to satisfy the boundary conditions (1.2), it must be

$$h(a) = h(b) = 0. \quad (1.6)$$

Then, since the corresponding increment of the functional equals

$$\Delta J = J[y + h] - J[y] = \int_a^b dx \left[f(x, y + h, y' + h') - f(x, y, y') \right],$$

it follows by Taylor's theorem that

$$\Delta J = \int_a^b dx \left[\frac{\partial f}{\partial y}(x, y, y')h + \frac{\partial f}{\partial y'}(x, y, y')h' \right] + O(h^2).$$

According to the definition 2, the variation of $J[y]$ is

$$\delta J = \int_a^b dx \left[\frac{\partial f}{\partial y}(x, y, y')h + \frac{\partial f}{\partial y'}(x, y, y')h' \right],$$

and it must be zero for all admissible h for y to be an extremum.

Integrating by parts

$$\int_a^b dx \frac{\partial f}{\partial y'} h' = \left(h \frac{\partial f}{\partial y'} \right) \Big|_a^b - \int_a^b dx h \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right)$$

and using the condition (1.6), we can write

$$\delta J = \int_a^b dx \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) h = 0, \quad \forall h$$

and finally, using the lemma L.1 we obtain the Euler-Lagrange differential equation

$$\frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) - \frac{\partial f}{\partial y} = 0. \quad \blacksquare$$

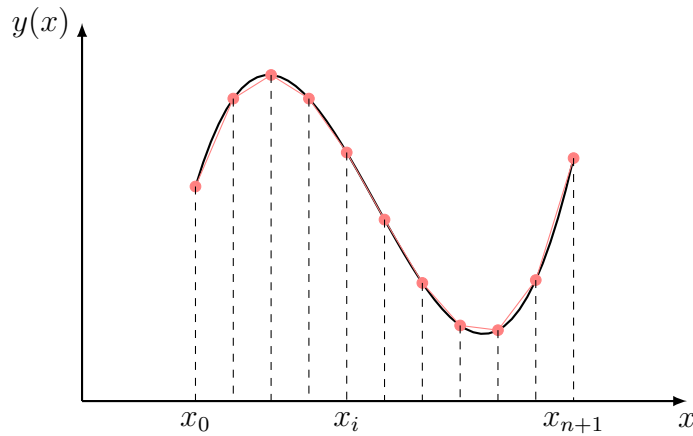


Figure 1.2 Subintervals.

Remark. To visualize the order of Taylor's expansion we may write

$$h(x) = \varepsilon \eta(x), \quad (1.7)$$

where ε is an infinitesimal parameter and $\eta(x)$ an arbitrary function.

Further, since we are interested in first order variations, neglecting terms smaller than ε in our Taylor's expansions, from now on we shall be slightly imprecise and write

$$\delta J \simeq J[y + h] - J[y] = J[y + \varepsilon \eta] - J[y].$$

1.3 Functional derivative

Discrete limit of a functional. Sometimes it is convenient to think about functionals as the limit of a suitable n -variables function. We shall apply this concept to introduce the functional derivative and look for a generalization.

Let's consider again the functional (1.1) subject to the conditions (1.2) and divide the interval $[a, b]$, in which $y(x)$ is defined, in $n + 1$ subintervals delimited by the points

$$a = x_0, x_1, \dots, x_n, x_{n+1} = b,$$

as shown in fig. 1.2.

Replacing the curve $y = y(x)$ by the polygonal line with vertices

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n), (x_{n+1}, y_{n+1}),$$

where $y_i \equiv y(x_i)$, it is possible to approximate the functional (1.1) by the sum

$$J(y_1, \dots, y_n) \equiv \sum_{i=0}^n f\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right) \Delta x, \quad (1.8)$$

where $\Delta x = x_{i+1} - x_i = (b - a)/(n + 1)$. Since $y_0 = a$ and $y_{n+1} = b$ are fixed, eq. (1.8) is a function of n -variables, namely y_1, \dots, y_n .

Let's take the derivative with respect to a particular y_k

$$\frac{\partial J(y_1, \dots, y_n)}{\partial y_k}$$

and evaluate the limit as $n \rightarrow \infty$.

Since y_k appears in only two terms of (1.8), corresponding to $i = k$ and $i = k - 1$, we find that

$$\begin{aligned} \frac{\partial J}{\partial y_k} &= \frac{\partial f}{\partial y}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right) \Delta x \\ &+ \frac{\partial f}{\partial y'}\left(x_{k-1}, y_{k-1}, \frac{y_k - y_{k-1}}{\Delta x}\right) - \frac{\partial f}{\partial y'}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right). \end{aligned} \quad (1.9)$$

We divide both sides by Δx , in order to avoid a nontrivial result,

$$\begin{aligned} \frac{1}{\Delta x} \frac{\partial J}{\partial y_k} &= \frac{\partial f}{\partial y}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right) \\ &- \frac{1}{\Delta x} \left[\frac{\partial f}{\partial y'}\left(x_{k-1}, y_{k-1}, \frac{y_k - y_{k-1}}{\Delta x}\right) + \frac{\partial f}{\partial y'}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right) \right]. \end{aligned} \quad (1.10)$$

As $\Delta x \rightarrow 0$, the expression (1.10) converges to

$$\frac{\delta J}{\delta y} \equiv \frac{\partial f}{\partial y}(x, y, y') - \frac{d}{dx} \frac{\partial f}{\partial y'}(x, y, y'), \quad (1.11)$$

called *functional* or *variational derivative* of the functional (1.1).

Recalling the Euler-Lagrange equation (1.4), it guarantees that the variational derivative of the functional under consideration vanishes at every point.

Differential in multivariate calculus. From what we've seen until now, we could infer a similarity between the extremum of a function and that of a functional. Indeed, in both cases the respective derivative must vanish. Since the intention is studying a functional about its extremum, it's useful to carry on this similarity

to define the functional derivative in a more general way. Before that, let's briefly review some aspects of the differential calculus for standard multivariate functions.

Let's consider a function $f : A \rightarrow \mathbb{R}$, where $A \subset \mathbb{R}^n$. Given an arbitrary increment $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{h}$, f is differentiable if exists a vector $\mathbf{m} \in \mathbb{R}^n$ such that

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \langle \mathbf{m}, \mathbf{h} \rangle + o(|\mathbf{h}|), \quad |\mathbf{h}| \rightarrow 0, \quad (1.12)$$

where $\langle \cdot, \cdot \rangle$ is the usual euclidean scalar product and $|\cdot|$ the usual norm. The linear term $\langle \mathbf{m}, \mathbf{h} \rangle$ is called differential of the function and is expressed by df . Setting $\mathbf{h} = |\mathbf{h}|\mathbf{h}_1$ and $\mathbf{e}_k = (0, \dots, 1, \dots, 0)$, where \mathbf{h}_1 is the versor of \mathbf{h} and $\mathbf{e}_1, \dots, \mathbf{e}_n$ the canonical base of \mathbb{R}^n , the directional derivative with respect to $\boldsymbol{\lambda}_1$, where $|\boldsymbol{\lambda}_1| = 1$, is defined as

$$\frac{\partial f}{\partial \boldsymbol{\lambda}_1} \equiv \lim_{\varepsilon \rightarrow 0} \frac{f(\mathbf{x} + \varepsilon \boldsymbol{\lambda}_1) - f(\mathbf{x})}{\varepsilon}. \quad (1.13)$$

If we set $\boldsymbol{\lambda}_1 = \mathbf{e}_i$, we obtain the i -th component of the gradient. We can now prove that

$$\mathbf{m} = \nabla f(\mathbf{x}). \quad (1.14)$$

Proof. Using the definition (1.13) with $\boldsymbol{\lambda}_1 = \mathbf{e}_k$ and by (1.12), we have

$$\frac{\partial f}{\partial x_k}(\mathbf{x}) = \lim_{\varepsilon \rightarrow 0} \frac{f(\mathbf{x} + \varepsilon \mathbf{e}_k) - f(\mathbf{x})}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{\langle \mathbf{m}, \varepsilon \mathbf{e}_k \rangle + o(|\varepsilon|)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} (m_k + o(1)) = m_k,$$

showing (1.14). ■

Another property of differential functions is

$$\frac{\partial f}{\partial \boldsymbol{\lambda}_1}(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \boldsymbol{\lambda}_1 \rangle. \quad (1.15)$$

Proof. Using (1.13), (1.12) and (1.14),

$$\begin{aligned} \frac{\partial f}{\partial \boldsymbol{\lambda}_1}(\mathbf{x}) &\equiv \lim_{\varepsilon \rightarrow 0} \frac{f(\mathbf{x} + \varepsilon \boldsymbol{\lambda}_1) - f(\mathbf{x})}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{\langle \mathbf{m}, \varepsilon \boldsymbol{\lambda}_1 \rangle + o(|\varepsilon|)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} (\langle \nabla f(\mathbf{x}), \boldsymbol{\lambda}_1 \rangle + o(1)) = \langle \nabla f(\mathbf{x}), \boldsymbol{\lambda}_1 \rangle, \end{aligned}$$

showing (1.15). ■

Combining (1.12), (1.14) and (1.15), for a differentiable function we can write

$$df = \langle \mathbf{m}, \mathbf{h} \rangle = \sum_{k=1}^n \frac{\partial f}{\partial x_k} h_k = |\mathbf{h}| \frac{\partial f}{\partial \mathbf{h}_1}(\mathbf{x}). \quad (1.16)$$

Functional derivative's general definition As we've seen, a functional $J[y]$ can be thought as a function of infinitely many variables $y(x)$. Therefore, we expect a natural generalization for (1.16). Let's try to guess its expression and an operative method to compute it. Comparing (1.3) and (1.12), we can notice that the variation of our functional is akin to the differential of a function, so an identity similar to (1.16) should hold for δJ . It's natural to replace the euclidean inner product with that of L^2 , i.e., the space of square-integrable functions. Considering a variation of the argument $y(x) \rightarrow y(x) + h(x)$, this would lead to

$$\delta J = \int dx \frac{\delta J}{\delta y} h(x), \quad (1.17)$$

that turns out to be a coherent definition for the functional derivative $\delta J/\delta y$. Therefore, the latter is akin to the gradient of a multivariate function, and it expresses how much the varied functional $J[y + h]$ differs from the original one, that is, $J[y]$. Writing the variation as in (1.7), namely $y(x) \rightarrow y(x) + \varepsilon\eta(x)$, from (1.16) we can identify $\int dx \frac{\delta J}{\delta y} \eta(x)$ with the directional derivative $\frac{\partial f}{\partial \mathbf{h}_1}(\mathbf{x})$, defined by (1.13). This leads to

$$\int dx \frac{\delta J}{\delta y} \eta(x) = \lim_{\varepsilon \rightarrow 0} \frac{J[y(x) + \varepsilon\eta(x)] - J[y(x)]}{\varepsilon} = \frac{d}{d\varepsilon} \left[J[y(x) + \varepsilon\eta(x)] \right]_{\varepsilon=0}, \quad (1.18)$$

that, again, turns out to be correct and provides a method to compute the variational derivative. It can be shown that ordinary derivatives' properties also apply for variational ones. For example, given two differentiable functionals $J[y]$, $K[y]$, with $y(x)$ and $z(u)$ two differentiable functions, and α and β two constants, we have

$$\frac{\alpha J + \beta K}{\delta y(x)} = \alpha \frac{\delta J}{\delta y} + \beta \frac{\delta K}{\delta y}, \quad (1.19a)$$

$$\frac{\delta JK}{\delta y} = \frac{\delta J}{\delta y} K + J \frac{\delta K}{\delta y}, \quad (1.19b)$$

$$\frac{\delta J[z(u)]}{\delta y(x)} = \frac{\delta J[z(u)]}{\delta z(y(x))} \frac{dz(u(x))}{dy(x)}. \quad (1.19c)$$

Return to Euler-Lagrange equation. Let's apply the new formalism to the functional (1.1) with conditions (1.2), to recover the Euler-Lagrange equation (1.4). It can be said that, in order for $y(x)$ to be an extremum of the functional $J[y]$, the variation δJ must vanish. As we shall see, in mechanics this is known as *least action principle*. Indeed, considering a variation $y(x) \rightarrow y(x) + h(x) = y(x) + \varepsilon\eta(x)$, by (1.17)

and (1.18) we have

$$\begin{aligned}
\delta J &= \int dx \frac{\delta J}{\delta y} h(x) = \varepsilon \int dx \frac{\delta J}{\delta y} \eta(x) = \varepsilon \lim_{\varepsilon \rightarrow 0} \frac{J[y(x) + \varepsilon \eta(x)] - J[y(x)]}{\varepsilon} \\
&= \varepsilon \frac{d}{d\varepsilon} \left[J[y(x) + \varepsilon \eta(x)] \right]_{\varepsilon=0} = \varepsilon \frac{d}{d\varepsilon} \left[\int_a^b dx f(x, y + \varepsilon \eta, y' + \varepsilon \eta') \right]_{\varepsilon=0} \\
&= \varepsilon \int_a^b dx \left[\frac{\partial f}{\partial y} \eta + \frac{\partial f}{\partial y'} \eta' \right] = \varepsilon \int_a^b dx \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right] \eta(x) + \frac{\partial f}{\partial y'} \varepsilon \eta(x) \\
&= \int_a^b dx \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right] h(x) = 0,
\end{aligned}$$

where we have integrated by parts and used (1.2). Using $\delta J = 0$ and the fundamental lemma L.1, we recover the Euler-Lagrange equation (1.4), as expected.

CLASSICAL PARTICLE THEORY

The special theory of relativity is corroborated by so many experiments that every physics theory should comply with it. In particular, we shall ensure the principle of relativity to be valid and develop a formalism to make relativistic invariance manifest. Since Lorentz transformations are those that yield the invariance of electromagnetism, we should start from there, looking for a covariant form for Maxwell's equations. Then, because of the strength of Lagrangian theory, and since the easiest way to require Lorentz invariance for a system is to impose that its action is a scalar, we should revise these aspects of classical mechanics, using the tools developed in the previous chapter.

2.1 Lagrangian formalism

Let's consider a classical system with n degrees of freedom, described by n -generalized coordinates $\mathbf{q}(t) \equiv (q_1(t), \dots, q_n(t))$ which correspond to a point in an n -dimensional space called *configuration space*. We'll call $\dot{\mathbf{q}}(t)$ generalized velocities.

As time flows, the point $\mathbf{q}(t)$ moves in the configuration space because of the change in the system's state, tracing out a curve.

We'll suppose that the system can be described through a function of \mathbf{q} and $\dot{\mathbf{q}}$, called *Lagrangian*

$$L = L(\mathbf{q}, \dot{\mathbf{q}}, t). \quad (2.1)$$

Higher order derivatives of \mathbf{q} are not necessary, since from the Lagrangian we derive the equations of motion, that are second order differential equations.

Defining the functional *action*,

$$S[\mathbf{q}] = \int_{t_a}^{t_b} dt L(\mathbf{q}, \dot{\mathbf{q}}, t), \quad (2.2)$$

we can apply the so-called *principle of least action*, or *Hamilton's principle*:

Principle (least action). *The motion of a mechanical system from $\mathbf{q}(t_a) = \mathbf{q}_a$ to $\mathbf{q}(t_b) = \mathbf{q}_b$ is such that the action has an extremum. Namely,*

$$\delta S = \delta \int_{t_a}^{t_b} dt L(\mathbf{q}, \dot{\mathbf{q}}, t) = 0. \quad (2.3)$$

In other words, as showed by (1.4), the system follows the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, \dots, n. \quad (2.4)$$

2.2 Special relativity and tensor notation

Metric Tensor. The special relativity can be summarized in two principles:

- *The principle of relativity:* the laws of physics are the same for all inertial observers. No preferred inertial system exists.
- *The principle of the constancy of the speed of light:* the speed of light in vacuum has the same value c in all inertial frames of reference.

These two principles lead to the invariance of the quantity

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad (2.5)$$

under Lorentz transformations. Using Einstein's notation, this quantity can be written as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \quad (2.6)$$

where $\eta_{\mu\nu}$ is the *metric tensor*

$$\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1) \quad (2.7)$$

and x^μ are the contravariant coordinates, defined by

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z). \quad (2.8)$$

From now on we'll observe the following *convention*: the Greek indices (μ, ν, \dots) can assume values from 0 to 3, while the Latin indices (i, j, \dots) from 1 to 3.

It's useful to define the covariant coordinates x_μ as

$$x_\mu = \eta_{\mu\nu} x^\nu \quad (2.9)$$

and the covariant metric tensor $\eta^{\mu\nu}$ such that

$$\eta^{\mu\nu} \eta_{\nu\rho} = \delta^\mu_\rho, \quad x^\mu = \eta^{\mu\nu} x_\nu, \quad (2.10)$$

so that we can recast eq. (2.6) as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = dx_\mu dx^\mu = \eta^{\mu\nu} dx_\mu dx_\nu.$$

Lorentz transformations. A point x^μ in Minkowski's space-time \mathcal{M} with metric η is called *event*. Two events are separated by a particular ds^2 , and, as shown in fig. 2.1, considering its sign we can classify three types of separations:

- $ds^2 > 0$: time-like separation.
- $ds^2 = 0$: light-like or null separation.
- $ds^2 < 0$: space-like separation.

We can define a *Lorentz transformation* as a linear transformation of the coordinates

$$x'^\mu = \Lambda^\mu_\nu x^\nu = \left(\frac{\partial x'^\mu}{\partial x^\nu} \right) x^\nu \quad (2.11)$$

such that the quantity $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ is invariant. Explicitly:

$$\eta_{\mu\nu} dx'^\mu dx'^\nu = \eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma dx^\rho dx^\sigma = \eta_{\rho\sigma} dx^\rho dx^\sigma.$$

By convention, considering Λ^μ_ν as a matrix, the first index represents its rows while the second one its columns. However, the usefulness of the notation developed is that we won't have to worry about matrices anymore.

In conclusion, a Lorentz transformation is a linear transformation such that

$$\eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma = \eta_{\rho\sigma}, \quad \eta^{\mu\nu} \Lambda^\rho_\mu \Lambda^\sigma_\nu = \eta^{\rho\sigma}. \quad (2.12)$$

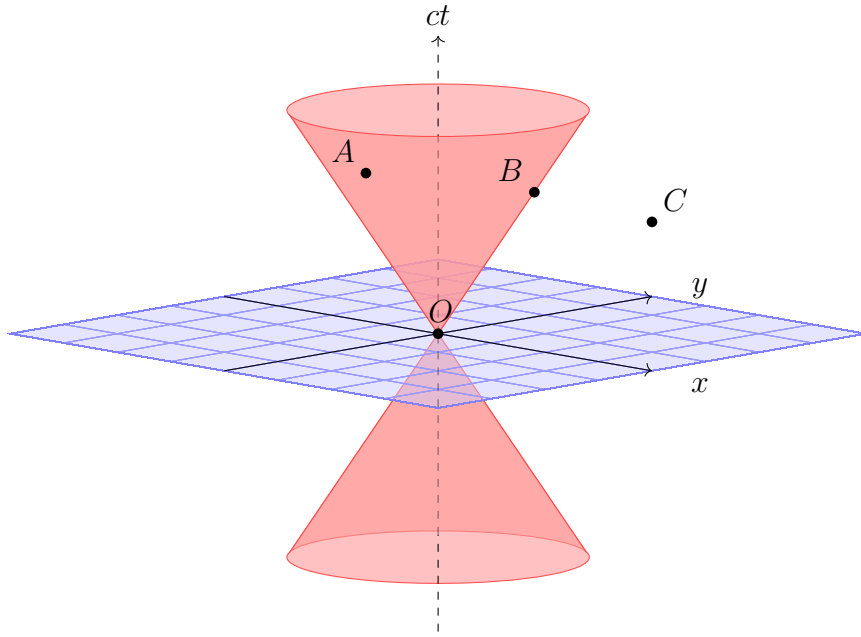


Figure 2.1 Lightcone in (2+1)-dimensions Minkowski's space-time. A and O are time-like separated, B and O are null separated and C and O are space-like separated.

In matrix notation, relation (2.12) reads

$$\Lambda^T \eta \Lambda = \eta,$$

from which we can show that

$$\det \Lambda = \pm 1. \quad (2.13)$$

Proof. Using Binet's theorem and that for a generic matrix A , $\det(A) = \det(A^T)$, we have

$$\det(\Lambda^T \eta \Lambda) = (\det(\Lambda))^2 \det(\eta) = \det(\eta),$$

the conclusion is straightforward. ■

Combining eq. (2.9), (2.10) and (2.11), we can find the Lorentz transformation for the covariant coordinates x_μ :

$$x'_\mu = \eta_{\mu\nu} x'^\nu = \eta_{\mu\nu} \Lambda^\nu_\rho x^\rho = \eta_{\mu\nu} \Lambda^\nu_\rho \eta^{\rho\sigma} x_\sigma, \quad (2.14)$$

where, defining

$$\Lambda_\mu^\sigma \equiv \eta_{\mu\nu} \Lambda^\nu_\rho \eta^{\rho\sigma}, \quad (2.15)$$

we can write the transformation (2.14) as

$$x'^{\mu} = \Lambda_{\nu}^{\mu} x^{\nu}. \quad (2.16)$$

To conclude, it's easy to prove that the matrix defined by eq. (2.15) is the inverse transpose of Λ , namely

$$\Lambda_{\mu}^{\tau} = (\Lambda^{-1,T})_{\mu}^{\tau}. \quad (2.17)$$

Proof. Let's multiply eq. (2.12) by $\eta^{\sigma\tau}$

$$\eta_{\mu\nu} \Lambda^{\mu}_{\rho} \Lambda^{\nu}_{\sigma} \eta^{\sigma\tau} = \eta_{\rho\sigma} \eta^{\sigma\tau},$$

and using eq. (2.15) and (2.10)

$$\Lambda^{\mu}_{\rho} \Lambda_{\mu}^{\tau} = \delta_{\rho}^{\tau},$$

from which it is immediate that eq. (2.17) must hold. ■

Tensor properties. Let's now consider the Minkowski's space-time \mathcal{M} and a generic Lorentz transformation:

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}, \quad x'_{\mu} = \Lambda_{\mu}^{\nu} x_{\nu}. \quad (2.18)$$

We can define the following quantities:

- *Scalars:* quantities $\Phi(x)$ that are invariant under a Lorentz transformation:

$$\Phi'(x') = \Phi(x).$$

- *Contravariant 4-vectors:* quantities V^{μ} that transform like the coordinates x^{μ} :

$$V'^{\mu}(x') = \Lambda^{\mu}_{\nu} V^{\nu}(x).$$

- *Covariant 4-vectors:* quantities V_{μ} that transform like the covariant coordinates x_{μ} :

$$V'_{\mu}(x') = \Lambda_{\mu}^{\nu} V_{\nu}(x).$$

It's immediate to show that a contravariant four-vector contracted with a covariant one yields a scalar:

$$V'^{\mu} \omega'_{\mu} = \Lambda^{\mu}_{\nu} \Lambda_{\mu}^{\rho} V^{\nu} \omega_{\rho} = \delta_{\nu}^{\rho} V^{\nu} \omega_{\rho} = V^{\nu} \omega_{\nu}.$$

- *Tensors*: a general (p, q) tensor is a quantity that transforms like:

$$T^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}(x') = \Lambda^{\mu_1}_{\alpha_1} \dots \Lambda^{\mu_p}_{\alpha_p} \Lambda_{\nu_1}^{\beta_1} \dots \Lambda_{\nu_q}^{\beta_q} T^{\alpha_1 \dots \alpha_p}_{\beta_1 \dots \beta_q}(x).$$

To each contravariant four-vector V^μ corresponds a covariant four-vector and vice versa,

$$\begin{aligned} V^\mu &= \eta^{\mu\nu} V_\nu, \\ V_\mu &= \eta_{\mu\nu} V^\nu. \end{aligned}$$

In general, to lower or raise an index of a tensor, it must be contracted with the metric tensor η . Following the convention (2.7), if we lower or raise the temporal index of a four-vector, i.e., the zeroth component, the sign remains the same, while it changes for a spacial index, i.e., an i -th component.

Given two four-vectors $A^\mu = (A^0, \mathbf{A})$ and $B^\mu = (B^0, \mathbf{B})$, their inner product is defined as

$$A \cdot B = A^\mu B_\mu = \eta_{\mu\nu} A^\mu B^\nu = A^0 B^0 - \mathbf{A} \cdot \mathbf{B}. \quad (2.20)$$

They are orthogonal if $A \cdot B = 0$. Further, the absolute value of a four-vector is

$$A \cdot A = A^\mu A_\mu = (A^0)^2 - \mathbf{A} \cdot \mathbf{A}, \quad (2.21)$$

from which we can distinguish three type of four-vectors, time-like if $A^\mu A_\mu > 0$, light-like if $A^\mu A_\mu = 0$ or space-like if $A^\mu A_\mu < 0$.

A $(2, 0)$ tensor $T^{\mu\nu}$ is said to be *symmetric* if $T^{\mu\nu} = T^{\nu\mu}$, while it's *antisymmetric* if $T^{\mu\nu} = -T^{\nu\mu}$. A generic $(2, 0)$ tensor $T^{\mu\nu}$ can be decomposed into a symmetric part $T^{(\mu\nu)}$ and an antisymmetric part $T^{[\mu\nu]}$:

$$T^{\mu\nu} = T^{(\mu\nu)} + T^{[\mu\nu]}, \quad (2.22a)$$

$$T^{(\mu\nu)} = \frac{1}{2}(T^{\mu\nu} + T^{\nu\mu}), \quad (2.22b)$$

$$T^{[\mu\nu]} = \frac{1}{2}(T^{\mu\nu} - T^{\nu\mu}). \quad (2.22c)$$

In general, the contraction of a symmetric tensor with an antisymmetric one vanishes.

Given an antisymmetric $(2, 0)$ tensor $F^{\mu\nu}$, its *dual* is defined as

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}, \quad (2.23)$$

where ε^{ijkl} is the *Levi-Civita symbol*, defined as

$$\varepsilon^{ijkl} = \begin{cases} +1 & \text{if } (i, j, k, l) \text{ is an even permutation of } (0, 1, 2, 3). \\ -1 & \text{if } (i, j, k, l) \text{ is an odd permutation of } (0, 1, 2, 3). \\ 0 & \text{otherwise.} \end{cases} \quad (2.24)$$

The dual tensor $\tilde{F}^{\mu\nu}$ is also antisymmetric.

From now on, we'll deal with tensorial equations, that equate tensors of the same rank. The benefit of this notation is that they would be manifestly Lorentz invariant.

Differentiation and integration. Let's define a covariant vector for differentiation in Minkowski's space-time \mathcal{M} , known as *four-gradient*,

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^i} \right) = \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right). \quad (2.25)$$

Since the coordinates x^μ behave as a contravariant four-vector, it's easy to show that ∂_μ is a covariant one.

Proof. Using $\Lambda^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu}$ and $(\Lambda^{-1})^\mu{}_\nu = \frac{\partial x^\mu}{\partial x'^\nu}$, that can be derived from the Lorentz transformation (2.11) and its inverse, and also eq. (2.17), we have

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu{}_\mu \frac{\partial}{\partial x^\nu} = (\Lambda^{-1,T})_\mu{}^\nu \frac{\partial}{\partial x^\nu} = \Lambda_\mu{}^\nu \frac{\partial}{\partial x^\nu},$$

showing that ∂_μ transforms like x_μ . ■

Applying the four-gradient to a scalar field $\phi(x)$, we obtain

$$\partial_\mu \phi(x) = \left(\frac{1}{c} \frac{\partial \phi}{\partial t}, \nabla \phi \right),$$

while applied to a vector field $V^\mu(x) = (V^0, \mathbf{V})$, it leads to the *four-divergence*, hence

$$\partial_\mu V^\mu = \frac{1}{c} \frac{\partial V^0}{\partial t} + \nabla \cdot \mathbf{V}.$$

Contracting ∂_μ with itself, yields the *d'Alembertian* operator:

$$\square \equiv \partial_\mu \partial^\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2.$$

Considering now the integration properties, the volume element in Minkowski's space-time is

$$d^4x = dx^0 dx^1 dx^2 dx^3 = c dt d^3x, \quad (2.26)$$

and it is Lorentz invariant.

Proof. Under a Lorentz transformation of the coordinates, $x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}$, the volume element (2.26) transforms as

$$d^4x' = |J(x)| d^4x,$$

where $J(x)$ is the Jacobian of the transformation, namely

$$J(x) = \det \left\{ \frac{\partial x'^{\mu}}{\partial x^{\nu}} \right\} = \det \Lambda.$$

Using eq. (2.13), we have

$$d^4x' = d^4x, \quad (2.27)$$

as claimed. ■

In analogy with \mathbb{R}^3 , a *hypersurface* σ in (3+1)-dimensional Minkowski's space-time is a 3-dimensional surface. Its element $d\sigma^{\mu}$ can be written as

$$d\sigma^{\mu} = n^{\mu} d\sigma,$$

where n^{μ} is a unit four-vector orthogonal to $d\sigma^{\mu}$. In particular, we are interested in *space-like hypersurfaces*, characterized by the property that their points are separated by a space-like distance:

$$x_1^{\mu}, x_2^{\mu} \in \sigma : (x_1^{\mu} - x_2^{\mu})(x_{1\mu} - x_{2\mu}) < 0,$$

implying that n^{μ} is time-like.

A space-like hyperplane is defined implicitly by the equation

$$n \cdot x = \lambda,$$

where n^{μ} is a time-like unit four-vector and λ a real parameter. Choosing the frame of reference in which $n_{\mu} = (1, 0, 0, 0)$, the previous relation is equivalent to

$$t = \frac{\lambda}{c},$$

that is \mathbb{R}^3 at time $t = \lambda/c$. In this particular case the hypersurface element is

$$d\sigma_\mu = n_\mu d\sigma = (dx^1 dx^2 dx^3, 0, 0, 0) = (d^3x, 0, 0, 0). \quad (2.28)$$

Finally, let's seek for a generalization of *Gauss theorem*. If we consider a 4-dimensional volume Ω and its boundary $\partial\Omega$,

$$\int_{\Omega} d^4x \partial_\mu V^\mu(x) = \int_{\partial\Omega} d\sigma_\mu V^\mu(x). \quad (2.29)$$

If certain properties are satisfied, the upper integral doesn't depend on the specific hypersurface chosen. Let's suppose that $V^\mu(x)$ has vanishing four-divergence, hence

$$\partial_\mu V^\mu = 0,$$

and the volume of integration Ω is delimited by two space-like hypersurfaces σ_1 and σ_2 , and a time-like hypersurface σ_∞ . Then, by (2.29) we can write

$$\int_{\Omega} d^4x \partial_\mu V^\mu(x) = 0 = \int_{\sigma_1} d\sigma_\mu V^\mu - \int_{\sigma_2} d\sigma_\mu V^\mu + \int_{\sigma_\infty} d\sigma_\mu V^\mu,$$

where the unit four-vectors n_1^μ and n_2^μ are oriented in the same direction. Further, let's suppose that V^μ approaches to zero fast enough, so that the integration on the time-like surface vanishes. We have

$$\int_{\sigma_1} d\sigma_\mu V^\mu = \int_{\sigma_2} d\sigma_\mu V^\mu = \int_{\sigma} d\sigma_\mu V^\mu, \quad (2.30)$$

meaning that the integral above is independent on the specific space-like hypersurface chosen. In particular, choosing the previously considered hypersurface with element (2.28), we have

$$\partial_\mu V^\mu = 0 \implies \int d^3x V^0(x) = \text{const in } t, \quad (2.31)$$

where we are integrating over \mathbb{R}^3 .

2.3 Charge and current densities

To introduce the charge and the current density of a system, let's start considering a system of N point-like particles, in which each particle has an electric charge e_n

and a position $\mathbf{x}_n(t)$ at time t . The *charge density* of the system is

$$\rho(\mathbf{x}, t) = \sum_{n=1}^N \rho_n(\mathbf{x}, t) = \sum_{n=1}^N e_n \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \quad (2.32)$$

where it has been used the Dirac's delta (cf. App. A.2) to localize each charge at its position.

The total charge of the system is

$$Q = \int_V d^3x \rho(\mathbf{x}, t) = \sum_n e_n \int_V d^3x \delta^3(\mathbf{x} - \mathbf{x}_n(t)) = \sum_n e_n,$$

where V is the volume containing the particles. In the last step we have used (A.7).

In general, we can define the *current density* of a particle as

$$\mathbf{j}_n(\mathbf{x}, t) = \rho_n(\mathbf{x}, t) \frac{d\mathbf{x}_n}{dt}(t), \quad (2.33)$$

so the current density of the system is

$$\mathbf{j}(\mathbf{x}, t) = \sum_n e_n \delta^3(\mathbf{x} - \mathbf{x}_n(t)) \frac{d\mathbf{x}_n}{dt}(t). \quad (2.34)$$

By substitution, it can be verified that (2.32) and (2.34) abide by the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \quad (2.35)$$

Let's now consider a generic system, that can be also continuous. The electric charge dq contained in an infinitesimal volume d^3x is

$$dq = \rho(\mathbf{x}, t) d^3x \quad (2.36)$$

and the current flowing through an infinitesimal oriented surface $d\boldsymbol{\sigma}$ is

$$dI = \mathbf{j}(\mathbf{x}, t) \cdot d\boldsymbol{\sigma}, \quad (2.37)$$

where $d\boldsymbol{\sigma} = d\sigma \mathbf{n}$, \mathbf{n} being a unit vector orthogonal to the surface.

The total charge enclosed in a volume V is

$$Q(t) = \int_V d^3x \rho(\mathbf{x}, t)$$

and the total current flowing through a surface S at time t is

$$I(t) = \int_S \mathbf{j}(\mathbf{x}, t) \cdot d\boldsymbol{\sigma}. \quad (2.38)$$

Again, the continuity equation (2.35) holds, indeed it's an experimental law.

It's now useful to exploit the formalism of sec. 2.2 to prove that ρ and \mathbf{j} can be combined into a four-vector, that will be useful to cast Maxwell's equations in a covariant form. The *four-current* is defined as

$$J^\mu = (J^0, \mathbf{J}) = (c\rho, \mathbf{j}) \quad (2.39)$$

or, equivalently, by (2.32), (2.34) and (2.39),

$$J^\mu = \sum_n e_n \delta^3(\mathbf{x} - \mathbf{x}_n(t)) \frac{dx_n^\mu}{dt}(t), \quad (2.40)$$

where $x_n^\mu = (ct, \mathbf{x}_n)$.

Let's introduce a formal integration over t' using Delta's definition (A.6)

$$\begin{aligned} J^\mu &= \int dt' \delta(t - t') \sum_n e_n \delta^3(\mathbf{x} - \mathbf{x}_n(t')) \frac{dx_n^\mu}{dt'} \\ &= c \int dt' \delta^4(x^\mu - x_n^\mu(t)) \sum_n e_n \frac{dx_n^\mu}{dt'}, \end{aligned}$$

where it has been used (A.13) and (A.11). It's convenient to change the integration variable to ds , since it is a scalar

$$\begin{aligned} J^\mu &= c \int ds \frac{dt'}{ds} \delta^4(x^\mu - x_n^\mu(t)) \sum_n e_n \frac{dx_n^\mu}{ds} \frac{ds}{dt'} \\ &= c \int ds \delta^4(x^\mu - x_n^\mu(t)) \sum_n e_n \frac{dx_n^\mu}{ds}. \end{aligned}$$

The last formulation makes evident that J^μ is a four-vector, since ds is a scalar, $\delta^4(x)$ is a scalar by (A.14), and x^μ is a four-vector.

Lastly, considering (2.39), the continuity equation (2.35) can be written as

$$\partial_\mu J^\mu = 0. \quad (2.41)$$

2.4 Electromagnetism

Lorentz transformations (cf. sec. 2.2) yield the invariance of Maxwell's equations. Therefore, the covariance of those equations is obvious. Here, our purpose is to make this covariance manifest, using the notation of sec. 2.2.

To begin with, *Maxwell's equations* in CGS units¹ are

$$\nabla \cdot \mathbf{E}(\mathbf{x}, t) = 4\pi\rho(\mathbf{x}, t), \quad (2.42a)$$

$$\nabla \wedge \mathbf{B}(\mathbf{x}, t) - \frac{1}{c} \frac{\partial \mathbf{E}(\mathbf{x}, t)}{\partial t} = \frac{4\pi}{c} \mathbf{j}(\mathbf{x}, t), \quad (2.42b)$$

$$\nabla \cdot \mathbf{B}(\mathbf{x}, t) = 0, \quad (2.42c)$$

$$\nabla \wedge \mathbf{E}(\mathbf{x}, t) + \frac{1}{c} \frac{\partial \mathbf{B}(\mathbf{x}, t)}{\partial t} = 0. \quad (2.42d)$$

The continuity equation (2.35), that, as already stated, must hold being an experimental fact, is contained in (2.42).

Proof. Let's derive eq. (2.42a) with respect to time,

$$\frac{\partial}{\partial t} \nabla \cdot \mathbf{E} = \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} = 4\pi \frac{\partial \rho}{\partial t}.$$

Taking the divergence of eq. (2.42b) and recalling that the divergence of a curl vanishes, we have:

$$\nabla \cdot \left(\nabla \wedge \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} \right) = -\frac{4\pi}{c} \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c} \nabla \cdot \mathbf{j}.$$

Using the previous relations, the conclusion is straightforward. ■

A general solution for the homogeneous equations (2.42c) and (2.42d) is provided by the *scalar* and *vector potentials*, ϕ and \mathbf{A} , defined as

$$\mathbf{B} = \nabla \wedge \mathbf{A}, \quad (2.43a)$$

$$\mathbf{E} = -\nabla\phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}. \quad (2.43b)$$

¹In CGS system we impose $\varepsilon_0 = \mu_0 = 1$.

Given the above-mentioned potentials, the inhomogeneous equations (2.42a) and (2.42b), describing their dynamics, become

$$\nabla^2 \phi + \frac{1}{c} \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) = -4\pi\rho, \quad (2.44a)$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \left(\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} \right) = -\frac{4\pi}{c} \mathbf{j}. \quad (2.44b)$$

Recalling that the divergence of a curl of a vector field and the curl of the gradient of a scalar field vanish and that $\nabla \wedge (\nabla \wedge \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$, everything can be easily verified.

Maxwell's equations (2.42) are invariant under *gauge transformations*, namely

$$\mathbf{A}(x) \rightarrow \mathbf{A}'(x) = \mathbf{A}(x) + \nabla \chi(x), \quad (2.45a)$$

$$\phi(x) \rightarrow \phi'(x) = \phi(x) - \frac{1}{c} \frac{\partial \chi}{\partial t}, \quad (2.45b)$$

meaning that the simultaneous transformations (2.45) yield

$$\mathbf{E} \rightarrow \mathbf{E}' = \mathbf{E},$$

$$\mathbf{B} \rightarrow \mathbf{B}' = \mathbf{B}.$$

χ is an arbitrary function, called *gauge function*. Moreover, notice that the transformations (2.45) don't involve a change in the system of reference, that is, there is no transformation over the coordinates x .

Since the sources ρ and \mathbf{j} can be gathered into a four-vector $J^\mu = (c\rho, \mathbf{j})$, as seen in sec. 2.3 by (2.39), we expect \mathbf{E} and \mathbf{B} , on the left side of eqs. (2.42), to be cast into a tensor, in order to write the equations in a covariant form. Indeed, it is possible through the *field-strength* or *electromagnetic tensor*, $F^{\mu\nu}$,

$$F^{\mu\nu} = -F^{\nu\mu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}, \quad (2.46)$$

that can also be written as

$$F^{0i} = -E_i \quad F^{ij} = -\varepsilon^{ijk} B_k.$$

Using (2.39) and (2.46), Maxwell's equations (2.42) become

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu, \quad (2.47a)$$

$$\partial_\mu \tilde{F}^{\mu\nu} = 0, \quad (2.47b)$$

where $\tilde{F}^{\mu\nu}$ is the antisymmetric dual field-strength tensor, defined, as in (2.23), by

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}. \quad (2.48)$$

In alternative, eq. (2.47b) is equivalent to

$$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0. \quad (2.49)$$

Using the field-strength tensor, we can derive two scalars,

$$F^{\mu\nu} F_{\mu\nu} = -2(E^2 - B^2), \quad (2.50a)$$

$$\tilde{F}^{\mu\nu} F_{\mu\nu} = -4\mathbf{E} \cdot \mathbf{B}, \quad (2.50b)$$

that we'll reclaim later.

As we have seen by (2.43), electromagnetism can be described using the potentials ϕ and \mathbf{A} . Since \mathbf{E} and \mathbf{B} are contained in the electromagnetic tensor $F^{\mu\nu}$, we expect ϕ and \mathbf{A} to form a four-vector that should have a relation with $F^{\mu\nu}$. It is called *four-potential*, and it's defined as

$$A^\mu = (\phi, \mathbf{A}). \quad (2.51)$$

From (2.43), (2.46) and (2.51) follows

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (2.52)$$

Using (2.52) in (2.47a) we have

$$\square A^\nu - \partial^\nu (\partial_\mu A^\mu) = \frac{4\pi}{c} J^\nu. \quad (2.53)$$

2.5 Gauge invariance

Making use of the four-potential (2.51), gauge transformations (2.45) read

$$A^\mu(x) \rightarrow A'^\mu(x) = A^\mu(x) - \partial^\mu \chi(x), \quad (2.54)$$

and (2.54) yield the invariance of the electromagnetic tensor (2.46),

$$F^{\mu\nu} \rightarrow F'^{\mu\nu} = \partial^\mu A'^\nu - \partial^\nu A'^\mu = F^{\mu\nu}. \quad (2.55)$$

This means that if $A^\mu(x)$ is a solution of (2.53), also $A'^\mu(x)$ is.

Thus, the electromagnetic tensor doesn't determine the four-potential uniquely, and then the independent components of A^μ must be up to three, not four. Using (2.54), we can impose a condition on A^μ , called *gauge fixing*. To simplify our computation, let's write down the *Lorenz gauge* condition,

$$\partial_\mu A^\mu = 0. \quad (2.56)$$

More than being Lorentz invariant, this gauge is useful since yields a simpler version of inhomogeneous Maxwell's equations (2.53), namely

$$\square A^\nu = \frac{4\pi}{c} J^\nu, \quad (2.57a)$$

$$\partial_\mu A^\mu = 0. \quad (2.57b)$$

However, it doesn't fix A^μ completely, since under a gauge transformation (2.54) on A^μ itself, with the constraint that the gauge function is a plane wave, the transformed potential A'^μ leaves (2.57) invariant. Explicitly, this *residual gauge freedom* is

$$A'^\mu(x) = A^\mu(x) - \partial^\mu \theta(x), \quad (2.58a)$$

$$\square \theta(x) = 0. \quad (2.58b)$$

Thus, A^μ has only two degrees of freedom, corresponding to the two polarizations of light.

2.6 Solution for Maxwell's equations

Let's consider a general source J^μ for the electromagnetic field. The solution of Maxwell's equations in Lorenz gauge, eq. (2.57), can be written as the sum of the homogeneous solution and a particular one:

$$A^\mu = A^\mu_{\text{homo}} + A^\mu_{\text{part}}.$$

While the homogeneous solution would lead to a plane wave, not particularly interesting for our purposes, the particular solution problem requires the definition of a Green function, an efficient method to tackle problems in quantum field theory, as we shall see in due time.

Therefore, let's define the *Green function* D for Maxwell's equations (2.57). It is a function of two point in Minkowski's space-time \mathcal{M} such that

$$D(t, \mathbf{x}; t', \mathbf{x}') : \quad \square_x D(t, \mathbf{x}; t', \mathbf{x}') = \delta^4(x - x'). \quad (2.59)$$

It represents the potential generated in x by a point-like source in x' . Since in \mathcal{M} there must be invariance under translations, D can only depend on the difference between x and x' , namely

$$D = D(x - x') = D(ct - ct', \mathbf{x} - \mathbf{x}').$$

Using the definition (2.59), eq. (2.57a) is satisfied by

$$A^\mu(t, \mathbf{x}) = \frac{4\pi}{c} \int d^4x' D(x - x') J^\mu(t', \mathbf{x}'). \quad (2.60)$$

Proof. Using (2.60), since only D depends on x ,

$$\begin{aligned} \square_x A^\mu &= \frac{4\pi}{c} \int d^4x' \square_x D(x - x') J^\mu(t', \mathbf{x}') \\ &= \frac{4\pi}{c} \int d^4x' \delta^4(x - x') J^\mu(t', \mathbf{x}') = \frac{4\pi}{c} J^\mu(t, \mathbf{x}), \end{aligned}$$

where we have used (2.59) and (A.13), obtaining (2.57a). ■

Obviously, in order for (2.60) to be a solution of (2.57), it must also satisfy the Lorenz gauge condition (2.57b). It can be verified, but we'll omit this computation. Let's seek to find an explicit expression for D .

To tackle the problem, let's operate in Fourier's space, writing, by (A.5),

$$D(x - x') = \frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha(x^\alpha - x'^\alpha)} \tilde{D}(k), \quad (2.61)$$

where $\tilde{D}(k)$ is the Fourier transform of D , defined by (A.4). Further, using (A.15), the delta can be written as

$$\delta^4(x - x') = \frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha(x^\alpha - x'^\alpha)}.$$

Substituting into the definition (2.59) yields

$$\square_x \left[\frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha(x^\alpha - x'^\alpha)} \tilde{D}(k) \right] = \frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha(x^\alpha - x'^\alpha)}.$$

Since the exponential is the only one that depends on x , the d'Alembertian acts only on it, yielding

$$\partial_\mu e^{-ik_\alpha(x^\alpha - x'^\alpha)} = -ik_\mu e^{-ik_\alpha(x^\alpha - x'^\alpha)}, \quad \partial^\mu \partial_\mu e^{-ik_\alpha(x^\alpha - x'^\alpha)} = -k^\mu k_\mu e^{-ik_\alpha(x^\alpha - x'^\alpha)}.$$

Then,

$$\int d^4k (-k_\mu k^\mu) e^{-ik_\alpha(x^\alpha - x'^\alpha)} \tilde{D}(k) = \int d^4k e^{-ik_\alpha(x^\alpha - x'^\alpha)}. \quad (2.62)$$

Because of the completeness of the Fourier transform, eq. (2.62) implies

$$(-k_\mu k^\mu) e^{-ik_\alpha(x^\alpha - x'^\alpha)} \tilde{D}(k) = e^{-ik_\alpha(x^\alpha - x'^\alpha)},$$

equivalent to

$$\tilde{D}(k) = \frac{1}{k^2 - k_0^2}, \quad (2.63)$$

where we have set $|\mathbf{k}| \equiv k$. Using (2.63) in (2.61) we have

$$D(x - x') = \frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha(x^\alpha - x'^\alpha)} \frac{1}{k^2 - k_0^2}.$$

Now, writing

$$d^4k = dk^0 d^3k \quad \text{and} \quad e^{-ik_\alpha(x^\alpha - x'^\alpha)} = e^{-ik_0(ct - ct')} e^{i\mathbf{k}(\mathbf{x} - \mathbf{x}')} ,$$

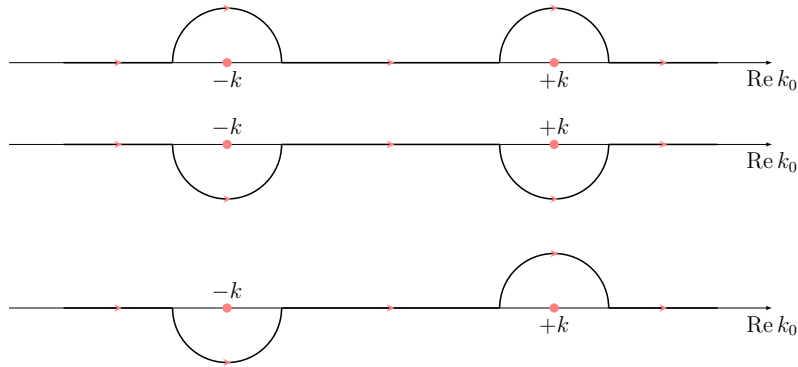


Figure 2.2 Possible deformations of the integration path to go around the poles on the real axis k_0 .

we obtain

$$D(x - x') = \frac{1}{(2\pi)^4} \int d^3k e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')} \int dk_0 \frac{e^{-ik_0(ct-ct')}}{k^2 - k_0^2}. \quad (2.64)$$

Notice that the index of the temporal component of a four-vector, in this case k^0 , can be lowered or raised without changing of sign, because of the chosen metric (2.7).

Let's first perform the integration over k_0 , defining for convenience

$$J_{\text{ret}}(k) = \int dk_0 \frac{e^{-ik_0(ct-ct')}}{k^2 - k_0^2}. \quad (2.65)$$

Inspecting (2.65), it has two poles on the real axis, for $k_0 = \pm k$. Thus, since the integrand is holomorphic, we can exploit the computation by the residue theorem (cf. sec. A.3), by deforming properly the integration path to go around the poles. However, it can be done in different ways, as depicted in fig. 2.2. Depending on the choice, we would obtain Green functions with different physical interpretation, so we should examine the implications in depth.

Let's remark that the Green function $D(x - x')$, as defined in (2.59), represents the potential measured in (t, \mathbf{x}) and generated by a point-like and instantaneous source positioned in (t', \mathbf{x}') . If we temporarily consider a point-like charge and we wish to measure its field at (t, \mathbf{x}) , because of causality, we expect the field at time t to depend on the behaviour of the charge at $t' < t$. Therefore, in general, we should impose the Green function (2.59) to vanish if $t' > t$, implementing the so-called *retarded Green function*, namely

$$D_{\text{ret}}(t - t'; \mathbf{x} - \mathbf{x}') = 0 \quad \text{if } t < t'. \quad (2.66)$$

In order to implement (2.66), while performing the integration (2.65), we must deform the integration path Γ_r as shown in fig. 2.3. Indeed, if $t < t'$, by Jordan's

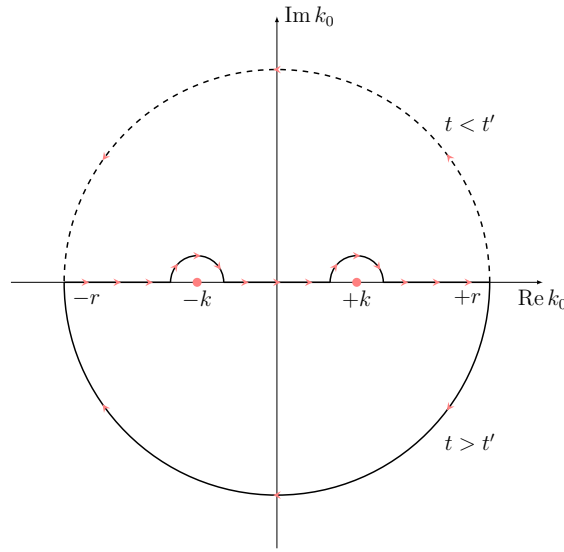


Figure 2.3 Deformation of the integration path to implement the retarded Green function. The path drawn in continuous line is Γ_r .

lemma L.2, we must close the integration path in the upper half-plane, as showed through the dashed line in fig. 2.3. Since there is no pole within this integration path, for Cauchy's theorem (A.16), the integral (2.65) vanishes, as we expect by (2.66). Conversely, if $t > t'$, in order to avoid the exponential in (2.65) to diverge and by Jordan's lemma, we must close the integration path in the lower half-plane, leading to the continuous line in fig. 2.3. In this case, we can use (A.28) and apply the residue theorem (A.23). Then, we can compute (2.65) as

$$\begin{aligned} J_{\text{ret}}(k) &= \lim_{r \rightarrow \infty} \oint_{\Gamma_r} dk_0 \frac{e^{-ick_0(t-t')}}{k^2 - k_0^2} \\ &= -2\pi i \sum_j \text{Res } f(k_0) e^{-ick_0(t-t')} \Big|_{k_0=k_{0j}}, \end{aligned} \quad (2.67)$$

where

$$f(k_0) = \frac{1}{k^2 - k_0^2},$$

k_{0j} are the poles of $f(k_0)$ within Γ_r and the minus sign is due to the orientation of Γ_r , that is clockwise, as shown in fig. 2.3. In this case, since the poles $k_0 = \pm k$ are simple

poles, we can compute the residue by (A.22), namely

$$\begin{aligned} \sum_j \text{Res } f(k_0) e^{-ick_0(t-t')} \Big|_{k_0=k_{0j}} &= \sum_j -\frac{e^{-ick_0(t-t')}}{2k_0} \Big|_{k_0=k_{0j}} \\ &= \frac{1}{2k} \left[e^{ick(t-t')} - e^{-ick(t-t')} \right] = \frac{i}{k} \sin [ck(t-t')]. \end{aligned} \quad (2.68)$$

Inserting (2.68) into (2.67) we obtain

$$J_{\text{ret}}(k) = \frac{2\pi}{k} \sin [ck(t-t')]. \quad (2.69)$$

Now, recalling the definition (2.65) and inserting (2.69) into (2.64) we have for the retarded Green function

$$D_{\text{ret}}(\mathbf{x} - \mathbf{x}') = (2\pi)^{-3} \int \frac{d^3k}{k} e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')} \sin [ck(t-t')] \theta(t-t'), \quad (2.70)$$

where the *Heaviside step function*

$$\theta(t-t') = \begin{cases} 1, & \text{if } t > t', \\ 0, & \text{if } t < t', \end{cases} \quad (2.71)$$

is used to implement the retarded condition (2.66).

To tackle the integral (2.70) it's convenient to use polar coordinates. Since in \mathcal{M} there is invariance under rotations, it's possible to rotate the system of reference in order for $\mathbf{x} - \mathbf{x}'$ to be parallel to the polar axis. Acting this way, we can write

$$\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}') = |\mathbf{k}| |\mathbf{x} - \mathbf{x}'| \cos \theta = kR \cos \theta,$$

where θ is the angle between the polar axis and \mathbf{k} , $k \equiv |\mathbf{k}|$ and $R \equiv |\mathbf{x} - \mathbf{x}'|$. Let's remark that \mathbf{x} is the observation point, where the field is measured, whilst \mathbf{x}' is the position of the source. Moreover, the volume element in polar coordinates is $d^3k = k^2 dk \sin \theta d\theta d\phi$. So, eq. (2.70) becomes

$$D_{\text{ret}} = (2\pi)^{-3} \int_0^\infty dk k \sin [ck(t-t')] \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta e^{ikR \cos \theta} \theta(t-t').$$

Performing the angular integrations by setting $\xi = \cos \theta$ we have

$$\begin{aligned}
 D_{\text{ret}}(x - x') &= (2\pi)^{-2} \int_0^\infty dk k \sin [ck(t - t')] \int_{-1}^{+1} d\xi e^{ikR\xi} \theta(t - t') \\
 &= (2\pi)^{-2} \int_0^\infty dk k \sin [ck(t - t')] \frac{e^{ikR} - e^{-ikR}}{ikR} \theta(t - t') \\
 &= \frac{2}{(2\pi)^2 R} \int_0^\infty dk \sin [ck(t - t')] \sin (kR) \theta(t - t') \\
 &= \frac{1}{(2\pi)^2 R} \int_{-\infty}^\infty dk \sin [ck(t - t')] \sin (kR) \theta(t - t'),
 \end{aligned}$$

where the last step is a consequence of the fact that the integrand is even in k . Using Euler's formula

$$D_{\text{ret}}(x - x') = \frac{1}{4(2\pi)^2 R} \int_{-\infty}^\infty dk \left[e^{-ick(t-t')} - e^{ick(t-t')} \right] \left[e^{ikR} - e^{-ikR} \right] \theta(t - t'),$$

and performing the multiplications, we have

$$\begin{aligned}
 D_{\text{ret}}(x - x') &= \frac{1}{4(2\pi)^2 R} \int_{-\infty}^\infty dk \left[e^{-ik[c(t-t')-R]} - e^{-ik[c(t-t')+R]} \right. \\
 &\quad \left. - e^{ik[c(t-t')+R]} + e^{ik[c(t-t')-R]} \right] \theta(t - t').
 \end{aligned} \tag{2.72}$$

Recalling the integral representation of the Dirac's delta (A.9), and the property (A.8), since by (2.71) $ik [c(t - t') + R] > 0$, the second and third term in the integral (2.72) vanish. Indeed, the Dirac's delta is different from zero only if its argument vanishes and since in this case it is strictly positive, it must be zero. Using again (A.9) for the remaining terms, we can write

$$D_{\text{ret}}(x - x') = \frac{1}{4\pi} \frac{\delta [c(t - t') - |\mathbf{x} - \mathbf{x}'|]}{|\mathbf{x} - \mathbf{x}'|} \theta(t - t'). \tag{2.73}$$

Now, recalling (2.60), the potential related to the retarded Green function (2.73) is

$$\begin{aligned}
 A^\mu_{\text{ret}}(t, \mathbf{x}) &= \frac{4\pi}{c} \int d^4x' D(x - x') J^\mu(t', \mathbf{x}') \\
 &= \frac{1}{c} \int d^4x' \frac{\delta [c(t - t') - R]}{R} J^\mu(t', \mathbf{x}') \theta(t - t').
 \end{aligned}$$

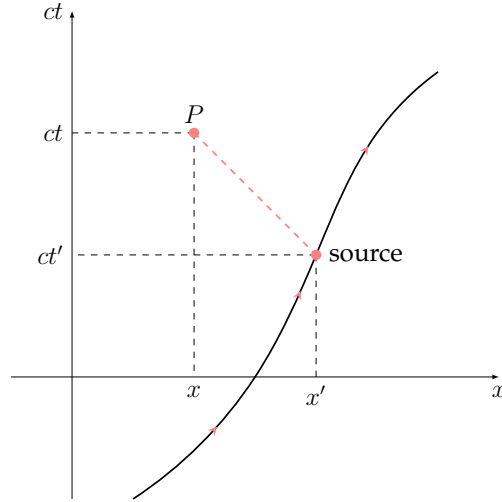


Figure 2.4 The field at time t is influenced by the source at time $t' = t - \frac{R}{c}$, where $R = |\mathbf{x} - \mathbf{x}'|$.

Writing $d^4x' = dx_0' d\mathbf{x}' = c dt' d^3x'$ and using (A.6) and (A.11) we have

$$A^\mu_{\text{ret}}(t, \mathbf{x}) = \frac{1}{c} \int d^3x' \frac{J^\mu \left(t - \frac{R}{c}, \mathbf{x}' \right)}{R}. \quad (2.74)$$

The retarded potential (2.74) shows that the field at time t is influenced by the source at time $t' = t - \frac{R}{c}$, where $R = |\mathbf{x} - \mathbf{x}'|$, as shown in fig. 2.4.

Now we would like to show the Lorentz invariance of (2.73). Using (A.12) we can write, interpreting the square as a Minkowski's square (2.21),

$$\begin{aligned} \delta[(x - x')^2] &= \delta[(x^\alpha - x'^\alpha)(x_\alpha - x'_\alpha)] = \delta[(x_0 - x_0')^2 - |\mathbf{x} - \mathbf{x}'|^2] \\ &= \delta[(x_0 - x_0')^2 - R^2] = \frac{1}{2R} [\delta(x_0 - x_0' + R) + \delta(x_0 - x_0' - R)]. \end{aligned} \quad (2.75)$$

Multiplying (2.75) by $\theta(x_0 - x_0')$ we have

$$\begin{aligned} \delta[(x - x')^2] \theta(x_0 - x_0') &= \frac{1}{2R} [\delta(x_0 - x_0' + R) + \delta(x_0 - x_0' - R)] \theta(x_0 - x_0') \\ &= \frac{1}{2R} [\delta(x_0 - x_0' - R)] \theta(x_0 - x_0'), \end{aligned} \quad (2.76)$$

where in the last step we have used that, since the Heaviside function requires $x_0 - x_0' > 0$, the argument $x_0 - x_0' + R$ of the delta can't vanish and then the delta must be zero. Substituting (2.76) into (2.73), we can write the retarded Green

function as

$$D_{\text{ret}}(x - x') = \frac{1}{2\pi} \delta [(x^\alpha - x'^\alpha)(x_\alpha - x'_\alpha)] \theta(x_0 - x'_0), \quad (2.77)$$

that, obviously, is nonzero only if $(x^\alpha - x'^\alpha)(x_\alpha - x'_\alpha) = 0$ or rather, if $x - x'$ is a light-like four-vector, as we expect (see fig. 2.4).

Furthermore, it can be shown that (2.74) satisfies the condition (2.57b), thus, it represents a particular solution for Maxwell's equations.

CLASSICAL FIELD THEORY

Let's think for a moment about the Einstein's principle of relativity. Because of the limited speed of propagation of information, we can't think about interaction between particles as instantaneous anymore. Instead, we should consider that if we change the position of a particle, the others are influenced only after a certain time interval. Therefore, the concept of a field mediating the interactions becomes a very powerful tool to describe physics. So, we should think about interactions in the following way: a particle interacts with a sort of field and then the latter interacts with another particle, and so on, but the interactions must be separated by a certain delay of time. Further, it's natural to have a unified formalism to describe physical entities, so we'll apply Lagrangian theory also for fields. Therefore, we'll seek to generalize the formalism to continuous systems and apply it to find some important equations, treating them as examples for now, but keeping in mind that the fields considered represent physical particles. In particular, we'll retrieve these equations to interpret the Klein-Gordon field as representing a particle with spin 0 and mass and the electromagnetic field the photon, with spin 1 and no mass. Finally, we'll seek to find Proca's equation, to describe particles with spin 1 and mass.

3.1 A first model: the rod

Trying to generalize Lagrangian formalism for continuous systems, we aim to show that it's possible to view one of them as the limit of a suitable discrete mechanical system.

We should start from a simplified model, studying a one dimensional elastic rod that can undergo small longitudinal vibrations. We can think about the rod as composed by an infinite number of massive particles with mass m , connected

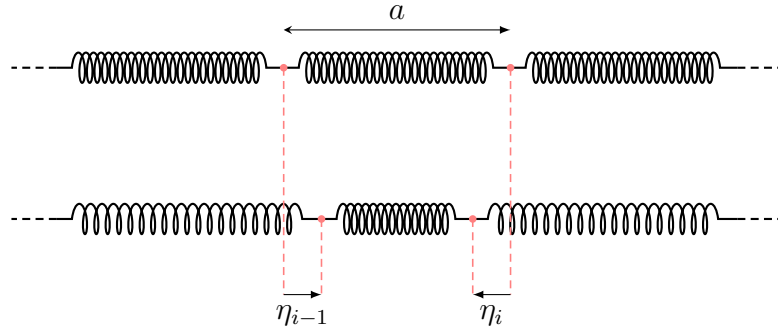


Figure 3.1 One dimensional elastic rod.

by massless identical springs with constant k and distanced a from each other, as shown in fig. 3.1. Let η_i be the displacement of the i -th particle from its equilibrium position. From classical mechanics, it is known that the kinetic and potential energy of the system can be written as

$$T = \frac{1}{2} \sum_i m \dot{\eta}_i^2, \quad V = \frac{1}{2} \sum_i k (\eta_{i+1} - \eta_i)^2,$$

respectively. Being the system conservative, the Lagrangian can be written as $L = T - V$, namely

$$L = \frac{1}{2} \sum_i [m \dot{\eta}_i^2 - k (\eta_{i+1} - \eta_i)^2] = \frac{1}{2} \sum_i a \left[\frac{m}{a} \dot{\eta}_i^2 - ka \left(\frac{\eta_{i+1} - \eta_i}{a} \right)^2 \right], \quad (3.1)$$

where the last step retrace what we have done in (1.10) dividing and multiplying by Δx , and will be useful while taking the limit as $a \rightarrow 0$.

The related Euler-Lagrange equations (2.4) for the Lagrangian (3.1) are

$$\frac{m}{a} \ddot{\eta}_i - ka \left(\frac{\eta_{i+1} - \eta_i}{a^2} \right) + ka \left(\frac{\eta_i - \eta_{i-1}}{a^2} \right) = 0. \quad (3.2)$$

Let's further suppose that the rod obeys to Hooke's law, so that the force F exerted on the rod is proportional to the elongation per unit length ξ

$$F = Y \xi, \quad (3.3)$$

where Y is Young's modulus. Given two contiguous points of our discrete system, the elongation per unit length can be written as

$$\xi = \frac{\eta_{i+1} - \eta_i}{a},$$

so that the force necessary to stretch the corresponding spring by this amount is

$$F = k(\eta_{i+1} - \eta_i) = ka\xi. \quad (3.4)$$

Let's consider the continuous limit. As $a \rightarrow 0$, $m/a \rightarrow \mu$, which is the mass per unit length of the rod. Comparing eq. (3.3) and (3.4), we can conclude that $ka \rightarrow Y$. Further, in eq. (3.1) we should replace η_i by $\eta(x)$,

$$\eta_i \rightarrow \eta(x), \quad (3.5)$$

where x is the position. We shall think about it as a merely continuous index, replacing i . Then, we can write

$$\frac{\eta_{i+1} - \eta_i}{a} = \frac{\eta(x+a) - \eta(x)}{a} \rightarrow \frac{d\eta}{dx}(x).$$

Gathering the previous considerations, the Lagrangian (3.1) for the continuous system becomes

$$L = \frac{1}{2} \int dx \left[\mu \dot{\eta}^2 - Y \left(\frac{d\eta}{dx} \right)^2 \right], \quad (3.6)$$

and the equation of motion (3.2) is

$$\mu \frac{d^2\eta}{dt^2} - Y \frac{d^2\eta}{dx^2} = 0. \quad (3.7)$$

Inspecting eq. (3.6), let's write the Lagrangian as

$$L = \int dx \mathcal{L}, \quad (3.8)$$

where \mathcal{L} is the *Lagrangian density*. In our example, it is

$$\mathcal{L} = \frac{1}{2} \left[\mu \left(\frac{d\eta}{dt} \right)^2 - Y \left(\frac{d\eta}{dx} \right)^2 \right], \quad (3.9)$$

where the symbols of the total derivative d/dt and d/dx are not ambiguous because of the independence between x and t .

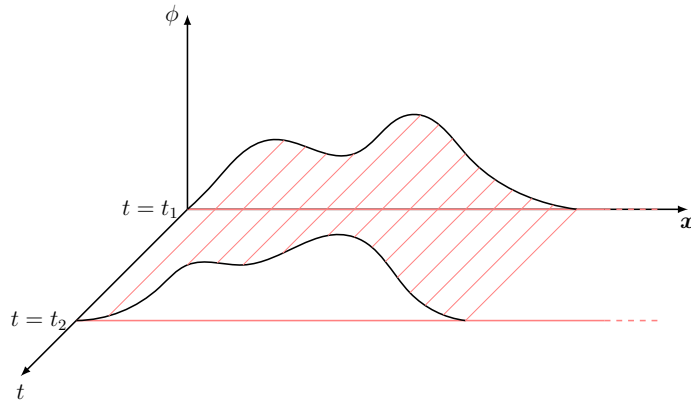


Figure 3.2 Integration domain Ω with boundary $\partial\Omega$.

3.2 From discrete to continuous systems

Let's start from the previous example to extrapolate a general formalism compatible with continuous fields.

Eq. (3.5) taught us that in classical field theory both space and time should be considered as parameters. From now on, we'll deal with the dynamics of $\phi(\mathbf{x}, t)$, which is a system with an infinite number of degrees of freedom. The following considerations can be easily generalized for a set of fields $\phi_a(\mathbf{x}, t)$, parametrized by the index a .

Let Ω be the domain in which the field is defined, as showed in fig. 3.2, where t can vary within the interval $t \in [t_1, t_2]$ and \mathbf{x} is defined in the subset V of the three-dimensional space, so that $\Omega = [t_1, t_2] \times V$. Its boundary will be $\partial\Omega$, consisting of the points belonging to the surface $S_V = \partial V$, at time t_1 or t_2 .

In analogy with classical particle theory (cf. eq. (2.1)), we shall suppose that the Lagrangian describing the continuous system can be cast in the form

$$L(\phi(\mathbf{x}, t), \dot{\phi}(\mathbf{x}, t)), \quad (3.10)$$

which depends on the values of the field and its time derivative, at a certain instant t in $[t_1, t_2]$ and at every point in V . Notice that the Lagrangian doesn't depend explicitly on \mathbf{x} and t , meaning that we are considering a closed system.

Furthermore, inspecting the example of the rod, let's try to apply the reasoning in reverse. The continuous system (the rod) has been discretized by dividing the 1-dimensional space of the coordinates into cells of volume a . The resultant Lagrangian (3.1) depended on the values of the approximants of the field at each cell.

Similarly, let's discretize the volume V into tiny cells of volume δV_i and call $\phi_i(t)$ the approximant of the field in the i -th cell (its mean value). Then, the discretized system will be described through a Lagrangian $L^*(\phi_i(t), \dot{\phi}_i)$ depending on infinitely many variables, and in the continuous limit it should be

$$L^*(\phi_i(t), \dot{\phi}_i) \xrightarrow{\delta V_i \rightarrow 0} L(\phi(\mathbf{x}, t), \dot{\phi}(\mathbf{x}, t)). \quad (3.11)$$

It's convenient to think about the Lagrangian (3.10) as a functional of the functions $\phi(\mathbf{x}, t)$ and $\dot{\phi}(\mathbf{x}, t)$ and carry out an independent variation of the field at each point \mathbf{x} , namely

$$\phi(\mathbf{x}) \rightarrow \phi(\mathbf{x}) + \delta\phi(\mathbf{x}).$$

Extending eq. (1.17) to a functional of two functions, we can write

$$\delta L[\phi, \dot{\phi}] = \int d^3x \left[\frac{\delta L}{\delta\phi(\mathbf{x}, t)} \delta\phi(\mathbf{x}, t) + \frac{\delta L}{\delta\dot{\phi}(\mathbf{x}, t)} \delta\dot{\phi}(\mathbf{x}, t) \right]. \quad (3.12)$$

On the other hand, since the Lagrangian of the discretized system, L^* , defined by eq. (3.11), is an ordinary function of the variables ϕ_i and $\dot{\phi}_i$, its variation can be written as

$$\begin{aligned} \delta L^*(\phi_i, \dot{\phi}_i) &= L^*(\phi_i + \delta\phi_i, \dot{\phi}_i + \delta\dot{\phi}_i) - L^*(\phi_i, \dot{\phi}_i) \\ &= \sum_i \left(\frac{\partial L^*}{\partial\phi_i(t)} \delta\phi_i(t) + \frac{\partial L^*}{\partial\dot{\phi}_i(t)} \delta\dot{\phi}_i(t) \right) \\ &= \sum_i \frac{1}{\delta V_i} \left(\frac{\partial L^*}{\partial\phi_i(t)} \delta\phi_i(t) + \frac{\partial L^*}{\partial\dot{\phi}_i(t)} \delta\dot{\phi}_i(t) \right) \delta V_i, \end{aligned} \quad (3.13)$$

where we have multiplied and divided by δV_i , in analogy with (1.10) and (3.1).

Comparing eq. (3.12) and (3.13), since variations at distinct points are independent of one another, we may expect the following identification

$$\frac{\delta L}{\delta\phi(\mathbf{x}, t)} \equiv \lim_{\delta V_i \rightarrow 0} \frac{1}{\delta V_i} \frac{\partial L^*}{\partial\phi_i(t)}, \quad (3.14a)$$

$$\frac{\delta L}{\delta\dot{\phi}(\mathbf{x}, t)} \equiv \lim_{\delta V_i \rightarrow 0} \frac{1}{\delta V_i} \frac{\partial L^*}{\partial\dot{\phi}_i(t)}, \quad (3.14b)$$

where \mathbf{x} belongs to the i -th cell.

Further, since the Euler-Lagrange equations for the discretized system are

$$\frac{\partial L}{\partial \phi_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}_i} = 0, \quad (3.15)$$

using (3.14) in (3.15), we expect

$$\frac{\delta L}{\delta \phi(\mathbf{x}, t)} - \frac{d}{dt} \frac{\delta L}{\delta \dot{\phi}(\mathbf{x}, t)} = 0. \quad (3.16)$$

Let's inspect the Lagrangian (3.1) of the discretized version of the rod. It can be written as the sum of the Lagrangians computed in each cell. It's reasonable to expect the same behaviour in the general case, since we don't want different points in space influence each other instantly, not to violate the propagation of information's speed limit. In addition, in the continuous limit of (3.1) the term $(\eta_{i+1} - \eta_i)/a$ approaches to the gradient of the field. We suppose there is a similar property for the general case. Summing up, we may write

$$\begin{aligned} L^*(\phi_i(t), \dot{\phi}_i(t)) &= \sum_i L_i^*(\phi_i(t), \nabla \phi_i(t), \dot{\phi}_i(t)) \\ &= \sum_i \frac{1}{\delta V_i} L_i^*(\phi_i(t), \nabla \phi_i(t), \dot{\phi}_i(t)) \delta V_i, \end{aligned}$$

where the last step will be useful while taking the continuous limit, which is

$$L(\phi(t), \dot{\phi}(t)) = \int_V d^3x \mathcal{L}(\phi(x^\mu), \nabla(\phi(x^\mu)), \dot{\phi}(x^\mu)), \quad (3.17)$$

where $x^\mu \equiv (ct, \mathbf{x})$ and \mathcal{L} is called *Lagrangian density*. For now, we'll write x^μ just for notation convenience, but the generalization on Minkowski's space is straightforward. The appearance of the field's gradient in the Lagrangian density will be useful to ensure its Lorentz invariance.

The action of the system is

$$S[\phi] = \int_{t_1}^{t_2} dt L(t) = \int_{t_1}^{t_2} dt \int_V d^3x \mathcal{L}(x^\mu) = \frac{1}{c} \int_\Omega d^4x^\mu \mathcal{L}(x^\mu),$$

where the Lagrangian density has been defined as

$$\mathcal{L}(\phi(x^\mu), \nabla \phi(x^\mu), \dot{\phi}(x^\mu)) \equiv \lim_{\delta V_i \rightarrow 0} \frac{1}{\delta V_i} L_i^*(\phi_i(t), \nabla \phi_i(t), \dot{\phi}_i(t)).$$

Let's finally consider arbitrary variations of the field $\phi(x^\mu)$ which vanish at the boundary $\partial\Omega$. Writing the Lagrangian as in (3.17), we can write its variation as

$$\begin{aligned} \delta L &= \int d^3x \left[\frac{\partial \mathcal{L}}{\partial \phi(\mathbf{x}, t)} \delta \phi(\mathbf{x}, t) + \frac{\partial \mathcal{L}}{\partial \nabla \phi(\mathbf{x}, t)} \delta \nabla \phi(\mathbf{x}, t) + \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\mathbf{x}, t)} \delta \dot{\phi}(\mathbf{x}, t) \right] \\ &= \int d^3x \left\{ \left[\frac{\partial \mathcal{L}}{\partial \phi(\mathbf{x}, t)} - \nabla \frac{\partial \mathcal{L}}{\partial \nabla \phi(\mathbf{x}, t)} \right] \delta \phi(\mathbf{x}, t) - \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\mathbf{x}, t)} \delta \dot{\phi}(\mathbf{x}, t) \right\}, \end{aligned} \quad (3.18)$$

and comparing the latter with (3.12), we can identify

$$\frac{\delta L}{\delta \phi(x^\mu)} = \left[\frac{\partial \mathcal{L}(x^\mu)}{\partial \phi(\mathbf{x}, t)} - \nabla \frac{\partial \mathcal{L}(x^\mu)}{\partial \nabla \phi(x^\mu)} \right], \quad (3.19a)$$

$$\frac{\delta L}{\delta \dot{\phi}(x^\mu)} = \frac{\partial \mathcal{L}(x^\mu)}{\partial \dot{\phi}(x^\mu)}. \quad (3.19b)$$

Using (3.19) in (3.16) we finally have the Euler-Lagrange equation for the continuous system, namely

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial^\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0. \quad (3.20)$$

3.3 Least action principle

Founded the Euler-Lagrange equation for a continuous system, we aim to show that it's possible to derive it from a variational principle, forcing the first variation of the action to vanish. From now on we'll call \mathcal{L} simply Lagrangian.

Let's suppose the system of interest is described through a Lagrangian

$$\mathcal{L}(\phi, \partial_\mu \phi, x^\mu), \quad (3.21)$$

which, this time, can be directly dependent on x^μ , opening the possibility for the field to interact with external sources. We suppose the domain is again Ω , with boundary $\partial\Omega$, as in fig. 3.2. Let's perform a variation both of the field ϕ and of the coordinates x^μ , which vanishes on the boundary $\partial\Omega$,

$$x^\mu \rightarrow x'^\mu = x^\mu + \delta x'^\mu, \quad (3.22a)$$

$$\phi(x) \rightarrow \phi'(x) = \phi(x) + \delta \phi(x), \quad (3.22b)$$

$$\delta \phi = 0, \quad \delta x^\mu = 0 \quad \text{on } \partial\Omega. \quad (3.22c)$$

Notice that we are considering a *functional variation* of ϕ , that is, ϕ and ϕ' are compared at the same point in spacetime x .

The action of the system is

$$S = \int_{\Omega} d^4x \mathcal{L}(\phi, \partial_{\mu}\phi, x^{\mu}), \quad (3.23)$$

and its variation, which follows from (3.22), is

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

Following the least action principle 2.1, we shall impose $\delta S = 0$.

As seen by (2.27), a consequence of a Lorentz transformation is the invariance of d^4x . However, if we consider a general transformation, we would have

$$d^4x' = |J(x)|d^4x = \det\left\{\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right\}d^4x \simeq \left[1 + \partial_{\mu}(\delta x^{\mu})\right]d^4x, \quad (3.24)$$

at first order in the variation δx^{μ} .

Proof. In general

$$d^4x' = |J(x)|d^4x,$$

where $J(x)$ is the Jacobian of the transformation $x \rightarrow x'$, namely

$$J(x) = \det\left\{\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right\}.$$

Since the transformation we relate to is (3.22a), we must compute

$$J(x) = \det(\delta_{\nu}^{\mu} + \partial_{\nu}\delta x^{\mu}).$$

Now, for a generic matrix A , we can compute the determinant as

$$\det\{A\} = e^{\text{Tr}\{\ln A\}}.$$

In this case, to first order in δx^{μ} , we have

$$\text{Tr}\{\ln(\delta_{\nu}^{\mu} + \partial_{\nu}\delta x^{\mu})\} \simeq \text{Tr}\{\partial_{\nu}\delta x^{\mu}\} = \partial_{\mu}\delta x^{\mu},$$

then, the Jacobian is

$$J(x) = e^{\partial_\mu \delta x^\mu} \simeq 1 + \partial_\mu \delta x^\mu,$$

as claimed. ■

Therefore, by (3.24), the variation of the action is

$$\delta S = \int_{\Omega} d^4x (\delta \mathcal{L} + \mathcal{L} \partial_\mu (\delta x^\mu)) = \int_{\Omega} d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta \phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta (\partial_\mu \phi) + \frac{\partial \mathcal{L}}{\partial x^\mu} \delta x^\mu + \mathcal{L} \partial_\mu (\delta x^\mu) \right].$$

Since $\delta (\partial_\mu \phi) = \partial_\mu (\delta \phi)$, we can write

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x^\mu} \delta x^\mu + \mathcal{L} \partial_\mu (\delta x^\mu) &= \partial_\mu (\mathcal{L} \delta x^\mu), \\ \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\mu (\delta \phi) &= \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta \phi \right] - \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right] \delta \phi. \end{aligned}$$

Gathering the previous considerations and using Gauss theorem we have

$$\delta S = \int_{\Omega} d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) \right] \delta \phi + \int_{\partial \Omega} d\sigma_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta \phi + \mathcal{L} \delta x^\mu \right], \quad (3.25)$$

which must vanish as supposed above.

The second integral vanishes because of (3.22c), while the first one, taking into account the fundamental lemma L.1, leads to

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial^\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0, \quad (3.26)$$

showing that it's possible to recover Euler-Lagrange equation from Hamilton's principle.

Let's finally highlight the properties a well written Lagrangian must satisfy. It must be real for the equations of motion to have a physical interpretation. In addition, in order for the theory to be Lorentz invariant, its action, or equivalently, its Lagrangian, must have the same property. Indeed, through (3.26), we would obtain covariant equations of motion.

3.4 Noether's theorem

Noether's theorem relates *symmetries* of the system to *conserved quantities*. Let's consider the following infinitesimal transformations of the coordinates and the field

$$x^\mu \rightarrow x'^\mu = x^\mu + \delta x^\mu, \quad (3.27a)$$

$$\phi(x^\mu) \rightarrow \phi'(x'^\mu) = \phi(x^\mu) + \Delta\phi(x^\mu), \quad (3.27b)$$

where $\Delta\phi(x^\mu)$ is the *total variation* of ϕ , not to be confused with the *functional variation* $\delta\phi(x^\mu) = \phi'(x^\mu) - \phi(x^\mu)$, computed at a fixed point of spacetime.

Let's remark that the theorem can be applied for continuous transformations, not discrete ones. To comprehend how the Lagrangian changes after (3.27a) and (3.27b), let's first find a relation, to first order in δx , between $\delta\phi$ and $\Delta\phi$. Combining eq. (3.22b) and (3.27b), we have

$$\Delta\phi = \phi'(x') - \phi(x) = \phi'(x') - \phi(x') + \phi(x') - \phi(x) \simeq \delta\phi + (\partial_\mu\phi)\delta x^\mu. \quad (3.28)$$

Furthermore, for later convenience, it's useful to define the *energy-momentum tensor*

$$\theta^\mu{}_\nu \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \delta^\mu_\nu\mathcal{L}. \quad (3.29)$$

Let's now consider (3.27a) and (3.27b), recalling that now Ω is an arbitrary hypersurface and there isn't the requirement $\delta x^\mu = 0$ and $\delta\phi = 0$ on $\partial\Omega$. By adding and subtracting a term from (3.25) yields

$$\begin{aligned} \delta S &= \int_\Omega d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \right) \right] \delta\phi \\ &+ \int_{\partial\Omega} d\sigma_\mu \left\{ \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} [\delta\phi + (\partial_\nu\phi)\delta x^\nu] - \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \delta^\mu_\nu\mathcal{L} \right] \delta x^\nu \right\}. \end{aligned}$$

Using (3.28) and (3.29), we have

$$\delta S = \int_\Omega d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \right) \right] \delta\phi + \int_{\partial\Omega} d\sigma_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \Delta\phi - \theta^\mu{}_\nu \delta x^\nu \right]. \quad (3.30)$$

Let's suppose the system has a *symmetry*, in particular that S is invariant under

$$\Delta x^\mu = \chi^\mu{}_\nu \delta\omega^\nu, \quad \Delta\phi = \Phi_\mu \delta\omega^\mu, \quad (3.31)$$

where $\delta\omega^\nu$ is an infinitesimal parameter, ν in general is a multiplet of indices, χ^μ_ν is a matrix and Φ_μ a set of scalars. Ignoring the case in which the action varies by a boundary term, we shall impose $\delta S = 0$. Besides, assuming that the transformed ϕ follows the Euler-Lagrange equation (3.26) and using (3.31) in (3.30), we can write

$$\int_{\partial\Omega} d\sigma_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \Phi_\nu - \theta^\mu_\rho \chi^\rho_\nu \right] \delta\omega^\nu = 0.$$

Inspecting the previous equation, since ω^ν are arbitrary independent variables, every term of the summation over ν must vanish. This leads to

$$\int_{\partial\Omega} d\sigma_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \Phi_\nu - \theta^\mu_\rho \chi^\rho_\nu \right] = 0. \quad (3.32)$$

Defining *Noether's current*

$$J^\mu_\nu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \Phi_\nu - \theta^\mu_\rho \chi^\rho_\nu, \quad (3.33)$$

and using Gauss theorem in (3.32) and the arbitrariness of Ω , we obtain

$$\partial_\mu J^\mu_\nu = 0, \quad (3.34)$$

that, evidently, is a *conservation equation* for the current J^μ_ν , that follows from the symmetry (3.31).

We are now in a position to define *Noether's charges* as

$$Q_\nu(\sigma) = \int_\sigma d\sigma_\mu J^\mu_\nu, \quad (3.35)$$

where σ is a space-like hypersurface that extends to infinity, covering all Minkowski's space-time. We'll assume that the field and its derivatives vanish rapidly towards space infinity. Because of the condition (3.34) and by (2.30), we can conclude that the integral above is independent on the specific space-like hypersurface we choose. Therefore, considering the surface with element (2.28) (cf. sec. 2.2), by (2.31) we have

$$Q_\nu(t) = \int d^3x J^0_\nu(x). \quad (3.36)$$

Integrating (3.34) over \mathbb{R}^3 we obtain

$$0 = \int d^3x \partial_\mu J^\mu_\nu = \int d^3x \partial_0 J^0_\nu + \int d^3x \partial_i J^i_\nu.$$

If we suppose that J is zero at spacial infinity, the last term vanishes as a consequence of 3-dimensional Gauss theorem. We then have

$$0 = \int d^3x \partial_0 J^0_\nu = \frac{1}{c} \frac{d}{dt} \int d^3x J^0_\nu = \frac{1}{c} \frac{dQ_\nu}{dt},$$

where in the last step we have used (3.36).

Definitely, Noether's theorem states that

$$\frac{dQ_\nu}{dt} = 0, \tag{3.37}$$

where Q_ν is the conserved quantity related to the symmetry (3.31), and J^0_ν can be interpreted as its density.

3.5 Scalar fields and Klein-Gordon Lagrangian

Let's apply the formalism developed to study the dynamics of a generic real scalar field $\phi(x)$. Recalling that the Lagrangian of the system must be real and Lorentz invariant, it can be written as

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi), \tag{3.38}$$

where the first part is the kinetic term, while the second one represents any potential depending on the field. Using Euler-Lagrange equation (3.26) yields

$$\square \phi + \frac{dV}{d\phi} = 0. \tag{3.39}$$

Furthermore, if the scalar field is complex, namely $\phi(x) = \phi_1 + i\phi_2$, the Lagrangian, that must be real, is

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - V(\phi^* \phi). \tag{3.40}$$

In this case ϕ_1 and ϕ_2 , or equivalently, ϕ and ϕ^* , must be treated independently, and by Euler-Lagrange equation (3.26) for ϕ and ϕ^* we have, respectively

$$\square\phi^* + \frac{\partial V}{\partial\phi} = 0, \quad (3.41a)$$

$$\square\phi + \frac{\partial V}{\partial\phi^*} = 0. \quad (3.41b)$$

The simplest case that can be studied is the *Klein-Gordon* field, whose potential is $V(\phi) = \frac{1}{2}\mu^2\phi^2$, with μ a real constant. Eq. (3.38) becomes

$$\mathcal{L}_{\text{KG}}^{\mathbb{R}} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\mu^2\phi^2, \quad (3.42)$$

and (3.39),

$$(\square + \mu^2)\phi = 0. \quad (3.43)$$

The generalization for a complex field, using (3.40) and (3.41) is straightforward, yielding

$$\mathcal{L}_{\text{KG}}^{\mathbb{C}} = \partial_\mu\phi^*\partial^\mu\phi - \mu^2\phi^*\phi, \quad (3.44)$$

and

$$(\square + \mu^2)\phi^* = 0, \quad (3.45a)$$

$$(\square + \mu^2)\phi = 0. \quad (3.45b)$$

If we look for a solution with a plane wave ansatz

$$\phi_p(x) \simeq e^{-ik_\alpha x^\alpha}, \quad (3.46)$$

with $k^\mu = (k^0, \mathbf{k})$ and $x^\mu = (x^0, \mathbf{x}) = (ct, \mathbf{x})$, by substitution in (3.45), we would obtain the following dispersion relation

$$-k_\alpha k^\alpha + \mu^2 = 0, \quad (3.47)$$

that has interesting implications in quantum field theory, as we shall see.

3.6 Lagrangian for electromagnetic field

Electromagnetic field in vacuum. Let's recall that in covariant form electromagnetism can be described through eq. (2.47a) and (2.47b), that, for clarity, are displayed below

$$\begin{aligned}\partial_\mu F^{\mu\nu} &= \frac{4\pi}{c} J^\nu, \\ \partial_\mu \tilde{F}^{\mu\nu} &= 0.\end{aligned}$$

Concentrating on the equation giving the dynamics of the field, eq. (2.47a), it can be written using the four-potential, through (2.52), as (2.53), namely

$$\begin{aligned}F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu, \\ \square A^\nu - \partial^\nu(\partial_\mu A^\mu) &= \frac{4\pi}{c} J^\nu.\end{aligned}$$

As turns out studying the quantization of the electromagnetic field, it's convenient to proceed using the four-potential A^μ . Therefore, the Lagrangian \mathcal{L} must contain the four-potential A^μ and its first derivatives $\partial_\mu A^\mu$. Further, electromagnetism is invariant under gauge transformations (cf. sec. 2.5), so we expect the Lagrangian to reflect this property. Finally, it must be Lorentz invariant. We should use these guidelines to guess a form for \mathcal{L} .

It's convenient to write the Lagrangian using $F^{\mu\nu}$, since, this way, it would contain A^ν and, by (2.55), it would be gauge invariant.

As seen in sec. 2.4, we can derive two scalars from the field-strength tensor $F^{\mu\nu}$, that is, $F^{\mu\nu} F_{\mu\nu}$ by eq. (2.50a) and $\tilde{F}^{\mu\nu} F_{\mu\nu}$ by eq. (2.50b). We could use one of these to write the Lagrangian, but which one? Since $\tilde{F}^{\mu\nu} F_{\mu\nu}$ is a four-divergence and so doesn't contribute to the equations of motion, $F^{\mu\nu} F_{\mu\nu}$ turns out to be the proper one.

Proof. Using the definition of the dual tensor (2.23) and eq. (2.52),

$$\begin{aligned}\tilde{F}^{\mu\nu} F_{\mu\nu} &= \frac{1}{2} \varepsilon^{\mu\nu\rho\tau} F_{\rho\tau} F_{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\rho\tau} (\partial_\mu A_\nu - \partial_\nu A_\mu) (\partial_\rho A_\tau - \partial_\tau A_\rho) \\ &= \frac{1}{2} \varepsilon^{\mu\nu\rho\tau} [\partial_\mu A_\nu \partial_\rho A_\tau - \partial_\nu A_\mu \partial_\rho A_\tau - \partial_\mu A_\nu \partial_\tau A_\rho + \partial_\nu A_\mu \partial_\tau A_\rho].\end{aligned}$$

Swapping two dummy indices in the parenthesis and exchanging the same indices on the Levi-Civita symbol (2.24), changing sign since the latter is completely antisymmetric, we can show that the four terms of the summation are identical. For

example, let's show that the second term is equal to the first one,

$$-\frac{1}{2}\varepsilon^{\mu\nu\rho\tau}\partial_\nu A_\mu\partial_\rho A_\tau = -\frac{1}{2}\varepsilon^{\nu\mu\rho\tau}\partial_\mu A_\nu\partial_\rho A_\tau = +\frac{1}{2}\varepsilon^{\mu\nu\rho\tau}\partial_\mu A_\nu\partial_\rho A_\tau.$$

Finally, we can write

$$\begin{aligned}\tilde{F}^{\mu\nu}F_{\mu\nu} &= 2\varepsilon^{\mu\nu\rho\tau}\partial_\mu A_\nu\partial_\rho A_\tau = 2\partial_\mu(\varepsilon^{\mu\nu\rho\tau}A_\nu\partial_\rho A_\tau) - 2\varepsilon^{\mu\nu\rho\tau}A_\nu\partial_\mu\partial_\rho A_\tau \\ &= 2\partial_\mu(\varepsilon^{\mu\nu\rho\tau}A_\nu\partial_\rho A_\tau),\end{aligned}$$

where in the last step, the second term of the sum vanishes because it represents a contraction in $\mu-\rho$ of a quantity that is antisymmetric in $\mu-\rho$ (i.e., $\varepsilon^{\mu\nu\rho\tau}$) and a quantity that is symmetric (i.e., $A_\nu\partial_\mu\partial_\rho A_\tau$). ■

Multiplied by a conventional constant, the Lagrangian of the electromagnetic field with no sources is

$$\mathcal{L}_{\text{em}}^0 = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu}. \quad (3.49)$$

Proof. Let's apply Euler-Lagrange equation (3.26) to (3.49). Since $\mathcal{L}_{\text{em}}^0$ contains only $F^{\mu\nu}$ which, by (2.52), contains only the derivatives of A^μ , we have

$$\frac{\partial\mathcal{L}_{\text{em}}^0}{\partial A_\nu} = 0.$$

Further, we have

$$\begin{aligned}\frac{\partial\mathcal{L}_{\text{em}}^0}{\partial(\partial_\mu A_\nu)} &= \frac{\partial\mathcal{L}_{\text{em}}^0}{\partial F_{\rho\tau}}\frac{\partial F_{\rho\tau}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{8\pi}F^{\rho\tau}(\delta_\rho^\mu\delta_\tau^\nu - \delta_\tau^\mu\delta_\rho^\nu) \\ &= -\frac{1}{8\pi}(F^{\mu\nu} - F^{\nu\mu}) = -\frac{1}{4\pi}F^{\mu\nu}.\end{aligned}$$

Applying Euler-Lagrange equation

$$\partial_\mu\frac{\partial\mathcal{L}_{\text{em}}^0}{\partial(\partial_\mu A_\nu)} = \frac{\partial\mathcal{L}_{\text{em}}^0}{\partial A_\nu}, \quad (3.50)$$

yields

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

that is Maxwell's equation (2.47a) with no sources. ■

Electromagnetic field with sources. We aim to add a term to $\mathcal{L}_{\text{em}}^0$ to represent the sources. In particular, we expect a term that leads to $(4\pi/c)J^\nu$ while applying

Euler-Lagrange equation without any influence on $\mathcal{L}_{\text{em}}^0$. This can be achieved by

$$\mathcal{L}_{\text{em}} = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}J^\mu A_\mu. \quad (3.51)$$

Applying (3.50) to (3.51), we would find (2.47a), as expected.

However, since (3.51) contains explicitly A^μ , performing a gauge transformation (2.54), \mathcal{L}_{em} will change,

$$\mathcal{L}'_{\text{em}} = -\frac{1}{16\pi}F'_{\mu\nu}F'^{\mu\nu} - \frac{1}{c}J^\mu A'_\mu = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}J^\mu A_\mu + \frac{1}{c}J^\mu \partial_\mu \chi = \mathcal{L}_{\text{em}} + \frac{1}{c}J^\mu \partial_\mu \chi.$$

Nevertheless, we can write

$$J^\mu \partial_\mu \chi = \partial_\mu (J^\mu \chi) - \chi \partial_\mu J^\mu = \partial_\mu (J^\mu \chi),$$

because of the continuity equation (2.41). Therefore, we have

$$\mathcal{L}'_{\text{em}} = \mathcal{L}_{\text{em}} + \frac{1}{c}\partial_\mu (J^\mu \chi),$$

and being the last term a four-divergence, it doesn't contribute to the equations of motion, that are, as a result, gauge invariant. Therefore, gauge invariance is related to the conservation of charge.

Towards Proca's equation. As known, the mediator of electromagnetic interactions is the *photon*, a particle with spin 1 and no mass. Let's consider the electromagnetic field in vacuum, represented by $\mathcal{L}_{\text{em}}^0$, eq. (3.49). Interpreting the field A^μ as representing the photon, we expect $\mathcal{L}_{\text{em}}^0$ not to contain any mass terms. On the other hand, as we shall see, the Klein-Gordon field describes a particle with spin 0 and mass. Inspecting eq. (3.42), we can guess that the quadratic term in ϕ represents the mass term of the underlying particle. This turns out to be correct. Therefore, if we wish to describe a particle with spin 1 and mass we could write the Lagrangian

$$\mathcal{L}_{\text{proca}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{\mu^2}{2}A_\rho A^\rho \quad (3.52)$$

and Euler-Lagrange equation (3.26) leads to the so-called *Proca equation*

$$\square A^\nu - \partial^\nu (\partial_\mu A^\mu) + \mu^2 A^\nu = 0. \quad (3.53)$$

Proof. Using eq. (2.52)

$$\mathcal{L}_{\text{proca}} = -\frac{1}{2}\partial_{\mu}A_{\nu}\partial^{\mu}A^{\nu} + \frac{1}{2}\partial_{\mu}A_{\nu}\partial^{\nu}A^{\mu} + \frac{\mu^2}{2}A_{\rho}A^{\rho},$$

so we have

$$\frac{\partial\mathcal{L}_{\text{proca}}}{\partial(\partial_{\mu}A_{\nu})} = -\partial^{\mu}A^{\nu} + \partial^{\nu}A^{\mu} = -F^{\mu\nu}, \quad \frac{\partial\mathcal{L}_{\text{proca}}}{\partial A_{\nu}} = \mu^2 A^{\rho}\delta_{\rho}^{\nu} = \mu^2 A^{\nu}.$$

Applying Euler-Lagrange equation

$$\partial_{\mu}\frac{\partial\mathcal{L}_{\text{proca}}}{\partial(\partial_{\mu}A_{\nu})} = \frac{\partial\mathcal{L}_{\text{proca}}}{\partial A_{\nu}},$$

we have

$$\partial_{\mu}F^{\mu\nu} = -\mu^2 A^{\nu}.$$

Using again (2.52) and changing sign yields

$$\partial_{\mu}(\partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}) = -\mu^2 A^{\nu},$$

the conclusion is straightforward. ■

An interesting property appears if we differentiate Proca's equation (3.53). We would obtain

$$\partial_{\nu}\partial_{\mu}F^{\mu\nu} + \mu^2\partial_{\nu}A^{\nu} = 0.$$

Since $\partial_{\nu}\partial_{\mu}$ is symmetric while $F^{\mu\nu}$ is antisymmetric, their contraction vanishes and remains

$$\partial_{\nu}A^{\nu} = 0, \tag{3.54}$$

a condition similar to (2.56). However, while in electromagnetism it was related to the specific choice of gauge fixing, here the condition must always be satisfied. Therefore, we can cast Proca's equation as

$$(\square + \mu^2)A^{\nu} = 0, \tag{3.55a}$$

$$\partial_{\nu}A^{\nu} = 0, \tag{3.55b}$$

equivalent to four Klein-Gordon fields (4.24) with the constraint (3.54).

RELATIVISTIC QUANTUM MECHANICS

As seen in chapter 3, it's natural to apply Lagrangian formalism, other than for particles, also for fields. In classical physics, where classical is meant to be non-quantum, particles and fields are two distinguished entities, the former describing localized objects and the latter extended physical quantities which allow interactions to happen. In quantum field theory this dualism is outdated and replaced by a unified physical entity: the quantum field.

This possibility arises trying to generalize Schrödinger equation to relativistic particles. Indeed, this process leads to states with negative energy, which make impossible the probabilistic interpretation that is so natural in quantum mechanics. A way to overcome these difficulties is to give up on the particle interpretation and try to describe a system of an infinite number of indistinguishable particles through a field. That said, our road will be to follow the first approach, the so-called first quantization, in the framework of relativistic quantum mechanics.

Therefore, we'll deal with wave equations describing the dynamics of quantum particles. In particular, we'll approach Klein-Gordon field as describing a massive particle with $s = 0$. The goal will be the scalar electrodynamics, which describes interactions between the previously described particles and the electromagnetic field. Then, we'll seek for a general method to obtain an interaction theory from a free one, leading to the gauge principle. Ultimately, we'll approach the latter critically.

4.1 Quantum mechanics in a nutshell

Let's revise some concepts about the foundation of quantum mechanics. Particles are described by wave functions $\psi(\mathbf{x}, t)$ obeying the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + U\psi. \quad (4.1)$$

In particular, particles' momentum \mathbf{p} and energy E are related to wave functions' angular frequencies ω and wave vectors \mathbf{k} by the Planck-Einstein relations

$$\begin{aligned} E &= \hbar\omega, \\ \mathbf{p} &= \hbar\mathbf{k}, \end{aligned}$$

that can be written in covariant form as

$$p^\mu = \hbar k^\mu, \quad p^\mu = (E/c, \mathbf{p}), \quad k^\mu = (\omega/c, \mathbf{k}). \quad (4.3)$$

More precisely, it's convenient to formulate the theory in Dirac's notation, where a state s of the system is codified by a normalized ket $|\psi\rangle$ defined in a Hilbert space \mathcal{H} , while its observables Q are represented by a selfadjoint¹ linear operator \hat{A} of \mathcal{H} . As known from linear algebra, being the operator selfadjoint, it always admits a generalized orthonormal basis of eigenkets, fulfilling the following requirements,

$$\hat{A} |\xi\rangle = |\xi\rangle x(\xi), \quad (4.4a)$$

$$\langle \xi | \eta \rangle = \delta(\xi, \eta), \quad (4.4b)$$

$$\int |\xi\rangle d\mu(\xi) \langle \xi| = \hat{1}, \quad (4.4c)$$

where $d\mu(\xi)$ is a suitable measure and $\delta(\xi, \eta)$ a generalized delta function, consisting of the product of Dirac's delta for continuous quantum numbers and the Kronecker's delta for discrete ones. Since the operator \hat{A} is Hermitian, its eigenvalues $x(\xi)$ are real.

In this context, Schrödinger's equation becomes

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (4.5)$$

¹A selfadjoint operator is a Hermitian one with the additional property of having an orthonormal basis of eigenkets.

Moreover, the result of a measurement of Q is not determined by s , however the probabilities of the possible results are. In particular, the probability of a possible outcome of a measurement of the observable within the numerical range χ is

$$p_{s,Q}(\chi) = \int_{x_\xi \in \chi} d\mu(\xi) |\langle \xi | \psi \rangle|^2, \quad (4.6)$$

where the kets $|\xi\rangle$ are a generalized orthonormal basis of eigenkets of \hat{A} and x_ξ are the corresponding eigenvalues. In addition, the expectation value of the results of the measurements of Q is given by

$$\langle Q \rangle_s = \int d\mu(\xi) x_\xi |\langle \xi | \psi \rangle|^2, \quad (4.7)$$

or, in general, for any real function $f(x)$,

$$\langle f(Q) \rangle_s = \int d\mu(\xi) f(x_\xi) |\langle \xi | \psi \rangle|^2. \quad (4.8)$$

Furthermore, a set of compatible observables Q_1, \dots, Q_n are represented by a set of commuting selfadjoint linear operators $\hat{A}_1 \dots \hat{A}_n$, where \hat{A}_i corresponds to Q_i , and they are said to be commuting if their commutator is vanishing, namely

$$[\hat{A}_i, \hat{A}_j] = \hat{A}_i \hat{A}_j - \hat{A}_j \hat{A}_i = 0. \quad (4.9)$$

The generalization of the previous relations is straightforward.

Moving on, the unique, up to a phase, choices of the orthonormal basis $|\xi\rangle$ associated with a complete set of commuting selfadjoint operators \hat{A}_i is called the \hat{A}_i representation. The representative set $\langle \xi | \phi \rangle$ of a ket $|\phi\rangle$ with respect to the basis is called the representative set of $|\phi\rangle$ in the representation. In this context, we are interested in the Schrödinger representation, consisting of the generalized orthonormal basis $|x\rangle$ of the position operator \hat{q} . Relations (4.4) become

$$\hat{q} |x\rangle = |x\rangle x, \quad (4.10a)$$

$$\langle x' | x \rangle = \delta(x' - x), \quad (4.10b)$$

$$\int |x\rangle d^3x \langle x| = \hat{1}. \quad (4.10c)$$

The representative $\langle \mathbf{x} | \psi \rangle$ of a ket $|\psi\rangle$ in the representation is the Schrödinger wave function $\psi(\mathbf{x})$ associated with $|\psi\rangle$, namely

$$\langle \mathbf{x} | \psi \rangle = \psi(\mathbf{x}). \quad (4.11)$$

In this representation, Schrödinger equation (4.5) becomes (4.1) and the position, momentum, angular momentum and Hamiltonian operators \hat{q} , \hat{p} , \hat{l} and \hat{H} act on the basis bras $\langle \mathbf{x} |$ as

$$\langle \mathbf{x} | \hat{q} = \mathbf{x} \langle \mathbf{x} |, \quad (4.12a)$$

$$\langle \mathbf{x} | \hat{p} = -i\hbar \nabla_{\mathbf{x}} \langle \mathbf{x} |, \quad (4.12b)$$

$$\langle \mathbf{x} | \hat{l} = -i\hbar \mathbf{x} \wedge \nabla_{\mathbf{x}} \langle \mathbf{x} |, \quad (4.12c)$$

$$\langle \mathbf{x} | \hat{H} = \left[-\frac{\hbar^2}{2m} \nabla_{\mathbf{x}}^2 + U(\mathbf{x}) \right] \langle \mathbf{x} |. \quad (4.12d)$$

In Schrödinger representation (4.10), by (4.11), eq. (4.6) becomes

$$p(\mathbf{x}, t) = \int d^3x |\langle \mathbf{x} | \psi \rangle|^2 = \int d^3x |\psi(\mathbf{x})|^2, \quad (4.13)$$

and can be interpreted as the probability of finding the particle described by ψ at position \mathbf{x} at time t . Therefore, we can define a probability density

$$\rho(\mathbf{x}, t) = \psi^* \psi. \quad (4.14)$$

In addition, we can define a probability current density $\mathbf{j}(\mathbf{x}, t)$ as

$$\mathbf{j}(\mathbf{x}, t) = \frac{\hbar}{2im} (\psi^* \nabla \psi - \psi \nabla \psi^*), \quad (4.15)$$

and show that ρ and \mathbf{j} follows the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \quad (4.16)$$

Proof. Supposing the potential is real and considering Schrödinger equation (4.1) multiplied by ψ^* and the conjugate equation multiplied by ψ ,

$$\begin{aligned} i\hbar\psi^*\frac{\partial\psi}{\partial t} &= -\frac{\hbar^2}{2m}\psi^*\nabla^2\psi + U\psi^*\psi, \\ -i\hbar\psi\frac{\partial\psi^*}{\partial t} &= -\frac{\hbar^2}{2m}\psi\nabla^2\psi^* + U\psi\psi^*. \end{aligned}$$

Subtracting them, we obtain

$$i\hbar\left(\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right) = -\frac{\hbar^2}{2m}(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*).$$

The first parenthesis, by (4.14), is equivalent to the time derivative of the probability density, while the second one can be written as

$$\nabla(\psi^*\nabla\psi - \psi\nabla\psi^*).$$

Using (4.15) the conclusion is straightforward. ■

This can be interpreted as the conservation of the probability, or rather, the particle being, at any time, somewhere in space. This leads to the conclusion that a non-relativistic quantum particle can't be created or destroyed, and we can see this if we take the non-relativistic limit of the energy,

$$E = mc^2\sqrt{1 + \frac{p^2}{m^2c^2}} = mc^2 + \frac{p^2}{2m} + \dots \quad (4.18)$$

Indeed, as $c \rightarrow \infty$, it would take an infinite amount of energy to create a particle.

4.2 From Schrödinger to Klein-Gordon equation

The easiest way to reach Schrödinger equation for a free quantum particle in Schrödinger representation is to start from its classical Hamiltonian,

$$H = \frac{p^2}{2m}, \quad (4.19)$$

and promote H and p to operators acting on the wave function ψ that describes the particle. This can be achieved performing the following formal substitutions

$$H \rightarrow \hat{H} = i\hbar \frac{\partial}{\partial t}, \quad (4.20a)$$

$$p \rightarrow \hat{p} = -i\hbar \nabla, \quad (4.20b)$$

leading to (4.1). The natural generalization of this method for a relativistic free particle would be starting from its Hamiltonian,

$$H = \sqrt{p^2 c^2 + m^2 c^4}, \quad (4.21)$$

and perform the substitutions (4.20), obtaining

$$i\hbar \frac{\partial \psi(\mathbf{x}, t)}{\partial t} = \sqrt{-\hbar^2 c^2 \nabla^2 + m^2 c^4} \psi(\mathbf{x}, t). \quad (4.22)$$

However, the previous equation contains the square root of a differential operator, leading to a non-local theory that is difficult to interpret and to handle. A simpler way to operate would be to remove the square root, starting from

$$H^2 = p^2 c^2 + m^2 c^4 \quad (4.23)$$

and then performing the substitutions (4.20). In this way, we would obtain the Klein-Gordon equation (3.43), namely

$$(\square + \mu^2)\psi(x) = 0, \quad (4.24)$$

where μ is a mass term, equivalent to

$$\mu = \frac{mc}{\hbar}. \quad (4.25)$$

Continuity equation. In analogy with Schrödinger equation, we wish to find a continuity equation like (4.16). Similarly as before, by subtracting eq. (4.24) multiplied by ψ^* and the conjugate equation multiplied by ψ , we obtain

$$\begin{aligned} \psi^*(\square + \mu^2)\psi - \psi(\square + \mu^2)\psi^* &= 0, \\ \psi^* \partial_\mu \partial^\mu \psi - \psi \partial_\mu \partial^\mu \psi^* &= \partial_\mu (\psi^* \partial^\mu \psi - \psi \partial^\mu \psi^*) = 0. \end{aligned}$$

If we define the current J^μ as

$$J^\mu = \frac{1}{2im}(\psi^* \partial^\mu \psi - \psi \partial^\mu \psi^*), \quad (4.27)$$

it would satisfy the continuity equation $\partial_\mu J^\mu = 0$. The time component of (4.27) should be the probability density ρ , as before, but this interpretation can't be correct, since it's not a positive definite quantity. Indeed, if we consider

$$J^0 = -\frac{i}{2m}(\psi^* \partial_0 \psi - \psi \partial_0 \psi^*),$$

since the Klein-Gordon equation (4.24) is a second-order differential equation in time, both ψ and $\partial_0 \psi$ can be chosen arbitrarily as initial conditions, so J^0 can also be negative.

Plane wave solution. As seen in sec. 3.5, the plane wave solution for Klein-Gordon equation leads to the dispersion relation (3.47). Since we are interested in relativistic quantum particles, using Planck-Einstein relations (4.3) we find that the plane wave is a solution of the equation if the mass-shell condition is satisfied, namely, it must be

$$E^2 = |\mathbf{p}|^2 c^2 + m^2 c^4 \implies E = \pm \sqrt{|\mathbf{p}|^2 c^2 + m^2 c^4} = \pm E_p. \quad (4.28)$$

Notice that, besides the expected solution with positive energy, there is the possibility for the energy to be negative. Since the model hasn't a lower boundary for energies, this leads to an interpretative problem. In quantum field theory this difficulty is overcome by reinterpreting the solution with $E = -E_p$ as representing antiparticles with positive energy.

Using again the Planck-Einstein relations (4.3), the plane wave solution (3.46) becomes

$$\phi_p(x) \sim e^{-\frac{i}{\hbar} p_\alpha x^\alpha} = e^{-\frac{i}{\hbar} E t + \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}}.$$

Using the two possible energies (4.28), we can index the plane wave solutions by the spacial momentum \mathbf{p} and the sign of the energy, namely

$$\begin{aligned} \phi_{\mathbf{p}}^+(x) &= e^{-\frac{i}{\hbar} E_p t + \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}}, \\ \phi_{\mathbf{p}}^-(x) &= e^{+\frac{i}{\hbar} E_p t - \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}}, \end{aligned}$$

where, by convention, in $\phi_{\mathbf{p}}^-(x)$ we have substituted $\mathbf{p} \rightarrow -\mathbf{p}$ to make it the complex conjugate of $\phi_{\mathbf{p}}^+(x)$.

Therefore, using a conventional normalization, we can write a general solution as a linear combination of the two plane waves, namely

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \left[a(\mathbf{p}) e^{-\frac{i}{\hbar} E_p t + \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} + b^*(\mathbf{p}) e^{+\frac{i}{\hbar} E_p t - \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} \right]. \quad (4.30)$$

The complex conjugate is

$$\phi^*(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \left[b(\mathbf{p}) e^{-\frac{i}{\hbar} E_p t + \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} + a^*(\mathbf{p}) e^{+\frac{i}{\hbar} E_p t - \frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} \right]. \quad (4.31)$$

Propagator. Let's now consider an inhomogeneous Klein-Gordon equation, namely

$$(\square + \mu^2)\phi(x) = -J(x), \quad (4.32)$$

where $J(x)$ is an arbitrary source and the minus sign is a convention. We can write a general solution by the Green function method, used in sec. 2.6 to solve Maxwell's equations. In particular, a Green function $G(x)$ for (4.32) is defined through

$$(\square + \mu^2)G(x) = -\delta^4(x), \quad (4.33)$$

and the solution for (4.32) can be written as

$$\phi(x) = \phi_0(x) + \int d^4y G(x-y)J(y), \quad (4.34)$$

where $\phi_0(x)$ is a solution of the homogeneous equation (4.24).

As already seen, there are many ways to implement a Green function to propagate information, depending on the boundary conditions chosen. In quantum field theory it would be more correct to take the *propagator* into account, which is an operator that propagates the quanta of a field. In this particular case, it can be defined as

$$\Delta(x-y) = \frac{i}{\hbar^2 c} G(x-y). \quad (4.35)$$

Without pursuing this argument further, we have to say that, in order to solve the negative energies' problem, we can't use the retarded Green function (2.66) anymore. Instead, we should implement the so-called *Feynman propagator*: it propagates the states with positive energy forward in time and those with negative energy backwards. This allows to reinterpret the latter as representing antiparticles with positive energy. To implement this prescription, we should proceed as in sec. 2.6, writing the

propagator in Fourier's space and choosing a suitable integration path to avoid the integrand's poles.

Using (A.5) we'd have

$$G(x) = \frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha x^\alpha} \tilde{G}(k).$$

Using (A.15) we can write the delta as

$$\delta^4(x) = \frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha x^\alpha},$$

and substituting into (4.33) leads to

$$\begin{aligned} (\square + \mu^2) \int d^4k e^{-ik_\alpha x^\alpha} \tilde{G}(k) &= - \int d^4k e^{-ik_\alpha x^\alpha}, \\ \int d^4k [k_\alpha k^\alpha - \mu^2] e^{-ik_\alpha x^\alpha} \tilde{G}(k) &= \int d^4k e^{-ik_\alpha x^\alpha}. \end{aligned}$$

Since the Fourier transform is complete, the last relation implies

$$\tilde{G}(k) = \frac{1}{k_\alpha k^\alpha - \mu^2},$$

leading to

$$G(x) = \frac{1}{(2\pi)^4} \int d^4k \frac{e^{-ik_\alpha x^\alpha}}{k_\alpha k^\alpha - \mu^2}.$$

Using (4.3) and (4.25), we can refer to p^μ rather than k^μ , and write

$$\begin{aligned} G(x-y) &= \frac{\hbar^2}{(2\pi)^4} \int d^4p \frac{e^{-\frac{i}{\hbar} p_\alpha (x^\alpha - y^\alpha)}}{p_\alpha p^\alpha - m^2 c^2} \\ &= \frac{\hbar^2}{(2\pi)^4} \int d^3p e^{\frac{i}{\hbar} \mathbf{p} \cdot (\mathbf{x} - \mathbf{y})} \int dp_0 \frac{e^{-\frac{i}{\hbar} p_0 (x_0 - y_0)}}{(p_0)^2 - |\mathbf{p}|^2 - m^2 c^2} \\ &= \frac{\hbar^2}{(2\pi)^4} \int d^3p e^{\frac{i}{\hbar} \mathbf{p} \cdot (\mathbf{x} - \mathbf{y})} \int dp_0 \frac{e^{-\frac{i}{\hbar} p_0 (x_0 - y_0)}}{\left(p_0 + \frac{E_p}{c}\right) \left(p_0 - \frac{E_p}{c}\right)}, \end{aligned} \quad (4.37)$$

where in the last step we have used (4.28). We can notice that there are two poles at $p_\alpha p^\alpha - m^2 c^2 = 0$, corresponding to particles that respect the mass-shell relation, called *real particles*. If $p_\alpha p^\alpha - m^2 c^2 \neq 0$ holds, the waves are thought to represent particles that can't propagate at large distances, called *virtual particles*.

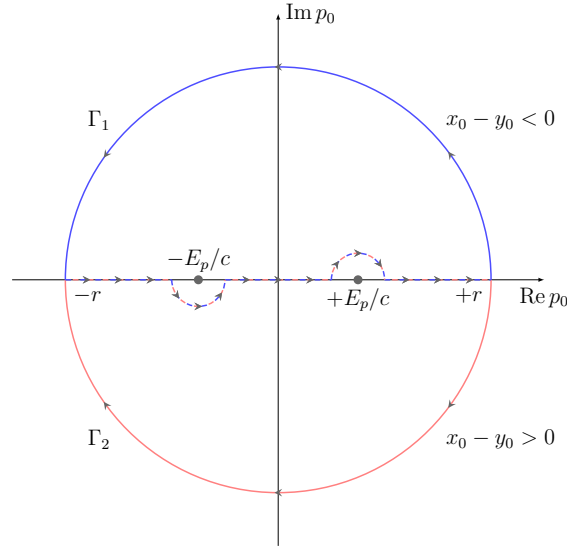


Figure 4.1 Deformation of the integration path to implement the Feynman propagator. The blue line, Γ_1 , is for $x_0 - y_0 < 0$ and allows propagating backward in time negative energies $p_0 = -E_p/c$. The red line, Γ_2 , is for $x_0 - y_0 > 0$ and allows propagating forward in time positive energies $p_0 = +E_p/c$.

Performing the integration over p_0 , we can implement the Feynman propagator choosing the integration path in fig. 4.1. Indeed, for $x_0 - y_0 < 0$, we must close the integration path in the upper half-plane to use Jordan's Lemma L.2, as showed by the blue line of fig. 4.1. Using (A.28) and the residue theorem (A.23), the final result is propagating the pole $p_0 = -E_p/c$ backward in time. Similarly, for $x_0 - y_0 > 0$, we must choose the red path of fig. 4.1, that entails the propagation in the future of the pole $p_0 = +E_p/c$.

Performing the computation,

$$\begin{aligned} I_0 &\equiv \int dp_0 \frac{e^{-\frac{i}{\hbar}p_0(x_0-y_0)}}{\left(p_0 + \frac{E_p}{c}\right)\left(p_0 - \frac{E_p}{c}\right)} \\ &= \frac{c}{2E_p} \int dp_0 e^{-\frac{i}{\hbar}p_0(x_0-y_0)} f(p_0), \end{aligned}$$

where

$$f(p_0) = \frac{1}{p_0 - \frac{E_p}{c}} - \frac{1}{p_0 + \frac{E_p}{c}}.$$

Moving on,

$$\begin{aligned}
 I_0 &= \frac{c}{2E_p} \left[\lim_{r \rightarrow \infty} \oint_{\Gamma_1} f(p_0) e^{-\frac{i}{\hbar} p_0 (x_0 - y_0)} \theta(y_0 - x_0) + \lim_{r \rightarrow \infty} \oint_{\Gamma_2} f(p_0) e^{-\frac{i}{\hbar} p_0 (x_0 - y_0)} \theta(x_0 - y_0) \right] \\
 &= \frac{i\pi c}{E_p} \left[\sum_{j_1} \text{Res } f(p_0) e^{-\frac{i}{\hbar} p_0 (x_0 - y_0)} \Big|_{p_0 = p_{0j_1}} \theta(y_0 - x_0) \right. \\
 &\quad \left. - \sum_{j_2} \text{Res } f(p_0) e^{-\frac{i}{\hbar} p_0 (x_0 - y_0)} \Big|_{p_0 = p_{0j_2}} \theta(x_0 - y_0) \right],
 \end{aligned}$$

where the minus sign is due to the orientation of Γ_2 and p_{0j_1} are the poles of $f(p_0)$ within Γ_1 and p_{0j_2} within Γ_2 . Using Cauchy's theorem A.16 and computing the residue with (A.22), we'd obtain

$$I_0 = -\frac{i\pi c}{E_p} \left[e^{-\frac{iE_p}{\hbar c} (y_0 - x_0)} \theta(y_0 - x_0) + e^{-\frac{iE_p}{\hbar c} (x_0 - y_0)} \theta(x_0 - y_0) \right] = -\frac{i\pi c}{E_p} e^{-\frac{iE_p}{\hbar c} |x_0 - y_0|}.$$

Inserting the last expression in (4.37), we can finally write the propagator (4.35) as

$$\Delta(x - y) = \int \frac{d^3 p}{(2\pi)^3} e^{i\mathbf{p} \cdot (\mathbf{x} - \mathbf{y})} \frac{e^{-\frac{iE_p}{\hbar c} |x_0 - y_0|}}{2E_p}. \quad (4.38)$$

Symmetries. The complex Klein-Gordon Lagrangian (3.44) is invariant under the following *global*² transformations,

$$\phi(x) \rightarrow \phi'(x) = e^{-i\alpha} \phi(x), \quad (4.39a)$$

$$\phi^*(x) \rightarrow \phi'^*(x) = e^{i\alpha} \phi^*(x), \quad (4.39b)$$

where α is a real constant. They are called *first-kind gauge transformations*. The infinitesimal form is

$$\delta\phi(x) = -i\alpha\phi(x), \quad (4.40a)$$

$$\delta\phi^*(x) = i\alpha\phi^*(x). \quad (4.40b)$$

In relation to notation (3.31) used for Noether's theorem, since this symmetry is *internal*, meaning that there is no transformation over the coordinates, we have

$$\chi = 0, \quad \Phi = -i\phi, \quad \Phi^* = i\phi^*.$$

²Global means that the transformation is carried out in the same way over the entire space, or equivalently, the parameter α is a constant.

Therefore, the Noether's current (3.33) is

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}(-i\phi) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^*)}(i\phi^*).$$

Performing the computation, inserting the Lagrangian (3.44) in the previous relation, we find

$$J^\mu = i(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*), \quad (4.41)$$

and the corresponding conserved charge by eq. (3.36) is

$$Q(t) = \int d^3x J^0(x) = \frac{i}{c} \int d^3x \left[\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t} \right].$$

4.3 Scalar electrodynamics

From now on, we shall assume $c = \hbar = 1$. As a consequence, it is $\mu = m$.

In quantum field theory the Klein-Gordon field represents a free particle with mass m and no spin. As we have just seen, if the field is complex, it admits a conserved current (4.41). Let's see that if we force the Lagrangian to be invariant under local transformations of the kind (4.39), this current allows an interaction with the electromagnetic field. Basically, we impose $\alpha = \alpha(x^\mu)$ to depend on the particular point of Minkowski's space-time. This way, we can write the so-called *second-kind gauge transformations* as

$$\phi(x) \rightarrow \phi'(x) = e^{-i\alpha(x)} \phi(x), \quad (4.42a)$$

$$\phi^*(x) \rightarrow \phi'^*(x) = e^{i\alpha(x)} \phi^*(x), \quad (4.42b)$$

or, infinitesimally,

$$\delta\phi(x) = -i\alpha(x)\phi(x), \quad (4.43a)$$

$$\delta\phi^*(x) = i\alpha(x)\phi^*(x). \quad (4.43b)$$

First, let's notice that $\partial_\mu \phi$ doesn't transform covariantly, meaning that its transformation is different from that of ϕ . Indeed, we have³

$$\partial_\mu \phi(x) \rightarrow \partial_\mu \phi'(x) = e^{-i\alpha(x)} \partial_\mu \phi - ie^{-i\alpha(x)} \partial_\mu \alpha(x) \phi(x).$$

³When there are no parentheses, the derivative ∂_μ is intended to be applied only on the contiguous term.

Consequently, the Lagrangian (3.44) is no longer invariant under the transformations (4.42). Indeed, it transforms as

$$\begin{aligned}
\mathcal{L} \rightarrow \mathcal{L}' &= \partial_\mu \phi'^* \partial^\mu \phi' - m^2 \phi'^* \phi' \\
&= \partial_\mu (e^{i\alpha(x)} \phi^*(x)) \partial^\mu (e^{-i\alpha(x)} \phi(x)) - m^2 \phi^*(x) \phi(x) \\
&= \mathcal{L} - i \partial^\mu \alpha \partial_\mu \phi^* \phi + i \partial_\mu \alpha \phi^* \partial^\mu \phi + \partial_\mu \alpha \partial^\mu \alpha \phi^* \phi \\
&= \mathcal{L} + \partial_\mu \alpha J^\mu + \partial_\mu \alpha \partial^\mu \alpha \phi^* \phi,
\end{aligned}$$

where we have used the definition (4.41).

As said before, we want the total Lagrangian to be invariant under gauge transformations (4.42). This is possible if we suppose the existence of a four-vector A_μ that couples to the current J^μ through a constant q , and transforms with the field as

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{q} \partial_\mu \alpha(x). \quad (4.44)$$

The term to add to \mathcal{L} will be

$$\mathcal{L}_1 = -q A_\mu J^\mu. \quad (4.45)$$

Under the simultaneous transformations (4.42) and (4.44), the current (4.41) turns into

$$J^\mu \rightarrow J'^\mu = J^\mu + 2 \partial^\mu \alpha \phi^* \phi,$$

and the Lagrangian (4.45) into

$$\mathcal{L}_1 \rightarrow \mathcal{L}'_1 = \mathcal{L}_1 - 2q A_\mu \partial^\mu \alpha \phi^* \phi - \partial_\mu \alpha J^\mu - 2 \partial_\mu \alpha \partial^\mu \alpha \phi^* \phi.$$

Therefore, the variation of $\mathcal{L} + \mathcal{L}_1$ will be

$$\mathcal{L} + \mathcal{L}_1 \rightarrow \mathcal{L}' + \mathcal{L}'_1 = \mathcal{L} + \mathcal{L}_1 - \partial_\mu \alpha \partial^\mu \alpha \phi^* \phi - 2q A_\mu \partial^\mu \alpha \phi^* \phi. \quad (4.46)$$

To make the total Lagrangian invariant under (4.42) and (4.44), we must add another term, that is

$$\mathcal{L}_2 = q^2 A_\mu A^\mu \phi^* \phi. \quad (4.47)$$

Proof. The Lagrangian \mathcal{L}_2 transforms as

$$\begin{aligned}\mathcal{L}_2 &\rightarrow \mathcal{L}'_2 = q^2 A'_\mu A'^\mu \phi'^* \phi' \\ &= q^2 \left(A_\mu + \frac{1}{q} \partial_\mu \alpha \right) \left(A^\mu + \frac{1}{q} \partial^\mu \alpha \right) \phi^* \phi \\ &= \mathcal{L}_2 + \partial_\mu \alpha \partial^\mu \alpha \phi^* \phi + 2q A_\mu \partial^\mu \alpha \phi^* \phi.\end{aligned}$$

Combined with (4.46), we have

$$\mathcal{L} + \mathcal{L}_1 + \mathcal{L}_2 \rightarrow \mathcal{L}' + \mathcal{L}'_1 + \mathcal{L}'_2 = \mathcal{L} + \mathcal{L}_1 + \mathcal{L}_2,$$

as claimed. ■

At this point the Lagrangian $\mathcal{L} + \mathcal{L}_1 + \mathcal{L}_2$ is invariant under the gauge transformations (4.42) and (4.44), but it's still reasonable to suppose that the field A_μ itself, other than interacting with the field's current J^μ , contributes directly on the total Lagrangian. Its contribution must be invariant under (4.44), and it's immediate to show that a correct term is given by

$$\mathcal{L}_3 = -\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu}, \quad F^{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A^\mu. \quad (4.48)$$

Summing up (3.44), (4.45), (4.47) and (4.48), the total Lagrangian

$$\mathcal{L}_{\text{tot}} = (\partial_\mu \phi + iq A_\mu \phi) (\partial^\mu \phi^* - iq A^\mu \phi^*) - m^2 \phi^* \phi - \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} \quad (4.49)$$

is invariant under (4.42) and (4.44). It's also known as *scalar electrodynamics Lagrangian*.

Let's analyze what we've found. The free Klein-Gordon Lagrangian (3.44) is invariant under global $U(1)$ transformations⁴, with the Noether's current (4.27) linked to it. Imposing an invariance under local $U(1)$ transformations leads to an interaction theory. Indeed, the Lagrangian (4.48) represents the free electromagnetic field (3.49), while (4.45) represents the coupling between the sources of the latter and the four-potential, as in (3.51). Therefore, the gauge potential A_μ can be interpreted as the electromagnetic vector potential (2.51), obeying the gauge transformation (2.54). Clearly, $\alpha(x) = q\chi$ represents the gauge function, while q is a coupling constant between the charged field ϕ and the gauge field A_μ . In quantum field theory it'll

⁴Group theory exceeds the purpose of this thesis, therefore we won't analyze it. A $U(1)$ transformation is, essentially, a rotation on a complex space.

be the charge of a particle excitation of the quantum field. Summarizing, we can interpret the electromagnetic field as a gauge field to be introduced with the aim to guarantee invariance under local $U(1)$ gauge transformations. Let's finally remark a very important aspect: gauge invariance requires that the gauge field is massless. Indeed, a hypothetical mass term for the gauge field would have the form

$$\mathcal{L}_{\text{mass}} = M^2 A_\mu A^\mu,$$

but this would make the total Lagrangian no longer invariant under the transformation (4.44).

A mathematical elegant way to derive the interacting theory (4.49) from the free one (3.44) consists of replacing the derivatives ∂_μ by the so-called *covariant derivatives*,

$$\partial_\mu \phi \rightarrow \mathcal{D}_\mu \phi = (\partial_\mu + iqA_\mu) \phi, \quad (4.50a)$$

$$\partial_\mu \phi^* \rightarrow \mathcal{D}_\mu \phi^* = (\partial_\mu - iqA_\mu) \phi^*. \quad (4.50b)$$

It's easy to show that they transform as the field.

Proof. Let's consider $\mathcal{D}_\mu \phi$ in (4.50a). Under (4.42) and (4.44), it transforms like

$$\begin{aligned} \mathcal{D}_\mu \phi &\rightarrow \mathcal{D}'_\mu \phi' = \partial_\mu \phi' + iqA'_\mu \phi' \\ &= e^{-i\alpha} [\partial_\mu \phi - i\partial_\mu \alpha \phi] + iq e^{-i\alpha} \left[A_\mu + \frac{1}{q} \partial_\mu \alpha \right] \phi \\ &= e^{-i\alpha} [\partial_\mu + iqA_\mu] \phi = e^{-i\alpha} \mathcal{D}_\mu \phi. \end{aligned}$$

Following step by step the computations with $\mathcal{D}_\mu \phi^*$ in (4.50b), we'd have

$$\mathcal{D}_\mu \phi^* \rightarrow e^{i\alpha} \mathcal{D}_\mu \phi^*,$$

showing what was claimed. ■

4.4 Gauge invariance as a principle

The method explained in the previous section is known as *gauge principle*. It consists of the replacement of the usual derivatives by the covariant ones, to obtain an interaction theory from a free one. The gauge field A^μ is interpreted as a new field with which the matter field interacts, and it has its own dynamics, that can be

highlighted by its relation to a stress-field tensor that obeys some field equations, like the electromagnetic tensor in the case studied.

Let's revise this principle critically, starting from *Michael Redhead's* suggestive prompt [18]:

“The gauge principle is generally regarded as the most fundamental cornerstone of modern theoretical physics. In my view its elucidation is the most pressing problem in current philosophy of physics.”

Let's retrace what we have done in the previous section with a slight change in notation, consisting of the addition of q somewhere, as we shall see. The superscript (m) will refer to the matter field, while (g) to the gauge field. The free Klein-Gordon Lagrangian

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi$$

is invariant under global, first-kind gauge transformation

$$\phi(x) \rightarrow \phi'(x) = e^{iq^{(m)}\alpha} \phi(x),$$

and, by Noether's theorem, there is an associated conserved current that obeys a continuity equation,

$$J^{\mu(m)} = iq^{(m)} [\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*],$$

$$\partial_\mu J^{\mu(m)} = 0.$$

The free Lagrangian is required to be invariant under local, second-kind gauge transformation

$$\phi(x) \rightarrow \phi'(x) = e^{iq^{(m)}\alpha(x)} \phi(x),$$

and it can be achieved by replacing the ordinary derivatives with the covariant ones,

$$\partial_\mu \rightarrow \mathcal{D}_\mu = \partial_\mu + iq^{(m)} A_\mu.$$

This procedure is equivalent to introduce a new vector potential $A_\mu(x)$ which transforms as

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{q^{(m)}} \partial_\mu \alpha(x),$$

and, as seen by (4.45), leads to an interaction term in the Lagrangian given by⁵

$$\mathcal{L}_{\text{int}}^{(m)} = -J_{\mu}^{(m)} A^{\mu}. \quad (4.52)$$

At this stage, we have done a merely internal change of coordinates. In order for the theory to have a physical relevance, we need to link A^{μ} to a field-strength tensor that obeys some field equations. This can be done by

$$F^{\mu\nu} = \partial^{\mu} A^{\nu} - \partial^{\nu} A^{\mu},$$

interpreting it as the electromagnetic tensor.

As remarked by *Holger Lyre* [11], we can't identify a priori A^{μ} and $F^{\mu\nu}$ as the electromagnetic potential and tensor. Indeed, let's consider Maxwell's electrodynamics. It can be described by the Maxwell's equations⁶

$$\begin{aligned} \partial_{\mu} F^{\mu\nu} &= 4\pi J^{\nu(g)}, \\ \partial_{\mu} \tilde{F}^{\mu\nu} &= 0, \end{aligned}$$

where, now, $F^{\mu\nu}$ is precisely the electromagnetic tensor. As we can see, to describe the dynamics of the field, we have to know its sources, represented by the four-current

$$J^{\mu(g)} = \rho^{(g)} \frac{dx^{\mu}}{ds},$$

where $\rho^{(g)}$ is the density of the field charge $q^{(g)}$ per volume element V and $\frac{dx^{\mu}}{ds}$ is the four-velocity of V (see sec. 2.3). As seen by (3.51), the coupling term between $J^{\mu(g)}$ and the four-potential A^{μ} is given by

$$\mathcal{L}_{\text{int}}^{(g)} = -J_{\mu}^{(g)} A^{\mu}. \quad (4.54)$$

Comparing (4.52) and (4.54), it's natural to make the identification

$$\mathcal{L}_{\text{int}} = \mathcal{L}_{\text{int}}^{(m)} = \mathcal{L}_{\text{int}}^{(g)}, \quad (4.55)$$

but it comprises a subtle implication. Indeed, this is possible only by the prescription

$$q^{(m)} = q^{(g)}, \quad (4.56)$$

⁵Notice that, now, the term $q^{(m)}$ is absorbed in the definition of J^{μ} .

⁶Remember the convention $c = \hbar = 1$.

which can be thought as a *generalized equivalence principle*, similar to the correspondence between inertial and gravitational mass in general relativity.

Indeed, let's suppose that (4.56) is not valid, namely

$$\frac{q^{(m)}}{q^{(g)}} \neq 1.$$

This would mean that different types of particles of equal electric charge would couple differently to the electromagnetic field. This isn't what we observe, since the experiments suggest that the coupling constant is universal.

In conclusion, both the gauge principle and the equivalence principle lead to the natural combination of the matter field and the interaction field, which can be now thought as the electromagnetic field without any ambiguity. This means that we can't think about the charge as a Klein-Gordon field's property, and, therefore, in an interacting system the division between source fields and fields mediating the interactions is somewhat artificial.

MATHEMATICAL APPENDIX

A.1 Fourier transform

A *Fourier transform* is a mathematical transformation that associates to a function

$$f : \mathbb{R}^n \rightarrow \mathbb{C} \quad (\text{A.1})$$

another function through the operation

$$(\mathcal{F}f)(\mathbf{k}) = \tilde{f}(\mathbf{k}) = \int_{\mathbb{R}^n} d^n x e^{i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x}), \quad (\text{A.2})$$

where $\boldsymbol{\xi} \cdot \mathbf{x} = \xi_1 x_1 + \cdots + \xi_n x_n$. Further, the *inverse Fourier transform* is defined as

$$(\bar{\mathcal{F}}f)(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} d^n k e^{-i\mathbf{k}\cdot\mathbf{x}} \tilde{f}(\mathbf{k}), \quad (\text{A.3})$$

and one is the inverse transformation of the other.

The generalization to Minkowski's space-time is straightforward. Indeed, considering the Minkowski's inner product, $k_\mu x^\mu$, defined by (2.20), we'd have

$$(\mathcal{F}f)(k) = \tilde{f}(k) = \int_{\mathcal{M}} d^4 x e^{ik_\alpha x^\alpha} f(x) \quad (\text{A.4})$$

for the Fourier transform and

$$(\bar{\mathcal{F}}f)(x) = \frac{1}{(2\pi)^4} \int_{\mathcal{M}} d^4 k e^{-ik_\alpha x^\alpha} \tilde{f}(k) \quad (\text{A.5})$$

for the inverse transform.

A.2 Dirac's delta

Being beyond our purposes, we won't dwell on the details of distributions' theory, but we'll formally define the *Dirac's delta* as a "function" $\delta(x - x')$ that, multiplied to a sufficiently regular function $f(x)$, and integrating over the real axis, yields

$$\int_{-\infty}^{+\infty} dx f(x) \delta(x - x') = f(x'). \quad (\text{A.6})$$

Integrating over the real axis,

$$\int_{-\infty}^{+\infty} dx \delta(x - x') = 1. \quad (\text{A.7})$$

It isn't necessary to integrate from $-\infty$ to $+\infty$, since we'd have the same result integrating over a finite interval that contains x' .

As follows, some useful properties.

- It is an even function,

$$\delta(x - x') = \delta(x' - x). \quad (\text{A.8})$$

- It can be thought as the inverse Fourier transform of the identity,

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dx e^{-ik(x-x')}. \quad (\text{A.9})$$

- The delta of a function $f(x)$ can be computed as

$$\delta[g(x)] = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|}, \quad (\text{A.10})$$

where x_i are simple zeroes of $g(x)$, namely $g(x_i) = 0$ and $g'(x_i) \neq 0$. In particular, given a constant a , we can write

$$\delta(ax) = \frac{\delta(x)}{|a|}. \quad (\text{A.11})$$

- Another useful relation following by (A.10) is

$$\delta(x^2 - a^2) = \frac{1}{2|a|} [\delta(x - a) + \delta(x + a)]. \quad (\text{A.12})$$

These relations can be easily generalized in an n -dimensional space and, in particular, we can define the Dirac's delta in Minkowski's space-time as

$$\delta^4(x^\mu - x'^\mu) = \delta(ct - ct')\delta(\mathbf{x} - \mathbf{x}'). \quad (\text{A.13})$$

The generalization of (A.7) obviously is

$$\int d^4x \delta^4(x) = 1, \quad (\text{A.14})$$

and since d^4x is a scalar by (2.27), this implies that $\delta^4(x)$ is Lorentz invariant.

Finally, eq. (A.9) becomes

$$\delta^4(x - x') = \frac{1}{(2\pi)^4} \int d^4k e^{-ik_\alpha(x^\alpha - x'^\alpha)}. \quad (\text{A.15})$$

A.3 Residue theorem

Let's first recall an important result for *holomorphic functions*.

Theorem T.2 (Cauchy's integral theorem). *Let $f(z)$ be a holomorphic function in an open, simply connected region D and γ a closed, simple, piecewise regular curve within D , then*

$$\oint_\gamma dz f(z) = 0. \quad (\text{A.16})$$

Let now $f(z)$ have an isolated singularity in $z = z_0$ and be holomorphic within a disk centered in z_0 , except for z_0 itself. The *Laurent series* of f about z_0 is

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \frac{a_{-1}}{z - z_0} + \frac{a_{-2}}{(z - z_0)^2} + \dots \quad (\text{A.17})$$

Because of the singularity in z_0 , at least one of the coefficient a_{-n} must not vanish. If a_{-n} is different from zero and all the successive coefficients vanish, namely

$$a_{n+1} = a_{n+2} = a_{n+3} = \dots = 0, \quad (\text{A.18})$$

then z_0 is said to be a *pole* of order n . If $n = 1$, that is, $a_{-1} \neq 0$ and the other coefficients with negative indices vanish, then z_0 is a *simple pole*.

Let $f(z)$ be holomorphic within an open region D except at a point z_0 in D , where f may have an isolated singularity. If γ is a closed simple curve, piecewise regular,

within D , that contains z_0 , then the *residue* of $f(z)$ at $z = z_0$ is

$$\operatorname{Res} f(z)|_{z=z_0} \equiv \frac{1}{2\pi i} \oint_{\gamma} dz f(z), \quad (\text{A.19})$$

where, for convention, γ is oriented counterclockwise, and the integral above doesn't depend on the particular choice of the curve. Remember that if the curve is oriented clockwise, there must be a minus sign.

It's easy to compute the residual if $z = z_0$ is a simple pole. Indeed, we'd have

$$\operatorname{Res} f(z)|_{z=z_0} = \lim_{z \rightarrow z_0} (z - z_0) f(z) \quad (\text{A.20})$$

or, equivalently, writing the function as

$$f(z) = \frac{p(z)}{q(z)}, \quad q(z_0) = 0, \quad (\text{A.21})$$

the residue is

$$\operatorname{Res} f(z)|_{z=z_0} = \frac{p(z_0)}{q'(z_0)}. \quad (\text{A.22})$$

Theorem T.3 (Residue Theorem). *Suppose $f(z)$ is holomorphic in a region D except for a finite set m of isolated singularities. Suppose also γ is a simple, closed curve in D , oriented counterclockwise, which contains the singularities. Then*

$$\oint_{\gamma} dz f(z) = 2\pi i \sum_{j=1}^m \operatorname{Res} f(z)|_{z=z_j}. \quad (\text{A.23})$$

We are interested in integrals of the form

$$I = \int_{-\infty}^{+\infty} dx f(x) e^{ix}, \quad (\text{A.24})$$

where $f(z)$ is holomorphic over the half-plane $\operatorname{Im} z \geq 0$, except a limited set of singularities.

We can write $I = \lim_{r \rightarrow \infty} I_r$, with

$$I = \int_{-r}^{+r} dx f(x) e^{ix}, \quad (\text{A.25})$$

and consider an integral over the complex plane, with a path of integration that coincide with the segment $[-r, +r]$ and a semicircle C_r with center O and

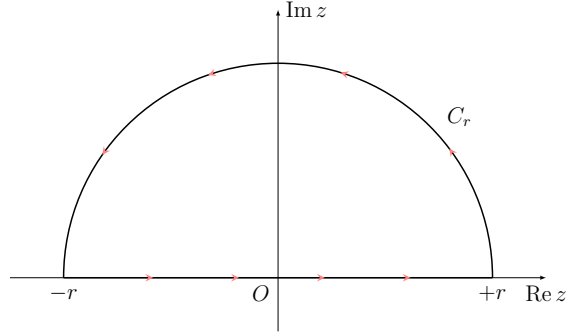


Figure A.1 Integration path.

radius r , as shown in fig. A.1. Then, if we set $\Gamma_r \equiv [-r, +r] \cup C_r$, we have

$$I = \oint_{\Gamma_r} dz f(z)e^{iz} = \int_{-r}^{+r} dx f(x)e^{ix} + \int_{C_r} dz f(z)e^{iz}, \quad (\text{A.26})$$

where (A.26) can be computed with residue theorem (A.23). If we can prove that

$$\int_{C_r} dz f(z)e^{iz} \xrightarrow{r \rightarrow \infty} 0, \quad (\text{A.27})$$

then we can compute (A.24) as

$$I = \lim_{r \rightarrow \infty} \oint_{\Gamma_r} dz f(z)e^{iz} = 2\pi i \sum_{\text{Im } z_k > 0} \text{Res } f(z)e^{iz} \Big|_{z=z_k}. \quad (\text{A.28})$$

Actually, eq. (A.27) is verified and is known as *Jordan's Lemma*.

Lemma L.2 (Jordan's lemma). *If $f(z)$ is defined on the superior half-plane $\text{Im } z > 0$ and if*

$$\lim_{|z| \rightarrow \infty} f(z) = 0 \quad (\text{A.29})$$

in a circular sector subtended by the arc C_r ,

$$C_r = \{z = re^{i\theta}, \quad \theta_1 \leq \theta \leq \theta_2\}, \quad (\text{A.30})$$

then

$$\int_{C_r} dz f(z)e^{iz} \xrightarrow{r \rightarrow \infty} 0. \quad (\text{A.31})$$

Similarly, if

$$\lim_{|z| \rightarrow 0} f(z) = 0, \quad (\text{A.32})$$

then

$$\int_{C_r} dz f(z) e^{iz} \xrightarrow{r \rightarrow 0} 0. \quad (\text{A.33})$$

If, instead of e^{iz} we would have e^{-iz} , we should have closed the integration path Γ_r in the inferior half-plane $\text{Im}_z < 0$, since in $\text{Im}_z > 0$ e^{-iz} diverges. Everything else is akin to what we have just seen.

REFERENCES

- [1] Drell S. D. and Bjorken J. D. *Relativistic Quantum Mechanics*. McGraw-Hill, 1964.
- [2] Jackson J. D. *Classical electrodynamics*. John Wiley and Sons, 1999.
- [3] Landau L. D. and Lifshitz E. M. *The classical theory of fields*. Addison Wesley, 1959.
- [4] Lurie D. *Particles and Fields*. John Wiley and Sons, 1968.
- [5] Tong D. *Quantum Field Theory*. Lecture Notes. University of Cambridge, 2006.
- [6] Giusti E. *Analisi Matematica 2*. Bollati Boringhieri, 2003.
- [7] Bastianelli F. *Relativistic Quantum Mechanics*. Lecture Notes. University of Bologna.
- [8] Ortolani F. *Appunti di Metodi Matematici*. Lecture Notes. University of Bologna, 2013.
- [9] Torre C. G. *Classical Field Theory*. Lecture Notes. Utah State University.
- [10] Goldstein H., Poole C., and Safko J. *Classical mechanics*. Addison Wesley, 2002.
- [11] Lyre H. "A Generalized Equivalence Principle". In: *International Journal of Modern Physics* 9 (2000).
- [12] Lyre H. "The Principles of Gauging". In: *The University of Chicago Press* 68.3 (2001).
- [13] Ryder L. H. *Quantum field theory*. Cambridge University Press, 1996.
- [14] Gelfand I.M. and Fomin S.V. *Calculus of variations*. Prentice Hall, 1963.
- [15] Hill E. L. "Hamilton's Principle and the Conservation Theorems of Mathematical Physics". In: *Reviews of Modern Physics* 23.3 (1951).
- [16] Landau L.D. and Lifshitz E.M. *Mechanics*. Pergamon Press, 1960.
- [17] Kot M. *A first course in the calculus of variations*. American Mathematical Society, 2014.
- [18] Redhead M. "The interpretation of gauge symmetry". In: *Symmetries in Physics: Philosophical Reflections*. Ed. by Brading K. and Castellani E. Cambridge University Press, 2003. Chap. 7.
- [19] Balbinot R. *Topics in Mathematical Methods and Models in Theoretical Physics*. Lecture Notes. University of Bologna.
- [20] Zucchini R. *Quantum Mechanics*. Lecture Notes. University of Bologna.
- [21] Barone V. *Relatività*. Bollati Boringhieri, 2004.

RINGRAZIAMENTI

A lungo ho pensato come poter scrivere i ringraziamenti e chi ringraziare. Che debba ringraziare Einstein e Schrödinger e una lunga lista di pensatori per avermi guidato ed entusiasmato ogni giorno? Forse sì, ma ad ora, che sono all'inizio di un percorso spero lungo, e dopo aver raggiunto il primo checkpoint, vorrei ringraziare chi ogni giorno mi ha ispirato e supportato nella vita quotidiana.

Pertanto, vorrei ringraziare la mia *famiglia*, che rappresenta la scintilla da cui ha iniziato ad ardere la passione per la conoscenza. Ringrazio mio *padre*, che mi ha trasmesso quel senso critico tanto scomodo delle volte quanto utile a non smarrirsi; mia *madre*, che, anche se da lontano, vive con me e attraverso i miei occhi l'avventura d'inseguire i propri sogni, supportandomi ogni giorno; e mio *fratello*, che spero non smetta mai di prendermi in giro per essere il prototipo di fisico tedioso, facendomi mettere i piedi per terra.

Ringrazio di cuore la mia *Cici*, che è sempre stata al mio fianco, e mi ha insegnato con l'affetto e la vicinanza ad accantonare stress e rabbia, a fare due passi indietro quando necessario, ad affrontare le cose con il sorriso anche quando non è spontaneo farlo. Mi hai ascoltato con entusiasmo parlare per ore di relatività e meccanica quantistica, dibattendo attivamente, mi hai aiutato a capire che tipo di tesi volessi scrivere, come scriverla e che tono darle, anche se dopo il liceo ti eri ripromessa di non sentire mai più parlare di fisica. E come se non bastasse, hai letto con me ogni riga di questo documento, correggendo ogni sciocco errore di un inglese che deve ancora maturare.

Ringrazio tutti i miei *compagni di corso*. Da solo non ce l'avrei mai fatta, e mi avete insegnato quanto sia bello collaborare e mettere insieme le forze per ottenere un risultato che nessuno potrebbe raggiungere individualmente. Ringrazio tutti i *Maggiovatovi*, con i quali mi sono confrontato perennemente. Mi avete insegnato come sia possibile parlare di fisica anche di fronte a una birra, e come esista altro nel mondo. Un ringraziamento particolare a *Feber*. Condividere energie, grinta e voglia di fare, ma anche insicurezze e intoppi, mi ha tenuto sempre in equilibrio.

Ringrazio *l'Ochetta*. Dall'inizio ho condiviso con te le difficoltà e le vittorie, i periodi di stress e quelli di euforia, ma ci hai sempre creduto. E *Coso*, sai di non poter mancare, non smettere mai di ricordarmi come l'amicizia possa trapassare ogni ostacolo. Ringrazio *Michele* e le nostre passeggiate senza sosta. Adoro quando camminiamo e parliamo e ci confrontiamo, dimenticandoci per un attimo del resto del mondo.

Un ringraziamento peculiare va a *Giò* e *Giuse*. Avete avuto la pazienza di cucinare in silenzio per farmi camminare senza sosta attorno a un tavolo e ripetere la teoria dello spin. Mi avete sopportato sia in configurazione di spin up che in spin down, e vi ringrazio. È stato bello ridere insieme e rendermi conto di quanto possa sembrare stravagante un tizio che parla di spazi a infinite dimensioni.

Ci sono molti che non ho nominato, ma sappiate che ognuno di voi mi ha trasmesso qualcosa. Anche solo una frase, uno stimolo, un'idea. Avrei potuto imparare la fisica da solo in una grotta, munito di soli libri e articoli, ma averlo fatto in un ambiente dinamico e stimolante è quanto di più prezioso mi sia capitato, ha reso possibile concludere questo percorso con molta più carica di quanta ne avessi all'inizio.