Contents

1 Introduction

According to the present knowledge of physics, the ultimate constituents of matter are *quarks* and *leptons*. Both of them are particles of spin 1/2 that interact by interchanging spin 1 particles, namely *photons, gluons,* W^+, W^- and Z^0. The existence of the Higgs spin 0 particle is presently under experimental investigation.
The issues of relativity and quantum mechanics, that are strictly necessary to understand atomic and subatomic world, have favored the development of local field theories in which, as we said, the interactions are mediated by the interchange of the (virtual) integer spin particles mentioned above. A general feature of these theories is that, in the field Lagrangian or Hamiltonian, the interaction term is *simply* added to the term that represents the free motion of the particles.
As for the free term of the matter, spin 1/2, particles, it gives rise to the Dirac equation, that represents the relativistic, quantum mechanical wave equation for these particles.
These arguments explain the great importance of Dirac equation for the study of particle physics at *fundamental* level. However, it is also strictly necessary to understand many important aspects of atomic physics, nuclear physics and of the phenomenological models for hadronic particles.

An introduction to this equation represents the objective of the present work that is mainly directed to students with good foundations in nonrelativistic quantum mechanics and some knowledge of special relativity and classical electrodynamics.
We shall not follow the historical development introduced by Dirac and adopted by many textbooks. In that case, the Lorentz transformation (boost) of the Dirac spinors is performed only in a second time, without clarifying sufficiently the connection between the mathematics and the physical meaning of that transformation.
In this paper the Dirac equation will be derived starting from the basic principles of special relativity and quantum mechanics, analyzing the transformation properties of the relativistic spinors.
This development will be carried out without entering into the mathematical details of the Lorentz group theory, but keeping the discussion at a more *physical* level only using the mathematical tools of linear vector algebra, as row by column matrix product and vector handling.

In our opinion this introductory approach is highly recommendable in order to stimulate the students to make independent investigations by using the powerful concept of *relativistic covariance*.

In a subsequent work we shall analyze in more detail the properties of Dirac equation and derive some relevant observable effects. To that work we shall also defer an introduction to the field theory formalism that is needed to give a complete physical description of subatomic world.

The subjects of the present work are examined in the following order.

In Subsection 1.1 we give some tedious but necessary explanations about the adopted notation.
In Section 2 we study some relevant aspects special relativity, focusing our attention on the properties of the Lorentz transformations.
Their fundamental properties are recalled in Subsection 2.1.
We briefly analyze, in Subsection 2.2, classical electrodynamics as a *relativistic* fields theory.
In Subsection 2.3 we examine the hyperbolic parametrization of the Lorentz transformations, introducing concepts and techniques that are widely applied in relativistic quantum mechanics for the construction of the boost operators.
Lorentz transformations in an arbitrary direction are given in subsection 2.4.
A very important point of this work is studied in Subsection 2.5, where the commutation rules of the Lorentz boost generators, rotation generators and parity transformation are derived.

In Section 3 we make use of the concepts of relativity to lay the foundations of relativistic quantum mechanics.
In Subsection 3.1 we discuss, as an example, the relativistic wave equation for a spin 0 particle.
In Subsection 3.2 we introduce the (quantum-mechanical) Dirac equation for spin 1/2 particles, starting from the commutation rules of the boost generators, rotation generators and parity transformation.
The properties of the Dirac Gamma matrices and their different representations are examined in Subsection 3.3.
Some relevant matrix elements of Dirac operators, as γ^5, are studied in Subsection 3.4.

Finally, plane wave solutions and the corresponding conserved current are found and discussed in Subsection 3.5.

The Appendix is devoted to study some useful properties of the Pauli matrices.

1.1 Notations and Conventions

We suggest the reader to read cursorily this Subsection and to go back to it when he finds some difficulty in understanding the other parts of the paper. First of all, the space time position of a particle is denoted as $x^\mu = (x^0, \mathbf{r})$ with $x^0 = ct$ and $\mathbf{r} = (x^1,\ x^2,\ x^3)$. To avoid confusion, we use this last notation *instead of* the standard one, that is $(x,\ y,\ z)$.
Greek letters of the "middle" part of the alphabet, as $\mu,\ \nu,\ \rho,\ \sigma, ...$ running from 0 to 3, are used to denote four-vector components. On the other hand the letters of the beginning of the Greek alphabet, as $\alpha,\ \beta,\ \delta, ...$ running from 1 to 3, denote three-vector components. This last notation with *upper indices* will be used extensively even though the corresponding quantity does not make part of a four-vector.
Repeated indices are always *summed*, unless otherwise explicitly stated.
For two three-vectors, say $\mathbf{a}$ and $\mathbf{b}$, the scalar product is denoted as

$$\mathbf{ab} = a^\alpha b^\alpha$$

If one of the two vectors is a set of the *three* Pauli (σ^δ) or Dirac (α^δ), (γ^δ) matrices, we use the notation

$$(\sigma\mathbf{a}) = \sigma^\delta a^\delta, \qquad (\alpha\mathbf{a}) = \alpha^\delta a^\delta, \qquad (\gamma\mathbf{a}) = \gamma^\delta a^\delta$$

Furthermore, the notation ∇ collectively indicates the derivatives with respect to the three components of the position vector $\mathbf{r}$.
Lower indices are only used for four-vectors and denote their *covariant components* as explained just after eq.(2.2). Invariant product of two four-vectors is introduced in eq.(2.3). For the unit vectors we use the standard notation

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|}$$

When a four-vector is used as an *argument* of a field or wave function, the Lorentz index $\mu,\ \nu,\ \rho,\ \sigma, ...$ is dropped and, more simply , we write

$$A^\mu(x), \qquad \psi(x)$$

where x represents collectively all the components of the four-vector x^μ.
In order to denote products of matrices and four-vectors, we arrange the components of a four-vector, say x^μ in a *column* vector, denoted as $[x]$. The corresponding *transposed* vector $[x]^T$ is a *row* vector. Standard Latin letters, without indices, are used to denote matrices. See, for example, eq.(2.6). We use this notation also for the set of the *four* Dirac matrices α^μ at the end of Subsection 3.2.

Four components Dirac spinors, introduced in Subsection 3.2, are handled according to the same rules of vector algebra. They are denoted by a Latin letter *without* parentheses.
We recall that the hermitic conjugate of the Dirac spinor u is a row spinor defined as:

$$u^\dagger = u^{*T}$$

For the *commutator* of two matrices (or operators), say Q, R, we use the notation

$$[Q, R] = QR - RQ$$

For the anticommutator we use curly brackets

$$\{Q, R\} = QR + RQ$$

2 Relativity

The principle of relativity, that was found by Galilei and Newton, states that it is possible to study physical phenomena from *different* inertial reference frames (RF) by means of the *same* physical laws. The hypothesis of an *absolute* reference frame is not allowed in physics.
Obviously, one has to transform the result of a measurement performed in a reference frame to another reference frame, primarily the measurements of time and space.
Requiring the speed of light c to be *independent* of the speed of the reference frame, as shown by th Michelson-Morley experiment, one obtains the *Lorentz transformations* that represent the formal foundation of Einstein's special relativity. The reader can find in ref.[1] a simple and satisfactory development of this point.

2.1 Fundamental Aspects of Lorentz Transformations

Considering a RF $\mathcal{S}'$ moving at velocity v along the x^1-axis with respect to $\mathcal{S}$, one has the standard Lorentz transformations

$$\begin{aligned} x'^0 &= \gamma(x^0 - \frac{v}{c}x^1) \\ x'^1 &= \gamma(-\frac{v}{c}x^0 + x^1) \\ x'^2 &= x^2 \\ x'^3 &= x^3 \end{aligned} \tag{2.1a}$$

where $x^0 = ct$, $(x^1, x^2, x^3) = \mathbf{r}$ and $\gamma = [1 - (v/c)^2]^{-1/2}$.

A thorough study of the subject of this Subsection, that consists in generalizing the previous equations, can be found in ref.[2]. In the present paper we highlight some specific aspects that are relevant for a quantum-mechanical description of elementary particles.

The Lorentz transformations of eq.(2.1a) can be syntetically written as

$$x'^\mu = L^\mu{}_\nu(v)x^\nu \tag{2.1b}$$

where the indices μ, ν take the values $0, 1, 2, 3$ and x^μ is denoted as *contravariant* four-vector.
By introducing the Minkowsky metric tensor

$$g_{\mu\nu} = g^{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{2.2}$$

one can construct *covariant* four-vectors $x_\mu = g_{\mu\nu}x^\nu$ and *invariant* quantities as *products* of covariant and contravariant four-vectors. For example, given two contravariant four-vectors, say $s^\mu = (s^0, \mathbf{s})$ and $l^\mu = (l^0, \mathbf{l})$, one can construct their covariant counterparts $s_\mu = (s^0, -\mathbf{s})$, $l_\mu = (l^0, -\mathbf{l})$ and the quantity

$$s_\mu l^\mu = s^\mu l_\mu = s_\mu g^{\mu\nu} l_\nu = s^\mu g_{\mu\nu} l^\nu = s^0 l^0 - \mathbf{sl} \tag{2.3}$$

that is *invariant* under Lorentz transformation:

$$s_\mu l^\mu = s'_\mu l'^\mu \tag{2.4}$$

In particular, the Lorentz transformation of eq.(2.1a) is obtained [1,2] by requiring the invariance of the propagation of a spherical light wave, that is the invariance of $x^\mu x_\mu = 0$.
The invariance equation (2.4) requires

$$g_{\mu\rho} L^\rho{}_\nu(v) L^\mu{}_\sigma(v) = g_{\nu\sigma} \tag{2.5}$$

In many cases it is very useful to work with standard linear algebra notation. Furthermore, at pedagogical level, this technique is very useful to introduce standard handling of Dirac spinors.
Identifying a four-vector x^μ with the column vector $[x]$, the invariant product of eq.(2.4) is written as

$$s^\mu g_{\mu\nu} l^\nu = [s]^T g[l] \tag{2.6}$$

where the upper symbol T denotes the operation of transposition. By means of this notation, eq.(2.5) reads

$$L(v) g L(v) = g \tag{2.7}$$

where we have used the important property, directly obtained from eq.(2.1a), that $L^T(v) = L(v)$. Also, a covariant four-vector x_μ is $[x_c] = g[x]$. Its transformation is

$$[x'_c] = gL(v)[x] = gL(v)gg[x] = gL(v)g[x_c] \tag{2.8}$$

Let us now multiply eq.(2.7) by g from the right, obtaining

$$L(v)gL(v)g = 1 \tag{2.9}$$

In consequence

$$gL(v)g = L^{-1}(v) \tag{2.10}$$

it means that the covariant four-vectors, look at eq.(2.8)!, transform with the *inverse* Lorentz transformations. By means of direct calculation or by using the principle of relativity one finds that

$$L^{-1}(v) = L(-v) \tag{2.11}$$

We recall some relevant physical quantities that are represented by (i.e. transform as) a four-vector. As previously discussed, we have the four-position (in time and space) of a particle denoted by x^μ.

We now define the four-vector that represents the energy and momentum of a particle.

Previously, we introduce the (invariant) rest mass of the particle. In the present work this quantity will be simply denoted as the *mass* m. We shall *never* make use of the so-called relativistic mass.

We also define the differential of the proper (invariant) time as

$$d\tau = \frac{1}{c}[dx_\mu dx^\mu]^{1/2} = \left[(dt)^2 - \frac{1}{c^2}(d\mathbf{r})^2\right]^{1/2} =$$

$$= dt\left[1 - (\frac{\mathbf{v}}{c})^2\right]^{1/2} = \frac{dt}{\gamma} \tag{2.12}$$

where the velocity

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}$$

represents the standard physical velocity of the particle measured by an observer in a given reference frame. Furthermore, the factor γ is a function of that velocity, of the form:

$$\gamma = \left[1 - (\frac{\mathbf{v}}{c})^2\right]^{-1/2}$$

The energy-momentum four-vector is obtained differentiating the four-position with respect to the proper time and multiplying the result by the mass m. One has

$$p^\mu = (\frac{E}{c}, \mathbf{p}) = m\frac{dx^\mu}{d\tau} = (mc\gamma, m\mathbf{v}\gamma) \tag{2.13}$$

In previous equation, E represents the energy of the particle and $\mathbf{p}$ its three-momentum. More explicitly, the energy is

$$E = mc^2\gamma$$

For small values of the velocity $|\mathbf{v}| << c$ one recovers the nonrelativistic limit, that is

$$E \simeq mc^2 + \frac{1}{2}m\mathbf{v}^2 + ... \tag{2.14a}$$

$$\mathbf{p} \simeq m\mathbf{v} + ... \tag{2.14b}$$

Note that the energy and momentum of a particle belong to the four-vector of eq.(2.13). In consequence, energy and momentum conservation can be written in a *manifestly covariant* form. For example, in a collision process in which one has a transition from an initial state (I) with N_I particles, to a final state (F) with N_F particles, the total energy and momentum conservation is written by means of the following four-vector equality

$$\sum_{i=1}^{N_I} p_i^\mu(I) = \sum_{i=1}^{N_F} p_i^\mu(F) \tag{2.15}$$

that holds in any reference frame. A complete discussion of the physical consequences of that equation and related matter is given in ref.[3]. Only recall that, at variance with nonrelativistic mechanics, mass is *not conserved*. In general, mass-energy transformations are represented by processes of creation and destruction of particles. As a special case, a scattering reaction is defined *elastic*, if all the particles of the final state remain *the same* (obviously, with the same mass) as those of the initial state.

Four-momentum conservation of eq.(2.15) is a very simple example. In general, a physical law written in a *manifestly covariant* form automatically fulfills the principle of relativity introduced at the beginning of this section. A physical law is written in a *manifestly covariant* form when it is written as an equality between two relativistic tensors of the same rank: two Lorentz invariants (scalars), two four-vectors, etc..

Going back to eq.(2.13) one can construct the following invariant

$$p^\mu p_\mu = \left(\frac{E}{c}\right)^2 - \mathbf{p}^2 = (mc)^2 \tag{2.16}$$

The second equality is obtained in the easiest way by calculating the invariant in the *rest frame* of the particle, where $p^\mu = (mc, \mathbf{0})$.
From the previous equation one can construct the Hamiltonian of a particle, that is the energy written as function of the momentum

$$E = [(\mathbf{p}c)^2 + (mc^2)^2]^{1/2} \tag{2.17}$$

that in the nonrelativistic limit reduces to

$$E \simeq mc^2 + \frac{\mathbf{p}^2}{2m} + ...$$

Note that in eq.(2.17) we have taken only the *positive* value of the square rooth. This choice is perfectly legitimate in a classical context, where the energy changes its value in a continous way. On the other hand negative energy solutions cannot be discarded when considering quantum-mechanical equations.

From eqs.(2.13) and (2.17), the velocity of a particle is

$$\mathbf{v} = \frac{\mathbf{p}}{E} \; , \qquad |\mathbf{v}| \le c$$

In the second relation, the equality is satified by *massless* particles. The constraint on velocity has a more general validity, as we shall see when revising electromagnetism: *a physical particle* cannot have a velocity greater than the speed of light c.
In this concern we observe that the word *information* has been also used but it should be clarified what does the word *information* really mean. Furthermore, the concepts and the requirements of quantum mechanics about physical states and measurements have not been taken into account when introducing that constraint.

For the study of both classical and quantum-mechanical (field) theories it is very important to determine the transformation properties of the derivative operator

$$\frac{\partial}{\partial x^\nu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial \mathbf{r}}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right)$$

The reader is suggested to derive them by using directly the chain rule. We propose here a simpler proof. Let us consider the invariant $x^\nu x_\nu = (x^0)^2 - \mathbf{r}^2$ and apply to it the derivative operator. One has

$$\frac{\partial}{\partial x^\mu} x_\nu x^\nu = 2x_\mu = 2(x^0, -\mathbf{r}) \tag{2.18a}$$

That is, the derivative with respect to the contravariant components gives, *and transforms as* , a covariant four-vector ($2x_\mu$ in the previous equation). Conversely, the derivative with respect to the covariant components *transforms as* a contravariant four-vector:

$$\frac{\partial}{\partial x_\mu} x_\nu x^\nu = 2x^\mu = 2(x^0, \mathbf{r}) \tag{2.18b}$$

For this reason the following notation is introduced

$$\frac{\partial}{\partial x^{\mu}} = \partial_{\mu} \tag{2.19a}$$

and

$$\frac{\partial}{\partial x_{\mu}} = \partial^{\mu} \tag{2.19b}$$

Straightforwardly one verifies that

$$\frac{\partial}{\partial x^{\mu}}\frac{\partial}{\partial x_{\mu}} = \partial_{\mu}\partial^{\mu} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \tag{2.20}$$

is an *invariant* operator.

2.2 Electromagnetism and Relativity

The elements that have been developed in the preceding Subsection will help us to understand the relativistic properties of classical electromagnetism.
In summary, electromagnetism is a *local* theory in which the interaction between charged particles is carried by the electromagnetic *field*, at light speed c. A complete analysis of this theory can be found, for example, in refs.[2,4].

With respect to interaction propagation, the reader should realize that Newton's theory of gravitational interaction is not compatible with special relativity. In fact the gravitational potential energy

$$V_g = -\frac{Gm_1m_2}{r}$$

depends *instantaneously* on the distance r between the two bodies. If one body, say the #1, changes its position or state, the potential energy, and in consequence, *the force* felt by the body #2 changes at the *same instant*, implying a transmission of the interaction at *infinite* velocity.
Note that, on the other hand, the expression of Coulomb potential energy, that is formally analogous to the Newton's gravitational one, holds exactly *exclusively in the static case*. According to classical electromagnetism, if the interacting particles are in motion, it represents *only approximatively* their

interaction. This approximation is considered good if their relative velocity is

$$|\mathbf{v}| << c$$

The fundamental quantity of electromagnetism is the vector potential *field* $A^\mu = (A^0, \mathbf{A})$.
A field is, by definition, a function of the time-space position x^ν. As done in most texbooks, in the following we shall drop the index ν of the argument, simply writing $A^\mu = A^\mu(x)$.
Synthetically, we recall that the Maxwell equations have the form

$$\partial_\nu \partial^\nu A^\mu = \frac{4\pi}{c} j^\mu \tag{2.21}$$

with the Lorentz invariant Gauge condition

$$\partial_\mu A^\mu = 0 \tag{2.22}$$

where we have introduced the *current density*

$$j^\mu = (c\rho(x), \mathbf{j}(x)) \tag{2.23}$$

Applying the derivative operator ∂_μ to eq.(2.21) and using eq.(2.22), one finds the current conservation equation, that is

$$\partial_\mu j^\mu = \frac{\partial}{\partial t}\rho(x) + \frac{\partial}{\partial \mathbf{r}}\mathbf{j}(x) = 0 \tag{2.24}$$

All the equations written above are *manifestly covariant* and the Lorentz transformations can be easily performed. If a solution of eqs.(2.21) and (2.22) is found in a reference frame $\mathcal{S}$, it is not necessary to solve the equations in the reference frame $\mathcal{S}'$, but simply one can transform the electromagnetic field:

$$A'^\mu(x') = L^\mu{}_\nu(v) A^\nu(x(x')) \tag{2.25}$$

In more detail, one has
(i) to tranform the field A^μ, mixing its components by means of $L^\mu{}_\nu(v)$, that is the first factor of the previous equation, but also
(ii) to express the argument x of the frame $\mathcal{S}$ as a function of x' measured in $\mathcal{S}'$, that is, recalling eq.(2.11)

$$x^\nu = L^\nu{}_\rho(-v) x'^\rho$$

We briefly define the last operation as *argument re-expression*.

The reader should note that such *double* transformation occours in the same way when a rotation is performed. In this case the space components $\mathbf{A}$ are mixed by the rotation matrix (for this reason the electromagnetic field is defined as a *vector* field) and the argument $\mathbf{r}$ must be expressed in terms of $\mathbf{r}'$ by means of the inverse rotation matrix. Under rotation, in the time component A^0, one only has the argument re-expression of $\mathbf{r}$.

In principle it is possible to construct a *scalar* field theory (even though there is no evidence of such theories at macroscopic level). In this case the field is represented by a one-component function $\phi(x)$. Both the Lorentz transformation and the rotations only affect the argument x in the same way as before, but no mixing can occur for the *single* component function ϕ. One has only to perform the *argument re-expression*.

We shall now explain with a physical relevant example the use of the transformation (2.25) for the electromagnetic field.
Let us consider a charged particle moving with velocity u along the x^1-axis. What is the field produced by this particle ?
We introduce a reference frame $\mathcal{S}$ in which the particle is at rest, while the observer in $\mathcal{S}'$ sees the particle moving with velocity u along x^1. The velocity of $\mathcal{S}'$ with respect to $\mathcal{S}$ is $v = -u$. The field in $\mathcal{S}$ is purely electrostatic, that is

$$A^0 = A^0(ct, \mathbf{r}) = \frac{q}{|\mathbf{r}|} \qquad (2.26a)$$

$$\mathbf{A} = \mathbf{A}(ct, \mathbf{r}) = \mathbf{0} \qquad (2.26b)$$

where q represents the charge of the particle. We find A'^μ by means of eq.(2.25). First, one has

$$A'^0 = \gamma A^0 \qquad (2.27a)$$

$$A'^1 = \frac{u}{c}\gamma A^0 \qquad (2.27b)$$

$$A'^\alpha = A^\alpha = 0 \qquad (2.27c)$$

with $\alpha = 2, 3$ and $\gamma = [1 - (u/c)^2]^{-1/2}$.
Now we express $|\mathbf{r}|$ in terms of $(ct', \mathbf{r}')$, that is we perform the *argument re-expression*.

By means of eqs.(2.1) and (2.11) one has

$$x^1 = \gamma(-ut' + x'^1)$$

$$x^\alpha = x'^\alpha$$

so that

$$|\mathbf{r}| = [\gamma^2(-ut' + x'^1)^2 + (x'^2)^2 + (x'^3)^2]^{1/2} \tag{2.28}$$

By means of the previous equation the final expression for the field of eqs.(2.27a,b) is

$$A'^0(ct', \mathbf{r}') = q\gamma[\gamma^2(-ut' + x'^1)^2 + (x'^2)^2 + (x'^3)^2]^{-1/2} \qquad (2.29a)$$

$$A'^1(ct', \mathbf{r}') = \frac{u}{c} A'^0(ct', \mathbf{r}') \qquad (2.29b)$$

This example has been chosen to explain the procedure for transforming a field function.
The field of eqs.(2.29a,b) can be directly derived in the frame $\mathcal{S}'$ by solving the Maxwell equations (2.21),(2.22) as done in refs.[2,4]. The technique of the Liénard Wiechert potentials can be used. But, as the reader should check, much more mathematical efforts are required.

2.3 The Hyperbolic Parametrization of the Lorentz Transformations

Going back to the Lorentz transformations of eq.(2.1) we note that the coefficients of the of the transformation matrix $L(v)$ are:

$$\gamma \ , \qquad\qquad -\frac{v}{c}\gamma$$

We square both terms (the minus sign disappears in the second one) and take the difference, obtaining

$$\gamma^2 - \left(\frac{v}{c}\gamma\right)^2 = 1$$

Recalling that the hyperbolic functions satisfy the relation

$$ch^2\omega - sh^2\omega = 1$$

one can choose the following parametrization

$$\gamma = ch\ \omega\ , \quad \frac{v}{c}\gamma = sh\ \omega \tag{2.30a}$$

with

$$\frac{v}{c} = th\ \omega \tag{2.30b}$$

that connects the *hyperbolic* parameter ω with the standard velocity v. We now show the reason why the parametrization $L(\omega)$ is very useful for the following developments.

Let us consider two subsequent Lorentz transformations along the x^1 axis with hyperbolic parameters η and ξ. The total transformation is given by the following product of Lorentz transformations, that, by using the vector algebra notation, is written in the form

$$[x'] = L(\eta)L(\xi)[x] \tag{2.31}$$

The reader can calculate explicitly $L(\eta)L(\xi)$ by means of standard rules for row by column matrix product, then recalling

$$sh(\eta + \xi) = sh\ \eta\ ch\ \xi + sh\ \xi\ ch\ \eta$$

$$ch(\eta + \xi) = ch\ \eta\ ch\ \xi + sh\ \xi\ sh\ \eta$$

one finds

$$L(\eta)L(\xi) = L(\eta + \xi) \tag{2.32}$$

Note that the previous result strictly depends on the chosen hyperbolic parametrization. Due to the relativistic *nonlinear* composition of velocities, considering two subsequent Lorentz transformations, with $v/c = th\ \eta$ and $w/c = th\ \xi$, one has, in contrast to eq.(2.32),

$$L(v)L(w) \neq L(v + w) \tag{2.33}$$

On the other hand, the composition of two Lorentz transformations with hyperbolic parametrization, as given in eq.(2.32), has the same form as the composition of two rotations around the same axis.

Eq.(2.32) is the clue for the following development.

By means of hyperbolic parametrization, we can now turn to express a finite Lorentz transformation in terms of the corresponding infinitesimal transformation.

Let us consider the case of *small* velocity, that is $v/c << 1$ or equivalently, for the hyperbolic parameter, $\omega \simeq 0$ (see eq.(2.30b)).
In particular, at first order in ω or in v/c, one has

$$ch\ \omega \simeq 1\ ,\quad sh\ \omega \simeq \omega \simeq \frac{v}{c}$$

and, in consequence

$$L(\omega) \simeq 1 + \omega K^1 \tag{2.34}$$

where 1 and K^1 respectively represent the identity matrix and the *generator* of the Lorentz transformation matrix along the x^1 axis. This second term is usually called *boost generator*.
Explictly, the matrix K^1 is easily obtained considering eqs.(2.1), (2.31a) and the above Taylor expansions of the hyperbolic functions. It has the form

$$K^1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = -\begin{bmatrix} \sigma^1 & 0 \\ 0 & 0 \end{bmatrix} \tag{2.35}$$

The second expression in the previous equation is given for pedagogical reasons, that is to familiarize the reader with *block* matrices.
In fact, the 4×4 matrix K^1 is written as a *block* matrix, in which each block is represented by a 2×2 matrix. In particular the upper left block is the Pauli matrix σ^1, in the other "0" blocks the four entries of each block are all vanishing. The properties of the Pauli matrices are studied in the Appendix. For their definition see eq.(A.1).
The reader should note the following two points:
(i) there is no direct connection between σ^1 of the previous equations and the quantum mechanical spin operator,
(ii) as for the row by column product of a block matrix, the same rules of standard matrices must be used.

In order to reconstuct the finite boost $L(\omega)$, (ω finite), we apply N times, with $N \to \infty$, the infinitesimal transformation of eq.(2.34).
The *linear* boost composition law of eq.(2.32) allows to derive the following equation

$$L(\omega) = \lim_{N\to\infty} \left(1 + \frac{\omega}{N} K^1\right)^N = \exp(\omega K^1) \tag{2.36a}$$

$$= 1 + (ch\ \omega - 1) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + sh\ \omega K^1 \qquad (2.36b)$$

The second equality of eq.(2.36a)is obtained by comparing the series expansion in powers of ω of the exponential, with $\left(1 + \frac{\omega}{N} K^1\right)^N$ for $N \to \infty$.
Eq.(2.36b) is derived working on that series expansion. One has the following rules for the powers of K^1

$$(K^1)^0 = 1 \ , \qquad (K^1)^{2n} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \ , \qquad (K^1)^{2n+1} = K^1$$

that, for example, can be derived from the corresponding properties of σ^1 by means of eq.(A.4).
The coefficients that multiply $(K^1)^{2n}$ and $(K^1)^{2n+1}$ can be *summed up*, giving $ch\ \omega - 1$ and $sh\ \omega$, respectively. One can straightforwardly check that eq.(2.36b) is equal to eq.(2.1) with the hyperbolic parametrization of eq.(2.30a). We have developed in some detail this example as a guide to construct the finite boost transformations for Dirac spinors in eqs.(3.17a,b).

We remind the reader that all the relevant properties of the boost transformation are contained in the infinitesimal form given in eq.(2.34) with the matrix boost generator of eq.(2.35). The finite expression of the boost is obtained by means of a standard mathematical procedure that does not add new physical information.

2.4 Lorentz Transformations in an Arbitrary Direction

In the previous developments we have considered Lorentz transformations along the x^1-axis. The transformations along the x^2- and x^3-axis are directly obtained interchanging the spatial variables. In this way (as done in ref.[4]) one obtains a sufficiently general treatment of relativistic problems. For completeness and to help the reader with the analysis of some textbooks (as for example ref.[2]) and research articles, we now study Lorentz transformations with an arbitrary boost velocity $\mathbf{v}$ direction. The comprehension of the other Sections of this work does not depend on this point. In consequence, the reader (if not interested) can go directly to eq.(2.41).
For definiteness we consider the time-space four-vector $x^\mu = (x^0, \mathbf{r})$, but the results hold for *any* four-vector.
The transformation equation (2.1) can be generalized in the following way

$$x'^{\,0} = \gamma(x^0 - \frac{\mathbf{rv}}{c}) \tag{2.37a}$$

$$\mathbf{r}'\hat{\mathbf{v}} = \gamma(-\frac{v}{c}x^0 + \mathbf{r}\hat{\mathbf{v}}) \tag{2.37b}$$

$$\mathbf{r}'_\perp = \mathbf{r}_\perp \tag{2.37c}$$

where the unit vector $\hat{\mathbf{v}}$ has been introduced so that $\mathbf{v} = v\hat{\mathbf{v}}$ with $v > 0$ and the notation $\mathbf{r}_\perp$ denotes the spatial components of $\mathbf{r}$ perpendicular to $\mathbf{v}$.
Eq.(2.37a) directly represents the Lorentz transformation of the time component of a four-vector for an arbitrary direction of the boost velocity.
Some handling is necessary for the spatial components of the four-vector. Starting from eqs.(2.37b,c) we now develop the transformation for $\mathbf{r}$. One can parametrize this transformation according to the following hypothesis

$$\mathbf{r}' = \mathbf{r} + f(v)\frac{1}{c^2}(\mathbf{rv})\mathbf{v} + g(v)\frac{1}{c}x^0\mathbf{v} \tag{2.38}$$

Note that it correctly reduces to the identity when $v = 0$ and automatically gives eq.(2.37c) for the perpendicular components of the four-vector.
Multiplying the previous equation by $\hat{\mathbf{v}}$ and comparing with eq(2.37b) one finds

$$g(v) = -\gamma \tag{2.39a}$$

$$f(v) = \frac{\gamma - 1}{(\frac{v}{c})^2} = \frac{\gamma^2}{\gamma + 1} \tag{2.39b}$$

where the last expression of eq.(2.39b) is obtained by using the standard definition of the factor γ.
Analogously to eq.(2.36a), the Lorentz transformation for the four-vector given by eqs.(2.37a) and (2.38) can be written in exponential form, as

$$L(\omega\hat{\mathbf{v}}) = \exp(\omega\hat{\mathbf{v}}\mathbf{K}) \tag{2.40}$$

with the same connection between ω and v as in eq.(2.30b). The matrices of the boost generator $\mathbf{K} = (K^1, K^2, K^3)$ are defined as

$$K^1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad K^2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad K^3 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \tag{2.41}$$

Note that K^1 had been already derived in eq.(2.35). Furthermore, K^2 and K^3 can be directly obtained performing the Lorentz transformation analogously to eq.(2.1) but along the x^2- and x^3-axis and repeating the procedure that leads to eq.(2.35).

2.5 The Commutation Rules of the Boost Generators

The most important property of the boost generators or, more precisely, of the matrices $\mathbf{K}$ given in eq.(2.41), is represented by their *commutation rules*. Let us consider an illustrative example. For generality, we shall denote the Lorentz transformation as *boost*, using the symbol B.
In a *first step*, we perform a boost along x^2 with a *small* velocity. At first order in the hyperbolic parameter ω^2 one has

$$B^2 \simeq 1 + \omega^2 K^2$$

Analogously, in a *second step*, we make a boost along x^1, with hyperbolic parameter ω^1, that is

$$B^1 \simeq 1 + \omega^1 K^1$$

The total boost, up to order $\omega^1\omega^2$, is

$$B^{12} = B^1 B^2 \simeq 1 + \omega^1 K^1 + \omega^2 K^2 + \omega^1 \omega^2 K^1 K^2 \tag{2.42a}$$

Note the important property that the product of two boosts is a Lorentz boost because it satisfies eq.(2.7), as it can be directly verified.
We now repeat the previous procedure *inverting* the order of the two boosts, obtaining

$$B^{21} = B^2 B^1 \simeq 1 + \omega^1 K^1 + \omega^2 K^2 + \omega^1 \omega^2 K^2 K^1 \tag{2.42b}$$

What is the difference between the two procedures ? Subtraction of eqs.(2.42a,b) gives

$$B^{12} - B^{21} \simeq \omega^1 \omega^2 [K^1, K^2] \tag{2.43}$$

where the standard notation for the commutator of the matrices K^1 and K^2 has been introduced. Explicit calculation gives

$$[K^1, K^2] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.44}$$

At this point two (connected) questions are in order. What is the meaning of the *noncommutativity* of the boost generators? Which physical quantity is represented by the commutator of the last equation?
To answer these questions it is necessary to recall some properties of the *rotations*.
They are initially defined in the three dimensional space. Let us rotate the vector $\mathbf{r}$ counterclockwise, around the x^3 axis, of the angle θ^3. For a small angle, at first order in θ^3, one obtains the rotated vector

$$\mathbf{r}' \simeq \mathbf{r} + \theta^3 \hat{\mathbf{k}} \times \mathbf{r} \tag{2.45a}$$

where $\hat{\mathbf{k}}$ represents the unit vector of the x^3 axis. One can put $\mathbf{r}$ and $\mathbf{r}'$ in the three component column vectors $[\mathbf{r}]$ and $[\mathbf{r}']$ so that the previous equation can be written with the vector algebra notation as

$$[\mathbf{r}'] \simeq \left(1 + \theta^3 s^3\right) [\mathbf{r}] \tag{2.45b}$$

where s^3 (see the next equation) represents the three-dimensional generator matrix of the rotations around the axis x^3. The same procedure can be repeated for the rotations around the axes x^1 and x^2. The generator matrices are

$$s^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \; s^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \; s^3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.46}$$

As it is well known the (previous) rotation generator matrices do not commute:

$$[s^\alpha, s^\beta] = \epsilon^{\alpha\beta\delta} s^\delta \tag{2.47}$$

where we have introduced the Levi-Civita antisymmetric tensor $\epsilon^{\alpha\beta\delta}$.
As for the *noncommutativity*, this situation is partially similar to the case of the boost generators shown in eq.(2.43) but, for the rotations, eq.(2.47) shows that, given two generator matrices, their commutator is proportional to the *third* matrix, while we have not yet identified the physical meaning of the matrix in the *r.h.s.* of eq.(2.44).

Pay attention ! In quantum mechanics, from eqs.(2.46), (2.47) we can introduce the *spin 1* operators as

$$j_1^\alpha = -i\hbar s^\alpha$$

satisfying the standard angular momentum commutation rules. Our s^α do not *directly* represent the three spin operators.

It is very important to note that the physical laws must be *invariant* under rotations. To make physics we assume that space is *isotropic*. As the hypothesis of an absolute reference frame must be refused, in the same way the idea of a *preferencial* direction in the space is not allowed by the conceptual foundations of physics.
Obviously, rotational invariance must be compatible with relativity. This fact is immediately evident recalling that the rotations *mix* the spatial components of a vector without changing the scalar product of two three-vectors, say $\mathbf{a}$ and $\mathbf{b}$: $\mathbf{a}'\mathbf{b}' = \mathbf{ab}$.
The time components of the corresponding four-vectors also remain unaltered: $a'^{\,0} = a^0$ and $b'^{\,0} = b^0$. Consider, as two relevant examples, the time and energy that represent the zero components of the position and momentum four-vectors, respectively.
It means that rotations satisfy the invariance equation (2.3) and, in consequence, they are fully compatible with relativity. In terms of 4×4 matrices, eq.(2.45b) is generalized as

$$[x'] \simeq \left(1 + \theta^3 S^3\right)[x] \tag{2.48}$$

The 4×4 generator matrices are defined in terms of 3×3 s^α as

$$S^\alpha = \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & s^\alpha & \\ 0 & & & \end{array}\right] \tag{2.49}$$

As it will be written in eq.(2.50b), the matrices S^α obviously satisfy the same commutation rules of eq.(2.47).

We can now verify that the *r.h.s.* of eq.(2.44) represents $\;-S^3$.
In general one has the following commutation rules

$$[K^\alpha, K^\beta] = -\epsilon^{\alpha\beta\delta} S^\delta \tag{2.50a}$$

For completeness, we also give

$$[S^\alpha, S^\beta] = \epsilon^{\alpha\beta\delta} S^\delta \tag{2.50b}$$

and

$$[S^\alpha, K^\beta] = \epsilon^{\alpha\beta\delta} K^\delta \qquad (2.50c)$$

where the last equation means that the boost generator $\mathbf{K}$ transforms as a vector under rotations.

As for the derivation of the Dirac equation that will be performed in the next section, we anticipate here that a set of K^α and S^α matrices (different from eqs.(2.41) and (2.46),(2.49)) will be found, that satisfy the *same* commutation rules of eqs.(2.50a-c). In mathematical terms, these new matrices are a different *representation* of the Lorentz group, allowing to satisfy in this way the relativistic invariance of the theory.

For the study of the Dirac equation, it is also necessary to introduce another invariance property related to a new, discrete, *space-time* transformation. It is the *parity transformation*, or *spatial inversion*, that changes the position three-vector $\mathbf{r}$ into $-\mathbf{r}$, leaving the time component unaltered. This definition shows that spatial inversion does not change the invariant product of two four-vectors and, in consequence, is *compatible* with relativity.
Parity is a *discrete* transformation that does not depend on any parameter. On the other hand, recall that rotations are *continous* tranformations, that *continously* depend on the rotation angle. Obviously, spatial inversion cannot be accomplished by means of rotations.

Note that, under parity transformation, ordinary, or *polar* vectors, as for example the momentum $\mathbf{p}$, do change sign in the same way as the position $\mathbf{r}$, while the *axial* vectors, as for example the orbital angular momentum $\mathbf{l} = \mathbf{r} \times \mathbf{p}$, do not change sign. On the other hand they transform in standard way under rotations.

Using the definition given above, parity transformation on the space-time position,

$$[x'] = \Pi[x]$$

is accomplished by means of the diagonal Minkowsky matrix. We can write

$$\Pi = g$$

that holds for the spatial inversion of *all* the four-vectors.

From the previous definition, one can easily verify the following *anticommutation* rule with the boost generators

$$\{\Pi, K^\alpha\} = 0 \tag{2.51a}$$

or equivalently

$$\Pi K^\alpha \Pi = -K^\alpha \tag{2.51b}$$

where we have used the standard property $\Pi^2 = 1$.
Furthermore

$$[\Pi, S^\alpha] = 0 \tag{2.52a}$$

or equivalently

$$\Pi S^\alpha \Pi = S^\alpha \tag{2.52b}$$

It shows that the rotation generators do not change sign under spatial inversion, that is they behave as an axial-vector.
The *determinant* of Lorentz boost and rotations is equal to $+1$, while for spatial inversion it is -1.

Note that eqs.(2.51a)-(2.52b) represent *general* properties of the parity transformation that do not depend on the tensor to which it is applied. They are derived, and hold, in the case of four-vectors, but they are also assumed to hold for the Dirac spinors. But, in this case, the following critical discussion is necessary.

In fact, after these formal developments, we can ask: being parity *compatible* with relativity, are the physical laws of nature *really* invariant under spatial inversion?
The situation is different with respect to rotations, that represent a *necessary* invariance for our understanding of nature.
Initially, parity was considered an invariance of physics, but in the fifties the situation changed. In fact, some experiments on beta decay showed that *weak interactions are not invariant* under spatial inversion. On the other hand, *gravitational, electromagnetic and strong (or nuclear) interactions* are parity invariant.
When deriving the Dirac equation, we shall require the fulfillment of parity invariance, having in mind the study of electromagnetic and strong interactions. In a following work we shall discuss the weakly-interacting neutrino equations, that are *not invariant* under parity transformation.

We conclude this section mentioning another discrete transformation, called *time reversal*, that consists in changing the sign of time: $t' = -t$. Classical laws of physics are invariant with respect to this change of the sense of direction of time. The action of time reversal on the space-time four vector is represented by the matrix $T = -\Pi = -g$.

At microscopic level, time reversal invariance is *exact* in strong and electromagnetic processes, but not in weak interactions. However, this violation is of different kind with respect to that of parity transformation.
We conclude pointing out that in the formalism of field theories the *product* of the three transformations : C (*Charge Conjugation*), P (*Parity*) and T (*Time Reversal*), is an exact invariance, as confirmed by the available experimental data.

3 Relativistic Quantum Wave Equations

In this Section we shall study the procedure to implement the principles of special relativity in the formalism of quantum mechanics in order to introduce the fundamental Dirac equation.
Previously, in Subsection 3.1 we shall analyze the general properties of the four-momentum operator in quantum mechanics and discuss at pedagogical level the Klein-Gordon equation for spinless particles.

3.1 Generalities and Spin 0 Equation

Let us firstly recall the Schrödinger equation for a free particle. In the coordinate representation it has the form

$$i\hbar \frac{\partial \psi(t, \mathbf{r})}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(t, \mathbf{r}) \tag{3.1}$$

It can be obtained by means of the following eqs.(3.2a-c), performing the *translation*, in terms of differential operators acting onto the wave function $\psi(t, \mathbf{r})$, of the standard nonrelativistic expression

$$E = \frac{\mathbf{p^2}}{2m}$$

It clearly shows that Schrödinger equation (3.1) is essentially nonrelativistic or, in other words, *not compatible* with Lorentz transformations.

As discussed in refs.[5,6], the fundamental relation that is used for the study of (relativistic) quantum mechanics associates the four-momentum of a particle to a space-time differential operator in the following form

$$p^\mu = i\hbar\partial^\mu \tag{3.2a}$$

that, as explained in Subsection 2.1, means

$$p^0 c = E = i\hbar\frac{\partial}{\partial t} \tag{3.2b}$$

and

$$\mathbf{p} = -i\hbar\nabla = -i\hbar\frac{\partial}{\partial \mathbf{r}} \tag{3.2c}$$

The reader may be surprised that at relativistic level the *same* relations hold as in nonrelativistic quantum mechanics. As a matter of fact, eqs.(3.2a-c) express *experimental* general properties of quantum waves, as given by the De Broglie hypothesis.
Furthermore, the connection with relativity is possible because $i\hbar\partial^\mu$ is a contravariant four-vector operator.
The easiest choice to write a relativistic wave equation consists in *translating* eq.(2.16) (*instead of* the nonrelativistic expression !) in terms of the space-time differential operators given by the previous equations. One has

$$-\hbar^2\partial_\mu\partial^\mu\psi(t,\mathbf{r}) = (mc)^2\psi(t,\mathbf{r}) \tag{3.3a}$$

or, more explicitly, multiplying by c^2

$$-(\hbar c)^2\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\psi(t,\mathbf{r}) = m^2c^4\psi(t,\mathbf{r}) \tag{3.3b}$$

Exactly as done for the electromagnetic field equations in Subsection 2.3, recalling the invariance of $\partial_\mu\partial^\mu$, one realizes that previous equation is *manifestly covariant.*

In order to make explicit calculations in atomic, nuclear and subnuclear physics, it is necessary to remember some numerical values (and the corresponding units !). We start considering the following quantities that appear in eq.(3.3b):

$$\hbar c = 197.327\ MeV\ fm$$

that is the Planck constant $\hbar$ multiplied by the speed of light c, expressed as an energy multiplied by a length. The energy is measured in MeV

$$1MeV = 10^6 eV = 1.6022 \times 10^{13} Joule$$

and the length in fm *(femtometers or Fermis)*

$$1fm = 10^{-15}m = 10^{-13}cm$$

Furthermore, the particle masses are conveniently expressed in terms of their rest energies. We give a few relevant examples

$$m_e c^2 = 0.511 \ MeV$$

for the electron

$$m_p c^2 = 938.27 \ MeV$$

for the proton, and

$$m_n c^2 = 939.57 \ MeV$$

for the neutron.
Also note that the operator ∂^μ is, dimensionally, a $length^{-1}$, that in our units gives fm^{-1}.

Going back to the formal aspects of eq.(3.3a,b), usually called Klein-Gordon equation, we note the two following aspects:

(i) Being based on the relativistic relation among energy, momentum and mass of eq.(2.16) with the De Broglie hypothesis of eqs.(3.2a-c), the *manifestly covariant* Klein-Gordon equation has a *general* validity, in the sense that the wave fuctions of *all* the relativistic free particles must satisfy that equation. As for the Dirac equation for spin 1/2 particles, see eq.(3.41) and the following discussion.

(ii) In the Klein-Gordon equation it does not appear the particle spin. Or, equivalently, the function $\psi(t, \mathbf{r})$ is a one-component or *scalar* field function that describes a spin 0 particle, as it happens in nonrelativistic quantum mechanics when spin is not included.
The effects of rotations and Lorentz boosts only consist in the *argument re-expression* discussed in Subsection 2.2. As explained in textbooks of quantum

mechanics, see for example ref.[7], the (infintesimal) rotations are performed by using the orbital angular momentum operator as generator.

The Klein-Gordon equation admits plane wave solutions, corresponding to eigenstates of the four-momentum $p^\mu = (\frac{E}{c}, \mathbf{p})$ in the form

$$\psi_p(t, \mathbf{r}) = N \exp\left[\frac{i}{\hbar}(-Et + \mathbf{pr})\right] \tag{3.4a}$$

$$= N \exp\left(-\frac{i}{\hbar} p_\mu x^\mu\right) = N \exp\left(-\frac{i}{\hbar}[p]^T g [x]\right) \tag{3.4b}$$

where N represents a normalization constant. The expression (3.4b) has been written using explicitly the Lorentz covariant notation.
The most relevant point here is that the energy eigenvalue E can assume both positive and negative values (we shall see that it holds true also for Dirac equation !) We have

$$E = p^0 c = \lambda \epsilon(\mathbf{p}) \tag{3.5a}$$

where

$$\epsilon(\mathbf{p}) = [(\mathbf{p}c)^2 + (mc^2)^2]^{1/2} \tag{3.5b}$$

and the *energy sign* $\lambda = +/-1$ have been introduced.
In quantum mechanics the $\lambda = -1$ solutions cannot be eliminated. They are strictly necessary to have a *complete* set of solutions of the wave equation. They can be correctly interpreted by means of charge conjugation in the framework of field theory, as done in most textbooks. Historically, starting from the work by Dirac, negative energy solutions lead to the very important discovery of the *antiparticles*, that have the same mass (and spin) but *opposite charge* with respect to the corresponding particles.
We shall not analyze this problem here but postpone it to a subsequent work.

As for the positive energy solutions, one can immediately check that in the nonrelativistic regime ($|\mathbf{p}|c << mc^2$) the Schrödinger limit is obtained.

As an illustrative exercise, it may be useful to perform a Lorentz boost in eq.(3.4b). Given that we are considering a scalar field, we have to make only the *argument re-expression*.

In this concern recall that, for positive energy, the wave function of eq.(3.4a,b) represents a particle state such that an observer in $\mathcal{S}$ measures the particle four-momentum p^μ.
In the reference frame $\mathcal{S}'$, for the space-time position one must use

$$[x] = L^{-1}[x']$$

(both the velocity and the hyperbolic parametrizations can be adopted and, for simplicity, no argument has been written in L^{-1}) and replace it in eq.(3.4b). In the argument of plane wave exponential one has

$$[p]^T g[x] = [p]^T g L^{-1}[x'] = [p]^T L g[x'] = [p']^T g[x']$$

where we have used $gL^{-1} = Lg$ from eq.(2.10) and also $[p]^T L = [p']^T$. In the previous result we recognize the invariance equation that, in standard notation, reads

$$p_\mu x^\mu = p'_\mu x'^\mu$$

Physically, it means that an observer in $\mathcal{S}'$ measures the particle transformed four-momentum p'^μ.

The Klein-Gordon equation admits a *conserved current*. We shall consider the form, related to a *transition process*, that is used in perturbation theory to calculate the corresponding probability amplitude.
To derive the conserved current one has to make the following three steps.

(i) Take eq.(3.3b) with a plane wave solution $\psi_{p_I}(t, \mathbf{r})$ for an initial state of four-momentum p_I.

(ii) Take eq.(3.3b) with a plane wave solution $\psi_{p_F}(t, \mathbf{r})$ for a final state of four-momentum p_F and make the complex conjugate.

(iii) Multiply the equation of step (i) by the complex conjugate $\psi^*_{p_F}(t, \mathbf{r})$ and the equation of step (ii) by $\psi_{p_I}(t, \mathbf{r})$. Then subtract these two equations, obtaining

$$[\partial^\mu \partial_\mu \psi^*_{p_F}(t, \mathbf{r})]\psi_{p_I}(t, \mathbf{r}) - \psi^*_{p_F}(t, \mathbf{r})\partial_\mu \partial^\mu \psi_{p_I}(t, \mathbf{r}) = 0 \qquad (3.6)$$

Note that the mass term has disappeared. The previous equation can be equivalently written as a *conservation equation* in the form

$$\partial_\mu J^\mu_{FI}(t, \mathbf{r}) = 0 \qquad (3.7)$$

where the *conserved current* is defined as (multiplying by the conventional factor $i\hbar$)

$$J^\mu_{FI}(t, \mathbf{r}) = i\hbar[\psi^*_{p_F}(t, \mathbf{r})\partial^\mu\psi_{p_I}(t, \mathbf{r}) - (\partial^\mu\psi^*_{p_F}(t, \mathbf{r}))\psi_{p_I}(t, \mathbf{r})] \tag{3.8a}$$

$$= (p^\mu_I + p^\mu_F)N_I N_F \exp\left(\frac{i}{\hbar}q_\mu x^\mu\right) \tag{3.8b}$$

In the last equation the four-momentum transfer $q^\mu = p^\mu_F - p^\mu_I$ of the transition process has been introduced.
The conserved current $J^\mu_{FI}(t, \mathbf{r})$ is *manifestly* a four-vector.

The latter eq.(3.8b), that is obtained by explicit use of the wave functions, is very interesting. The first term $(p^\mu_I + p^\mu_F)$ represents the so-called four-vector *vertex factor*.
Applying to eq.(3.8b) the derivative operator ∂_μ one verifies that current conservation relies on the following *kinematic* property of the vertex factor

$$q_\mu(p^\mu_I + p^\mu_F) = {p_F}^\mu p_{F\mu} - {p_I}^\mu p_{I\mu} = 0 \tag{3.9}$$

that is *automatically* satisfied because the mass of the particle remains *the same* in the initial and final state.

As for the general properties of the current given in eqs.(3.8a,b) we find that in the *static case*, i.e. $p_F = p_I$, the time component J^0_{II} is *negative* if negative energy states ($\lambda = -1$) are considered. It means that one cannot attach to J^0_{II} the meaning of *probability density* as it was done with the Schrödinger equation. For this reason we do not discuss in more detail the plane wave normalization constant N.
Again, a complete interpretation of the Klein-Gordon equation and of its current is obtained in the context of field theory.

3.2 Spin 1/2 Dirac Equation

In nonrelativistic quantum-mechanics a spin 1/2 particle is described by a *two-component* spinor ϕ. The spinor rotation is performed by mixing its components. At first order in the rotation angle θ^α, one has

$$\phi' \simeq (1 - \frac{i}{2}\theta^\alpha\sigma^\alpha)\phi \tag{3.10}$$

where the three Pauli matrices σ^α have been introduced. Their properties are studied in the Appendix.
What is important to note here is that the matrix operators $S^\alpha_{[2]} = -\frac{i}{2}\sigma^\alpha$ play the same rôle in realizing the rotations as the matrices S^α defined in eq.(2.49). For this reason, their commutation rules are the same as those given in eq.(2.50b). Also, the *spin* or *intrinsic angular momentum* operator is defined [7] multiplying by $\hbar/2$ the Pauli matrices σ^α.
Formally, we have introduced the two-dimensional representation of the rotation group (the three-dimensional representation corresponds to spin 1, etc.).
Finally, the spatial *argument* of the spinor ϕ (not written expicitly in eq.(3.10)) is rotated with the same rules previously discussed for the arguments of the field functions, that is one has to perform the *argument re-expression*.
In quantum mechanics, the generator of these rotations is the orbital angular momentum operator $\mathbf{l} = \mathbf{r} \times \mathbf{p}$, so that the total angular momentum is given by the three generators of the total rotation (on the spinor and on the argument), in the form

$$j^\alpha = l^\alpha + \frac{\hbar}{2}\sigma^\alpha$$

We can now try to introduce relativity. We shall follow a strategy similar to that of refs.[5,8], but avoiding many unessential (at this level) mathematical details.
First, we note that for a particle *at rest*, the relativistic theory must coincide with the previous nonrelativistic treatment.
Second, we make the following question: can we find a set of three 2×2 boost matrices (acting on the two-component spinors) that satisfy, with the $S^\alpha_{[2]} = -\frac{i}{2}\sigma^\alpha$ replacing the S^α, the same commutation rules as the K^α in eqs.(2.50a-c)?
The answer is *yes*. A simple *inspection* of eqs.(2.50a-c) and use of the standard property of the Pauli matrices given in eq.(A.2) show that the matrices $K^\alpha_{[2]} = \frac{\tau}{2}\sigma^\alpha$ satisfy those commutation rules.
Eq.(2.50a) requires $\tau^2 = 1$, while eq.(2.50c) does not give any new constraint on the parameter τ, that, in consequence can be chosen equivalently as
$\tau = +/-1$.

However, a serious problem arises when trying to introduce the parity transformation matrix. It must satisfy, both the anticommutation rule with the boost generators as in eq.(2.51a) and the commutation rule with the rotation generators as in eq.(2.52a). In our 2×2 case, boost generators and rotation generators are proportional to the Pauli matrices, so there is no matrix that satisfies *at the same time* the two rules [8].

In consequence, we can construct a two-dimensional theory for spin 1/2 particles that is invariant under Lorentz transformations but not under parity transformations.

On the other hand, the first objective that we want to reach is the study of the electromagnetic interactions of the electrons in atomic physics and in scattering processes. To this aim we need an equation that is invariant under spatial inversion.

A parity noninvariant equation for spin 1/2 particles, based on the transformation properties outlined above, will be used for the study of the *neutrinos* that are created, destroyed and in general interact only by means of *weak interactions* that are not invariant under spatial inversions.

In order to construct a set of matrices for spin 1/2 particles satisfying both Lorentz and parity commutation rules, we make the two following steps:
(i) we consider matrices with *larger* dimension;
(ii) we exploit the sign ambiguity of τ in the boost generator.
More precisely, it is sufficient to introduce the following 4×4 *block* matrices

$$K^{\delta}_{[D]} = \frac{1}{2} \begin{bmatrix} \sigma^{\delta} & 0 \\ 0 & -\sigma^{\delta} \end{bmatrix} = \frac{1}{2}\alpha^{\delta} \tag{3.11}$$

where we have taken $\tau = +1$ and $\tau = -1$ in the upper and lower diagonal block, respectively.

Important note: the previous equation represents the definition of the three matrices α^{δ}. We use the greek letter δ (instead of α) as spatial index to avoid confusion between the indices and the matrices.

With no difficulty, for the spinor rotations we introduce

$$\Sigma^{\delta} = \begin{bmatrix} \sigma^{\delta} & 0 \\ 0 & \sigma^{\delta} \end{bmatrix} \tag{3.12a}$$

so that

$$S^{\delta}_{[D]} = -\frac{i}{2}\Sigma^{\delta} \tag{3.12b}$$

Note that, taking into account the discussion for the transformation of the two-dimensional spinors with $S^\delta_{[2]}$ and $K^\delta_{[2]}$, the commutation rules of eqs.(2.50a-c) for Lorentz transformations and rotations are automatically satified by the block diagonal matrices $K^\delta_{[D]}$, $S^\delta_{[D]}$ introduced above.

For the spatial inversion, we find the 4×4 block matrix

$$\Pi_{[D]} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \beta \tag{3.13}$$

with the property

$$\Pi_{[D]} = \Pi^\dagger_{[D]} = \Pi^{-1}_{[D]} \tag{3.14}$$

It satisfies the anticommutation with the boost generators of eq.(2.51a,b), that means

$$\{\Pi_{[D]}, K^\delta_{[D]}\} = \{\beta, \alpha^\delta\} = 0 \tag{3.15}$$

The specific form of $\Pi_{[D]}$ straightforwardly satisfies also the rules (2.52a,b). In technical words, we have obtained a *representation* of the Lorentz group, including parity, for spin 1/2 particles.

Introducing explicitly the four component Dirac spinor u, its boost transformation is written in the form

$$u' = B_{[D]}(\omega) u \tag{3.16}$$

The (infinitesimal) form of $B_{[D]}(\omega)$ at first order in ω is

$$B_{[D]}(\omega) \simeq 1 - \frac{1}{2}\, \omega \; (\alpha \hat{\mathbf{v}}) \tag{3.17a}$$

where, as usual, $\hat{\mathbf{v}}$ represents the unity vector of the boost velocity. (For simplicity, we do not write it explicitly in $B_{[D]}(\omega)$.)
The *finite* transformation is obtained in the same way as in eqs.(2.36a,b) and (2.40) but using the properties of the Pauli matrices, as it is shown in detail in eqs.(A.17),(A.18) and in the following discussion in the Appendix. One has

$$B_{[D]}(\omega) = \exp[-\frac{\omega}{2}(\alpha\hat{\mathbf{v}})] = ch(\frac{\omega}{2}) - (\alpha\hat{\mathbf{v}}) sh(\frac{\omega}{2}) \tag{3.17b}$$

On the other hand, the spinor rotations are obtained by replacing the σ^δ with the 4×4 matrices Σ^δ in eq.(3.10).

Furthermore, when changing the reference frame, one has always to perform the *argument re-expression* in the Dirac spinors u.
We note that, while the rotations are represented by a *unitary operator*, the Lorentz boost are *not*. More precisely, $B_{[D]}(\omega)$ is a *antiunitary operator*, that is

$$B^{\dagger}_{[D]}(\omega) = B_{[D]}(\omega) \tag{3.18}$$

A unitary, but infinite dimensional (or nonlocal) representation of the boost for spin 1/2 particles can be obtained. This problem will be studied in a different work.

The next task is to construct *matrix elements* (in the sense of vector algebra and not of quantum mechanics, because no spatial integration is performed) of the form $u_b{}^{\dagger} M u_a$, that, when boosting u_a and u_b, transform as Lorentz scalar and Lorentz four-vectors. The case of pseudoscalars and axial-vectors will be studied in Subsection 3.4.
We shall keep using the word *matrix elements* throughout this work, but in most textbooks they are commonly denoted as *Dirac covariant bilinear* quantities.

Given a generic 4×4 matrix M, by means of eq.(3.17a) the transformation of the matrix element up to first order in ω, is

$$u_b'^{\dagger} M u_a' \simeq u_b{}^{\dagger} M u_a - \frac{1}{2}\ \omega\ \hat{v}^{\delta} u_b{}^{\dagger} \{\alpha^{\delta}, M\} u_a \tag{3.19}$$

The Lorentz scalar matrix element is easily determined by means of a matrix M_s that anticommutes with the α^{δ} so that the second term in the *r.h.s.* of eq.(3.19) is vanishing. Simply recalling eqs.(3.15) and (3.13) one has

$$M_s = \beta = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{3.20}$$

where we are using the definition of the β Dirac matrix given in eq.(3.13).

As for the four-vector matrix element, one needs four matrices M_v^{μ}. To find their form in a simple way, let us consider a boost along the x^1-axis, that in eq.(3.19) means $\hat{\mathbf{v}} = (1, 0, 0)$. By means of eq.(3.19), to recover the four-vector Lorentz transformation (see eqs.(2.1) and (2.34)), one needs

$$\frac{1}{2}\{\alpha^1, M_v^0\} = M_v^1 \tag{3.21}$$

for the transformation of M_v^0, and

$$\frac{1}{2}\{\alpha^1, M_v^1\} = M_v^0 \tag{3.22}$$

for the transformation of M_v^1.
The solution is easily found calculating the anticommutators of the Dirac matrices α^δ by means of the anticommutators of the Pauli matrices of eq.(A.3). One has

$$M_v^0 = \alpha^0 = 1, \quad M_v^1 = \alpha^1 \tag{3.23a}$$

and the solution for all the components is

$$M_v^\mu = \alpha^\mu = (1, \alpha^1, \alpha^2, \alpha^3) \tag{3.23b}$$

Pay attention: $\alpha^0 = 1$ is not introduced in most textbooks.

We can resume the previous equations, also for finite Lorentz boosts, as

$$B_{[D]}(\omega)\beta B_{[D]}(\omega) = \beta \tag{3.24a}$$

or, equivalently

$$B_{[D]}(\omega)\beta = \beta B_{[D]}^{-1}(\omega) \tag{3.24b}$$

for the scalar matrix elements, and

$$B_{[D]}(\omega)\alpha^\mu B_{[D]}(\omega) = L^\mu{}_\nu(\omega)\alpha^\nu \tag{3.25}$$

for the four-vector ones.

The previous developments, recalling the expression of the four-momentum operator given in eqs.(3.2a-c), allow to write a *linear* covariant wave equation in the form

$$i\hbar c\ \partial_\mu \alpha^\mu \psi(x) = mc^2\ \beta\psi(x) \tag{3.26}$$

that is the Dirac equation, where m is the particle mass and $\psi(x) = \psi(t, \mathbf{r})$ is a four component Dirac spinor representing the particle wave function.
Intuitively, the covariance of the Dirac equation can be proven multiplying the previous equation from the left by a generic hermitic conjugate Dirac spinor. In the *l.h.s.* one has a Lorentz scalar given by the product of the (contravariant) four-vector martix element of α^μ with the (covariant) operator $i\hbar c\ \partial_\mu$. In the *r.h.s.* one has the Lorentz scalar directly given by the matrix element of β.

More formally, we can prove the covariance of the Dirac equation in the following way. We write the same equation in $\mathcal{S}'$ and show that is equivalent to the (original) equation in $\mathcal{S}$. We have

$$i\hbar c\,\partial\,'_{\mu}\alpha^{\mu}\psi'(x') = mc^2\,\beta\psi'(x') \tag{3.27}$$

The spinor in $\mathcal{S}'$ is related to the spinor in $\mathcal{S}$ by means of eq.(3.16):

$$\psi'(x') = B_{[D]}(\omega)\psi(x'(x)) \tag{3.28}$$

We replace the last expression in eq.(3.27) and multiply from the left that equation by $B_{[D]}(\omega)$. In the *r.h.s.*, by means of eq.(3.24a) one directly obtains $\beta\psi$. In the *l.h.s.*, one has to consider eq.(3.25), transforming the equation in the form

$$i\hbar c\,\partial\,'_{\mu}L^{\mu}{}_{\nu}(\omega)\alpha^{\nu}\psi(x'(x)) = mc^2\,\beta\psi(x'(x))$$

We can use the more synthetic vector algebra notation, writing

$$\partial\,'_{\mu}L^{\mu}{}_{\nu}\alpha^{\nu} = [\partial\,']^T gL[\alpha] = [\partial]^T g[\alpha] = \partial_{\mu}\alpha^{\mu} \tag{3.29}$$

where in the second equality we have used $gL = L^{-1}g$.
In this way we have shown the equivalence of eq.(3.27), written in $\mathcal{S}'$, with the original equation (3.26), written in $\mathcal{S}$.

3.3 The Gamma Dirac Matrices and the Standard Representation

The physical content of the Dirac equation is completely contained in eq.(3.26) and in the related transformation properties. However, to work in a more direct way with Dirac equation and its applications, some more developments are necessary.
First, we introduce the *Dirac adjoint* spinor that is preferably used (instead of the hermitic conjugate) to calculate matrix elements. It is defined as

$$\bar{u} = u^{\dagger}\beta \tag{3.30}$$

Its transformation law is straightforwardly obtained in the form

$$\bar{u}' = u'^{\dagger}\beta = u^{\dagger}B_{[D]}(\omega)\beta = \bar{u}B^{-1}_{[D]}(\omega) \tag{3.31}$$

where eq.(3.24b) has been used. As it must be for a representation of the Lorentz boost, $B^{-1}_{[D]}(\omega)$ is obtained inverting the direction of the boost velocity

$$B^{-1}_{[D]}(\omega) = ch(\frac{\omega}{2}) + (\alpha\hat{\mathbf{v}})sh(\frac{\omega}{2}) \simeq 1 + \frac{1}{2}\,\omega\ (\alpha\hat{\mathbf{v}}) \tag{3.32}$$

As an exercise, the reader can check that $B_{[D]}(\omega)B^{-1}_{[D]}(\omega) = 1$ by using the properties of the α^δ matrices.
Note that in the previous results there is no new physical content. We can represent the Lorentz scalar (*invariant*) as

$$u^\dagger_b \beta u_a = \bar{u}_b u_a \tag{3.33}$$

In fact, we have learned in eq.(3.31) that $\bar{u}$ transforms with $B^{-1}_{[D]}(\omega)$.

We introduce the Dirac matrices γ^μ defined as

$$\gamma^\mu = \beta\alpha^\mu \tag{3.34a}$$

Recalling that $\beta^2 = 1$, one has

$$\alpha^\mu = \beta\gamma^\mu \tag{3.34b}$$

The four-vector matrix element can be written as

$$u^\dagger_b \alpha^\mu u_a = \bar{u}_b \gamma^\mu u_a \tag{3.35}$$

and the Dirac equation (3.26) takes the usual form

$$i\hbar c\ \partial_\mu \gamma^\mu \psi(x) = mc^2\ \psi(x) \tag{3.36}$$

For clarity we give the explicit expression of the γ^μ :

$$\gamma^0 = \beta = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \gamma^\delta = \begin{bmatrix} 0 & -\sigma^\delta \\ \sigma^\delta & 0 \end{bmatrix} \tag{3.37}$$

As it will be discussed in the following, this is the so-called *spinorial representation* of the Dirac matrices.
Starting from the anticommuation rules of the α^μ one finds the following *fundamental !* anticommutation rules of the γ^μ

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{3.38}$$

Furthermore, one easily verifies that γ^0 is hermitic while the γ^δ are anti-hermitic:

$$\gamma^{\mu\dagger} = g^{\mu\mu}\gamma^\mu = \gamma^0\gamma^\mu\gamma^0 \tag{3.39}$$

Note that in $g^{\mu\mu}$ the index μ not summed; the last equality is obtained by standard use of eq.(3.38). Furthermore, the previous equation also holds in the *standard representation* of the Dirac matrices that will be introduced in the following.

We can now easily examine the *usual* procedure that is adopted to introduce the Dirac equation. Consider, for example, refs.[6,9]. The differential wave equation for a spin 1/2 particle is assumed to be *linear* with respect to the four-momentum operator introduced in eq.(3.2a,c) and to the particle mass. According to this hypothesis, the equation is written as

$$i\hbar c\ \partial_\mu\Gamma^\mu\psi(x) = mc^2\ \psi(x) \tag{3.40}$$

where the Γ^μ are four adimensional matrices to be determined.
Then, one multiplies by $i\hbar c\ \partial_\mu\Gamma^\mu$ and, by using the same eq.(3.40), obtains in the *l.h.s.* another factor mc^2. The equation takes the form

$$-(\hbar c)^2\partial_\nu\Gamma^\nu\partial_\mu\Gamma^\mu\psi(x) = (mc^2)^2\ \psi(x) \tag{3.41}$$

As we said in Subsection 3.1, the wave function of *any* relativistic particle must satisfy the Klein-Gordon equation (3.3a,b). This property must be verified also in our case. To this aim, we make the following algebraic manipulation

$$\partial_\nu\Gamma^\nu\partial_\mu\Gamma^\nu = \frac{1}{2}\ \partial_\mu\partial_\nu(\Gamma^\nu\Gamma^\nu + \Gamma^\mu\Gamma^\nu)$$

It shows that the Γ^μ must satisfy the anticommutation rules of eq.(3.38). The lowest dimension for which it is possible is 4 and we can identify the Γ^μ with the γ^μ of eq.(3.37) that have been derived by means of relativistic transformation properties.
In any case, (we repeat) the previous development is useful to show that the solutions of the Dirac equation are also solutions of the Klein-Gordon one. We can expect that also Dirac equation admits negative energy solutions.

We now face a different problem. In Subsection 3.2 we have seen that the relevant point for the covariance of the Dirac equation is represented by the anticommutation rules of the α^δ and β matrices. The same is true for the γ^μ. In other words, their specific form is not important, provided that the anticommutation rules are fulfilled. We now look for another representation, different from eq.(3.37), and more useful for practical calculations. We construct this new representation starting from a specific solution of Dirac equation (3.26) or (3.36).

Let us consider a particle at rest, that is, in a three-momentum eigenstate with $\mathbf{p} = 0$. The spatial components $\frac{\partial}{\partial \mathbf{r}}$ of the derivative operator, when applied to the corresponding wave function, give zero. The Dirac equation reduces to

$$i\hbar \frac{\partial \psi(x)}{\partial t} = mc^2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \psi(x) \tag{3.42a}$$

We can split the Dirac spinor into two two-component spinors

$$\psi = \begin{bmatrix} \eta \\ \xi \end{bmatrix}$$

So that eq.(3.42a) is written as a system of coupled equations:

$$i\hbar \frac{\partial \eta}{\partial t} = mc^2 \xi$$

$$i\hbar \frac{\partial \xi}{\partial t} = mc^2 \eta \tag{3.42b}$$

We can sum and subtract these two equations introducing the new two-component spinors

$$\varphi = \frac{1}{\sqrt{2}}(\xi + \eta)$$

$$\chi = \frac{1}{\sqrt{2}}(\xi - \eta) \tag{3.43}$$

(the factor $\frac{1}{\sqrt{2}}$ guarantees that normalization of the new Dirac spinor is not changed). One finds

$$i\hbar \frac{\partial \varphi}{\partial t} = mc^2 \varphi$$

$$i\hbar \frac{\partial \chi}{\partial t} = -mc^2 \chi \tag{3.44}$$

These equations are equivalent to eq.(3.42b) but they are *decoupled.* Technically, we have diagonalized the *r.h.s.* rest frame Hamiltonian of eq.(3.42a).

The solutions are easily found:

$$\psi_+ = \begin{bmatrix} \varphi \\ 0 \end{bmatrix}$$

with *positive energy* $E = +mc^2$,

$$\psi_- = \begin{bmatrix} 0 \\ \chi \end{bmatrix}$$

with *negative energy* $E = -mc^2$. The presence of two energy values represents a general property of relativistic wave equations.
The advantage of the solutions $\psi_{+/-}$ of eq.(3.44) is that *only one* two-component spinor is nonvanishing while the other is zero. In the positive energy case, the nonvanishing spinor can be identified with the nonrelativistic one. Furthermore, when considering a positive enery particle with small (nonrelativistic) velocity, we can expect the lower components of ψ_+ to be (not zero but) small with respect to the upper ones.
For these reasons we apply that transformation to a generic Dirac spinor, not only in the case $\mathbf{p} = \mathbf{0}$.
More formally, we perform the transformation of eq.(3.43) by introducing the following matrix

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3.45}$$

that satisfies

$$U^\dagger = U^{-1} = U$$

We multiply from the left the Dirac equation (3.36) by U and insert $UU = 1$ between the γ^μ and ψ . In this way we transform the Dirac wave function and, at the same time, the Dirac matrices obtaining

$$\gamma^\mu_{st} = U\gamma^\mu U \tag{3.46}$$

where the γ^μ_{st} are the Dirac matrices in the *standard representation*, while the γ^μ of eq.(3.37) have been given in the so-called *spinorial representation*. In most physical problem (specially if a connection with nonrelativistic physics is wanted) the *standard representation* is adopted. Generally the index *"st"* is not explicitly written. In the following of the present work we shall also adopt this convention.

Note that, due to the property of U given above, if two matrices in the spinorial representation satisfy an (anti)commutation rule, the corresponding matrices in the standard representation also satify the same rule.

In particular, this property holds for the anticommutation rule of eq.(3.38) of the γ^μ. In the standard representation they have the form

$$\gamma^0 = \beta = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \gamma^\delta = \begin{bmatrix} 0 & \sigma^\delta \\ -\sigma^\delta & 0 \end{bmatrix} \tag{3.47}$$

The hermitic conjugate satify the same eq.(3.39). As for the α^μ, by using eq.(3.34b), one has

$$\alpha^\mu_{st} = U\alpha^\mu U = U\beta U U\gamma^\mu U = \gamma^0_{st}\gamma^\mu_{st} \tag{3.48}$$

Explicitly, *without writing the index "st"*, they are

$$\alpha^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \alpha^\delta = \begin{bmatrix} 0 & \sigma^\delta \\ \sigma^\delta & 0 \end{bmatrix} \tag{3.49}$$

Note that the spin Σ^δ matrices of eq.(3.12a) keep the same form in the spinorial and standard representation.
In consequence, one can define $K^\delta_{[D]} = \frac{1}{2}\alpha^\delta$ and $S^\delta_{[D]} = -\frac{i}{2}\Sigma^\delta$ by using the standard representation for the α^δ (and the Σ^δ): the boost and rotation generators commutation rules are equivalently fulfilled. Furthermore, the expression of the boost operator is the same as in eq.(3.17a,b), with the α^δ written in the standard representation.

3.4 Parity Transformations and the Matrix γ^5

There is a fifth matrix that anticommutes with the other γ^μ. It is γ^5:

$$\{\gamma^\mu, \gamma^5\} = 0 \tag{3.50}$$

In the spinorial and standard representations, one has, respectively

$$\gamma^5_{sp} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \gamma^5_{st} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \tag{3.51}$$

Note that $\gamma^{5\dagger} = \gamma^5$ and $(\gamma^5)^2 = 1$.

Furthermore, we use the definition of ref.[5], but, as done in many texbooks, γ^5 can be defined multiplying eq.(3.51) by -1. All its properties remain unchanged. Pay attention to which definition is used !

To understand the physical meaning of the matrix elements of γ^5, it is useful to go back to Dirac spinor parity transformation. As shown in eq.(3.13), this transformation is $u' = \beta u$ being $\beta = \gamma^0$. Let us consider the parity transformation for Lorentz scalar and four-vector matrix elements. Standard use of the γ^μ anticommutation rule (3.38) gives

$$\bar{u_b}' u_a' = \bar{u}_b u_a \tag{3.52a}$$

and

$$\bar{u_b}' \gamma^0 u_a' = \bar{u}_b \gamma^0 u_a \tag{3.52b}$$

$$\bar{u_b}' \gamma^\delta u_a' = -\bar{u}_b \gamma^\delta u_a \tag{3.52c}$$

These results have an easy physical interpretation: a Lorentz scalar and a time component of a four-vector (for example a charge density) do not change sign under spatial inversion, while the spatial components of a four-vector (for example a current density) do change sign.

Let us now consider the following matrix element

$$\bar{u}_b \gamma^5 u_a$$

The Lorentz boost are studied by means of eq.(3.19) taking $M_{ps} = \gamma^0 \gamma^5$. Standard use of eqs.(3.50) (3.38) and (3.34a,b) show that

$$\{\alpha^\delta, \gamma^0 \gamma^5\} = 0 \tag{3.53}$$

so that we can conclude that our matrix element is *invariant* under Lorentz transformations. The same can be shown for rotations using the generator of eq.(3.12a,b).
But, what happens with spatial inversion ? We have

$$\bar{u_b}' \gamma^5 u_a' = \bar{u_b}' \gamma^0 \gamma^5 \gamma^0 u_a' = -\bar{u}_b u_a \tag{3.54}$$

It means that our matrix element changes sign under parity transformation. It is a *pseudo-scalar* quantity.

In terms of elementary quantities, a pseudo-scalar is given by the product of an axial vector (see the discussion of subsect 2.5) with a standard vector, for example the spin with the three-momentum: $\mathbf{sp}$. (It is not possible to use the orbital angular momentum instead of spin because one has $\mathbf{lp} = 0$, identically).

We now consider the following matrix element

$$\bar{u}_b \gamma^5 \gamma^\mu u_a$$

Standard handling (that is left as an exercise) with the γ^μ and γ^5 shows that, under Lorentz boosts and rotations, it transforms as a four-vector, but, under spatial inversion, one has

$$\bar{u_b}' \gamma^5 \gamma^0 u_a{}' = \bar{u}_b \gamma^0 \gamma^5 \gamma^0 \gamma^0 u_a = -\bar{u}_b \gamma^5 \gamma^0 u_a \qquad (3.55a)$$

and

$$\bar{u_b}' \gamma^5 \gamma^\delta u_a{}' = \bar{u}_b \gamma^0 \gamma^5 \gamma^\delta \gamma^0 u_a = +\bar{u}_b \gamma^5 \gamma^\delta u_a \qquad (3.55b)$$

We have an *axial four-vector*. Its time component changes sign, while the space components do not.

3.5 Plane Wave Solutions and the Conserved Dirac Current

In this last Subsection we shall find the plane wave solutions of the Dirac equation for a noninteracting particle, and, as in the case of the Klein-Gordon equation, we shall determine the conserved current.

At this point the equations become very large and it is necessary to find a strategy to simplify the calculations and avoid to lose the physical meaning of the developments. For this reason, most textbooks adopt the system of units in which

$$\hbar = c = 1$$

In any part of the calculations one can go back to the standard units recalling the following dimensional equalities

$$[\hbar] = [E]\ [T], \qquad [c] = [L]\ [T]^{-1}$$

and use the numerical values given in Subsection 3.1.

In this way, Dirac equation (3.36) is written in the form

$$[i\partial_\mu \gamma^\mu - m]\ \psi(x) = 0 \tag{3.56}$$

Let us make the hypothesis that the wave function $\psi(x)$ can be factorized in plane wave exponential, identical to that of the Klein-Gordon equation given in eqs.(3.4a,b), and a Dirac spinor *not depending* on the four-vector x. Also using eq.(3.5a) for positive and negative energy, being λ the energy sign, we can write

$$\psi_{\lambda \mathbf{p}\sigma}(x) = u(\lambda, \mathbf{p}, \sigma) \exp\left[i(-\lambda\epsilon(\mathbf{p})t + \mathbf{pr})\right] \tag{3.57}$$

The spin label σ of the Dirac spinor (not to be confused with the Pauli matrices) will be discussed in the following.
Applying the space-time derivative operator to the previous equation one has

$$i\partial_\mu \psi_{\lambda \mathbf{p}\sigma}(x) = (\lambda\epsilon(\mathbf{p}), -\mathbf{p}) u(\lambda, \mathbf{p}, \sigma) \exp\left[i(-\lambda\epsilon(\mathbf{p})t + \mathbf{pr})\right] \tag{3.58}$$

where the the minus sign in $-\mathbf{p}$ is due to the use of covariant components of the operator $i\partial_\mu$.
We insert the last result in the Dirac equation (3.56). Cancelling the exponential factor, it remains the following matrix equation for the Dirac spinor:

$$[\lambda\epsilon(\mathbf{p})\gamma^0 - (\mathbf{p}\gamma) - m] u(\lambda, \mathbf{p}, \sigma) = 0 \tag{3.59}$$

As in eq.(3.43), we write the four component Dirac spinor in terms of two two-component ones:

$$u(\lambda, \mathbf{p}, \sigma) = \begin{bmatrix} \varphi \\ \chi \end{bmatrix} \tag{3.60}$$

where $\varphi,\ \chi$ are respectively defined as *upper* and *lower* components of the spinor. For brevity we do not write the indices $\lambda,\ \mathbf{p},\ \sigma$ in φ and χ.
Using the γ^μ in the standard representation of eq.(3.47), we can write eq.(3.59) in the form:

$$(\lambda\epsilon(\mathbf{p}) - m)\varphi - (\mathbf{p}\sigma)\chi = 0 \tag{3.61a}$$

$$(\lambda\epsilon(\mathbf{p}) + m)\chi - (\mathbf{p}\sigma)\varphi = 0 \tag{3.61b}$$

Considering positive energy states, that is $\lambda = +1$, we obtain the lower components χ_+ in terms of φ_+ by means of eq.(3.61b):

$$\chi_+ = \frac{(\mathbf{p}\sigma)}{\epsilon(\mathbf{p}) + m}\ \varphi_+ \tag{3.62a}$$

In this case it is not possible to write φ_+ in terms of χ_+ using eq.(3.61a) because, with $\lambda = +1$, the factor $\lambda\epsilon(\mathbf{p}) - m$ is vanishing for $\mathbf{p} = 0$.
Conversely, for negative energy states, that is $\lambda = -1$, from eq.(3.61a) we obtain the upper components:

$$\varphi_- = -\frac{(\mathbf{p}\sigma)}{\epsilon(\mathbf{p}) + m}\, \chi_- \tag{3.62b}$$

In this way we have found the plane wave solutions of Dirac equation for a noninteracting particle. The two-component spinors φ_+, χ_- can be chosen (but it is not the only possible choice), as those of the nonrelativistic theory. Denoting them as w_σ, with the property $w^\dagger_{\sigma'} w_\sigma = \delta_{\sigma'\sigma}$, one has explicitly

$$w_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad w_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

for spin *up* and *down*, respectively.
In consequence the Dirac spinors $u(\lambda, \mathbf{p}, \sigma)$ can be put in the form

$$u(+1, \mathbf{p}, \sigma) = N \begin{bmatrix} w_\sigma \\ \frac{(\mathbf{p}\sigma)}{\epsilon(\mathbf{p})+m}\, w_\sigma \end{bmatrix} \tag{3.63a}$$

and

$$u(-1, \mathbf{p}, \sigma) = N \begin{bmatrix} -\frac{(\mathbf{p}\sigma)}{\epsilon(\mathbf{p})+m}\, w_\sigma \\ w_\sigma \end{bmatrix} \tag{3.63b}$$

We point out that, in general, the spin label σ of w_σ *does not represent* the spin eigenvalue in a fixed direction, for example the x^3 axis. This property holds true *only* for a particle at rest. In this case the previous solutions coincide with the solutions of eq.(3.44).
General properties of spin and angular momentum for Dirac equation will be studied in a subsequent work.

The Dirac spinors of eqs.(3.63a,b) can be also conveniently written as

$$u(\lambda, \mathbf{p}, \sigma) = N\ u(\lambda, \mathbf{p}) w_\sigma \tag{3.64}$$

with

$$u(+1, \mathbf{p}) = N \begin{bmatrix} 1 \\ \frac{(\mathbf{p}\sigma)}{\epsilon(\mathbf{p})+m} \end{bmatrix} \tag{3.65a}$$

and

$$u(-1,\mathbf{p}) = N \begin{bmatrix} -\frac{(\mathbf{p}\sigma)}{\epsilon(\mathbf{p})+m} \\ 1 \end{bmatrix} \qquad (3.65b)$$

where the $u(\lambda,\mathbf{p})$ represent 4×2 matrices. They must be applied onto the two-component (column) spinors w_σ, giving as result the four component Dirac (column) spinors of eqs.(3.63a,b).
Note that, in contrast to the nonrelativistic case, the Dirac spinors *depend* on the momentum of the particle.

We now discuss the normalization factor N. In nonrelativistic theory, the plane wave of a spin 1/2 particle is "normalized" as

$$\psi_{\mathbf{p}\sigma}(x) = \frac{1}{\sqrt{V}} w_\sigma \exp\left[i(-Et+\mathbf{pr})\right]$$

where V represents the (macroscopic) volume where the particle stays. The probabilty of finding the particle in this volume is set equal to one. However, V is a *fictitious* quantity that always disappears when physical (observable) quantities are calculated. In consequence, for the sake of simplicity, one can put $V=1$. In this way, one has

$$\psi^\dagger_{\mathbf{p}\sigma'}(x)\psi_{\mathbf{p}\sigma}(x) = \delta_{\sigma\sigma'}$$

A similar result can be obtained for the Dirac equation plane waves, putting in eqs.(3.63a)-(3.65b)

$$N = N^{nc} = \sqrt{\frac{\epsilon(\mathbf{p})+m}{2\epsilon(\mathbf{p})}} \qquad (3.66)$$

where nc stands for *not covariant* . In fact this normalization cannot be directly used for the calculation of covariant amplitudes. With this noncovariant normalization, the Dirac wave function satifies the following normalization equation that is analogous to the nonrelativistic one

$$\psi^\dagger_{\lambda'\mathbf{p}\sigma'}(x)\psi_{\lambda\mathbf{p}\sigma}(x) = \delta_{\lambda\lambda'}\delta_{\sigma\sigma'} \qquad (3.67)$$

As an exercise, verify this result and that of eq.(3.69), by using eq.(A.7) for the products of $(\sigma\mathbf{p})$. Also use the identity

$$\mathbf{p}^2 = [\epsilon(\mathbf{p})]^2 - m^2 = (\epsilon(\mathbf{p})+m)(\epsilon(\mathbf{p})-m)$$

The *covariant normalization* is obtained taking

$$N = N^{cov} = \sqrt{\frac{\epsilon(\mathbf{p}) + m}{2m}} = \sqrt{\frac{\epsilon(\mathbf{p})}{m}}\ N^{nc} \tag{3.68}$$

By using this normalization one has

$$\bar{u}(\lambda', \mathbf{p}, \sigma')u(\lambda, \mathbf{p}, \sigma) = (-1)^{\lambda}\delta_{\lambda\lambda'}\delta_{\sigma\sigma'} \tag{3.69}$$

that, also recalling eq.(3.52a), represents an explicitly Lorentz invariant condition.

In many textbooks a slightly different covariant normalization in used, that is

$$N^{cov'} = N^{cov}\sqrt{2m}$$

so that a factor $2m$ appears in the *r.h.s.* of eq.(3.69).
When reading a book or an article for the study of a specific problem, pay attention to which normalization is really used !

For further developments it is also introduced the spinor corresponding to negative energy, *negative momentum* $-\mathbf{p}$ (and spin label σ). From eq.(3.63b) or (3.65b) one has

$$u(-1, -\mathbf{p}) = N \begin{bmatrix} \frac{(\mathbf{p}\sigma)}{\epsilon(\mathbf{p})+m} \\ 1 \end{bmatrix} \tag{3.70}$$

Note that

$$u(-1, -\mathbf{p}) = -\gamma^5 u(+1, \mathbf{p}) \tag{3.71}$$

That spinor is standardly applied to w_σ, as in eqs.(3.63b) and (3.64).

We conclude this section studying the transition current associated to the Dirac equation in the same way as we studied that of the Klein-Gordon equation in eqs.(3.6)-(3.8b).
First, one has to write the Dirac equation for the *adjoint* wave function

$$\bar{\psi}(x) = \psi^{\dagger}(x)\gamma^0$$

To this aim, take the Dirac equation (3.56) and calculate the hermitic conjugate. By using eq.(3.39), one finds

$$-i\partial_\mu\psi^{\dagger}(x)\gamma^0\gamma^\mu\gamma^0 - \psi^{\dagger}(x)m = 0 \tag{3.72a}$$

Multiplying this equation from the right by $-\gamma^0$ one obtains

$$i\partial_\mu \bar{\psi}(x)\gamma^\mu + \bar{\psi}(x)m = 0 \tag{3.72b}$$

that is the searched equation.
As done for the Klein-Gordon equation we obtain the conserved current by means of the following three steps.
(i) Take eq.(3.56) with a plane wave, initial state, solution $\psi_I(x)$ corresponding to energy sign λ_I, three-momentum $\mathbf{p}_I$ and spin label σ_I.
(ii) Analogously, take eq.(3.72b) with a plane wave, final state, solution $\bar{\psi}_F(x)$.
(iii) Multiply the equation of step (i) by $\bar{\psi}_F(x)$ and the equation of step (ii) by $\psi_I(x)$. Then *sum* these two equations (note that the scalar mass term disappears), obtaining

$$\partial_\mu J^\mu_{FI}(x) = 0 \tag{3.73}$$

where the Dirac *conserved current* is

$$J^\mu_{FI}(x) = \bar{\psi}_F(x)\gamma^\mu \psi_I(x) \tag{3.74a}$$

$$= \bar{u}(\lambda_F, \mathbf{p}_F, \sigma_F)\gamma^\mu u(\lambda_I, \mathbf{p}_I, \sigma_I)\exp(iq_\mu x^\mu) \tag{3.74b}$$

with the four-momentum tranfer $q^\mu = p^\mu_F - p^\mu_I$. The four-vector character of the Dirac current is manifestly shown by the previous equation.
The Dirac four-vector vertex is

$$\bar{u}_F\gamma^\mu u_I = \bar{u}(\lambda_F, \mathbf{p}_F, \sigma_F)\gamma^\mu u(\lambda_I, \mathbf{p}_I, \sigma_I) \tag{3.75}$$

Due to current conservation it satisfies, analogously to eq.(3.9),

$$q_\mu \bar{u}_F \gamma^\mu u_I = 0 \tag{3.76}$$

Note that in the static case the current density (differently from the Klein Gordon equation) is a positive quantity both for positive and negative enery states, as shown explictly by the second equality of the following equation:

$$J^0_{II} = \bar{\psi}_I(x)\gamma^0\psi_I(x) = \psi^\dagger_I(x)\psi_I(x) > 0 \tag{3.77}$$

This property allows to attach (for some specific problems) a probabilistic interpretation to that quantity and to consider $\psi(x)$ as a wave function in the same sense of nonrelativistic quantum mechanics. However, the presence of negative energy solutions requires, in general, the introduction of the field theory formalism.

The vertex of eq.(3.75) at first glance looks very different with respect to that of the Klein-Gordon equation $(p_F^\mu + p_I^\mu)$ given in eq.(3.8b). The so-called Gordon decomposition, with some algebra on the Dirac matrices, shows that it can be written in a form that is more similar to the Klein-Gordon one. This procedure will be analyzed in a subsequent work.
For the moment, using the properties of the Pauli matrices, the reader can show that

$$\bar{u}(\lambda, \mathbf{p}, \sigma')\gamma^\mu u(\lambda, \mathbf{p}, \sigma) = \frac{p^\mu}{m}\delta_{\sigma\sigma'} \tag{3.78}$$

with $p^\mu = (\epsilon(\mathbf{p}), \mathbf{p})$. The covariant normalization of eq.(3.68) has been used.

We conclude this work noting that, at this point, the reader should be able to use the main tools related to Dirac equation, being also familiarized with the issues of relativity in quantum mechanical theories.
More formal details and calculations of physical observables can be found in many textbooks and will be studied in a subsequent work.

4 Appendix. Properties of the Pauli Matrices

The three Pauli matrices are defined as follows

$$\sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{A.1}$$

they are 2×2, traceless, hermitic ($\sigma^{\alpha\dagger} = \sigma^\alpha$) matrices. The Pauli matrices fulfill the the following commutation rules

$$[\sigma^\alpha, \sigma^\beta] = 2i\epsilon^{\alpha\beta\gamma}\sigma^\gamma \tag{A.2}$$

One defines the *spin*, that is the intrinsic angular momentum operator, multiplying the σ^α by $\hbar/2$.
By means of this definition, the spin satisfies the standard angular momentum commutation rules, that are

$$[j^\alpha, j^\beta] = i\hbar\epsilon^{\alpha\beta\gamma} j^\gamma$$

Independently, the Pauli matrices fulfill the anticommutation rules

$$\{\sigma^\alpha, \sigma^\beta\} = 2\delta^{\alpha\beta} \tag{A.3}$$

Summing up eqs.(A.2) and (A.3) and dividing by two, one obtains the very useful relation

$$\sigma^\alpha \sigma^\beta = \delta^{\alpha\beta} + i\epsilon^{\alpha\beta\gamma}\sigma^\gamma \tag{A.4}$$

Obviously *only two* of eqs.(A.2), (A.3) and (A.4) are independent.

Given the three-vectors **a** and **b**, let us multiply the previous expression by a^α and b^β, summing over the components. One obtains

$$(\sigma\mathbf{a})(\sigma\mathbf{b}) = \mathbf{ab} + i(\sigma\mathbf{a} \times \mathbf{b}) \tag{A.5}$$

Note that $(\sigma\mathbf{a})$ represents the following matrix

$$(\sigma\mathbf{a}) = \begin{bmatrix} a^3 & a^1 - ia^2 \\ a^1 + ia^2 & -a^3 \end{bmatrix} \tag{A.6}$$

and analogously for $(\sigma\mathbf{b})$ and $(\sigma\mathbf{a} \times \mathbf{b})$.
In eq.(A.5), if $\mathbf{b} = \mathbf{a}$, the vector product is vanishing, so that one has

$$(\sigma\mathbf{a})^2 = \mathbf{a}^2 \tag{A.7}$$

Starting from this equality we can calculate the *function* $f(\sigma\mathbf{a})$.
To this aim we recall that, if a function $f(x)$ of a standard variable x has the Taylor expansion

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \tag{A.8}$$

the (same) function of the *matrix* $(\sigma\mathbf{a})$ is *defined* as follows

$$f(\sigma\mathbf{a}) = \sum_{n=0}^{\infty} c_n (\sigma\mathbf{a})^n \tag{A.9}$$

The result is obviously a 2×2 matrix.
Incidentally, the previous definition, that makes use of the Taylor expansion in powers of the *argument* matrix, is a general one: it holds not only for $(\sigma\mathbf{a})$

but also if the argument of the function is a matrix of *any* dimension or if it is a linear operator. In the present case, the powers $(\sigma\mathbf{a})^n$ in eq.(A.9) can be calculated by means of eq.(A.7). We also use $(\sigma\mathbf{a})^0 = 1$.
We make here some algebraic developments to obtain a "closed" expression for eq.(A.9).
First, let us write separately the even and the odd powers in the expansion (A.8):

$$f(x) = \sum_{m=0}^{\infty} c_{2m} x^{2m} + \sum_{l=0}^{\infty} c_{2l+1} x^{2l+1} \qquad (A.10)$$

Do the same for $f(-x)$:

$$f(-x) = \sum_{m=0}^{\infty} c_{2m} x^{2m} - \sum_{l=0}^{\infty} c_{2l+1} x^{2l+1} \qquad (A.11)$$

So that, summing and subtracting the last two equations, one has:

$$\frac{1}{2}[f(x) + f(-x)] = \sum_{m=0}^{\infty} c_{2m} x^{2m} \qquad (A.12)$$

$$\frac{1}{2}[f(x) - f(-x)] = \sum_{l=0}^{\infty} c_{2l+1} x^{2l+1} \qquad (A.13)$$

Let us now go back to eq.(A.9), introducing the unit vector $\hat{\mathbf{a}}$ and the absolute value (*positive* !) $|\mathbf{a}|$, by means of the standard relation

$$\mathbf{a} = |\mathbf{a}|\hat{\mathbf{a}} \qquad (A.14)$$

Furthermore, by means of eq.(A.7), one has

$$(\sigma\mathbf{a})^{2m} = (\mathbf{a}^2)^m = |\mathbf{a}|^{2m} \qquad (A.15)$$

$$(\sigma\mathbf{a})^{2l+1} = (\sigma\mathbf{a})(\mathbf{a}^2)^l = (\sigma\hat{\mathbf{a}})|\mathbf{a}|^{2l+1} \qquad (A.16)$$

In consequence, writing separately the even and odd powers in eq.(A.9), and using eqs.(A.12,13), one obtains

$$f(\sigma\mathbf{a}) = \sum_{m=0}^{\infty} c_{2m}|\mathbf{a}|^{2m} + (\sigma\hat{\mathbf{a}})\sum_{l=0}^{\infty} c_{2l+1}|\mathbf{a}|^{2l+1} =$$

$$= \frac{1}{2}[f(|\mathbf{a}|) + f(-|\mathbf{a}|)] + \frac{1}{2}[f(|\mathbf{a}|) - f(-|\mathbf{a}|)](\sigma\hat{\mathbf{a}}) \qquad (A.17)$$

In order to derive the second equality of eq.(3.17b), being the α^δ, defined in eq.(3.11) as block *diagonal* matrices, one can procede separately for the two blocks. Let us consider first the upper left block. By means of the previous equation, one has

$$\exp\left[-\frac{\omega}{2}(\sigma\hat{\mathbf{v}})\right] =$$
$$= \frac{1}{2}\left[\exp(|\frac{\omega}{2}|) + \exp(-|\frac{\omega}{2}|)\right] - \frac{1}{2}\left[\exp(|\frac{\omega}{2}|) - \exp(-|\frac{\omega}{2}|)\right] sgn(\omega)(\sigma\hat{\mathbf{v}}) =$$
$$= ch(\frac{\omega}{2}) - (\sigma\hat{\mathbf{v}})sh(\frac{\omega}{2}) \qquad (A.18)$$

In the previous equation $sgn(\omega)$ gives the sign of ω. Also, we have used $\hat{\mathbf{a}} = -sgn(\omega)\hat{\mathbf{v}}$ and $|\mathbf{a}| = |\omega|$ in eq.(A.17).
As for the lower right block, one easily obtain the result that is analogous to the previous one, but with a plus sign in front of the second term. Recalling the form of the α^δ matrices, one obtains the final result of eq.(3.17b).

The reader can now look at this development in a slightly different way. Recalling the form of the α^δ of eq.(3.11), the powers of $(\alpha\mathbf{a})$, satisfy the *same* relations as eqs.(A.15) and (A.16) for the powers of $(\sigma\mathbf{a})$. In consequence, one can repeat the calculations of eqs.(A.17) and (A.18) simply replacing the σ^δ with the α^δ, obtaining directly eq.(3.17b).
Furthermore, in this way, one realizes that the result remains the same also in the standard representation and only depends on the anticommutation rules of the α^δ matrices.

References

[1] C. Kittel, W.D. Knight, M.A. Ruderman, *Mechanics*, in Berkeley Physics Course, (Mc Graw-Hill Education, New York 1965).

[2] J.D. Jackson, *Classical Electrodynamics*, Second Edition, (John Wiley and Sons, New York 1975).

[3] R. Hagedorn, *Relativistic Kinematics*, (W. A. Benjamin, New York 1963).

[4] L.D. Landau, E.M. Lifshits, *The Classical Theory of Fields*, in Course of theoretical Physics, Vol.2, Fourth Edition, (Elsevier, Butterworth Heinemann, 1980).

[5] E. M. Lifshitz, L. P. Pitaevskii, V. B. Berestetskii, *Quantum Electrodynamics*, in Course of theoretical Physics, Vol.4, Second Edition, (Elsevier, Butterworth Heinemann, 1982).

[6] J.D. Bjorken, S.D. Drell, *Relativistic Quantum Mechanics*, (McGraw-Hill College, New York 1965).

[7] J.J. Sakurai, *Modern Quantum Mechanics*, (Addison Wesley Publishing Co., USA, 1994).

[8] N. Cabibbo *Relatività - Teoria di Dirac*, Class notes of the Università di Roma La Sapienza, Dipartimento di Fisica, 2003, in http://chimera.roma1.infn.it/NICOLA/poincare.pdf

[9] P.A.M. Dirac, *The principles of Quantum Mechanics*, Fourth Edition, (Oxford University Press, USA, 1982).