Quantum Field Theory

Lecture Notes

Petr Jizba

September 21, 2021

Preface

Should authors feel compelled to justify the writing of yet another lecture notes on Quantum Field Theory? In an overpopulated world, should parents feel compelled to justify bringing forth yet another child? Perhaps not! But an act of creation is also an act of love, and a love story can always be happily shared. These notes originated from a series of lectures on Quantum Filed Theory delivered at the Faculty of Nuclear Science and Physical Engineering, Czech Technical University in Prague, over the period from 2019 to 2020. During the writing, I have attempted to maintain a cohesive self-contained content. The material is discussed in sufficient detail to enable the students to follow every step, but some crucial theoretical aspects are not covered such as the non-perturbative aspects of Yang–Mills gauge theories or quantum field theory of gravity. Still it is hoped that these notes will serve as a useful introduction to Quantum Field Theory.

A working knowledge of basic quantum mechanics and related mathematical formalisms, e.g., Hilbert spaces and operators, is required to understand the contents of these lecture notes. Nevertheless, I have attempted to recall necessary definitions throughout the chapters and the numerous notes.

I would like to express my gratitude to Doctors V. Zatloukal and J. Kňap for their diligent reading of the manuscript and constructive criticisms. Also special thanks go to M. Blasone, G. Vitiello and H. Kleinert for teaching me non-perturbative techniques, as well as to the students of QFT I and II courses for their patience and their numerous suggestions. Finally these notes would not have seen the light of day had it not been for the heroic efforts of three modern day scribes and illuminators, Georgy Ponimatkin, David Grund and Diana Mária Krupová to whom I am deeply grateful.

Books

There are many books on Quantum Field Theory, most are rather long. All those listed below are worth looking at. They provide a wealth of a complemental material for these lecture notes.

- E.M. Peskin and D.V Schroeder, *An Introduction to Quantum Field Theory*, (Addison-Wesley Publishing Co., 1996).

  Provides a good introduction with an extensive discussion of gauge theories including QCD and various applications.

- M. Srednicky, *Quantum Field Theory*, (Cambridge University Press, 2007).
  Represents a comprehensive modern book organised by considering spin-0, spin-1 2 and spin-1 fields in turn.

- S. Weinberg, *The Quantum Theory of Fields, vol. I Foundations* and *vol. II Modern Applications*, (Cambridge University Press, 1995,1996).
  Written by a Nobel Laureate, contains lots of details which are not covered elsewhere, perhaps a little idiosyncratic and less introductory than the above.

- Z. Zinn-Justin, *Quantum Field Theory and Critical Phenomenam*, (Oxford University Press, 2002).
  Book devotes a large fraction to applications to critical phenomena in statistical physics but covers gauge theories at some length as well, not really an introductory book.

- C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, (McGraw-Hill International Book Co., 1980).
  At one time the standard book, containing a lot of detailed calculations but the treatment of non abelian gauge theories is a bit cursory and somewhat dated.

Detailed Contents

Part I

Relativistic Quantum Mechanics

Non-Relativistic Wave Function | 1

1.1 Transformations

In quantum mechanics the physical state of a particle is represented by wave function $\psi(\boldsymbol{x}, t)$. In the Schrödinger picture we have $\psi(\boldsymbol{x}, t) = \langle \boldsymbol{x}|\psi(t)\rangle$, i.e. time evolution in this picture is contained in the *state vectors* and corresponding *base vectors* are time independent (basis is rigid). On the other hand, in the Heisenberg picture we have $\psi(\boldsymbol{x}, t) = \langle \boldsymbol{x}, t|\psi\rangle$, i.e. *state vectors* are time independent and *base vectors* evolve (basis evolves in time). Consequently, in the Schrödinger picture the dynamics is given by the Schrödinger equation, which prescribes evolution of state vectors, while in the Heisenberg picture the evolution is given by the Heisenberg equation which prescribes the evolution of the complete set of observables (i.e., Hermitian operators). The later, in turn, imply the time dependent base vectors (via corresponding eigenfunctions).

Choice of the representation is in quantum mechanics simply matter of convenience (though work with Schrödinger equation is often simpler due to its linearity) since these representations are unitarily equivalent. This is essence of the so-called Stone–von Neumann uniqueness theorem. We will see that the Stone–von Neumann theorem is typically broken in quantum theories with infinitely many degrees of freedom. This will have important consequences for the entire structure of Quantum Field Theory (e.g., renormalization, non-trivial vacuum condensates, etc.).

Let us now recall behaviour of the quantum-mechanical wave function under two important transformations, namely **rotation**

$$\psi(\boldsymbol{x}) \stackrel{\boldsymbol{R}}{\rightarrow} \psi_{\boldsymbol{R}}(\boldsymbol{x}) = \psi(\boldsymbol{R}^{-1}\boldsymbol{x}), \tag{1.1}$$

and **translation**

$$\psi(\boldsymbol{x}) \stackrel{\boldsymbol{a}}{\rightarrow} \psi_{\boldsymbol{a}}(\boldsymbol{x}) = \psi(\boldsymbol{x} - \boldsymbol{a}). \tag{1.2}$$

Here we have omitted time argument as it is immaterial for this discussion. Time argument will, however, be important in the next chapter where relativistic transformation of state vectors will be considered.

In quantum mechanics are symmetry transformations for (compact) groups implemented via unitary operations:

$$\begin{aligned} \psi_{\boldsymbol{a}}(\boldsymbol{x}) &= U_{\boldsymbol{a}}(\boldsymbol{a})\psi(\boldsymbol{x}) = \psi(\boldsymbol{x} - \boldsymbol{a}), \quad (1.3)\\ \psi_{\boldsymbol{R}}(\boldsymbol{x}) &= U_{\boldsymbol{R}}(\theta)\psi(\boldsymbol{x}) = \psi(\boldsymbol{R}^{-1}(\theta)\boldsymbol{x}). \quad (1.4) \end{aligned}$$

In addition, if particles have spin (or other internal indices associated with them) i.e.

$$\psi(\boldsymbol{x}) \rightarrow \psi_{\alpha}(\boldsymbol{x}), \quad \alpha \in \boldsymbol{I}, \tag{1.5}$$

then such wave functions will change under rotation according to

$$\psi_\alpha(\boldsymbol{x}) \xrightarrow{\boldsymbol{R}} \psi_{\boldsymbol{R},\alpha}(\boldsymbol{x}) = U_{\boldsymbol{R}}(\theta)_{\alpha\beta}\psi_\beta(\boldsymbol{x}) = D_{\alpha\beta}(\boldsymbol{R})\psi_\beta(\boldsymbol{R}^{-1}\boldsymbol{x}), \tag{1.6}$$

here $D(\boldsymbol{R})$ is an appropriate representation of the group element $\boldsymbol{R} \in SO(3)$, which acts on internal indices. It is important to recognize that the (unitary) rotation operator $U_{\boldsymbol{R}}(\theta)$ not only shifts ψ from $\boldsymbol{x}$ to $\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{x}$, but also rotates the "direction" of ψ. In Section 2.15. we will prove that the rotation in the space of internal indices will give rise to *spin* while the rotation of the wave function from one spatial position to another will give rise to *orbital angular momentum*.

Example 1.1.1 For example, consider spin-$\frac{1}{2}$ particle, in this case wave function index takes values $\alpha = \pm\frac{1}{2}$. From quantum mechanics we know, that our transformation can be written as

$$D(\boldsymbol{R}) = e^{-i\theta \boldsymbol{n}\boldsymbol{s}},$$

which is a matrix that acts on Pauli spinors. Here θ is the angle of rotation, $\boldsymbol{n}$ is axis of the rotation, and $\boldsymbol{s} = \frac{1}{2}\boldsymbol{\sigma} = \frac{1}{2}(\sigma_1, \sigma_2, \sigma_3)$ is the vector of Pauli matrices, which are given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In this representation $s_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is diagonal, and so α has interpretation of being the (eigen)value of s_3.

The corresponding rotation operator $U_{\boldsymbol{R}}(\theta)$ reads

$$U_{\boldsymbol{R}}(\theta) = e^{-i\theta \boldsymbol{n}(\boldsymbol{s}+\boldsymbol{L})},$$

where $\boldsymbol{L}$ is the orbital angular momentum, i.e.

$$\boldsymbol{L} = x \times \boldsymbol{p} = \boldsymbol{x} \times (-i\hbar\boldsymbol{\nabla}).$$

It should be notes that $\boldsymbol{s} + \boldsymbol{L}$ is the total angular momentum of the particle (i.e. quantity that is conserved).

Exercises: Transformations in 3D

Exercise 1.1 Wave-function transforms under change of coordinates as scalar field: $\psi'(\boldsymbol{x}') = \psi(\boldsymbol{x})$. Verify invariance of the timeless Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{x}) - E\right]\psi(\boldsymbol{x}) = 0,$$

under a) rotations, b) translations. What conditions on the potential $V(\boldsymbol{x})$ need to be met?

Rotation group $SO(3)$

Generators of rotations around axes x_j, $j = 1, 2, 3$, are the skew-symmetric matrices

$$\mathbb{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathbb{T}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbb{T}_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Finite rotation by an angle φ around the axis in direction of a unit vector $\boldsymbol{n}$, $\boldsymbol{\varphi} \equiv \varphi \boldsymbol{n}$, is given by the matrix

$$\mathbb{R}(\boldsymbol{\varphi}) = e^{\varphi_j \mathbb{T}_j},$$

where the sum over j is implicitly assumed.

Exercise 1.2 Show that $\mathbb{R}(\boldsymbol{\varphi}) \in SO(3)$ (the group of real 3×3 orthogonal matrices with det = 1).

[**Hint:** Use the identity $\det e^{\mathbb{A}} = e^{\mathrm{Tr}\,\mathbb{A}}$.]

Exercise 1.3 Verify the identities

$$[\mathbb{T}_i, \mathbb{T}_j] = -\varepsilon_{ijk} \mathbb{T}_k .$$

[**Hint:** Note that the matrix elements $(\mathbb{T}_i)_{jk} = \varepsilon_{ijk}$.]

Group $SU(2)$ and algebra $su(2)$

Lie group $SU(2)$ is the set of matrices

$$SU(2) = \{\mathbb{U} \in \mathbb{C}^{2,2} \mid \mathbb{U}^\dagger = \mathbb{U}^{-1}, \det \mathbb{U} = 1\}.$$

Lie algebra $su(2)$ is the set of matrices

$$su(2) = \{\mathbb{A} \in \mathbb{C}^{2,2} \mid \mathbb{A}^\dagger = -\mathbb{A}, \mathrm{Tr}\,\mathbb{A} = 0\}.$$

(So that $e^{\mathbb{A}} \in SU(2)$.) The Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

give rise to a standard basis $\{i\tau^1, i\tau^2, i\tau^3\}$ of $su(2)$, where $\tau^i = \sigma^i/2$.

Exercise 1.4 Show that

a) $(\sigma^1)^2 = (\sigma^2)^2 = (\sigma^3)^2 = \mathbb{I}$,

b) $[\sigma^i, \sigma^j] = 2i\varepsilon_{ijk}\sigma^k$,

c) $\{\sigma^i, \sigma^j\} = \sigma^i\sigma^j + \sigma^j\sigma^i = 2\delta^{ij}\mathbb{I}$.

Exercise 1.5 Calculate the (Casimir) operators:

a) $\mathbb{T}_j \mathbb{T}_j$ for algebra $so(3)$,

b) $i\tau^j i\tau^j$ for algebra $su(2)$.

(Sum over j is implicitly assumed.)

Exercise 1.6 Expand $e^{\frac{i}{2}\varphi_j\sigma^j}$ into a series to find

$$e^{\frac{i}{2}\varphi_j\sigma^j} = \left(\cos\frac{\varphi}{2}\right)\mathbb{I} + i\left(\sin\frac{\varphi}{2}\right)\frac{\varphi_j\sigma^j}{\varphi}, \qquad \varphi \equiv \sqrt{\varphi_j\varphi_j}.$$

Exercise 1.7 Show that for infinitesimal transformations, $|\boldsymbol{\varphi}| \ll 1$,

$$\sigma^j (e^{\varphi_i \mathbb{T}_i})_{jk} a_k = e^{\frac{i}{2}\sigma^i \varphi_i} a_k \sigma^k e^{-\frac{i}{2}\sigma^j \varphi_j} ,$$

where $(a_1, a_2, a_3) \in \mathbb{R}^3$.

Exercise 1.8 Show that $\mathbb{U}(\vec{\varphi}) \equiv e^{\frac{i}{2}\sigma^j \varphi_j}$ can be cast in the form

$$\mathbb{U}(\vec{\varphi}) = e^{i\sigma^3 \frac{\beta_1+\beta_2}{2}} e^{i\sigma^2 \gamma} e^{i\sigma^3 \frac{\beta_1-\beta_2}{2}} ,$$

and find relations between the parameters $(\varphi_1, \varphi_2, \varphi_3)$ and $(\beta_1, \beta_2, \gamma)$.

[**Hint:** Use the result of Exercise 1.6.]

2 Klein–Gordon equation

2.1 Relativistic Conventions

Here and throughout we will assume so-called *natural units*, in which $c = \hbar = 1$. In this system of units, E, p have units of lenght^{-1} = time^{-1}. So, as for their units *time* and *space* are considered on equal footing.

Also the following relativistic conventions will be used:

- Space-time 4-vector will be denoted by $x^\mu = (x^0, \boldsymbol{x}) = (t, \boldsymbol{x})$.
- 4-momentum will be denoted by $p^\mu = (p^0, \boldsymbol{p}) = (E, \boldsymbol{p})$.
- Scalar product is given by

$$a \cdot b = \boldsymbol{g}_{\mu\nu} a^\mu b^\nu = a^\mu b_\mu = a_\mu b^\mu \,, \tag{2.1}$$

where

$$\boldsymbol{g}_{\mu\nu} = \boldsymbol{g}^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \mathrm{diag}(1, -1, -1, -1) \,, \tag{2.2}$$

is a metric tensor, with $\boldsymbol{g}^{\mu\nu}$ being inverse to the $\boldsymbol{g}_{\mu\nu}$. We can immediately derive a simple relation between metric tensors and Kronecker delta:

$$\boldsymbol{g}_{\mu\nu} \boldsymbol{g}^{\nu\sigma} = \delta_\mu{}^\sigma = \delta_\mu^\sigma \,. \tag{2.3}$$

A *Lorentz transformation* $\boldsymbol{L}$ maps 4-vectors according to the linear relation

$$x^\mu \xrightarrow{\boldsymbol{L}} x'^\mu = \boldsymbol{L}^\mu{}_\nu x^\nu \,. \tag{2.4}$$

Here $\boldsymbol{L}^\mu{}_\nu \in SO(1,3)$ is element of the so-called *Lorentz group*. If we define the inverse transform as

$$x^\nu = \boldsymbol{L}_\mu{}^\nu x'^\mu \,, \tag{2.5}$$

we see that $\boldsymbol{L}_\mu{}^\nu$ and $\boldsymbol{L}^\mu{}_\nu$ are inverse to each other. From the fact that Lorentz transformation should preserve scalar product of 4-vectors, i.e. $a' \cdot b' = a \cdot b$, the following relations must hold

$$\boldsymbol{g}_{\mu\mu'} \boldsymbol{L}^\mu{}_\nu \boldsymbol{L}^{\mu'}{}_{\nu'} = \boldsymbol{g}_{\nu\nu'} \,, \tag{2.6}$$

$$\boldsymbol{g}^{\mu\mu'} \boldsymbol{L}_\mu{}^\nu \boldsymbol{L}_{\mu'}{}^{\nu'} = \boldsymbol{g}^{\nu\nu'} \,. \tag{2.7}$$

By taking determinant of both sides of (2.6) we arrive at the fact that

$$\det{}^2 \boldsymbol{L} = 1 \,. \tag{2.8}$$

Hence we can divide Lorentz transformations into two classes — *proper*, for which $\det \boldsymbol{L} = 1$ and *improper*, for which $\det \boldsymbol{L} = -1$.

Differentials

We define differential as a scalar operator given by

$$d \equiv dx^{\mu} \frac{\partial}{\partial x^{\mu}} . \tag{2.9}$$

In relativistic notation we define $\partial_{\mu} \equiv \frac{\partial}{\partial x^{\mu}}$, then differential can be then written as

$$d = dx^{\mu} \partial_{\mu} . \tag{2.10}$$

Such definition of d is Lorentz invariant (as can be shown from the fact that $\partial'_{\mu} = \boldsymbol{L}_{\mu}{}^{\nu} \partial_{\nu}$). We define covariant 4-gradient operator to be

$$\partial_{\mu} = \left(\frac{\partial}{\partial x_0}, \boldsymbol{\nabla} \right) , \tag{2.11}$$

and its contravariant counterpart as

$$\partial^{\mu} = \left(\frac{\partial}{\partial x_0}, -\boldsymbol{\nabla} \right) . \tag{2.12}$$

Using those we can define the *d'Alembert operator* (or simply *d'Alembertian*) as

$$\boldsymbol{g}^{\mu\nu} \partial_{\mu} \partial_{\nu} = \partial_{\mu} \partial^{\mu} = \left(\frac{\partial^2}{\partial x_0^2} - \nabla^2 \right) \equiv \Box . \tag{2.13}$$

In the context of special relativity it is common to denote the metric tensor $\boldsymbol{g}^{\mu\nu}$ as $\boldsymbol{\eta}^{\mu\nu}$ or simply $\eta^{\mu\nu}$.

2.2 Structure of Lorentz Transformation

We begin our study of Lorentz transformations by taking infinitesimal limit of such transformation

$$\boldsymbol{L}^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} , \tag{2.14}$$

which can be thought of as infinitesimal deformation from identical transformation (here $||\omega|| \ll 1$). Starting from the fact that $\boldsymbol{L}_{\mu}{}^{\alpha}$ and $\boldsymbol{L}^{\mu}{}_{\alpha}$ are inverse to each other, we can write

$$\boldsymbol{L}_{\mu}{}^{\nu'} \boldsymbol{L}^{\mu}{}_{\alpha} = \delta^{\nu'}{}_{\alpha} . \tag{2.15}$$

By using (2.14) we arrive to the following equation

$$(\delta_{\mu}{}^{\nu'} + \omega_{\mu}{}^{\nu'})(\delta^{\mu}{}_{\alpha} + \omega^{\mu}{}_{\alpha}) = \delta^{\nu'}{}_{\alpha} . \tag{2.16}$$

If we restrict ourselves only to the first order in ω

$$\delta^{\nu'}_{\alpha} + \omega_{\alpha}{}^{\nu'} + \omega^{\nu'}{}_{\alpha} = \delta^{\nu'}{}_{\alpha} . \tag{2.17}$$

By subtracting Kronecker δ's from both sides and lowering all indices one gets

$$\omega_{\alpha\nu'} + \omega_{\nu'\alpha} = 0 . \tag{2.18}$$

This statement implies that ω is a 4×4 antisymmetric (or skew-symmetric) matrix, which has 6 independent parameters in the case of infinitesimal Lorentz transformation. This fact also holds for finite Lorentz transformations.

Properties of Lie groups (some preliminaries)

The transformation laws of continuous groups (Lie groups) such as rotation or Lorentz group are typically conveniently expressed in an infinitesimal form. By combining successive infinitesimal transformations it is always possible to reconstruct from these the finite transformation laws. This is a consequence of the fact that exponential function e^x can always be obtained by a product of many small-x approximations. In particular, consider $e^{\delta\alpha X} \approx 1+\delta\alpha X$, where $\delta\alpha = \alpha/N,\ N \gg 1$. By taking successive applications of N such infinitesimal transformations we obtain

$$(1+\alpha X/N)(1+\alpha X/N)\times\ldots\times(1+\alpha X/N) \;=\; (1+\alpha X/N)^N\,,$$

which in the limit of large N tends to $e^{\alpha X}$. This can also be extended to more parameters α_i. In such a case one should substitute αX with $\sum_i \alpha_i X_i$. Here X_i are the so-called *group generators*. The finite group transformation is then given by $L(\alpha) = e^{\sum_i \alpha_i X_i}$. One can recover the group generators from a generic group element $L(\alpha)$ by taking $\left.\frac{\partial L(\alpha)}{\partial\alpha_i}\right|_{\alpha=0} = X_i$.

When we pass from infinitesimal to finite transformation, the generic group element will read

$$\boldsymbol{L}^{\rho}{}_{\tau} \;=\; \left(e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}}\right)^{\rho}{}_{\tau}\,. \qquad (2.19)$$

We can find $\boldsymbol{M}^{\mu\nu}$ by comparing expression (2.19) for $\|\omega_{\mu\nu}\| \ll 1$ ($\omega_{\mu\nu} = -\omega_{\nu\mu}$) with the infinitesimal form of $\boldsymbol{L}^{\rho}{}_{\tau}$ given by (2.14). This yields

$$\begin{aligned}
\boldsymbol{L}^{\rho}{}_{\tau} \;\underset{\|\omega_{\mu\nu}\|\ll 1}{=}\;& \delta^{\rho}{}_{\tau} - \frac{i}{4}(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau}\omega_{\mu\nu} \;=\; \delta^{\rho}{}_{\tau} + \omega^{\rho}{}_{\tau} \\
=\;& \delta^{\rho}{}_{\tau} + \eta^{\rho\mu}\eta_{\tau}{}^{\nu}\omega_{\mu\nu} \;=\; \delta^{\rho}{}_{\tau} + \frac{1}{2}\eta^{\rho\mu}\delta_{\tau}{}^{\nu}(\omega_{\mu\nu} - \omega_{\nu\mu}) \\
=\;& \delta^{\rho}{}_{\tau} + \frac{1}{2}(\eta^{\rho\mu}\delta^{\nu}_{\tau} - \eta^{\rho\nu}\delta^{\mu}_{\tau})\omega_{\mu\nu}\,. \qquad (2.20)
\end{aligned}$$

From this we have

$$(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau} \;=\; 2i\left(\eta^{\rho\mu}\delta^{\nu}_{\tau} - \eta^{\rho\nu}\delta^{\mu}_{\tau}\right)\,. \qquad (2.21)$$

2.3 Relativistic Wave Equations

A spinless relativistic particle can be described in terms of a scalar wave function $\phi(\boldsymbol{x}, t)$. This wave function cannot posses any internal index,

Properties of Lie groups (some preliminaries)

The transformation laws of continuous groups (Lie groups) such as rotation or Lorentz group are typically conveniently expressed in an infinitesimal form. By combining successive infinitesimal transformations it is always possible to reconstruct from these the finite transformation laws. This is a consequence of the fact that exponential function e^x can always be obtained by a product of many small-x approximations. In particular, consider $e^{\delta\alpha X} \approx 1 + \delta\alpha X$, where $\delta\alpha = \alpha/N$, $N \gg 1$. By taking successive applications of N such infinitesimal transformations we obtain

$$(1 + \alpha X/N)(1 + \alpha X/N) \times \ldots \times (1 + \alpha X/N) \; = \; (1 + \alpha X/N)^N \, ,$$

which in the limit of large N tends to $e^{\alpha X}$. This can also be extended to more parameters α_i. In such a case one should substitute αX with $\sum_i \alpha_i X_i$. Here X_i are the so-called *group generators*. The finite group transformation is then given by $L(\alpha) = e^{\sum_i \alpha_i X_i}$. One can recover the group generators from a generic group element $L(\alpha)$ by taking $\left.\frac{\partial L(\alpha)}{\partial \alpha_i}\right|_{\alpha=0} = X_i$.

When we pass from infinitesimal to finite transformation, the generic group element will read

$$\boldsymbol{L}^{\rho}{}_{\tau} \; = \; \left(e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}}\right)^{\rho}{}_{\tau} \, . \tag{2.17}$$

We can find $\boldsymbol{M}^{\mu\nu}$ by comparing expression (2.17) for $\|\omega_{\mu\nu}\| \ll 1$ ($\omega_{\mu\nu} = -\omega_{\nu\mu}$) with the infinitesimal form of $\boldsymbol{L}^{\rho}{}_{\tau}$ given by (2.14). This yields

$$\begin{aligned}
\boldsymbol{L}^{\rho}{}_{\tau} \; \underset{\|\omega_{\mu\nu}\| \ll 1}{=} \; & \delta^{\rho}{}_{\tau} - \frac{i}{4}(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau}\omega_{\mu\nu} \; = \; \delta^{\rho}{}_{\tau} + \omega^{\rho}{}_{\tau} \\
= \; & \delta^{\rho}{}_{\tau} + \eta^{\rho\mu}\eta_{\tau}{}^{\nu}\omega_{\mu\nu} \; = \; \delta^{\rho}{}_{\tau} + \frac{1}{2}\eta^{\rho\mu}\delta_{\tau}{}^{\nu}(\omega_{\mu\nu} - \omega_{\nu\mu}) \\
= \; & \delta^{\rho}{}_{\tau} + \frac{1}{2}(\eta^{\rho\mu}\delta^{\nu}_{\tau} - \eta^{\rho\nu}\delta^{\mu}_{\tau})\omega_{\mu\nu} \, .
\end{aligned} \tag{2.18}$$

From this we have

$$(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau} \; = \; 2i\left(\eta^{\rho\mu}\delta^{\nu}_{\tau} - \eta^{\rho\nu}\delta^{\mu}_{\tau}\right) \, . \tag{2.19}$$

2.3 Relativistic Wave Equations

A spinless relativistic particle can be described in terms of a scalar wave function $\phi(\boldsymbol{x}, t)$. This wave function cannot posses any internal index, which would otherwise bear information about other degrees of freedom, such as spin. Relativistic particles satisfy the energy-momentum dispersion relation

$$E \; = \; \sqrt{m^2 + \boldsymbol{p}^2} \, . \tag{2.20}$$

In classical relativity we do not consider negative sign in the dispersion relation.

Recall that $p^\mu = (E, \boldsymbol{p})$ and that there exists a relativistic invariant given by

$$p^\mu p_\mu = p_0^2 - \boldsymbol{p}^2 = m^2 . \tag{2.21}$$

In the formalism of first quantization, quantum mechanics is brought about by identifying operators with dynamical quantities

$$\boldsymbol{p} \to -i\boldsymbol{\nabla}, \quad E \to i\frac{\partial}{\partial t} . \tag{2.22}$$

Applying this prescription to the relativistic invariant (2.21) we arrive at the following equation

$$\left(-\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}^2\right)\phi(x) = m^2\phi(x) . \tag{2.23}$$

From the fact that $\partial_\mu = \left(\frac{\partial}{\partial t}, \boldsymbol{\nabla}\right)$ we can equivalently rewrite this equation as

$$\partial^\mu \partial_\mu \phi = \Box\phi = -m^2\phi . \tag{2.24}$$

Finally, we arrive at the relativistic wave equation known as the Klein–Gordon equation, given by

Klein–Gordon equation just reflects energy dispersion relation (similarly as Schrödinger equation) so, all relativistic wave functions should satisfy this equation.

$$(\Box + m^2)\phi(x) = 0 . \tag{2.25}$$

If we accept this equation and seek solution of a definite energy and momentum, we get

$$\phi(x) \propto e^{-ipx} = e^{-iEt + i\boldsymbol{p}\cdot\boldsymbol{x}} = e^{-ip_0 x_0 + i\boldsymbol{p}\cdot\boldsymbol{x}} . \tag{2.26}$$

Adopting $\partial_\mu \phi = -ip_\mu \phi$ we get that $\Box\phi = -p^2\phi$ and then

$$(-p^2 + m^2)\phi = 0 . \tag{2.27}$$

Both positive and negative energy solutions are relevant in relativistic quantum theory!

So if $\phi \neq 0$ we have condition that $p^2 = m^2$ and hence

$$E = \pm\sqrt{\boldsymbol{p}^2 + m^2} . \tag{2.28}$$

Why can't we directly quantize relativistic energy relation?

A question may rise, why can't we directly quantize dispersion relation $\omega_{\boldsymbol{p}} = E_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2}$ using fact that $\boldsymbol{p} \to -i\boldsymbol{\nabla}$? To make sense to such a function of operator we have to interpret it in terms of its Taylor expansion:

$$H_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2} = m\left(1 + \frac{\boldsymbol{p}^2}{m^2}\right)^{\frac{1}{2}} = m + \frac{\boldsymbol{p}^2}{2m} - \frac{\boldsymbol{p}^4}{8m^3} + \cdots$$

Unfortunately, in this way we can not form covariant wave equation, i.e. if we formed a coordinate space representation of a state vector $|\psi\rangle$, the resulting wave equation would have *one* time derivative and *infinite* series of increasing spatial derivatives. There is no way to put time and space on an "equal footing". Nonetheless, let us go ahead and try to build a wave equation

$$i\frac{\partial}{\partial t}\langle \boldsymbol{x}|\psi(t)\rangle = \langle \boldsymbol{x}|H_{\boldsymbol{p}}|\psi(t)\rangle .$$

The matrix element $\langle \boldsymbol{x}|H_{\boldsymbol{p}}|\psi(t)\rangle$ is proportional to the infinite sum of $\langle \boldsymbol{x}|\boldsymbol{p}^n|\psi(t)\rangle = (-i)^n \frac{\partial^n}{\partial x^n}\langle \boldsymbol{x}|\psi(t)\rangle$ terms. This in turn renders wave function to be *non-local*, since it must reach further and further away from the region near $\boldsymbol{x}$ in order to evaluate the time derivative. Indeed, while the left-hand side can be written as

$$\lim_{\Delta t\to 0}\frac{\langle \boldsymbol{x}|\psi(t+\Delta t)\rangle - \langle \boldsymbol{x}|\psi(t)\rangle}{\Delta t},$$

a typical term on the right-hand side, i.e., term $(-i)^n \frac{\partial^n}{\partial x^n}\langle \boldsymbol{x}|\psi(t)\rangle$ has the form (for simplicity we consider $\boldsymbol{x}$ to be one-dimensional)

$$\lim_{\Delta x\to 0}\frac{(-i)^n}{(\Delta x)^n}\sum_{k=0}^{n}(-1)^k\begin{pmatrix} n \\ k \end{pmatrix}\langle x + \left(\frac{n}{2}-k\right)\Delta x|\psi(t)\rangle .$$

Here we use the so-called *central difference relation* for n-th derivative.

So, on the right-hand side we need all possible integer multiples of Δx. Eventually, the causality will be violated for any spatially localized function $\langle \boldsymbol{x}|\psi(t)\rangle$ since for understanding physics in the interval Δt we need to know physics in the interval $\bar{\Delta}x = k\Delta x$ (k is an arbitrary integer), which for sufficiently large k certainly satisfies $\Delta x^\mu \Delta x_\mu = (\Delta x^0)^2 - (\bar{\Delta}\boldsymbol{x})^2 < 0$, i.e., we require space-like separated events. Because of that we must abandon this approach and work with square of $H_{\boldsymbol{p}}$, (i.e., ω_p^2) instead. This will remove the problem of the square root, but will introduce a different problem — negative energies. This will still prove to be more useful way to proceed.

Let us look at non-relativistic limit of Klein–Gordon equation. A mode with $E = m + \varepsilon$ would oscillate in time as $\phi \propto e^{-iEt}$. In the non-relativistic regime ε is much smaller than the rest mass m. We can factor-out the fast-oscillating part of the ϕ away and rewrite it as

Here $\varepsilon = \frac{\boldsymbol{p}^2}{2m} + O(\boldsymbol{p}^4/m^3)$.

$$\phi(x) = \phi(\boldsymbol{x},t) = e^{-imt}\varphi(\boldsymbol{x},t) . \tag{2.29}$$

Field φ is oscillating much more slowly that e^{-imt} in time. By inserting this into the Klein–Gordon equation and using the fact that

$$\frac{\partial}{\partial t}e^{-imt}(\cdots) = e^{-imt}\left(-im + \frac{\partial}{\partial t}\right)(\cdots) , \tag{2.30}$$

we obtain

$$
\begin{aligned}
&\left(\frac{\partial}{\partial t} - \nabla^2 + m^2\right)\phi(x) \\
&\quad = \frac{\partial}{\partial t}\left[e^{-imt}\left(-im + \frac{\partial}{\partial t}\right)\varphi\right] - e^{-imt}\nabla^2\varphi(\boldsymbol{x},t) + m^2 e^{-imt}\varphi \\
&\quad = e^{-imt}\left(-im + \frac{\partial}{\partial t}\right)\left(-im + \frac{\partial}{\partial t}\right)\varphi - e^{-imt}\nabla^2\varphi + m^2 e^{-imt}\varphi \\
&\quad = e^{-imt}\left[\left(-m^2 - 2im\frac{\partial}{\partial t} + \frac{\partial^2}{\partial t^2}\right)\varphi - \nabla^2\varphi + m^2\varphi\right] = 0\,. \qquad (2.31)
\end{aligned}
$$

By the way, the Klein-Gordon equation was actually discovered before Schrödinger equation by Erwin Schrödinger himself.

Dropping $\frac{\partial^2\varphi}{\partial t^2}$ as small compared to $-imt\frac{\partial\varphi}{\partial t}$ we find that

$$
i\frac{\partial}{\partial t}\varphi = -\frac{\nabla^2}{2m}\varphi\,, \qquad (2.32)
$$

which is nothing but the Schrödinger equation for a free particle.

Let us focus on general solution to the Klein-Gordon equation, $\phi(x)$. With the help of Fourier decomposition we can write

$$
\phi(x) = \int \frac{\mathrm{d}^4 p}{(2\pi)^3} e^{-ipx}\,\widetilde{\phi}(p)\,. \qquad (2.33)
$$

To find the solution we will solve the Klein-Gordon equation in momentum space, which yields

$$
(p^2 - m^2)\widetilde{\phi}(p) = 0\,. \qquad (2.34)
$$

Equations of this form are solved by the Dirac δ-functions. In particular, the solution in momentum space reads

Here we use the well known property of Dirac δ-function, namely that $\delta(f(x)) = \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|}$, where x_i are roots of f.

$$
\begin{aligned}
\widetilde{\phi}(p) &= f(p)\delta(p^2 - m^2) \\
&= \frac{f(p)\delta(p_0 + \sqrt{\boldsymbol{p}^2 + m^2})}{2\sqrt{\boldsymbol{p}^2 + m^2}} + \frac{f(p)\delta(p_0 - \sqrt{\boldsymbol{p}^2 + m^2})}{2\sqrt{\boldsymbol{p}^2 + m^2}}\,. \qquad (2.35)
\end{aligned}
$$

Using this knowledge and denoting $\omega_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2}$, we can write the full solution as

$$
\begin{aligned}
\phi(x) &= \int \frac{\mathrm{d}^4 p}{(2\pi)^3}\frac{1}{2\omega_{\boldsymbol{p}}} e^{-ipx}[f(p)\delta(p_0 + \omega_{\boldsymbol{p}}) + f(p)\delta(p_0 - \omega_{\boldsymbol{p}})] \\
&= \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3}\frac{1}{2\omega_{\boldsymbol{p}}}[e^{-i\omega_{\boldsymbol{p}}t + i\boldsymbol{p}\cdot\boldsymbol{x}} f(\omega_{\boldsymbol{p}}, \boldsymbol{p}) + e^{i\omega_{\boldsymbol{p}}t + i\boldsymbol{p}\cdot\boldsymbol{x}} f(-\omega_{\boldsymbol{p}}, \boldsymbol{p})] \\
&= \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3}\frac{1}{2\omega_{\boldsymbol{p}}}[e^{-ipx}\underbrace{f(\omega_{\boldsymbol{p}}, \boldsymbol{p})}_{f(\boldsymbol{p})} + e^{ipx}\underbrace{f(-\omega_{\boldsymbol{p}}, -\boldsymbol{p})}_{g(\boldsymbol{p})}]\,. \qquad (2.36)
\end{aligned}
$$

Here $p^\mu = (\omega_{\boldsymbol{p}}, \boldsymbol{p})$.

So, we see that a general solution of the Klein–Gordon equation is a superposition of positive and negative energy eigenstate solutions.

If we want to interpret $\phi(x)$ as a wave function, we have to find a non-negative norm, which is conserved by time evolution and is Lorentz invariant. Let us define the norm of $\phi(x)$ to be

$$||\phi||^2 = (\phi|\phi) = i\int d^3\boldsymbol{x}\left[\phi^*\frac{\partial\phi}{\partial x^0} - \left(\frac{\partial\phi}{\partial x^0}\right)^*\phi\right]. \qquad (2.37)$$

This is, in a sense, a natural candidate for the norm. The naturalness of this choice comes from the analogy with quantum mechanics — continuity equation, which defines the probability density.

We know that each 4-current should have the form $J_\mu = (\rho, \boldsymbol{J})$ and should be conserved (after equations of the motion are taken into account), i.e., $\partial^\mu J_\mu = 0$. To this end we consider the 4-current

$$J_\mu(x) = \frac{i}{2m}\left[\phi^*\partial_\mu\phi - (\partial_\mu\phi)^*\phi\right], \qquad (2.38)$$

(factor $1/2m$ is only a convention that ensures a correct non-relativistic limit, see Eq. (2.41)). Eq. (2.38) can be equivalently rewritten as

$$\begin{aligned} \boldsymbol{J}(x) &= \frac{i}{2m}\left[\phi^*\boldsymbol{\nabla}\phi - (\boldsymbol{\nabla}\phi)^*\phi\right], \\ \rho(x) &= \frac{i}{2m}\left[\phi^*\partial_0\phi - (\partial_0\phi)^*\phi\right]. \end{aligned} \qquad (2.39)$$

Let us now compute $\partial_\mu J^\mu = \partial^\mu J_\mu$:

$$\begin{aligned} \partial^\mu J_\mu(x) &= i\left[\partial^\mu(\phi^*\partial_\mu\phi) - \partial^\mu(\phi\partial_\mu\phi^*)\right] \\ &= i[(\partial_\mu\phi^*)(\partial^\mu\phi) + \underbrace{\phi^*\partial^2\phi}_{-m^2\phi^*\phi} - (\partial_\mu\phi)(\partial^\mu\phi^*) - \underbrace{\phi\partial^2\phi^*}_{-m^2\phi\phi^*}] \\ &= 0. \end{aligned} \qquad (2.40)$$

The existence and explicit form of the conserved currents will be discussed in connection with Noether's theorem in Section **??**.

So, the current J_μ is conserved and can be used to prove time-independence of the norm (as in ordinary quantum mechanics).

Current in non-relativistic limit

In non-relativistic limit, we assume that $\phi(x) = e^{-imt}\varphi(\boldsymbol{x}, t)$ where $\varphi(\boldsymbol{x}, t)$ is supposed to be a non relativistic wave function. By inserting the aforementioned form of $\phi(x)$ to the explicit form of J_μ we obtain:

$$\begin{aligned} \boldsymbol{J}_{NR}(x) &= \frac{i}{2m}\left[\varphi^*\boldsymbol{\nabla}\varphi - (\boldsymbol{\nabla}\varphi)^*\varphi\right], \\ \rho_{NR}(x) &= \frac{i}{2m}\left[(-im)\varphi\varphi^* + \varphi^*\partial_0\varphi - (im)\varphi\varphi^* - (\partial_0\varphi)^*\varphi\right] \\ &= \frac{i}{2m}\left[(-i2m)\varphi\varphi^*\right] = \varphi\varphi * . \end{aligned} \qquad (2.41)$$

Here we have neglected $\partial_t\varphi$ in comparison to $-im\varphi$. Eq. (2.41) is the well know form of Schrödinger's conserved probability current and charge (i.e., probability density).

With the conserved current we can now show that the norm is time independent, indeed

$$\underbrace{\frac{\partial}{\partial t}\int_V \mathrm{d}^3\boldsymbol{x}\,\rho}_{\substack{\text{Change in total}\\ \text{probability inside } V}} = \int_V \mathrm{d}^3\boldsymbol{x}\,\boldsymbol{\nabla}\cdot\boldsymbol{J} = \underbrace{\int_{\partial V}\mathrm{d}\boldsymbol{S}\cdot\boldsymbol{J}}_{\substack{\text{Flux of } \boldsymbol{J}\\ \text{through the } \partial V = S^2(R)}} \to 0\,. \tag{2.42}$$

Convergence issue

We want to show that $\int_V \mathrm{d}^3x\,\rho$ is finite, so that ρ could (potentially) represent a density of probability. Since in the non relativistic case $\rho_N = \phi^*\phi = \varphi\varphi*$, we know that the integral (in spherical coordinates) $\int \mathrm{d}\Omega\,\mathrm{d}r r^2|\phi|^2 < \infty$. Since our fields shout for large r behave as $\phi \lesssim \frac{1}{r^{3/2+\varepsilon}}$. Our current then has to behave at large r like $\boldsymbol{J} \sim \phi\boldsymbol{\nabla}\phi \lesssim \frac{1}{r^{4+\varepsilon}}$. Since $\partial V = S^2(R) \sim R^2$ and $\boldsymbol{J} \lesssim \frac{1}{R^{4+\varepsilon}}$, we have that the total integral of $\lim_{R\to\infty}\int_{\partial V}\mathrm{d}\boldsymbol{S}\cdot\boldsymbol{J} = 0$.

Let us now show that our norm is relativistically invariant. To this end we can write:

$$\begin{aligned}||\phi||^2 &= \int \mathrm{d}^3\boldsymbol{x}\rho(x) = \int \mathrm{d}^4x\,\underbrace{\delta(x^0)}_{\frac{\partial}{\partial x^0}\theta(x^0)}\,\rho(x)\\ &= \int \mathrm{d}^4x\,J^\alpha\frac{\partial}{\partial x^\alpha}\theta(n^\beta x_\beta)\,.\end{aligned} \tag{2.43}$$

Here $n^\beta = (1,\ 0,\ 0,\ 0)$.

The effect of a Lorentz transformation on $||\phi||^2$ is then evidently simply to change n^μ. So, in this connection we define another wave function norm, $||\widetilde{\phi}||^2$ as

Time-like vectors have not only dominant time component with respect to space-like components, but in addition, if they are related via Lorentz transform with the vector $n^\mu = (1,\ 0,\ 0,\ 0)$ the have the zero component positive for all *orthochronous Lorentz transformations*.

$$||\widetilde{\phi}||^2 = \int \mathrm{d}^4x J^\alpha\frac{\partial}{\partial x^\alpha}\theta(n'^\beta x_\beta)\,. \tag{2.44}$$

Here n' is a generic time-like 4-vector obtained from n via Lorentz transformation, i.e., $n'^\beta = \boldsymbol{L}^\beta{}_\gamma n^\gamma$. Taking difference between two norms we obtain

$$||\phi||^2 - ||\widetilde{\phi}||^2 = \int \mathrm{d}^4x J^\alpha\frac{\partial}{\partial x^\alpha}\left[\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\right], \tag{2.45}$$

Because $\partial_\alpha J^\alpha = 0$ (as seen before) we can rewrite expression inside the integral as

$$||\phi||^2 - ||\widetilde{\phi}||^2 = \int \mathrm{d}^4x\frac{\partial}{\partial x^\alpha}\left[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\right]. \tag{2.46}$$

By using 4-dimensional version of Gaussian theorem we obtain that

$$||\phi||^2 - ||\widetilde{\phi}||^2 = \int \mathrm{d}S_\alpha\left[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\right]. \tag{2.47}$$

To show that this is zero, consider two possibilities:

1. J^α can be presumed to vanish if $|\boldsymbol{x}| \to \infty$ with fixed t.
2. $\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)$ vanishes for $|t| \to \infty$ with $\boldsymbol{x}$ fixed.

Hence, the difference is zero and the norm is relativistically invariant (in fact Lorentz scalar) as needed.

Let us now return to the general solution

$$\phi(x) = \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_p} \left[f(\boldsymbol{p}) e^{-ipx} + g(\boldsymbol{p}) e^{ipx} \right] , \tag{2.48}$$

and explore, what its norm would look like:

$$\begin{aligned} ||\phi||^2 &= (\phi, \phi) \\ & i \int \mathrm{d}^3\boldsymbol{x} \left\{ \left[\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_p} \left(f^*(\boldsymbol{p}) e^{ipx} + g^*(\boldsymbol{p}) e^{-ipx} \right) \right. \right. \\ & \times \left. \int \frac{\mathrm{d}^3\boldsymbol{q}}{(2\pi)^3 2\omega_q} \left(f(\boldsymbol{q})(-i\omega_q) e^{-iqx} + g(\boldsymbol{q})(i\omega_q) e^{iqx} \right) \right] \\ & - \left[\int \frac{\mathrm{d}^3\boldsymbol{q}}{(2\pi)^3 2\omega_q} \left(f(\boldsymbol{q}) e^{-iqx} + g(\boldsymbol{q}) e^{iqx} \right) \right. \\ & \times \left. \left. \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_q} \left(f^*(\boldsymbol{p})(i\omega_p) e^{ipx} + g^*(\boldsymbol{p})(-i\omega_p) e^{-ipx} \right) \right] \right\} . \end{aligned} \tag{2.49}$$

Since our norm is time independent, elements of type $e^{\pm i(\omega_p + \omega_q)t}$ must cancel, and only terms of the type $e^{\pm i(\omega_p - \omega_q)t}$ should be considered. Continuing

$$\begin{aligned} ||\phi||^2 &= i \int \mathrm{d}^3\boldsymbol{x} \int \frac{\mathrm{d}^3\boldsymbol{p}\,\mathrm{d}^3\boldsymbol{q}}{(2\pi)^6 4\omega_p \omega_q} \left[\left(f^*(\boldsymbol{p}) f(\boldsymbol{q})(-i\omega_q) e^{ix(p-q)} \right. \right. \\ &+ \left. g^*(\boldsymbol{p}) g(\boldsymbol{q})(i\omega_q) e^{-ix(p-q)} \right) \\ &- \left(f(\boldsymbol{q}) f^*(\boldsymbol{p})(i\omega_p) e^{ix(p-q)} \right. \\ &+ \left. \left. g(\boldsymbol{q}) g^*(\boldsymbol{p})(-i\omega_p) e^{-ix(p-q)} \right) \right] , \end{aligned} \tag{2.50}$$

and hence finally we have

$$||\phi||^2 = \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_p} \left[|f(\boldsymbol{p})|^2 - |g(\boldsymbol{p})|^2 \right] . \tag{2.51}$$

This norm is not generally positive definite! However, if we restrict our attention to positive energy only, i.e., $g(\boldsymbol{p}) = 0$, then $||\phi||^2$ is positive definite.

Note

In a similar spirit as for norm, we can be define the general scalar product between two states as

$$(\psi, \phi) = i \int \mathrm{d}^3\boldsymbol{x} \left[\psi^* \partial_0 \phi - \phi (\partial_0 \psi)^* \right] .$$

The reader is encouraged to check that this scalar product is indeed time independent and Lorentz invariant.

We thus see that ρ cannot represent probability density, it may well be considered (while satisfying continuity equation) as the density of charge (or any other conserved quantity).

Apart from semidefinite probability density, there is a second problem with Klein–Gordon equation. In particular, any plane wave function, i.e.

$$\phi(x) = N e^{\pm i(\omega_{\boldsymbol{p}} t - \boldsymbol{p}\cdot\boldsymbol{x})}, \tag{2.52}$$

satisfies Klein–Gordon equation, provided that $E^2 = \omega_{\boldsymbol{p}}^2 = \boldsymbol{p}^2 + m^2$. Thus negative energies $E = -\sqrt{\boldsymbol{p}^2 + m^2}$ are on the same footing as the physical ones $E = \sqrt{\boldsymbol{p}^2 + m^2}$. This leads to a problem — energy spectrum is unbounded from below.

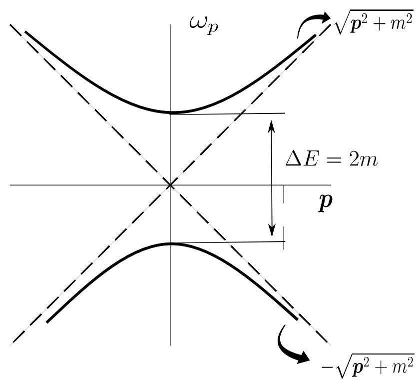


Figure 2.1: Energy spectrum of a free quantum relativistic particle.

Of course, even in classical physics, the relativistic relation $E^2 = \boldsymbol{p}^2 + m^2$ has two solutions $E = \pm\sqrt{\boldsymbol{p}^2 + m^2}$. However, in classical physics we can simply assume that the only physical particles are those with $E \geq 0$. This is because the positive-energy solutions have $E > mc^2$, while the negative ones have $E \leq -mc^2$. Hence there is a finite gap between them and in classical (non-quantum) physics there does not exist any continuous process that can take a particle from positive to negative energy.

P.A.M. Dirac, Proc. Roy. Soc. A117, 610 (1928).

In relativistic quantum mechanics the problem is more pressing. As Dirac pointed out in his 1928 paper, the interaction of electrons with radiation can produce transition, in which a positive energy electrons falls into a negative energy state, with the energy carried off by two or more photons.

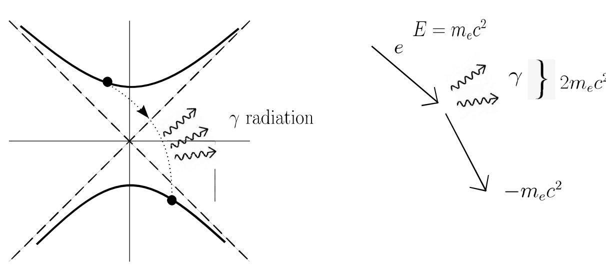


Figure 2.2: Spontaneous emission from excited state. For instance, an electron in rest could emit two photon quanta with total energy $2m_e c^2$ and hence end up in the negative energy level.

This brings about a problem. If we have a quantum particle whose state satisfies the Klein–Gordon equation we could, in principle, extract an arbitrary amount of energy from it (in the form radiated photons). This, in turn, would lead to the perpetuum mobile of the first kind. In addition, when particle reaches the negative energy states there is nothing that would prevent it to decay to even lower energy state. Consequently, the matter (together with us) would be unstable!

Exercises: Lorentz transformations

Lorentz group

Lorentz transformations $\boldsymbol{L}$ are defined by the property

$$g_{\mu\nu}\boldsymbol{L}^{\mu}{}_{\rho}\boldsymbol{L}^{\nu}{}_{\sigma} \;=\; g_{\rho\sigma}\,,$$

where $g_{\mu\nu} = g^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$ is the Minkowski metric.

Their exponential form reads

$$\boldsymbol{L} \;=\; \exp\left(-\frac{i}{4}\omega_{\mu\nu}\boldsymbol{M}^{\mu\nu}\right) \quad , \quad (\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\sigma} \;=\; 2i(g^{\mu\rho}\delta^{\nu}_{\sigma} - g^{\nu\rho}\delta^{\mu}_{\sigma})\,,$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$, $\boldsymbol{M}^{\mu\nu} = -\boldsymbol{M}^{\nu\mu}$.

Exercise 2.1 Find explicit "tabular" form of the matrix $(\boldsymbol{M}^{01})^{\rho}{}_{\sigma}$ (the generator of boosts in x^1-direction).

Exercise 2.2 Verify the commutation relations of the Lorentz algebra $so(1,3)$

$$[\boldsymbol{M}^{\mu\nu},\boldsymbol{M}^{\rho\sigma}] \;=\; -2i(g^{\mu\rho}\boldsymbol{M}^{\nu\sigma} - g^{\mu\sigma}\boldsymbol{M}^{\nu\rho} + g^{\nu\sigma}\boldsymbol{M}^{\mu\rho} - g^{\nu\rho}\boldsymbol{M}^{\mu\sigma})\,.$$

Klein–Gordon equation

Exercise 2.3 Show that the Klein-Gordon equation ($\hbar = c = 1$)

$$(\partial_\mu\partial^\mu \;+\; m^2)\phi(x) \;=\; 0\,, \qquad x \equiv (x^0,x^1,x^2,x^3)\,, \qquad \partial_\mu \;\equiv\; \frac{\partial}{\partial x^\mu}\,,$$

is invariant under Lorentz transformations $x^{\mu\prime} = L^{\mu}{}_{\nu}x^{\nu}$.

Scalar product between two states is defined

$$(\psi,\phi) \;=\; i\int d^3x[\psi^*\partial_0\phi - \phi\,\partial_0\psi^*]\,.$$

Exercise 2.4 Show that (ψ,ϕ) is time-independent

[**Hint:** Make use of the fact that $\psi(x)$ and $\phi(x)$ are solutions of the Klein-Gordon equation.]

Exercise 2.5 Show that (ψ,ϕ) is relativistically invariant.

[**Hint:** Write $(\psi,\phi) = i\int d^4x(\psi^*\partial_0\phi - \phi\,\partial_0\psi^*)\delta(x^0)$, and use $\delta(x^0) = \partial_0\theta(x^0)$.]

3 Dirac Equation — introduction

Klein-Gordon equation is a second order differential equation in time, which can be recognized as a reason why the norm is not positive definite. Dirac sought an equation, that would remedy this "difficulty". It turned out, that by appropriately "linearizing" relativistic wave equation, Dirac arrived (by coincidence) on the wave equation for electron, which indeed provides positive definite probability density. Since the spin is involved, the wave function is not anymore Lorentz scalar (recall Pauli-Schrödinger wave equation, which has as a solution two-component spinor wave function that is not scalar with respect to Galileo group).

Dirac had two goals:

1. Equation for wave function that is linear and first order in time derivative. Relativistic invariance then suggests that the equation will also be of first order in spatial derivatives.
2. Positive definite norm.

Assume that this equation has the form

$$\left(i\gamma^0 \frac{\partial}{\partial x^0} + i\boldsymbol{\gamma}\cdot\boldsymbol{\nabla}\right)\psi(x) = m\psi(x)\,, \tag{3.1}$$

which might be written in a shorthand notation

$$\left(i\gamma_\mu \partial^\mu - m\right)\psi(x) = 0\,. \tag{3.2}$$

By defining Feynman's slash notation $\partial\!\!\!/ = \gamma_\mu\partial^\mu$ this equation reduces to

$$\left(i\partial\!\!\!/ - m\right)\psi(x) = 0\,. \tag{3.3}$$

Here $\{\gamma^\mu\} = \{\gamma^0, \boldsymbol{\gamma}\}$ are some unspecified numbers or matrices. We require that $\psi(x)$ should also satisfy Klein-Gordon equation, since Klein–Gordon equation just states that $p_\mu p^\mu = m^2$. Multiplying (3.2) by $(i\gamma_\nu\partial^\nu + m)$ we get

$$0 = \left(i\gamma_\nu\partial^\nu + m\right)\left(i\gamma_\mu\partial^\mu - m\right)\psi(x) = \left(-\gamma_\mu\gamma_\nu\partial^\mu\partial^\nu - m^2\right)\psi(x)\,. \tag{3.4}$$

We can rewrite $\gamma_\mu\gamma_\nu\partial^\mu\partial^\nu$ as

$$\gamma_\mu\gamma_\nu\partial^\mu\partial^\nu = \frac{1}{2}\gamma_\mu\gamma_\nu\partial^\mu\partial^\nu + \frac{1}{2}\gamma_\nu\gamma_\mu\partial^\nu\partial^\mu = \frac{1}{2}\{\gamma_\mu, \gamma_\nu\}\partial^\mu\partial^\nu\,, \tag{3.5}$$

where $\{\gamma_\mu, \gamma_\nu\} = \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu$. Here $\{A, B\}$ is a symmetric combination of A and B. This operation is called *anti-commutator*.

To obtain Klein–Gordon equation we must impose condition

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}\,, \quad \{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\,. \tag{3.6}$$

Because $\gamma^\mu\gamma^\nu\partial^\mu\partial^\nu = \partial_\nu\partial^\nu$ we get $(\Box + m^2)\psi(x) = 0$.

Dirac's derivation

Dirac started with the following ansatz:

$$i\frac{\partial\psi}{\partial t} = \left(\frac{1}{i}\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m\right)\psi = H_D\psi\,, \tag{3.7}$$

where H_D is Dirac's Hamiltonian, which should be Hermitian (and hence $\boldsymbol{\alpha}$ and β are Hermitian). Klein–Gordon equation implies that $\{\alpha_i, \alpha_k\} = 0$, $\{\alpha_i, \beta\} = 0$ for $i \neq k$ and $\alpha_i^2 = \beta^2 = \mathbb{1}$.

By rewriting Dirac equation (3.2) explicitly as

$$\left(i\gamma^0\partial_0 + \gamma^i\partial_i - m\right)\psi = 0\,, \tag{3.8}$$

and multiplying it by the inverse of γ^0 we get

$$\left[i(\gamma^0)^{-1}\gamma^0\partial_0 + i(\gamma^0)^{-1}\gamma^i\partial_i - (\gamma^0)^{-1}m\right]\psi = 0\,, \tag{3.9}$$

which is equivalent to

$$i\partial_0\psi = \left[\frac{1}{i}(\gamma^0)^{-1}\gamma^i\partial_i + (\gamma^0)^{-1}m\right]\psi. \tag{3.10}$$

Consequently we see that $\boldsymbol{\alpha} = (\gamma^0)^{-1}\boldsymbol{\gamma}$ and $\beta = (\gamma^0)^{-1}$. Because $\{\gamma^0, \gamma^0\} = 2$, we see that $\gamma^0 = (\gamma^0)^{-1} = (\gamma^0)^\dagger$. From anti-commutation relation for γ^0 and γ^i we have that $\gamma^0\gamma^i = -\gamma^i\gamma^0$ and from Hermiticity we also have that $\gamma^0\gamma^i = (\gamma^0\gamma^i)^\dagger = \gamma^{i\dagger}\gamma^{0\dagger}$, from which we see that $\gamma^0\gamma^i\gamma^0 = \gamma^{i\dagger} = -\gamma^i(\gamma^0)^2 = -\gamma^i$.

Relation

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{1}\,, \tag{3.11}$$

is known as *Clifford algebra* $\mathrm{CL}_{1,3}(\mathbb{R})$ or simply *Dirac algebra*.

Now we can ask ourselves, what is the smallest dimension of γ^μ in 4-dimensional space. In fact, matrices α^i and β have eigenvalues equal to ± 1. For $i \neq j$ we have

$$\begin{aligned} \det\left(\alpha^i\alpha^j\right) &= \det\left(-\alpha^j\alpha^i\right) = (-1)^d\det\left(\alpha^j\alpha^i\right)\,, \\ \det\left(\alpha^i\beta\right) &= (-1)^d\det\left(\beta\alpha^i\right)\,. \end{aligned} \tag{3.12}$$

So, the dimension of α^i, $i = 1, 2, 3$ and β must be even. Since for $d = 2$ there exists only 3 anti-commuting Hermitian matrices — Pauli matrices, we have $d \geq 4$. There are many representations of $\mathrm{CL}_{1,3}(\mathbb{R})$ with $d > 4$ (altought they are rarely used in practice). An explicit representation with $d = 4$ is provided by matrices

Pauli matrices by themselves generate Clifford algebra $\mathrm{CL}_{0,3}(\mathbb{R})$ via relation $\{\sigma_a, \sigma_b\} = 2\delta_{ab}\mathbf{1}$. This algebra is known as Pauli algebra.

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{pmatrix}, \quad \boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}. \tag{3.13}$$

Here $\boldsymbol{\sigma}$ are Pauli matrices. This representation is known as *Dirac representation* and is useful when discussing non-relativistic limit of the theory. A useful technical trick, for calculating gamma matrices in

Dirac's representation (but also in other representations) is based on properties of tensor product $\otimes$ on matrices. Because

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \cdot (\boldsymbol{C} \otimes \boldsymbol{D}) \ = \ (\boldsymbol{A} \cdot \boldsymbol{C}) \otimes (\boldsymbol{B} \cdot \boldsymbol{D}) \,, \tag{3.14}$$

and

$$\gamma^0 \ = \ \sigma^3 \otimes \mathbb{1}, \qquad \boldsymbol{\gamma} \ = \ i\sigma^2 \otimes \boldsymbol{\sigma} \,, \tag{3.15}$$

and

$$\sigma^i \sigma^j \ = \ \delta^{ij} \ + \ i\varepsilon^{ijk}\sigma^k \,, \tag{3.16}$$

we can immediately check that γ matrices in Dirac's representation satisfy the defining Dirac algebra (3.11).

3.1 Lorentz Invariance of Dirac Equation

Recall that a non-relativistic particle with spin has a wave function $\psi_\alpha(x)$ (Weyl spinor) which transforms under rotation $\boldsymbol{R}$ as

$$\psi_\alpha(\boldsymbol{x}) \ \stackrel{\boldsymbol{R}}{\rightarrow} \ D_{\alpha\beta}(\boldsymbol{R})\psi_\beta(\boldsymbol{R}^{-1}\boldsymbol{x}) \,. \tag{3.17}$$

In a similar way, the Dirac wave function under a Lorentz transformation $\boldsymbol{L}$ transforms as

$$\psi(x) \ \stackrel{\boldsymbol{L}}{\rightarrow} \ \psi_{\boldsymbol{L}}(x) \ = \ S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x) \,, \tag{3.18}$$

where $S(\boldsymbol{L})$ is an appropriate representation of the Lorentz group, that acts in the vector space, in which the Dirac wave function takes its values. From this we see that $S(\boldsymbol{L})$ should be a 4×4 matrix.

In order to show that Dirac's equation is Lorentz invariant we need to show that

$$\left(i\gamma^\mu \partial_\mu - m\right) \psi_{\boldsymbol{L}}(x) \ = \ 0 \,, \tag{3.19}$$

provided

$$\left(i\gamma^\mu \partial_\mu - m\right) \psi(x) \ = \ 0 \,, \tag{3.20}$$

is satisfied. To show this we rewrite the left-hand-side (LHS) of (3.19) as

$$\begin{aligned} \text{LHS} \ &= \ \left(i\gamma^\mu \partial_\mu - m\right) S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x) \\ &= \ S(\boldsymbol{L}) \left[iS^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) - m\right] \psi(\boldsymbol{L}^{-1}x) \\ &= \ S(\boldsymbol{L}) \left[iS^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) - m\right] \psi(x') \,, \end{aligned} \tag{3.21}$$

where $x' = \boldsymbol{L}^{-1}x$. If we can find an appropriate matrix $S(\boldsymbol{L})$ such that

$$S^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) \ = \ \gamma^\mu \partial'_\mu \,, \tag{3.22}$$

Since now $x'^\mu = (\boldsymbol{L}^{-1})^\mu{}_\nu x^\nu$ we have that $x^\mu = \boldsymbol{L}^\mu{}_\nu x'^\nu$ and thus $\partial'_\mu = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = \boldsymbol{L}^\nu{}_\mu \frac{\partial}{\partial x^\nu}$.

then

$$\left(i\gamma^\mu \partial_\mu - m\right) \psi_{\boldsymbol{L}}(x) \ = \ S(\boldsymbol{L})(i\gamma^\mu \partial'_\mu - m)\psi(x') \,. \tag{3.23}$$

Since ψ is a Dirac wave function, then

$$\left(i\gamma^\mu \partial'_\mu - m\right) \psi(x') \ = \ 0 \quad \Rightarrow \quad \left(i\gamma^\mu \partial_\mu - m\right) \psi_{\boldsymbol{L}}(x) \ = \ 0 \,. \tag{3.24}$$

Consequently, also $\psi_{\boldsymbol{L}}$ is Dirac's wave function and Dirac's equation is relativistically invariant provided such $S(\boldsymbol{L})$ exists. To find $S(\boldsymbol{L})$ we have to explore properties of Lorentz group in more detail.

Lorentz group

From the condition (2.6) we know that there exist 2 kinds of Lorentz transformations — proper (those with $\det\{\boldsymbol{L}\} = 1$) and improper (those with $\det\{\boldsymbol{L}\} = -1$). We can expand this classification even further by realizing that

$$1 = \eta^{00} = \boldsymbol{L}^0{}_\nu \boldsymbol{L}^0{}_{\nu'} \eta^{\nu\nu'} = (\boldsymbol{L}^0{}_0)^2 - \sum_{i=1}^{3} (\boldsymbol{L}^0{}_i)^2 . \tag{3.25}$$

Rewriting this we arrive at the condition

$$(\boldsymbol{L}^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\boldsymbol{L}^0{}_i)^2 . \tag{3.26}$$

Lorentz transformations for which $\boldsymbol{L}^0{}_0 \geq 1$ are called *orthochronous* transformations, those with $\boldsymbol{L}^0{}_0 \leq -1$ are called *non-orthochronous*. We can not switch between those four types of Lorentz transformations using continuous process. Transitions between these disconnected parts of the Lorentz group can be done only via discrete transformations such as *parity* or *time reversal* (see Chapter 6.1).

Let us recall (Chapter 2.2.) that Lorentz group has 6 independent parameters $\omega_{\mu\nu} = -\omega_{\nu\mu}$ that define it. In terms of these we can construct finite Lorentz transformations as

$$\boldsymbol{L}^\rho{}_\tau = \left(e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}}\right)^\rho{}_\tau . \tag{3.27}$$

Here $M^{\mu\nu}$ are so-called generators of Lorentz transformation, which are fixed by comparison with the infinitesimal transformation (i.e., when $\|\omega_{\mu\nu}\| \ll 1$). In Chapter 2.2 we have found that

$$(\boldsymbol{M}^{\mu\nu})^\rho{}_\tau = 2i\left(g^{\rho\mu}\delta^\nu_\tau - g^{\rho\nu}\delta^\mu_\tau\right) . \tag{3.28}$$

Rotation group

To proceed, we start with a familiar example — *rotation group*. Elements of a rotation group are defined by

$$\boldsymbol{R}_{ij} = \left(e^{-i\theta \boldsymbol{n}_k \boldsymbol{J}_k}\right)_{ij} = \left(e^{-i\omega_k \boldsymbol{J}_k}\right)_{ij} , \tag{3.29}$$

where $\boldsymbol{n}$ is a unit vector along the axis of rotation. Group generators J_i satisfy the usual angular momentum commutation relations

$$\left[\boldsymbol{J}_i, \boldsymbol{J}_j\right] = i\varepsilon_{ijk}\boldsymbol{J}_k . \tag{3.30}$$

Recall from quantum mechanics that *vector operators* are sets of 3 operators, say $V_i(\boldsymbol{x})$ that transform according to

$$U(\boldsymbol{R})^\dagger V_i(\boldsymbol{x}) U(\boldsymbol{R}) \;=\; \sum_j \boldsymbol{R}_{ij} V_j(\boldsymbol{x})\,. \tag{3.31}$$

Here $U(\boldsymbol{R})$ is a representation of rotation group which acts on state space (e.g. $L^2(\mathbb{R})$) and $\boldsymbol{R}_{ij}$ is a representation of rotation group that acts on the operator indices.

For infinitesimal rotations one obtains

$$(1 + i\omega_k \mathbb{J}_k)\boldsymbol{V}_i(1 - i\omega_k \mathbb{J}_k) \;=\; [1 - i\omega_k(\boldsymbol{J}_k)_{ij}]\boldsymbol{V}_j\,. \tag{3.32}$$

Here $\mathbb{J}_k$ are operators of angular momentum, acting on corresponding state space (e.g. $L^2(\mathbb{R})$) and $(\boldsymbol{J}_k)_{ij}$ is a vector representation of angular momentum in 3D, which is known to be

$$\boldsymbol{J}_1 = i\begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{pmatrix},\quad \boldsymbol{J}_2 = i\begin{pmatrix} 0 & 0 & 1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{pmatrix},\quad \boldsymbol{J}_3 = i\begin{pmatrix} 0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}. \tag{3.33}$$

This can succinctly be written as: $(\boldsymbol{J}_j)_{ik} = i\varepsilon_{ijk}$.

After algebraic manipulations we obtain from (3.32)

$$i\omega_k\,[\mathbb{J}_k, \boldsymbol{V}_i] \;=\; -i\omega_k(\boldsymbol{J}_k)_{ij}\boldsymbol{V}_j\,, \tag{3.34}$$

which then implies that

$$[\mathbb{J}_k, \boldsymbol{V}_i] = -(\boldsymbol{J}_k)_{ij}\boldsymbol{V}_j \;=\; -i\varepsilon_{ikj}\boldsymbol{V}_j \;=\; i\varepsilon_{kij}\boldsymbol{V}_j\,. \tag{3.35}$$

This relation represents an algebraic condition for vector operators. As a byproduct we see that the generators $\mathbb{J}_k$ themselves are vector operators.

Adjoint representation

Representations where elements (group generators) A_i of algebra $[A_i, A_j] = c_{ikj}A_k$ are defined via structure constants c_{ikj} so that $(A_i)_{jk} = c_{ikj}$ are called *adjoint* or *regular* representations.

Lorentz group

From the representation of the generators (3.28) we can deduce the commutation relations determining the Lie algebra of Lorentz group. In particular, we find that

$$\left[\boldsymbol{M}^{\mu\nu}, \boldsymbol{M}^{\alpha\beta}\right] \;=\; 2i\left\{\boldsymbol{g}^{\mu\beta}\boldsymbol{M}^{\nu\alpha} + \boldsymbol{g}^{\nu\alpha}\boldsymbol{M}^{\mu\beta} - \boldsymbol{g}^{\mu\alpha}\boldsymbol{M}^{\nu\beta} - \boldsymbol{g}^{\nu\beta}M^{\mu\alpha}\right\}. \tag{3.36}$$

For $\boldsymbol{M}^{ij}$, $i, j = 1, 2, 3$ we have

$$\begin{aligned} \left[\boldsymbol{M}^{12}, \boldsymbol{M}^{13}\right] &= 2i\boldsymbol{M}^{23}\,,\\ \left[\boldsymbol{M}^{23}, \boldsymbol{M}^{12}\right] &= -2i\boldsymbol{M}^{31}\,, \end{aligned} \tag{3.37}$$

Dirac's representation (but also in other representations) is based on properties of tensor product $\otimes$ on matrices. Because

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \cdot (\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{A} \cdot \boldsymbol{C}) \otimes (\boldsymbol{B} \cdot \boldsymbol{D}), \tag{3.14}$$

and

$$\gamma^0 = \sigma^3 \otimes \mathbb{1}, \quad \boldsymbol{\gamma} = i\sigma^2 \otimes \boldsymbol{\sigma}, \tag{3.15}$$

and

$$\sigma^i \sigma^j = \delta^{ij} + i\varepsilon^{ijk}\sigma^k, \tag{3.16}$$

we can immediately check that γ matrices in Dirac's representation satisfy the defining Dirac algebra (3.11).

3.1 Lorentz Invariance of Dirac Equation

Recall that a non-relativistic particle with spin has a wave function $\psi_\alpha(x)$ (Weyl spinor) which transforms under rotation $\boldsymbol{R}$ as

$$\psi_\alpha(\boldsymbol{x}) \xrightarrow{\boldsymbol{R}} D_{\alpha\beta}(\boldsymbol{R})\psi_\beta(\boldsymbol{R}^{-1}\boldsymbol{x}). \tag{3.17}$$

In a similar way, the Dirac wave function under a Lorentz transformation $\boldsymbol{L}$ transforms as

$$\psi(x) \xrightarrow{\boldsymbol{L}} \psi_{\boldsymbol{L}}(x) = S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x), \tag{3.18}$$

where $S(\boldsymbol{L})$ is an appropriate representation of the Lorentz group, that acts in the vector space, in which the Dirac wave function takes its values. From this we see that $S(\boldsymbol{L})$ should be a 4×4 matrix.

In order to show that Dirac's equation is Lorentz invariant we need to show that

$$\left(i\gamma^\mu \partial_\mu - m\right)\psi_{\boldsymbol{L}}(x) = 0, \tag{3.19}$$

provided

$$\left(i\gamma^\mu \partial_\mu - m\right)\psi(x) = 0, \tag{3.20}$$

is satisfied. To show this we rewrite the left-hand-side (LHS) of (3.19) as

$$\begin{aligned} \text{LHS} &= \left(i\gamma^\mu \partial_\mu - m\right) S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x) \\ &= S(\boldsymbol{L}) \left[iS^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) - m\right] \psi(\boldsymbol{L}^{-1}x) \\ &= S(\boldsymbol{L}) \left[iS^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) - m\right] \psi(x'), \end{aligned} \tag{3.21}$$

where $x' = \boldsymbol{L}^{-1}x$. If we can find an appropriate matrix $S(\boldsymbol{L})$ such that

$$S^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) = \gamma^\mu \partial'_\mu, \tag{3.22}$$

Since now $x'^\mu = (\boldsymbol{L}^{-1})^\mu{}_\nu x^\nu$ we have that $x^\mu = \boldsymbol{L}^\mu{}_\nu x'^\nu$ and thus $\partial'_\mu = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = \boldsymbol{L}^\nu{}_\mu \frac{\partial}{\partial x^\nu}$.

then

$$\left(i\gamma^\mu \partial_\mu - m\right)\psi_{\boldsymbol{L}}(x) = S(\boldsymbol{L})(i\gamma^\mu \partial'_\mu - m)\psi(x'). \tag{3.23}$$

Since ψ is a Dirac wave function, then

$$\left(i\gamma^\mu \partial'_\mu - m\right)\psi(x') = 0 \quad \Rightarrow \quad \left(i\gamma^\mu \partial_\mu - m\right)\psi_{\boldsymbol{L}}(x) = 0. \tag{3.24}$$

Consequently, also $\psi_{\boldsymbol{L}}$ is Dirac's wave function and Dirac's equation is relativistically invariant provided such $S(\boldsymbol{L})$ exists. To find $S(\boldsymbol{L})$ we have to explore properties of Lorentz group in more detail.

Lorentz group

From the condition (2.6) we know that there exist 2 kinds of Lorentz transformations — proper (those with $\det\{\boldsymbol{L}\} = 1$) and improper (those with $\det\{\boldsymbol{L}\} = -1$). We can expand this classification even further by realizing that

$$1 \;=\; \eta^{00} \;=\; \boldsymbol{L}^0{}_\nu \boldsymbol{L}^0{}_{\nu'} \eta^{\nu\nu'} \;=\; (\boldsymbol{L}^0{}_0)^2 - \sum_{i=1}^{3} (\boldsymbol{L}^0{}_i)^2 \,. \qquad (3.25)$$

Rewriting this we arrive at the condition

$$(\boldsymbol{L}^0{}_0)^2 \;=\; 1 \;+\; \sum_{i=1}^{3} (\boldsymbol{L}^0{}_i)^2 \,. \qquad (3.26)$$

Lorentz transformations for which $\boldsymbol{L}^0{}_0 \geq 1$ are called *orthochronous* transformations, those with $\boldsymbol{L}^0{}_0 \leq -1$ are called *non-orthochronous*. We can not switch between those four types of Lorentz transformations using continuous process. Transitions between these disconnected parts of the Lorentz group can be done only via discrete transformations such as *parity* or *time reversal* (see Chapter 6.1).

Let us recall (Chapter 2.2.) that Lorentz group has 6 independent parameters $\omega_{\mu\nu} = -\omega_{\nu\mu}$ that define it. In terms of these we can construct finite Lorentz transformations as

$$\boldsymbol{L}^\rho{}_\tau = \left(e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}}\right)^\rho{}_\tau \,. \qquad (3.27)$$

Here $M^{\mu\nu}$ are so-called generators of Lorentz transformation, which are fixed by comparison with the infinitesimal transformation (i.e., when $\|\omega_{\mu\nu}\| \ll 1$). In Chapter 2.2 we have found that

$$(\boldsymbol{M}^{\mu\nu})^\rho{}_\tau \;=\; 2i\left(g^{\rho\mu}\delta^\nu_\tau - g^{\rho\nu}\delta^\mu_\tau\right) \,. \qquad (3.28)$$

Rotation group

To proceed, we start with a familiar example — *rotation group*. Elements of a rotation group are defined by

$$\boldsymbol{R}_{ij} \;=\; \left(e^{-i\theta \boldsymbol{n}_k \boldsymbol{J}_k}\right)_{ij} \;=\; \left(e^{-i\omega_k \boldsymbol{J}_k}\right)_{ij} \,, \qquad (3.29)$$

where $\boldsymbol{n}$ is a unit vector along the axis of rotation. Group generators J_i satisfy the usual angular momentum commutation relations

$$\left[\boldsymbol{J}_i, \boldsymbol{J}_j\right] \;=\; i\varepsilon_{ijk}\boldsymbol{J}_k \,. \qquad (3.30)$$

Recall from quantum mechanics that *vector operators* are sets of 3 operators, say $V_i(\boldsymbol{x})$ that transform according to

$$U(\boldsymbol{R})^\dagger V_i(\boldsymbol{x}) U(\boldsymbol{R}) = \sum_j \boldsymbol{R}_{ij} V_j(\boldsymbol{x}) \,. \tag{3.31}$$

Here $U(\boldsymbol{R})$ is a representation of rotation group which acts on state space (e.g. $L^2(\mathbb{R})$) and $\boldsymbol{R}_{ij}$ is a representation of rotation group that acts on the operator indices.

For infinitesimal rotations one obtains

$$(1 + i\omega_k \mathbb{J}_k) \boldsymbol{V}_i (1 - i\omega_k \mathbb{J}_k) = [\delta_{ij} - i\omega_k (\boldsymbol{J}_k)_{ij}] \boldsymbol{V}_j \,. \tag{3.32}$$

Here $\mathbb{J}_k$ are operators of angular momentum, acting on corresponding state space (e.g. $L^2(\mathbb{R})$) and $(\boldsymbol{J}_k)_{ij}$ is a vector representation of angular momentum in 3D, which is known to be

$$\boldsymbol{J}_1 = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{J}_2 = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{J}_3 = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.33}$$

This can succinctly be written as: $(\boldsymbol{J}_j)_{ik} = i\varepsilon_{ijk}$.

Adjoint representation

Representations where elements (group generators) A_i of algebra $[A_i, A_j] = c_{ikj} A_k$ are defined via structure constants c_{ikj} so that $(A_i)_{jk} = c_{ikj}$ are called *adjoint* or *regular* representations.

After algebraic manipulations we obtain from (3.32)

$$i\omega_k \, [\mathbb{J}_k, \boldsymbol{V}_i] = -i\omega_k (\boldsymbol{J}_k)_{ij} \boldsymbol{V}_j \,, \tag{3.34}$$

which then implies that

$$[\mathbb{J}_k, \boldsymbol{V}_i] = -(\boldsymbol{J}_k)_{ij} \boldsymbol{V}_j = -i\varepsilon_{ikj} \boldsymbol{V}_j = i\varepsilon_{kij} \boldsymbol{V}_j \,. \tag{3.35}$$

This relation represents an algebraic condition for vector operators. As a byproduct we see that the generators $\mathbb{J}_k$ themselves are vector operators.

Lorentz group

From the representation of the generators (3.28) we can deduce the commutation relations determining the Lie algebra of Lorentz group. In particular, we find that

$$\left[\boldsymbol{M}^{\mu\nu}, \boldsymbol{M}^{\alpha\beta}\right] = 2i\left(\eta^{\mu\beta}\boldsymbol{M}^{\nu\alpha} + \eta^{\nu\alpha}\boldsymbol{M}^{\mu\beta} - \eta^{\mu\alpha}\boldsymbol{M}^{\nu\beta} - \eta^{\nu\beta}M^{\mu\alpha}\right). \tag{3.36}$$

Sometimes one choses instead of generators $\boldsymbol{M}^{\mu\nu}$ generators $\tilde{\boldsymbol{M}}^{\mu\nu} = 1/2\boldsymbol{M}^{\mu\nu}$. In such a case the factor 2 on the RHS of (3.36) is not present.

For $\boldsymbol{M}^{ij}$, $i, j = 1, 2, 3$ we have

$$\begin{aligned} \left[\boldsymbol{M}^{12}, \boldsymbol{M}^{13}\right] &= 2i\boldsymbol{M}^{23}\,, \\ \left[\boldsymbol{M}^{23}, \boldsymbol{M}^{12}\right] &= -2i\boldsymbol{M}^{31}\,, \end{aligned} \tag{3.37}$$

etc. Defining

Clearly $\boldsymbol{M}^{jk} = \boldsymbol{M}_{jk}$.

$$\boldsymbol{J}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{M}^{jk} \quad \Leftrightarrow \quad \boldsymbol{M}^{jk} = 2\varepsilon^{jki}\boldsymbol{J}_i\,, \tag{3.38}$$

then from (3.37) we have

$$\begin{aligned} [\boldsymbol{J}_3,(-\boldsymbol{J}_2)] &= i\boldsymbol{J}_1 \quad \Leftrightarrow \quad [\boldsymbol{J}_2,\boldsymbol{J}_3] = i\boldsymbol{J}_1\,, \\ [\boldsymbol{J}_1,\boldsymbol{J}_3] &= -i\boldsymbol{J}_2\,. \end{aligned} \tag{3.39}$$

It can be easily checked that we can generally write

$$[\boldsymbol{J}_i,\boldsymbol{J}_j] = i\varepsilon_{ijk}\boldsymbol{J}_k\,. \tag{3.40}$$

From this we see that $\boldsymbol{J}_i$ are generators of rotations, since they close the familiar Lie algebra for group of rotations in 3-dimensional space, i.e., $SO(3)$.

Similarly we can define

$$\boldsymbol{M}_{i0} = 2\boldsymbol{K}_i = -\boldsymbol{M}^{i0} = \boldsymbol{M}^{0i}\,. \tag{3.41}$$

From this we see, for instance, that

$$[\boldsymbol{M}^{01},\boldsymbol{M}^{02}] = -2i\boldsymbol{M}^{12} \quad \Rightarrow \quad [\boldsymbol{K}_1,\boldsymbol{K}_2] = -i\boldsymbol{J}_3\,. \tag{3.42}$$

It can again be checked that one generally has

$$[\boldsymbol{K}_i,\boldsymbol{K}_j] = -i\varepsilon_{ijk}\boldsymbol{J}_k \quad \Leftrightarrow \quad [\boldsymbol{K}^i,\boldsymbol{K}^j] = -i\varepsilon^{ijk}\boldsymbol{J}_k\,. \tag{3.43}$$

Here $\boldsymbol{K}_i$ are the so-called *generators of boosts* in i-th direction. A precise meaning of this terminology will be clarified in Chapter 4.2.

To close the algebra we also need commutators of the type $[\boldsymbol{J},\boldsymbol{K}]$. It can be verified that

$$[\boldsymbol{J}_i,\boldsymbol{K}_j] = i\varepsilon_{ijk}\boldsymbol{K}_k\,. \tag{3.44}$$

So, the Lie algebra (3.36) can be equivalently rewritten as

$$\begin{aligned} [\boldsymbol{J}_i,\boldsymbol{J}_j] &= i\varepsilon_{ijk}\boldsymbol{J}_k\,, \\ [\boldsymbol{J}_i,\boldsymbol{K}_j] &= i\varepsilon_{ijk}\boldsymbol{K}_k\,, \\ [\boldsymbol{K}_i,\boldsymbol{K}_j] &= -i\varepsilon_{ijk}\boldsymbol{J}_k\,. \end{aligned} \tag{3.45}$$

These commutation relations define $so(3,1)$ algebra. Note that the first commutation relation implies that algebra $so(3) \sim su(2)$ is a subalgebra of $su(3,1)$ algebra. On the other hand, boosts do not form a subalgebra of $su(3,1)$, so we need both boosts and rotations to form closed algebra.

Group of Lorentz transformations — terminology

Defining or *fundamental* representation of $SO(3,1)$ group is given by etc. Defining

Clearly $\boldsymbol{M}^{jk} = \boldsymbol{M}_{jk}$.

$$\boldsymbol{J}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{M}^{jk} \quad \Leftrightarrow \quad \boldsymbol{M}^{jk} = 2\varepsilon^{jki}\boldsymbol{J}_i\,, \tag{3.38}$$

then from (3.37) we have

$$\begin{aligned} [\boldsymbol{J}_3, (-\boldsymbol{J}_2)] &= i\boldsymbol{J}_1 \quad \Leftrightarrow \quad [\boldsymbol{J}_2, \boldsymbol{J}_3] = i\boldsymbol{J}_1\,, \\ [\boldsymbol{J}_1, \boldsymbol{J}_3] &= -i\boldsymbol{J}_2\,. \end{aligned} \tag{3.39}$$

It can be easily checked that we can generally write

$$[\boldsymbol{J}_i, \boldsymbol{J}_j] = i\varepsilon_{ijk}\boldsymbol{J}_k\,. \tag{3.40}$$

From this we see that $\boldsymbol{J}_i$ are generators of rotations, since they close the familiar Lie algebra for group of rotations in 3-dimensional space, i.e., $SO(3)$.

Similarly we can define

$$\boldsymbol{M}_{i0} = 2\boldsymbol{K}_i = -\boldsymbol{M}^{i0} = \boldsymbol{M}^{0i}\,. \tag{3.41}$$

From this we see, for instance, that

$$[\boldsymbol{M}^{01}, \boldsymbol{M}^{02}] = -2i\boldsymbol{M}^{12} \quad \Rightarrow \quad [\boldsymbol{K}_1, \boldsymbol{K}_2] = -i\boldsymbol{J}_3\,. \tag{3.42}$$

It can again be checked that one generally has

$$[\boldsymbol{K}_i, \boldsymbol{K}_j] = -i\varepsilon_{ijk}\boldsymbol{J}_k \quad \Leftrightarrow \quad [\boldsymbol{K}^i, \boldsymbol{K}^j] = -i\varepsilon^{ijk}\boldsymbol{J}_k\,. \tag{3.43}$$

Here $\boldsymbol{K}_i$ are the so-called *generators of boosts* in i-th direction. A precise meaning of this terminology will be clarified in Chapter 4.2.

To close the algebra we also need commutators of the type $[\boldsymbol{J}, \boldsymbol{K}]$. It can be verified that

$$[\boldsymbol{J}_i, \boldsymbol{K}_j] = i\varepsilon_{ijk}\boldsymbol{K}_k\,. \tag{3.44}$$

So, the Lie algebra (3.36) can be equivalently rewritten as

$$\begin{aligned} [\boldsymbol{J}_i, \boldsymbol{J}_j] &= i\varepsilon_{ijk}\boldsymbol{J}_k\,, \\ [\boldsymbol{J}_i, \boldsymbol{K}_j] &= i\varepsilon_{ijk}\boldsymbol{K}_k\,, \\ [\boldsymbol{K}_i, \boldsymbol{K}_j] &= -i\varepsilon_{ijk}\boldsymbol{J}_k\,. \end{aligned} \tag{3.45}$$

These commutation relations define $so(3,1)$ algebra. Note that the first commutation relation implies that algebra $so(3) \sim su(2)$ is a subalgebra of $su(3,1)$ algebra. On the other hand, boosts do not form a subalgebra of $su(3,1)$, so we need both boosts and rotations to form closed algebra.

Group of Lorentz transformations — terminology

Defining or *fundamental* representation of $SO(3,1)$ group is given by matrices satisfying the defining relation

$$x^\mathsf{T} x = (\boldsymbol{L}x)^\mathsf{T}(\boldsymbol{L}x) = x^\mathsf{T}\boldsymbol{L}^\mathsf{T}\boldsymbol{L}x = \text{invariant}.$$

From this we can explicitly write that $\boldsymbol{L}^\mathsf{T}\boldsymbol{L} = \mathbb{1}$ (i.e., $\boldsymbol{L}_\mu{}^\nu \boldsymbol{L}^\mu{}_\alpha = \delta^\nu{}_\alpha$) and hence $\boldsymbol{L}^\mathsf{T} = \boldsymbol{L}^{-1}$. From this we can see that $\boldsymbol{L}$ are orthogonal matrices preserving the spacetime distance $x_0^2 - \sum_{i=1}^3 x_i^2$. Those matrices are of $O(3,1)$ type and since $\det \boldsymbol{L} = 1$ for proper transformations, we stress this extra fact in "S" in the $SO(3,1)$ group name.

Commutation relations can be diagonalized via transformation

$$\boldsymbol{N}_i = \frac{1}{2}(\boldsymbol{J}_i + i\boldsymbol{K}_i), \quad \boldsymbol{N}_i^\dagger = \frac{1}{2}(\boldsymbol{J}_i - i\boldsymbol{K}_i). \tag{3.46}$$

This diagonalization does not provide Hermitian generators!

From this follows that

$$[\boldsymbol{N}_i, \boldsymbol{N}_j^\dagger] = 0, \quad [\boldsymbol{N}_i, \boldsymbol{N}_j] = i\varepsilon_{ijk}\boldsymbol{N}_k, \quad [\boldsymbol{N}_i^\dagger, \boldsymbol{N}_j^\dagger] = i\varepsilon_{ijk}\boldsymbol{N}_k^\dagger. \tag{3.47}$$

The relation $[\boldsymbol{N}_i, \boldsymbol{N}_j] = i\varepsilon_{ijk}\boldsymbol{N}_k$ (and the same for $\boldsymbol{N}^\dagger$) closes the $su(2)$ algebra. Hence, we can view $so(3,1)$ algebra as being isomorphic to $su(2) \oplus su(2) \sim sl(2,\mathbb{C})$.

The Lorentz group is thus isomorphic to the group $SU(2) \otimes SU(2)$ or more explicitly, $SU(2)_{\boldsymbol{N}} \otimes SU(2)_{\boldsymbol{N}^\dagger}$.

Finite dimensional representations of the Lorentz group Lie algebra are easily obtained from those of $su(2)$. As in the theory of angular momenta we can find for $su(2)$ algebra the *Casimir operator* that quantifies the algebra (and hence group) representation. For $SU(2)_{\boldsymbol{N}}$ and $SU(2)_{\boldsymbol{N}^\dagger}$ the latter can be defined as $\sum_{i=1}^3 \boldsymbol{N}_i^2 = n(n+1)$ and $\sum_{i=1}^3 \boldsymbol{N}_i^{\dagger 2} = m(m+1)$. Constants n and m describe nothing but the size of the angular momenta (or spin). The representation of $so(3,1)$ can then be denoted with the pair (n,m). Note, in particular, that the transformation with respect to spatial parity is given by

$$\underbrace{\boldsymbol{J}_i \xrightarrow{P} \boldsymbol{J}_i}_{\text{Pseudovector}} \quad \text{and} \quad \underbrace{\boldsymbol{K}_i \xrightarrow{P} -\boldsymbol{K}_i}_{\text{Vector}}, \tag{3.48}$$

and hence

$$\boldsymbol{N}_i \overset{P}{\leftrightarrow} \boldsymbol{N}_i^\dagger \quad \Rightarrow \quad (n,m) \overset{P}{\leftrightarrow} (m,n). \tag{3.49}$$

Generally representations of Lorentz group need not to be parity invariant (for example parity is violated in weak interactions). In addition, since $\boldsymbol{J}_i = \boldsymbol{N}_i + \boldsymbol{N}_i^\dagger$, we can identify the spin of a given representation $(n+m)$, e.g., spin-0 particle is described (i.e, its wave function is) in the representation $(0,0)$, spin-1/2 particle can be in (parity non-invariant) representations $(1/2,0)$ or $(0,1/2)$ while spin-1 particle can be in representations $\underbrace{(1/2,1/2)}_{\text{Parity invariant}}$ or $\underbrace{(1,0) \text{ or } (0,1)}_{\text{Parity non-invariant}}$.

Parity violation was experimentally observed in week interactions (namely in the pion decay) in 1956 by the Chinese American physicist Chien-Shiung Wu.

Spin 1/2 representations

Representation $(1/2,0)$ is known as a *left-handed spinor* (handedness is a convention), while the representation $(0,1/2)$ is known as a *right-handed spinor*. Corresponding wave functions are typically 2-component objects that are called *Weyl spinors*.

When parity is relevant, one considers the linear combination $(0, 1/2) \oplus (1/2, 0)$, which yields 4-component wave function known as a *Dirac spinor* or *bispinor*.

Lorentz Invariance of Dirac Equation Continued

Our goal is to show that the Dirac equation transforms covariantly under Lorentz transformations. So, equivalently, we require that the Lorentz transformed wave function

$$\psi(x) \xrightarrow{\boldsymbol{L}} \psi_{\boldsymbol{L}}(x) \ = \ S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x)\,, \tag{3.50}$$

should satisfy Dirac's equation provided $\psi(\boldsymbol{x})$ does. Here $S(\boldsymbol{L})$ is a representation of the Lorentz group that acts only on the indices of wave function ψ. As such, it must satisfy the group composition law $S(\boldsymbol{L}_1)S(\boldsymbol{L}_2) = S(\boldsymbol{L}_1\boldsymbol{L}_2)$. On the other hand, Dirac's equation transforms covariantly provided that

So, in particular $S(\boldsymbol{L})S(\boldsymbol{L}^{-1}) = S(\mathbb{1}) = \mathbb{1}$ and hence $S(\boldsymbol{L}^{-1}) = S^{-1}(\boldsymbol{L})$.

$$S^{-1}(\boldsymbol{L})\gamma^{\mu}S(\boldsymbol{L}) \ = \ \boldsymbol{L}^{\mu}{}_{\nu}\gamma^{\nu}\,. \tag{3.51}$$

It can be easily checked that this condition is compatible with the group composition law and so $S(L)$ is a representation of the Lorentz group. Question now stand how look the corresponding generators and to what representation they belong to. This information can be obtained by considering an infinitisemal Lorentz transformation. In particular, in such a case we can write

Again we are dealing with connected part of the Lorentz group that contains the unit element.

$$\boldsymbol{L} \ = \ \mathbb{1} \ - \ \frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}\,, \tag{3.52}$$

and correspondingly

$$S(\boldsymbol{L}) \ = \ \mathbb{1} \ - \ \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}\,. \tag{3.53}$$

Here $\boldsymbol{\sigma}^{\mu\nu}$ are the generators of the Lorentz group in the representation that is appropriate to Dirac space (space of Dirac's wave functions — bispinors). Analogously

$$S^{-1}(\boldsymbol{L}) \ = \ S(\boldsymbol{L}^{-1}) \ = \ \mathbb{1} \ + \ \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}\,. \tag{3.54}$$

Inserting this into (3.51) we obtain that

$$\left(\mathbb{1} + \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\boldsymbol{\omega}_{\mu\nu}\right)\gamma^{\rho}\left(\mathbb{1} - \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\boldsymbol{\omega}_{\mu\nu}\right) \ = \ \left(\delta^{\rho}_{\tau} \ + \ \omega^{\rho}{}_{\tau}\right)\gamma^{\tau}\,, \tag{3.55}$$

and thus

$$\frac{i}{4}[\boldsymbol{\sigma}^{\mu\nu}\boldsymbol{\omega}_{\mu\nu}, \gamma^{\rho}] \ = \ \omega^{\rho}{}_{\tau}\gamma^{\tau}\,. \tag{3.56}$$

By writing

$$\omega^{\rho}{}_{\tau}\gamma^{\tau} = \omega_{\mu\nu}\eta^{\mu\rho}\gamma^{\nu} = \frac{1}{2}\omega_{\mu\nu}\left(\eta^{\mu\rho}\gamma^{\nu} - \eta^{\nu\rho}\gamma^{\mu}\right) , \tag{3.57}$$

Eq. (3.56) can be rewritten as

$$\frac{i}{4}[\boldsymbol{\sigma}^{\mu\nu}, \gamma^{\rho}] = \frac{1}{2}\left(\eta^{\mu\rho}\gamma^{\nu} - \eta^{\nu\rho}\gamma^{\mu}\right) . \tag{3.58}$$

This should determine the Dirac representation generators $\boldsymbol{\sigma}^{\mu\nu}$.

This condition is satisfied if $\boldsymbol{\sigma}^{\mu\nu} = \frac{i}{2}[\gamma^{\mu}, \gamma^{\nu}]$. Indeed, by using the well known identity $[\boldsymbol{AB}, \boldsymbol{C}] = \boldsymbol{A}\{\boldsymbol{B}, \boldsymbol{C}\} - \{\boldsymbol{A}, \boldsymbol{C}\}\boldsymbol{B}$ we can write

$$\begin{aligned}
&\frac{i}{4}\left(\frac{i}{2}[\gamma^{\mu}\gamma^{\nu}, \gamma^{\rho}] - \frac{i}{2}[\gamma^{\nu}\gamma^{\mu}, \gamma^{\rho}]\right) \\
&\quad = -\frac{1}{8}\left(\gamma^{\mu}\{\gamma^{\nu}, \gamma^{\rho}\} - \{\gamma^{\mu}, \gamma^{\rho}\}\gamma^{\nu} - \gamma^{\nu}\{\gamma^{\mu}, \gamma^{\rho}\} + \{\gamma^{\nu}, \gamma^{\rho}\}\gamma^{\mu}\right) \\
&\quad = -\frac{1}{8}\left(2\gamma^{\mu}\eta^{\nu\rho} - 2\eta^{\mu\rho}\gamma^{\nu} - 2\gamma^{\nu}\eta^{\mu\rho} + 2\eta^{\nu\rho}\gamma^{\mu}\right) \\
&\quad = \frac{1}{2}\left(\gamma^{\nu}\eta^{\mu\rho} - \gamma^{\mu}\eta^{\nu\rho}\right) .
\end{aligned} \tag{3.59}$$

It can also be checked that generators $\boldsymbol{\sigma}^{\mu\nu} = \frac{i}{2}[\gamma^{\mu}, \gamma^{\nu}]$ satisfy the correct Lorentz group algebra

$$[\boldsymbol{\sigma}^{\mu\nu}, \boldsymbol{\sigma}^{\alpha\beta}] = 2i\left(\eta^{\mu\beta}\boldsymbol{\sigma}^{\nu\alpha} + \eta^{\nu\alpha}\boldsymbol{\sigma}^{\mu\beta} - \eta^{\mu\alpha}\boldsymbol{\sigma}^{\nu\beta} - \eta^{\nu\beta}\boldsymbol{\sigma}^{\mu\alpha}\right) . \tag{3.60}$$

Finally we can write that the finite Lorentz transformations $S(\boldsymbol{L})$ have the form

$$S(\boldsymbol{L}) = e^{-\frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}} \quad \text{with} \quad \boldsymbol{L} = e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}} . \tag{3.61}$$

This closes our proof of the Lorentz covariance of the Dirac equation.

Exercises: Dirac equation — basics

Dirac algebra of γ-matrices

Dirac matrices γ^{μ} are 4×4 matrices satisfying the anti-commutation relations

$$\{\gamma^{\mu}, \gamma^{\nu}\} = 2g^{\mu\nu}\mathbb{1} .$$

In their standard (or Dirac) representation they have the form

$$\gamma^{0} = \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{pmatrix} = \sigma^{3}\otimes\mathbb{1}, \qquad \gamma^{j} = \begin{pmatrix} \mathbf{0} & \sigma^{j} \\ -\sigma^{j} & \mathbf{0} \end{pmatrix} = i\sigma^{2}\otimes\sigma^{j} .$$

The tensor (or Kronecker) product $\mathbb{A}\otimes\mathbb{B}$, of an $n\times n'$ matrix $\mathbb{A}$, and an $m\times m'$ matrix $\mathbb{B}$, is an $(nm)\times(n'm')$ matrix defined by

$$\mathbb{A}\otimes\mathbb{B} = \begin{pmatrix} A_{11}\mathbb{B} & \dots & A_{1n'}\mathbb{B} \\ \vdots & & \vdots \\ A_{n1}\mathbb{B} & \dots & A_{nn'}\mathbb{B} \end{pmatrix} .$$

The tensor product has following properties

$$
\begin{aligned}
&1) \quad (\alpha \mathbb{A}_1 + \mathbb{A}_2) \otimes \mathbb{B} = \alpha \mathbb{A}_1 \otimes \mathbb{B} + \mathbb{A}_2 \otimes \mathbb{B} \\
&2) \quad (\mathbb{A} \otimes \mathbb{B}) \otimes \mathbb{C} = \mathbb{A} \otimes (\mathbb{B} \otimes \mathbb{C}) \\
&3) \quad \text{(in general)} \quad \mathbb{A} \otimes \mathbb{B} \neq \mathbb{B} \otimes \mathbb{A} \\
&4) \quad (\mathbb{A} \otimes \mathbb{B})(\mathbb{C} \otimes \mathbb{D}) = (\mathbb{A}\mathbb{C}) \otimes (\mathbb{B}\mathbb{D}) \\
&5) \quad (\mathbb{A} \otimes \mathbb{B})^{-1} = \mathbb{A}^{-1} \otimes \mathbb{B}^{-1} \\
&6) \quad (\mathbb{A} \otimes \mathbb{B})^{\dagger} = \mathbb{A}^{\dagger} \otimes \mathbb{B}^{\dagger} .
\end{aligned}
$$

Exercise 3.1 Verify that the matrices γ^μ in Dirac representation satisfy Clifford algebra (3.11).

Identities in the following exercises can be proven using the defining relations $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}$, i.e., they hold in any representation of Dirac matrices.

Exercise 3.2 Show that:

$$
\begin{aligned}
&1) \quad \gamma^\mu \gamma_\mu = 4\mathbb{1} \\
&2) \quad \gamma^\mu \gamma^\nu \gamma_\mu = -2\gamma^\nu \\
&3) \quad \gamma^\mu \gamma^\nu \gamma^\rho \gamma_\mu = 4g^{\nu\rho}\ \mathbb{1} .
\end{aligned}
$$

One defines the matrix $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, which in the standard representation reads

$$
\gamma^5 = \begin{pmatrix} \mathbf{0} & \mathbb{1} \\ \mathbb{1} & \mathbf{0} \end{pmatrix} = \sigma^1 \otimes \mathbb{1} .
$$

Exercise 3.3 Show that γ^5 satisfies: a) $(\gamma^5)^2 = \mathbb{1}$, and b) $\gamma^\mu\gamma^5 = -\gamma^5\gamma^\mu$ (anti-commutes with all γ-matrices).

Exercise 3.4 Verify the following 'trace' identities:

$$
\begin{aligned}
&1) \quad \mathrm{Tr}(\gamma^\mu \gamma^\nu) = 4g^{\mu\nu} \\
&2) \quad \mathrm{Tr}(\gamma^\mu) = 0 \\
&3) \quad \mathrm{Tr}(\gamma^{\mu_1} \ldots \gamma^{\mu_{2n+1}}) = 0 \quad (\forall n) \\
&4) \quad \mathrm{Tr}(\gamma^5) = 0 \\
&5) \quad \mathrm{Tr}(\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma) = 4(g^{\mu\nu} g^{\rho\sigma} - g^{\mu\rho} g^{\nu\sigma} + g^{\mu\sigma} g^{\nu\rho}) \\
&6) \quad \text{if} \quad \not{p} = \gamma^\mu p_\mu \quad \text{then} \quad \mathrm{Tr}(\not{p}\not{q}) = 4p^\mu q_\mu .
\end{aligned}
$$

Exercise 3.5 Show that

$$
\det(\not{p} - m\,\mathbb{1}) = \det(\not{p} + m\mathbb{1}) ,
$$

and deduce that matrices $\not{p} \pm m\mathbb{1}$ are singular for four-momenta satisfying $p_\mu p^\mu = m^2$.

Note that if some matrices γ^μ (e.g., the Dirac representation matrices) satisfy Eq. (3.1), then for a unitary matrix $\mathbb{U}$

$$\{\mathbb{U}\gamma^\mu\mathbb{U}^\dagger, \mathbb{U}\gamma^\nu\mathbb{U}^\dagger\} = \mathbb{U}\{\gamma^\mu, \gamma^\nu\}\mathbb{U}^\dagger = \mathbb{U}\,2\eta^{\mu\nu}\,\mathbb{U}^\dagger = 2\eta^{\mu\nu}.$$

That is, matrices $\tilde{\gamma}^\mu = \mathbb{U}\gamma^\mu\mathbb{U}^\dagger$ constitute another representation of the Dirac algebra.

Problem 3.1 The Weyl (or chiral) representation of γ-matrices, γ^μ_W, is obtained from the standard Dirac representation γ^μ_D, via the transformation

$$\gamma^\mu_W = U\gamma^\mu_D U^\dagger, \qquad U = \frac{1}{\sqrt{2}}(\mathbb{1} + \gamma^5_D\gamma^0_D).$$

Determine the matrices $\gamma^0_W, \gamma^1_W, \gamma^2_W, \gamma^3_W$, and γ^5_W.

Lorentz group and spin representation

The Lorentz group

$$O(1,3) = \{\boldsymbol{L} \in \mathbb{R}^{4,4} \mid \eta_{\mu\nu}\boldsymbol{L}^\mu{}_\rho\boldsymbol{L}^\nu{}_\sigma = \eta_{\rho\sigma}\},$$

has 4 connected components, which differ by $\det\boldsymbol{L} = \pm1$, and $\operatorname{sign}(L^0{}_0) = \pm1$.

The component with $\det\boldsymbol{L} = +1$ and $\operatorname{sign}(L^0{}_0) = +1$ (*proper orthochronous* transformations) is denoted $SO^+(1,3)$, and its elements can be cast in the exponential form, Eq. (2.3). One can 'move' between the various connected components with the help of *parity* (or spatial inversion) $\boldsymbol{P}$, and *time reversal* $\boldsymbol{T}$:

$$\boldsymbol{P} = \operatorname{diag}(1,-1,-1,-1) \qquad \boldsymbol{T} = \operatorname{diag}(-1,1,1,1).$$

Denote

$$\boldsymbol{J}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{M}^{jk}, \qquad \boldsymbol{K}_i = \frac{1}{2}\boldsymbol{M}^{0i},$$

the rotation generators, and the boost generators, respectively. It holds that

$$[\boldsymbol{J}_i, \boldsymbol{J}_j] = i\varepsilon_{ijk}\boldsymbol{J}_k,$$

$$[\boldsymbol{K}_i, \boldsymbol{K}_j] = -i\varepsilon_{ijk}\boldsymbol{J}_k,$$

$$[\boldsymbol{J}_i, \boldsymbol{K}_j] = i\varepsilon_{ijk}\boldsymbol{K}_k.$$

Exercise 3.6 In this exercise we split the Lorentz algebra $so(1,3)$ is split into two independent algebras $su(2)$. Introduce

$$\boldsymbol{N}_i = \frac{1}{2}(\boldsymbol{J}_i + i\boldsymbol{K}_i), \qquad \boldsymbol{N}^\dagger_i = \frac{1}{2}(\boldsymbol{J}_i - i\boldsymbol{K}_i),$$

and show that

$$[\boldsymbol{N}_i, \boldsymbol{N}_j] = i\varepsilon_{ijk}\boldsymbol{N}_k,$$

$$[\boldsymbol{N}^\dagger_i, \boldsymbol{N}^\dagger_j] = i\varepsilon_{ijk}\boldsymbol{N}^\dagger_k,$$

$$[\boldsymbol{N}_i, \boldsymbol{N}^\dagger_j] = 0.$$

Exercise 3.7 Let

$$[\mathbb{A}_a, \mathbb{A}_b] = c_{abc}\mathbb{A}_c, \qquad a,b,c = 1,\ldots,n,$$

be the commutation relations of a Lie algebra of matrices $\{\mathbb{A}_1,\ldots,\mathbb{A}_n\}$. The *adjoint representation* (of this Lie algebra) is formed by the matrices

$$(\mathbb{C}_a)_{bc} = -c_{abc}.$$

Verify that the matrices $\mathbb{C}_a$ obey the same commutation relations as the matrices $\mathbb{A}_a$. Determine the adjoint representation of the algebra a) $so(3)$, b) $su(2)$.

[**Hint:** Use the Jacobi identity $[[\mathbb{A},\mathbb{B}],\mathbb{C}] + [[\mathbb{C},\mathbb{A}],\mathbb{B}] + [[\mathbb{B},\mathbb{C}],\mathbb{A}] = 0$.]

Spin representation of Lorentz algebra

By the requirement of invariance of the Dirac equation under Lorentz transformations $\boldsymbol{L}$, when the Dirac wave-function transforms as

$$\psi'(x') = S(\boldsymbol{L})\psi(x)\,,$$

we obtain the condition

$$S(L)^{-1}\gamma^\mu S(L) = L^\mu_{\ \nu}\gamma^\nu\,.$$

Its infinitesimal form, for

$$L \approx \mathbb{1} - \frac{i}{4}\omega_{\mu\nu}M^{\mu\nu}\,, \qquad S(\boldsymbol{L}) \approx \mathbf{1} - \frac{i}{4}\omega_{\mu\nu}\boldsymbol{\sigma}^{\mu\nu}\,,$$

reads

$$\frac{i}{4}[\sigma^{\mu\nu},\gamma^\rho] = \frac{1}{2}(\eta^{\mu\rho}\gamma^\nu - \eta^{\nu\rho}\gamma^\mu)\,. \tag{3.62}$$

Exercise 3.8 Show that

$$\boldsymbol{\sigma}^{\mu\nu} = \frac{i}{2}[\gamma^\mu,\gamma^\nu]\,.$$

satisfies Eq. (3.62).

Exercise 3.9 Show that $\boldsymbol{\sigma}^{\mu\nu} = \frac{i}{2}[\gamma^\mu,\gamma^\nu]$ satisfy the commutation relations of the Lorentz algebra, i.e.

$$[\boldsymbol{\sigma}^{\mu\nu},\boldsymbol{\sigma}^{\alpha\beta}] = -2i(\eta^{\mu\alpha}\boldsymbol{\sigma}^{\nu\beta} - \eta^{\mu\beta}\boldsymbol{\sigma}^{\nu\alpha} + \eta^{\nu\beta}\boldsymbol{\sigma}^{\mu\alpha} - \eta^{\nu\alpha}\boldsymbol{\sigma}^{\mu\beta})\,.$$

Dirac Equation — technical developments I

4

4.1 Dirac Bilinears

Dirac bilinears are relevant for construction of observables in quantum field theory but they will also help us to construct other quantities of interest, e.g., probability current.

First, general Dirac's wave function has form

$$\psi(x) \;=\; \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}, \quad \psi^\dagger(x) \;=\; (\psi_1^*(x), \psi_2^*(x), \psi_3^*(x), \psi_4^*(x)) \,. \tag{4.1}$$

We define *spinor adjoint* to ψ as

$$\overline{\psi}(x) \;=\; \psi^\dagger(x)\gamma^0 \,. \tag{4.2}$$

Let us first see how $\overline{\psi}(x)$ transforms under the Lorentz transformation. To this end we use two simple facts, namely

$$\begin{aligned} \psi(x) \;&\xrightarrow{\boldsymbol{L}}\; \psi_{\boldsymbol{L}}(x) \;=\; S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x)\,, \\ \psi^\dagger(x) \;&\xrightarrow{\boldsymbol{L}}\; \psi^\dagger_{\boldsymbol{L}}(x) \;=\; \psi^\dagger(\boldsymbol{L}^{-1}x)S^\dagger(\boldsymbol{L})\,. \end{aligned} \tag{4.3}$$

Now, we can multiply from right the second equation by γ^0 , i.e.

$$\overline{\psi}(x) \;=\; \psi^\dagger(x)\gamma^0 \xrightarrow{\boldsymbol{L}} \psi^\dagger_{\boldsymbol{L}}(x)\gamma^0 \;=\; \psi^\dagger(\boldsymbol{L}^{-1}x)S^\dagger(\boldsymbol{L})\gamma^0 \,. \tag{4.4}$$

We now use the fact that

$$S(\boldsymbol{L}) \;=\; e^{-\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}} \;\;\Rightarrow\;\; S^\dagger(\boldsymbol{L}) \;=\; e^{\frac{i}{4}(\sigma^{\mu\nu})^\dagger\omega_{\mu\nu}}\,, \tag{4.5}$$

where $(\sigma^{\mu\nu})^\dagger$ is

$$(\sigma^{\mu\nu})^\dagger \;=\; \left(\frac{i}{2}[\gamma^\mu, \gamma^\nu]\right)^\dagger \;=\; -\frac{i}{2}[\gamma^{\nu\dagger}, \gamma^{\mu\dagger}] \;=\; \frac{i}{2}[\gamma^{\mu\dagger}, \gamma^{\nu\dagger}]\,. \tag{4.6}$$

How do we "rotate" $\gamma^{\nu\dagger}$ to γ^ν? We already know that be by writing Dirac equation in Schrödinger like form, i.e.

$$i\partial_t\psi \;=\; -i\gamma^0\gamma^i\partial_i\psi + \gamma^0 m\psi \;\equiv\; \boldsymbol{H}_D\psi\,, \tag{4.7}$$

($\boldsymbol{H}_D$ is a Dirac Hamiltonian) we get from the presumed hermiticity of $\boldsymbol{H}_D$ that

$$\begin{aligned} \gamma^0 \;&=\; \gamma^{0\dagger}\,, \\ \gamma^{i\dagger} \;&=\; -\gamma^i \;=\; -\gamma^0\gamma^0\gamma^i \;=\; \gamma^0\gamma^i\gamma^0\,. \end{aligned} \tag{4.8}$$

4 Dirac Equation — technical developments I

4.1 Dirac Bilinears

Dirac bilinears are relevant for construction of observables in quantum field theory but they will also help us to construct other quantities of interest, e.g., probability current.

First, general Dirac's wave function has form

$$\psi(x) \;=\; \begin{pmatrix}\psi_1(x)\\ \psi_2(x)\\ \psi_3(x)\\ \psi_4(x)\end{pmatrix}, \quad \psi^\dagger(x) \;=\; (\psi_1^*(x), \psi_2^*(x), \psi_3^*(x), \psi_4^*(x))\,. \tag{4.1}$$

We define *spinor adjoint* to ψ as

$$\overline{\psi}(x) \;=\; \psi^\dagger(x)\gamma^0\,. \tag{4.2}$$

Let us first see how $\overline{\psi}(x)$ transforms under the Lorentz transformation. To this end we use two simple facts, namely

$$\begin{aligned}\psi(x) &\xrightarrow{\boldsymbol{L}} \psi_{\boldsymbol{L}}(x) \;=\; S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x)\,,\\ \psi^\dagger(x) &\xrightarrow{\boldsymbol{L}} \psi^\dagger_{\boldsymbol{L}}(x) \;=\; \psi^\dagger(\boldsymbol{L}^{-1}x)S^\dagger(\boldsymbol{L})\,.\end{aligned} \tag{4.3}$$

Now, we can multiply from right the second equation by γ^0 , i.e.

$$\overline{\psi}(x) \;=\; \psi^\dagger(x)\gamma^0 \xrightarrow{\boldsymbol{L}} \psi^\dagger_{\boldsymbol{L}}(x)\gamma^0 \;=\; \psi^\dagger(\boldsymbol{L}^{-1}x)S^\dagger(\boldsymbol{L})\gamma^0\,. \tag{4.4}$$

We now use the fact that

$$S(\boldsymbol{L}) \;=\; e^{-\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}} \quad\Rightarrow\quad S^\dagger(\boldsymbol{L}) \;=\; e^{\frac{i}{4}(\sigma^{\mu\nu})^\dagger\omega_{\mu\nu}}\,, \tag{4.5}$$

where $(\sigma^{\mu\nu})^\dagger$ is

$$(\sigma^{\mu\nu})^\dagger \;=\; \left(\frac{i}{2}[\gamma^\mu,\gamma^\nu]\right)^\dagger \;=\; -\frac{i}{2}[\gamma^{\nu\dagger},\gamma^{\mu\dagger}] \;=\; \frac{i}{2}[\gamma^{\mu\dagger},\gamma^{\nu\dagger}]\,. \tag{4.6}$$

How do we "rotate" $\gamma^{\nu\dagger}$ to γ^ν? We already know that be by writing Dirac equation in Schrödinger like form, i.e.

$$i\partial_t\psi \;=\; -i\gamma^0\gamma^i\partial_i\psi + \gamma^0 m\psi \;\equiv\; H_D\psi\,, \tag{4.7}$$

(H_D is a Dirac Hamiltonian) we get from the presumed hermiticity of H_D that

$$\begin{aligned}\gamma^0 &\;=\; \gamma^{0\dagger}\,,\\ \gamma^{i\dagger} &\;=\; -\gamma^i \;=\; -\gamma^0\gamma^0\gamma^i \;=\; \gamma^0\gamma^i\gamma^0\,.\end{aligned} \tag{4.8}$$

Those identities are valid irrespective of chosen representation and a particular example of γ-matrices that satisfy these conditions is provided by the oldest representation of Dirac matrices (that is due to Dirac himself), i.e.

$$\begin{aligned} \gamma^0 &= \sigma^3 \otimes \mathbb{1} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \\ \gamma^i &= i\sigma^2 \otimes \sigma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \end{aligned} \tag{4.9}$$

Some properties of γ-matrices

Here we summaries some important properties of γ-matrices.

- $\gamma^0 = (\gamma^0)^{-1}$, $(\gamma^0)^2 = \mathbb{1}$ and $(\gamma^0)^\dagger = \gamma^0$
- $(\gamma^i)^\dagger = -\gamma^i$
- $\gamma^0(\gamma^0)^\dagger\gamma^0 = \gamma^0\gamma^0\gamma^0 = \gamma^0$
- $\gamma^0(\gamma^i)^\dagger\gamma^0 = -\gamma^0\gamma^i\gamma^0 = \gamma^i$
- $\gamma^0(\gamma^\mu)^\dagger\gamma^0 = \gamma^\mu$
- $\gamma^0(\boldsymbol{\sigma}^{\mu\nu})^\dagger\gamma^0 = \gamma^0\frac{1}{2}[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger]\gamma^0 = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) = \boldsymbol{\sigma}^{\mu\nu}$

Now consider

$$\gamma^0 S(\boldsymbol{L})^\dagger\gamma^0 = e^{\frac{i}{4}\omega_{\mu\nu}\gamma^0(\boldsymbol{\sigma}^{\mu\nu})^\dagger\gamma^0} = e^{\frac{i}{4}\omega_{\mu\nu}\boldsymbol{\sigma}^{\mu\nu}} = (S(\boldsymbol{L}))^{-1}. \tag{4.10}$$

So $\gamma^0 S(\boldsymbol{L})^\dagger\gamma^0 = S^{-1}(\boldsymbol{L}) = S(\boldsymbol{L}^{-1})$. From this it follows that

$$\begin{aligned} \overline{\psi}(x) = \psi^\dagger(x)\gamma^0 &\xrightarrow{\boldsymbol{L}} \psi^\dagger(\boldsymbol{L}^{-1}x)S^\dagger(\boldsymbol{L})\gamma^0 \\ &= \psi^\dagger(\boldsymbol{L}^{-1}x)\gamma^0\gamma^0 S(\boldsymbol{L})^\dagger\gamma^0 \\ &= \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L}). \end{aligned} \tag{4.11}$$

So finally we have the following transformation rules

$$\begin{aligned} \psi(x) \xrightarrow{\boldsymbol{L}} \psi_{\boldsymbol{L}}(x) &= S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x), \\ \overline{\psi}(x) \xrightarrow{\boldsymbol{L}} \overline{\psi}_{\boldsymbol{L}}(x) &= \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L}). \end{aligned} \tag{4.12}$$

These relations are key in forming bilinears.

Classification of bilinears

First, we begin with *scalar bilinears*

$$\begin{aligned}\overline{\psi}(x)\psi(x) \xrightarrow{\boldsymbol{L}} \overline{\psi}_{\boldsymbol{L}}(x)\psi_{\boldsymbol{L}}(x) &= \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L})S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x) \\ &= \overline{\psi}(\boldsymbol{L}^{-1}x)\psi(\boldsymbol{L}^{-1}x)\,. \end{aligned} \tag{4.13}$$

By defining $s(x) \equiv \overline{\psi}(x)\psi(x)$ we can rewrite (4.13) as

$$s(x) \xrightarrow{\boldsymbol{L}} s_{\boldsymbol{L}}(x) \;=\; s(\boldsymbol{L}^{-1}x)\,. \tag{4.14}$$

This is transformation property of *scalar fields*. Next, the *vector fields* (or *vector currents*) are defined as

$$J^{\mu}(x) \;=\; \overline{\psi}(x)\gamma^{\mu}\psi(x)\,. \tag{4.15}$$

Transforming this current we obtain

$$\begin{aligned} J^{\mu}(x) \xrightarrow{\boldsymbol{L}} J^{\mu}_{\boldsymbol{L}}(x) &= \overline{\psi}_{\boldsymbol{L}}(x)\gamma^{\mu}\psi_{\boldsymbol{L}}(x) \\ &= \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L})\gamma^{\mu}S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x)\,. \end{aligned} \tag{4.16}$$

By recalling that the following relation for γ-matrices holds

$$S^{-1}(\boldsymbol{L})\gamma^{\mu}S(\boldsymbol{L}) \;=\; \boldsymbol{L}^{\mu}{}_{\nu}\gamma^{\nu}\,, \tag{4.17}$$

we get

$$\begin{aligned} J^{\mu}_{\boldsymbol{L}}(x) &= \overline{\psi}(\boldsymbol{L}^{-1}x)\boldsymbol{L}^{\mu}{}_{\nu}\gamma^{\nu}\psi(\boldsymbol{L}^{-1}x) \\ &= \boldsymbol{L}^{\mu}{}_{\nu}\overline{\psi}(\boldsymbol{L}^{-1}x)\gamma^{\nu}\psi(\boldsymbol{L}^{-1}x)\,. \end{aligned} \tag{4.18}$$

Thus

$$J^{\mu}_{\boldsymbol{L}}(x) \;=\; \boldsymbol{L}^{\mu}{}_{\nu}\overline{\psi}(\boldsymbol{L}^{-1}x)\gamma^{\nu}\psi(\boldsymbol{L}^{-1}x) \;=\; \boldsymbol{L}^{\mu}{}_{\nu}J^{\nu}(\boldsymbol{L}^{-1}x)\,. \tag{4.19}$$

This is indeed the correct transformation law for a vector field.

In order to discuss *pseudoscalars* and *pseudovectors*, we will introduce a new γ-matrix, namely

$$\gamma^{5} \;=\; i\gamma^{0}\gamma^{1}\gamma^{2}\gamma^{3}\,, \tag{4.20}$$

which in the Dirac representation has the form

$$\gamma^{5} \;=\; \sigma^{1}\otimes\mathbb{1} \;=\; \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{4.21}$$

Basic properties of γ^5

- $(\gamma^5)^{\dagger} \;=\; \gamma^5$
- $(\gamma^5)^2 \;=\; \mathbb{1}$
- $\{\gamma^5, \gamma^{\mu}\} \;=\; 0$ for $\mu = 0, 1, 2, 3$

Now, the bilinear

$$P(x) \;=\; \overline{\psi}(x)\gamma^5\psi(x) \tag{4.22}$$

is a *pseudoscalar*, i.e. under standard (proper, orthochronous) Lorentz transformations it behaves like a scalar but changes its sign under the parity transformation. To see this, let us realize that

$$\gamma^5 \;=\; -i\gamma^1\gamma^0\gamma^2\gamma^3 \;=\; i\gamma^1\gamma^0\gamma^3\gamma^2 \;=\; \frac{i}{4!}\varepsilon_{\mu\nu\sigma\tau}\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\tau\,, \tag{4.23}$$

and also that (all involved indices are distinct)

$$\varepsilon^{\mu\nu\sigma\tau}\gamma^5 \;=\; i\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\tau\,. \tag{4.24}$$

Here $\varepsilon_{\mu\nu\sigma\tau}$ is a permutation symbol (4D analogue of Levi–Civita symbol):

$$\begin{aligned} \varepsilon_{\mu\nu\sigma\tau} &= 1 \quad \text{if} \quad (\mu\nu\sigma\tau) \quad \text{is even permutation of} \quad (0123)\,, \\ \varepsilon_{\mu\nu\sigma\tau} &= -1 \quad \text{if} \quad (\mu\nu\sigma\tau) \quad \text{is odd permutation of} \quad (0123)\,, \\ \varepsilon_{\mu\nu\sigma\tau} &= 0 \quad \text{therwise}\,. \end{aligned} \tag{4.25}$$

Continuing from (4.22) we have

$$\begin{aligned} P(x) \;=\; \overline{\psi}(x)\gamma^5\psi(x) \;&\xrightarrow{\boldsymbol{L}}\; \overline{\psi}_{\boldsymbol{L}}(x)\gamma^5\psi_{\boldsymbol{L}}(x) \\ &=\; \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L})\gamma^5 S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x)\,. \end{aligned} \tag{4.26}$$

We can rewrite term $S^{-1}(\boldsymbol{L})\gamma^5 S(\boldsymbol{L})$ as follows:

$$\begin{aligned} S^{-1}(\boldsymbol{L})\gamma^5 S(\boldsymbol{L}) \;&=\; \frac{i}{4!}\varepsilon_{\mu\nu\sigma\tau}S^{-1}(\boldsymbol{L})\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\tau S(\boldsymbol{L}) \\ &=\; \frac{i}{4}\varepsilon_{\mu\nu\sigma\tau}(S^{-1}(\boldsymbol{L})\gamma^\mu S(\boldsymbol{L}))\ldots(S^{-1}(\boldsymbol{L})\gamma^\tau S(\boldsymbol{L})) \\ &=\; \frac{i}{4!}\varepsilon_{\mu\nu\sigma\tau}\boldsymbol{L}^\mu{}_{\mu'}\boldsymbol{L}^\nu{}_{\nu'}\boldsymbol{L}^\sigma{}_{\sigma'}\boldsymbol{L}^\tau{}_{\tau'}\gamma^{\mu'}\gamma^{\nu'}\gamma^{\sigma'}\gamma^{\tau'} \\ &=\; \gamma^5(\det\boldsymbol{L})\,. \end{aligned} \tag{4.27}$$

Here we have used the fact that

$$\begin{aligned} \det\boldsymbol{A} \;&=\; \sum_{\{J_i\}}\varepsilon_{J_1,\ldots,J_n}A_{1J_1}\ldots A_{nJ_n} \\ &=\; \frac{1}{n!}\sum_{\{J_i\}}\sum_{\{k_i\}}\varepsilon_{J_1,\ldots,J_n}\varepsilon_{K_1,\ldots,K_n}A_{K_1,J_1}\ldots A_{K_n,J_n}\,. \end{aligned} \tag{4.28}$$

Consequently

$$P(x) \;=\; \overline{\psi}(x)\gamma^5\psi(x) \;\xrightarrow{\boldsymbol{L}}\; \overline{\psi}_{\boldsymbol{L}}(x)\gamma^5\psi_{\boldsymbol{L}}(x) \;=\; (\det\boldsymbol{L})P(\boldsymbol{L}^{-1}x)\,, \tag{4.29}$$

and the function $P(x)$ is a Lorentz scalar for all proper Lorentz transformations ($\det\boldsymbol{L} = 1$).

Notice that for the Lorentz transformations involving parity reversal

$$\boldsymbol{L}^P \;=\; \mathrm{diag}(1,-1,-1,-1)\,, \tag{4.30}$$

the transformation also changes sign as $\det\boldsymbol{L}^P = -1$. Complete set of bilinears is given in the following table. Note that all bilinears present in Tab. 4.1 have the form $\overline{\psi}(x)\Gamma_i\psi(x)$, where Γ_i is one of 16 possible

Bilinear	Transformation properties
$\overline{\psi}(x)\psi(x)$	Scalar
$\overline{\psi}(x)\gamma^5\psi(x)$	Pseudoscalar
$\overline{\psi}(x)\gamma^\mu\gamma^5\psi(x)$	Pseudovector field
$\overline{\psi}(x)\gamma^\mu\psi(x)$	Vector field
$\overline{\psi}(x)[\gamma^\mu,\gamma^\nu]\psi(x)$	Antisymmetric tensor field

Table 4.1: List of Dirac's bilinears.

matrices: $\mathbb{1}$, γ^μ, $[\gamma^\mu,\gamma^\nu]$, γ^5 and $\gamma^5\gamma^\mu$. There is no way how to build a non-trivial *symmetric* tensor out of bilinears.

Basic properties of Γ_i matrices

- apart from $\Gamma_1 = \mathbb{1}$ we have that $\mathrm{Tr}(\Gamma_i) = 0$
- apart from $\Gamma_1 = \mathbb{1}$ we have that $\Gamma_i\Gamma_j = -\Gamma_j\Gamma_i$
- $\Gamma_i^2 = \pm\mathbb{1}$
- $\mathrm{Tr}(\Gamma_i\Gamma_j) = 0$ for $i \neq j$ and $\mathrm{Tr}(\Gamma_i^2) = \pm 4$
- $\sum_{i=1}^{16} \alpha_i\Gamma_i = 0$ iff all $\alpha_i = 0$

So, Γ_i matrices form a basis in the space of 4×4 matrices.

4.2 Current for a Dirac wave function

Main motivation of Dirac was to have consistent probability current with positive definite probability density. There is now a natural candidate for a probability current, namely

$$J_\mu(x) = \overline{\psi}(x)\gamma_\mu\psi(x)\,. \tag{4.31}$$

By analogy with non-relativistic quantum mechanics one can define the norm via current as

$$(\psi,\psi) = ||\psi||^2 = \int \mathrm{d}^3x J_0(x) = \int \mathrm{d}V n^\mu J_\mu(x)\,. \tag{4.32}$$

The last integral is over a space-like hyperplane orthogonal to the (time-like) 4-vector n^μ . We want to show the following:

- $||\psi||^2$ is time independent.
- $||\psi||^2$ is a Lorentz invariant non-negative norm.

In order to prove time independence of $||\psi||^2$ we show first that $\partial_\mu J^\mu(x) = 0$. To this end we need to compute

$$\partial_\mu\left(\overline{\psi}(x)\gamma^\mu\psi(x)\right) = \left(\partial_\mu\overline{\psi}(x)\right)\gamma^\mu\psi(x) + \overline{\psi}(x)\gamma^\mu\left(\partial_\mu\psi(x)\right)\,. \tag{4.33}$$

Recalling that $\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) = 0$ we have

$$\gamma^\mu\partial_\mu\psi(x) = -im\psi(x)\,. \tag{4.34}$$

bilinears is given in the following table. Note that all bilinears present in Tab. 4.1 have the form $\overline{\psi}(x)\Gamma_i\psi(x)$, where Γ_i is one of 16 possible

Bilinear	Transformation properties
$\overline{\psi}(x)\psi(x)$	Scalar
$\overline{\psi}(x)\gamma^5\psi(x)$	Pseudoscalar
$\overline{\psi}(x)\gamma^\mu\gamma^5\psi(x)$	Pseudovector field
$\overline{\psi}(x)\gamma^\mu\psi(x)$	Vector field
$\overline{\psi}(x)[\gamma^\mu,\gamma^\nu]\psi(x)$	Antisymmetric tensor field

Table 4.1: List of Dirac's bilinears.

matrices: $\mathbb{1}$, γ^μ, $[\gamma^\mu,\gamma^\nu]$, γ^5 and $\gamma^5\gamma^\mu$. There is no way how to build a non-trivial *symmetric* tensor out of bilinears.

Basic properties of Γ_i matrices

- apart from $\Gamma_1 = \mathbb{1}$ we have that $\mathrm{Tr}(\Gamma_i) = 0$
- apart from $\Gamma_1 = \mathbb{1}$ we have that $\Gamma_i\Gamma_j = -\Gamma_j\Gamma_i$
- $\Gamma_i^2 = \pm\mathbb{1}$
- $\mathrm{Tr}(\Gamma_i\Gamma_j) = 0$ for $i \neq j$ and $\mathrm{Tr}(\Gamma_i^2) = \pm 4$
- $\sum_{i=1}^{16} \alpha_i\Gamma_i = 0$ iff all $\alpha_i = 0$

So, Γ_i matrices form a basis in the space of 4×4 matrices.

4.2 Current for a Dirac wave function

Main motivation of Dirac was to have consistent probability current with positive definite probability density. There is now a natural candidate for a probability current, namely

$$J_\mu(x) = \overline{\psi}(x)\gamma_\mu\psi(x)\,. \tag{4.31}$$

By analogy with non-relativistic quantum mechanics one can define the norm via current as

$$(\psi,\psi) = ||\psi||^2 = \int \mathrm{d}^3x J_0(x) = \int \mathrm{d}V n^\mu J_\mu(x)\,. \tag{4.32}$$

The last integral is over a space-like hyperplane orthogonal to the (time-like) 4-vector n^μ . We want to show the following:

- $||\psi||^2$ is time independent.
- $||\psi||^2$ is a Lorentz invariant non-negative norm.

In order to prove time independence of $||\psi||^2$ we show first that $\partial_\mu J^\mu(x) = 0$. To this end we need to compute

$$\partial_\mu\left(\overline{\psi}(x)\gamma^\mu\psi(x)\right) = \left(\partial_\mu\overline{\psi}(x)\right)\gamma^\mu\psi(x) + \overline{\psi}(x)\gamma^\mu\left(\partial_\mu\psi(x)\right)\,. \tag{4.33}$$

Recalling that $\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) = 0$ we have

$$\gamma^\mu\partial_\mu\psi(x) = -im\psi(x)\,. \tag{4.34}$$

For the adjoint wave function $\overline{\psi}$ we obtain an equation of motion by taking first the hermitian conjugation of Dirac's equation, which yields

$$\left[\left(i\gamma^\mu\partial_\mu - m\right)\psi(x)\right]^\dagger = 0 \quad\Leftrightarrow\quad i\partial_\mu\psi^\dagger(x)\,(\gamma^\mu)^\dagger + m\psi^\dagger(x) = 0\,. \tag{4.35}$$

Taking advantage of the fact that

$$\gamma^0(\gamma^\mu)^\dagger\gamma^0 = \gamma^\mu \quad\Leftrightarrow\quad (\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0\,, \tag{4.36}$$

we can multiply (4.35) by γ^0 from right and write

$$\begin{aligned} 0 &= \left(i\partial_\mu\psi^\dagger(x)(\gamma^\mu)^\dagger + m\psi^\dagger(x)\right)\gamma^0 \\ &= \left(i\partial_\mu\psi^\dagger(x)\gamma^0\gamma^\mu + m\psi^\dagger(x)\gamma^0\right) \\ &= \left(i\partial_\mu\overline{\psi}(x)\gamma^\mu + m\overline{\psi}(x)\right) \\ &= \overline{\psi}(x)\left(i\gamma^\mu\overleftarrow{\partial_\mu} + m\right)\,. \end{aligned} \tag{4.37}$$

From this we see that

$$\left(\partial_\mu\overline{\psi}(x)\right)\gamma^\mu = im\overline{\psi}\,. \tag{4.38}$$

Plugging this into the (4.33) we obtain

$$\begin{aligned} &\left(\partial_\mu\overline{\psi}(x)\right)\gamma^\mu\psi(x) + \overline{\psi}(x)\gamma^\mu\left(\partial_\mu\psi(x)\right) \\ &\qquad = im\overline{\psi}(x)\psi(x) - im\overline{\psi}(x)\psi(x) = 0\,. \end{aligned} \tag{4.39}$$

So, current J^μ is indeed conserved.

Note also that there is yet another possible candidate for probability current, namely the axial vector current

$$J^\mu(x) = \overline{\psi}(x)\gamma^\mu\gamma^5\psi(x)\,. \tag{4.40}$$

It can be, however, shown that in this case

$$\partial_\mu J^\mu(x) = 2im\overline{\psi}(x)\gamma^5\psi(x)\,. \tag{4.41}$$

So, this is conserved only when $m = 0$. Since in the standard model of particle physics there are no fundamental spin 1/2 particles that would have $m = 0$ (at observational energy scales), we will discard this type of current from our reasoning.

Let us now return back to the norm. Choose $n^\mu = (1,\ 0,\ 0,\ 0)$, then

$$\begin{aligned} (\psi,\psi) &= \int d^3\boldsymbol{x}\,J_0(x) = \int d^3\boldsymbol{x}\,\overline{\psi}(x)\gamma^0\psi(x) \\ &= \int d^3\boldsymbol{x}\,\psi^\dagger(x)\gamma^0\gamma^0\psi(x) = \int d^3\boldsymbol{x}\,\psi^\dagger(x)\psi(x) \geq 0\,. \end{aligned} \tag{4.42}$$

From this we can see, that the would-be probability density $\rho = J_0$ is positive definite. In addition, the norm is also time independent, since

$$-\underbrace{\frac{\partial}{\partial t}\int_V d^3\boldsymbol{x}\,\rho}_{\substack{\text{Change in total}\\ \text{probability inside } V}} = \int_V d^3\boldsymbol{x}\boldsymbol{\nabla}\cdot\boldsymbol{J} = \underbrace{\int_{\partial V} d\boldsymbol{S}\cdot\boldsymbol{J}}_{\substack{\text{Flux of } \boldsymbol{J}\\ \text{through the } \partial V}} \to 0\,. \tag{4.43}$$

Here we used the same argument as in the Klein–Gordon particle case.

Finally, in order to see that the norm is relativistically invariant we can proceed as in the Klein–Gordon case, i.e., we first rewrite the norm as

$$\begin{aligned} ||\psi||^2 &= \int d^3\boldsymbol{x}\rho(x) = \int d^4x\,\underbrace{\delta(x_0)}_{\frac{\partial}{\partial x^0}\theta(x_0)}\,\rho(x) \\ &= \int d^4x J^\alpha \frac{\partial}{\partial x^\alpha}\theta(n^\beta x_\beta)\,. \end{aligned} \tag{4.44}$$

where $n^\mu = (1,\ 0,\ 0,\ 0)$. When $||\psi||^2$ is Lorentz invariant it should equal to

$$||\widetilde{\psi}||^2 = \int d^4x J^\alpha \frac{\partial}{\partial x^\alpha}\theta(n'^\beta x_\beta)\,, \tag{4.45}$$

where n'^μ is a unit time-like vector obtained from n^μ via proper orthochronous Lorentz transformation.

Equivalence between $||\widetilde{\psi}||^2$ and $|\psi||^2$ cane be explored by looking at their difference, i.e.

$$\begin{aligned} ||\psi||^2 - ||\widetilde{\psi}||^2 &= \int d^4x J^\alpha \frac{\partial}{\partial x^\alpha}\Big(\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\Big) \\ &\overset{\partial_\alpha J^\alpha = 0}{=} \int d^4x \frac{\partial}{\partial x^\alpha}\Big[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\Big] \\ &= \int dS_\alpha \Big[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\Big]\,. \end{aligned} \tag{4.46}$$

To show that this is zero it is enough to realize that:

1. J^α vanishes if $|\boldsymbol{x}| \to \infty$ with fixed t.
2. $\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)$ vanishes for $|t| \to \infty$ with $\boldsymbol{x}$ fixed (for orthochronous transformations).

Hence, the difference (4.46) is zero and the norm is relativistically invariant. Again, we can define related scalar product as

$$(\psi_1, \psi_2) = \int dV n_\mu J^{\mu(1,2)}(x)\,, \tag{4.47}$$

where $J^{\mu(1,2)}(x) = \overline{\psi}_1(x)\gamma^\mu\psi_2(x)$.

Plane Wave Solutions of Dirac Equation

We know that because $\psi(x)$ satisfies Dirac equation, it must also satisfy Klein–Gordon equation. In particular, one may expect that wave functions of definite energy and momentum will be described by plane waves (i.e., de Broglie monochromatic waves) of the form

$$\begin{aligned} \psi_p^+(x) &= u(p)e^{-ipx} = u(p)e^{-i\omega_{\boldsymbol{p}}t+i\boldsymbol{p}\cdot\boldsymbol{x}} \quad \text{(positive energy)}, \\ \psi_p^-(x) &= v(p)e^{ipx} = v(p)e^{i\omega_{\boldsymbol{p}}t-i\boldsymbol{p}\cdot\boldsymbol{x}} \quad \text{(negative energy)}, \end{aligned} \tag{4.48}$$

where $p_0 = \omega_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2} > 0$. For the positive-energy plane wave $i\partial_\mu \to p_\mu$ and given that $\left(i\gamma^\mu\partial_\mu - m\right)\psi_p^+ = 0$ we obtain

$$\left(\gamma^\mu p_\mu - m\right)u(p) = 0 \quad \Leftrightarrow \quad \left(\not{p} - m\right)u(p) = 0, \tag{4.49}$$

Similarly for the negative-energy solution we have

$$\left(\gamma^\mu p_\mu + m\right)v(p) = 0 \quad \Leftrightarrow \quad \left(\not{p} + m\right)v(p) = 0. \tag{4.50}$$

In order to have a non-trivial solutions of Eqs. (4.49)-(4.50) we need to show that $\det\left(\not{p} - m\right) = 0$ and $\det\left(\not{p} + m\right) = 0$. To this end we use the following trick:

$$\begin{aligned} \det\left(\gamma^\mu p_\mu - m\right) &= \det\left(\gamma^5\gamma^5\left(\gamma^\mu p_\mu - m\right)\right) \\ &= \det\left(\gamma^5\left(\gamma^\mu p_\mu - m\right)\gamma^5\right) \\ &= \det\left(\left(-\gamma^\mu p_\mu - m\right)\gamma^5\gamma^5\right) \\ &= \det\left(\gamma^\mu p_\mu + m\right). \end{aligned} \tag{4.51}$$

The second equation follows from property of determinants and the third from anticommutativity of γ^5 with γ^μ. From this we see that

$$\begin{aligned} \det\left[\left(\gamma^\mu p_\mu - m\right)\left(\gamma^\mu p_\mu + m\right)\right] &= \det\left(\gamma^\nu\gamma^\mu p_\mu p_\nu - m^2\right) \\ &= \det{}^2\left(\gamma^\mu p_\mu \pm m\right). \end{aligned} \tag{4.52}$$

Using properties of γ-matrices we can further write that

$$\gamma^\nu\gamma^\mu p_\mu p_\nu - m^2 = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\, p_\mu p_\nu - m^2 = p^2 - m^2, \tag{4.53}$$

and hence

$$\det{}^2\left(\gamma^\mu p_\mu \pm m\right) = (p^2 - m^2)^4 = 0. \tag{4.54}$$

Consequently, plane-wave solutions with non-trivial amplitudes are solutions of Dirac's equation.

Positive Energy Solutions

From the Dirac representation of γ-matrices we find that

$$\left(\gamma^\mu p_\mu - m\right) = \begin{pmatrix} (E-m)\mathbb{1} & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E+m)\mathbb{1} \end{pmatrix}. \tag{4.55}$$

Consider $u(p)$ in the form

$$u(p) = \begin{pmatrix} \chi \\ \varphi \end{pmatrix}, \tag{4.56}$$

where χ and φ are both 2-component columns. From this we can rewrite equation (4.49) as a system of two coupled equations

$$(E-m)\chi - \boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi = 0, \tag{4.57}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi - (E+m)\varphi = 0. \tag{4.58}$$

The *on-mass-shell* condition $p^2 = m^2$, in fact, ensures, that these two equations are equivalent. Indeed, multiplying (4.57) by $(E+m)$ we obtain

$$\underbrace{(E^2-m^2)}_{\boldsymbol{p}^2}\chi - \boldsymbol{\sigma}\cdot\boldsymbol{p}(E+m)\varphi = 0. \tag{4.59}$$

Similarly, by multiplying (4.58) from left by $\boldsymbol{\sigma}\cdot\boldsymbol{p}$ we get

$$\sigma^i\sigma^j p^i p^j \chi - \boldsymbol{\sigma}\cdot\boldsymbol{p}(E+m)\varphi = 0. \tag{4.60}$$

Using the fact that $\sigma^i\sigma^j p^i p^j$ can be rewritten as $\frac{1}{2}\{\sigma^i, \sigma^j\}p^i p^j = \delta^{ij}p^i p^j = \boldsymbol{p}^2$ we get that those two equations are indeed equivalent and any of the two can be used. In particular, from the second equation we obtain that

$$\varphi = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi, \tag{4.61}$$

which implies that

$$u(p) \propto \begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix}. \tag{4.62}$$

For our future convenience we fix the normalization of $u(p)$ so that

$$u(p) = \sqrt{E+m}\begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix}. \tag{4.63}$$

Other Normalizations

Often the normalization is chosen differently, namely

$$u(p) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{E+m}{2m}}\chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{\sqrt{2m(E+m)}}\chi \end{pmatrix}.$$

This gives nicer normalization for $\bar{u}u$ and $\bar{v}v$.

There is another, more physical way of solving $(\not{p}-m)\,u(p) = 0$ and $(\not{p}+m)\,v(p) = 0$ equations. Let assume that $m \neq 0$. In the rest frame of the particle $p^\mu = (m, \boldsymbol{0})$, which implies that the aforementioned Dirac equations reduce to

$$\left(\gamma^0 - \mathbb{1}\right) u(m, \mathbf{0}) \;=\; 0\,, \tag{4.64}$$

$$\left(\gamma^0 + \mathbb{1}\right) v(m, \mathbf{0}) \;=\; 0\,. \tag{4.65}$$

There are clearly 2 linearly independent solutions for both u and v, namely

$$u^{(1)}(m, \mathbf{0}) \;=\; \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad u^{(2)}(m, \mathbf{0}) \;=\; \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$v^{(1)}(m, \mathbf{0}) \;=\; \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v^{(2)}(m, \mathbf{0}) \;=\; \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{4.66}$$

We could now boost these solutions from a rest up to a velocity $|\boldsymbol{v}| = \frac{|\boldsymbol{p}|}{p_0}$ by a pure Lorentz transformation $S(\boldsymbol{L}) = e^{-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu}}$.

3-velocity in relativity

Note that $p^\mu = \left(\frac{E}{c}, \boldsymbol{p}\right) = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}(c, \boldsymbol{v})$. So, by setting $c = 1$ we have that $|\boldsymbol{v}| = \frac{|\boldsymbol{p}|}{p_0}$.

Inner workings of this will be explicitly seen in section that deals with boost transformations.

There exists also a simpler passage to the solution that employs a trivial fact that

$$\left(\not{p} - m\right)\left(\not{p} + m\right) \;=\; p^2 - m^2 \;=\; \left(\not{p} + m\right)\left(\not{p} - m\right) \;=\; 0\,. \tag{4.67}$$

So, in particular

$$\left(\not{p} - m\right)\left(\not{p} + m\right) u^{(\lambda)} \;=\; 0, \qquad \lambda = 1, 2\,. \tag{4.68}$$

This implies that the positive-energy solution is of the form

$$\begin{aligned} \left(\not{p} + m\right) u^{(\lambda)}(m, \mathbf{0}) \;&=\; \begin{pmatrix} E + m & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E - m) \end{pmatrix} u^{(\lambda)}(m, \mathbf{0}) \\ &=\; \begin{pmatrix} (E + m)\chi^{(\lambda)} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p}\,\chi^{(\lambda)} \end{pmatrix}. \end{aligned} \tag{4.69}$$

Here we employed the fact that

$$u^{(\lambda)}(m, \mathbf{0}) \;=\; \begin{pmatrix} \chi^{(\lambda)}(m, \mathbf{0}) \\ 0 \\ 0 \end{pmatrix}, \tag{4.70}$$

and that $\chi^{(\lambda)}$ is either $\begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$.

Negative energy solution is obtained similarly, namely it is of the form

$$\begin{aligned}(\not{p}-m)\, v^{(\lambda)}(m,\mathbf{0}) &= \begin{pmatrix}E-m & -\boldsymbol{\sigma}\cdot\boldsymbol{p}\\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E+m)\end{pmatrix} v^{(\lambda)}(m,\mathbf{0})\\ &= \begin{pmatrix}-\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\chi^{(\lambda)}\\ -(E+m)\chi^{(\lambda)}\end{pmatrix}. \qquad (4.71)\end{aligned}$$

Where in analogy with positive energy solution, we write

$$v^{(\lambda)}(m,\mathbf{0}) = \begin{pmatrix}0\\0\\ \chi^{(\lambda)}(m,\mathbf{0})\end{pmatrix}, \qquad (4.72)$$

and $\chi^{(\lambda)}$ is again $\begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$. Relation with our former normalization is obtained when we write

$$\begin{aligned}u^{(\lambda)}(p) &= \frac{\not{p}+m}{\sqrt{E+m}}u^{(\lambda)}(m,\mathbf{0}),\\ v^{(\lambda)}(p) &= \frac{-\not{p}+m}{\sqrt{E+m}}v^{(\lambda)}(m,\mathbf{0}). \qquad (4.73)\end{aligned}$$

Non-relativistic limit and relation to the Schrödinger equation

Recall that positive and negative energy solutions to the massive Dirac equation have the form

$$u_\lambda = \sqrt{E+m}\begin{pmatrix}\chi_\lambda\\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda\end{pmatrix} \equiv \begin{pmatrix}u_L\\u_S\end{pmatrix}. \qquad (4.74)$$

Note that $u_S = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}u_L$. Similarly for negative energy solution we can write

$$v_\lambda = \sqrt{E+m}\begin{pmatrix}\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\varphi_\lambda\\ \varphi_\lambda\end{pmatrix} \equiv \begin{pmatrix}v_S\\v_L\end{pmatrix}, \qquad (4.75)$$

and $v_S = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}v_L$. In the non-relativistic limit $|\boldsymbol{p}| \ll m$ and so $u_S \ll u_L$ and $v_S \ll v_L$. The subscript "S" refers to the so-called *small component* and the subscript "L" to the *large component*.

The positive energy solutions satisfy

$$\begin{pmatrix}m\mathbb{1} & \boldsymbol{\sigma}\cdot\boldsymbol{p}\\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -m\mathbb{1}\end{pmatrix}\begin{pmatrix}u_L\\u_S\end{pmatrix} = E\begin{pmatrix}u_L\\u_S\end{pmatrix}. \qquad (4.76)$$

This equation can be rewritten as

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\,u_S = (E-m)u_L, \qquad (4.77)$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\,u_L = (E+m)u_S. \qquad (4.78)$$

and that $\chi^{(\lambda)}$ is either $\begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$.

Negative energy solution is obtained similarly, namely it is of the form

$$\begin{aligned}(\not p - m)\, v^{(\lambda)}(m,\mathbf{0}) &= \begin{pmatrix}E-m & -\boldsymbol{\sigma}\cdot\boldsymbol{p}\\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E+m)\end{pmatrix} v^{(\lambda)}(m,\mathbf{0})\\ &= \begin{pmatrix}-\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\chi^{(\lambda)}\\ -(E+m)\chi^{(\lambda)}\end{pmatrix}. \end{aligned} \tag{4.71}$$

Where in analogy with positive energy solution, we write

$$v^{(\lambda)}(m,\mathbf{0}) = \begin{pmatrix}0\\0\\ \chi^{(\lambda)}(m,\mathbf{0})\end{pmatrix}, \tag{4.72}$$

and $\chi^{(\lambda)}$ is again $\begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$. Relation with our former normalization is obtained when we write

$$\begin{aligned}u^{(\lambda)}(p) &= \frac{\not p + m}{\sqrt{E+m}} u^{(\lambda)}(m,\mathbf{0}),\\ v^{(\lambda)}(p) &= \frac{-\not p + m}{\sqrt{E+m}} v^{(\lambda)}(m,\mathbf{0}).\end{aligned} \tag{4.73}$$

Non-relativistic limit and relation to the Schrödinger equation

Recall that positive and negative energy solutions to the massive Dirac equation have the form

$$u_\lambda = \sqrt{E+m}\begin{pmatrix}\chi^{(\lambda)}\\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi^{(\lambda)}\end{pmatrix} \equiv \begin{pmatrix}u_L\\ u_S\end{pmatrix}. \tag{4.74}$$

Note that $u_S = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}u_L$. Similarly for negative energy solution we can write

$$v_\lambda = \sqrt{E+m}\begin{pmatrix}\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\varphi^{(\lambda)}\\ \varphi^{(\lambda)}\end{pmatrix} \equiv \begin{pmatrix}v_S\\ v_L\end{pmatrix}, \tag{4.75}$$

and $v_S = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}v_L$. In the non-relativistic limit $|\boldsymbol{p}| \ll m$ and so $|u_S| \ll |u_L|$ and $|v_S| \ll |v_L|$. The subscript "S" refers to the so-called *small component* and the subscript "L" to the *large component*.

The positive energy solutions satisfy

$$\begin{pmatrix}m\mathbb{1} & \boldsymbol{\sigma}\cdot\boldsymbol{p}\\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -m\mathbb{1}\end{pmatrix}\begin{pmatrix}u_L\\ u_S\end{pmatrix} = E\begin{pmatrix}u_L\\ u_S\end{pmatrix}. \tag{4.76}$$

This equation can be rewritten as

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\, u_S = (E-m)u_L, \tag{4.77}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\, u_L = (E+m)u_S. \tag{4.78}$$

By substituting (4.78) to (4.77) we get

$$\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{p})(\boldsymbol{\sigma}\cdot\boldsymbol{p})}{E+m}u_L = (E-m)u_L . \tag{4.79}$$

We now use the fact that $(\boldsymbol{\sigma}\cdot\boldsymbol{p})(\boldsymbol{\sigma}\cdot\boldsymbol{p}) = \boldsymbol{p}^2$ and that $|\boldsymbol{p}| \ll m$ implies $E \approx m$. With this we can rewrite (4.79) as

$$\frac{\boldsymbol{p}^2}{2m}u_L = (E-m)u_L = E_{NR}u_L , \tag{4.80}$$

where E_{NR} is the non-relativist energy spectrum. This is the usual Schrödinger equation for a free non-relativistic particle. To understand the 2-component nature of u_L we need to understand better how $u(p)$ transforms under the Lorentz group.

Applications of Lorentz Transformations on Dirac Wave Functions

We will begin with *rotations of Dirac wave functions*. We know that that the group of rotations $\{\boldsymbol{R}\}$ is a subgroup of the group of Lorentz transformations $\{\boldsymbol{L}\}$. We have seen that a generic element of the Lorentz group has the form $\boldsymbol{L} = e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}}$. In particular, generators of rotations are described by $\boldsymbol{M}^{ij}$ $(i,j = 1,2,3)$. Connection between $\boldsymbol{M}^{ij}$ and $\boldsymbol{J}_i$ is established via relation

Recall that $[\boldsymbol{J}_i,\boldsymbol{J}_j] = i\varepsilon_{ijk}\boldsymbol{J}_k$ — i.e., they close a familiar algebra of rotations $su(2) \sim so(3)$.

$$\boldsymbol{J}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{M}^{jk} \quad \text{or equivalently} \quad \boldsymbol{M}^{jk} = 2\varepsilon^{jki}\boldsymbol{J}_i . \tag{4.81}$$

In particular, when $\boldsymbol{M}^{\mu\nu}$ acts on 4-vectors (i.e., it is in a fundamental representation) then

$$(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau} = 2i\left[\eta^{\mu\rho}\eta^{\nu}{}_{\tau} - \eta^{\nu\rho}\eta^{\mu}{}_{\tau}\right] . \tag{4.82}$$

This gives us

$$(\boldsymbol{J}_i)^{\rho}{}_{\tau} = \frac{i}{2}\varepsilon_{ijk}\left[\eta^{\rho j}\eta^{k}{}_{\tau} - \eta^{\rho k}\eta^{j}{}_{\tau}\right] . \tag{4.83}$$

So, for instance, for a third component we can write

$$(\boldsymbol{J}_3)^{\rho}{}_{\tau} = \frac{i}{2}\varepsilon_{3jk}\left[\eta^{\rho j}\eta^{k}{}_{\tau} - \eta^{\rho k}\eta^{j}{}_{\tau}\right] . \tag{4.84}$$

From this we see, that j has can be either 1 or 2, and k either 2 or 1, respectively. Hence, non-trivial contributions to $(\boldsymbol{J}_3)^{\rho}{}_{\tau}$ come only from components $\rho = 1,2$, $\tau = 2,1$. Namely

Upper index labels rows, while the lower one labels columns.

$$\begin{aligned}(\boldsymbol{J}_3)^{1}{}_{2} &= \frac{i}{2}\varepsilon_{312}\eta^{11}\eta^{2}_{2} - \frac{i}{2}\varepsilon_{321}\eta^{11}\eta^{2}_{2} = -i ,\\ (\boldsymbol{J}_3)^{2}{}_{1} &= \frac{i}{2}\varepsilon_{321}\eta^{22}\eta^{1}_{1} - \frac{i}{2}\varepsilon_{312}\eta^{22}\eta^{1}_{1} = i .\end{aligned} \tag{4.85}$$

Thus we can finally write that $\boldsymbol{J}_3$ has an explicit form

$$\boldsymbol{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{4.86}$$

In a similar way we obtain

$$\boldsymbol{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \quad \boldsymbol{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix} . \tag{4.87}$$

It is easy to see that the generators $\boldsymbol{J}_i$ can be written in the compact form

$$(\boldsymbol{J}_i)^{\mu}{}_{\nu} = -i \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \varepsilon_{ijk} \end{array} \right) = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & (\boldsymbol{J}_i)_{jk} \end{array} \right) , \tag{4.88}$$

where $(\boldsymbol{J}_i)_{jk}$ are angular momentum generators in *adjoint representation*.

On the other hand, we know that a general element of the subgroup of rotations has the form

$$R = e^{-\frac{i}{4}\omega_{ij}\boldsymbol{M}^{ij}} = e^{-\frac{i}{2}\varepsilon^{ijk}\omega_{ij}\boldsymbol{J}_k} = e^{-i\boldsymbol{\theta}^k \boldsymbol{J}_k} . \tag{4.89}$$

Here we have defined $\boldsymbol{\theta}^k = \frac{1}{2}\varepsilon^{ijk}\omega_{ij}$. So, for instance the rotation around z-axis should have the form

$$\boldsymbol{R}_3 = e^{-i\boldsymbol{\theta}\boldsymbol{J}_3} . \tag{4.90}$$

This can be computed explicitly. Let us first denote the central 2×2 matrix in $\boldsymbol{J}_3$, i.e.

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \tag{4.91}$$

as $\tilde{\boldsymbol{J}}_3$. Using the fact that $\tilde{\boldsymbol{J}}_3$ is *involutory matrix* , i.e., $(\tilde{\boldsymbol{J}}_3)^2 = \mathbf{1}$, we can write

$$
\begin{aligned}
e^{-i\theta J_3} &= \begin{pmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{0} & \sum_{n=0}^{\infty} \frac{(-i\theta)^n}{n!} \tilde{J}_3^n & \mathbf{0} \\ 0 & \mathbf{0} & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{0} & \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!}(\theta)^{2n}\mathbf{1}_{2\times 2} - i\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}(\theta)^{2n+1}\tilde{J}_3 & \mathbf{0} \\ 0 & \mathbf{0} & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{0} & \cos\theta \mathbf{1}_{2\times 2} - i\sin\theta \tilde{J}_3 & \mathbf{0} \\ 0 & \mathbf{0} & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
\end{aligned}
\tag{4.92}
$$

which describes a rotation in the $x-y$ plane (i.e., rotation around z-axis) by angle θ. Along the same lines one can show that a rotation along arbitrary axis specified by the unit 3-vector $\boldsymbol{n}$ is done by

One can indeed check that $\boldsymbol{J}_k = i\left.\frac{\partial R}{\partial \theta^k}\right|_{\boldsymbol{\theta}=0}$.

$$
\boldsymbol{R} = e^{-i\theta^i \boldsymbol{J}_i}. \tag{4.93}
$$

Note that we can equivalently write that

$$
\boldsymbol{\theta}^i \boldsymbol{J}_i = \theta \boldsymbol{n}^i \boldsymbol{J}_i, \tag{4.94}
$$

where θ is rotation angle around axis $\boldsymbol{n}$. Since $\boldsymbol{\theta}$ corresponds to respective components of the rotational angle we have succeeded to relate the abstract parameters ω_{ij} with the physical parameters $\boldsymbol{\theta}_i$ (cf. $\boldsymbol{\theta}^k = \frac{1}{2}\varepsilon^{ijk}\boldsymbol{\omega}_{ij}$).

The corresponding representation of rotations that acts on Dirac wave functions can be constructed from generators σ^{ij} by realizing the following parallelism

$$
\begin{aligned}
\boldsymbol{L} &= e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}} \quad \leftrightarrow \quad S(\boldsymbol{L}) = e^{-\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}}, \\
\boldsymbol{R} &= e^{-\frac{i}{4}\boldsymbol{M}^{ij}\omega_{ij}} = e^{-i\theta \boldsymbol{n}^i \boldsymbol{J}_i} \quad \leftrightarrow \quad S(\boldsymbol{R}) = e^{-\frac{i}{4}\sigma^{ij}\omega_{ij}} = e^{-i\theta \boldsymbol{n}^i \hat{\boldsymbol{\sigma}}_i}.
\end{aligned}
\tag{4.95}
$$

Here $\hat{\boldsymbol{\sigma}}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{\sigma}^{jk} \Leftrightarrow \boldsymbol{\sigma}^{jk} = 2\varepsilon^{ijk}\hat{\boldsymbol{\sigma}}_i$. Let us find explicit form of $\hat{\boldsymbol{\sigma}}_i$. Starting from the definition of $\boldsymbol{\sigma}^{ij}$, we get in Dirac's representation

$$
\begin{aligned}
\boldsymbol{\sigma}^{ij} &= \frac{i}{2}\left[\gamma^i, \gamma^j\right] = i\gamma^i\gamma^j \\
&= i\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} = i\begin{pmatrix} -\sigma^i\sigma^j & 0 \\ 0 & -\sigma^i\sigma^j \end{pmatrix} \\
&= \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}\varepsilon^{ijk}.
\end{aligned}
\tag{4.96}
$$

write

$$
\begin{aligned}
e^{-i\theta \boldsymbol{J}_3} &= \begin{pmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{0} & \sum_{n=0}^{\infty} \frac{(-i\theta)^n}{n!} \tilde{\boldsymbol{J}}_3^n & \mathbf{0} \\ 0 & \mathbf{0} & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{0} & \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} (\theta)^{2n} \mathbf{1}_{2\times 2} - i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (\theta)^{2n+1} \tilde{\boldsymbol{J}}_3 & \mathbf{0} \\ 0 & \mathbf{0} & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{0} & \cos\theta \mathbf{1}_{2\times 2} - i \sin\theta \tilde{\boldsymbol{J}}_3 & \mathbf{0} \\ 0 & \mathbf{0} & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (4.92)
\end{aligned}
$$

which describes a rotation in the $x - y$ plane (i.e., rotation around z-axis) by angle θ. Along the same lines one can show that a rotation along arbitrary axis specified by the unit 3-vector $\boldsymbol{n}$ is done by

One can easily check that $\boldsymbol{J}_3 = i \left. \frac{\partial \boldsymbol{R}_3}{\partial \theta} \right|_{\theta=0} = i \left. \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\theta}^3} \right|_{\boldsymbol{\theta}=0}$, and more generally $\boldsymbol{J}_k = i \left. \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\theta}^k} \right|_{\boldsymbol{\theta}=0}$.

$$
\boldsymbol{R} = e^{-i\boldsymbol{\theta}^i \boldsymbol{J}_i} . \qquad (4.93)
$$

Note that we can equivalently write that

$$
\boldsymbol{\theta}^i \boldsymbol{J}_i = \theta \boldsymbol{n}^i \boldsymbol{J}_i , \qquad (4.94)
$$

where θ is rotation angle around axis $\boldsymbol{n}$. Since $\boldsymbol{\theta}$ corresponds to respective components of the rotational angle we have succeeded to relate the abstract parameters ω_{ij} with the physical parameters $\boldsymbol{\theta}_i$ (cf. $\boldsymbol{\theta}^k = \frac{1}{2}\varepsilon^{ijk}\boldsymbol{\omega}_{ij}$).

The corresponding representation of rotations that acts on Dirac wave functions can be constructed from generators σ^{ij} by realizing the following parallelism

$$
\begin{aligned}
\boldsymbol{L} &= e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}} \leftrightarrow S(\boldsymbol{L}) = e^{-\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}} , \\
\boldsymbol{R} &= e^{-\frac{i}{4}\boldsymbol{M}^{ij}\omega_{ij}} = e^{-i\theta \boldsymbol{n}^i \boldsymbol{J}_i} \leftrightarrow S(\boldsymbol{R}) = e^{-\frac{i}{4}\boldsymbol{\sigma}^{ij}\omega_{ij}} = e^{-i\theta \boldsymbol{n}^i \hat{\boldsymbol{\sigma}}_i} . \qquad (4.95)
\end{aligned}
$$

Here $\hat{\boldsymbol{\sigma}}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{\sigma}^{jk} \Leftrightarrow \boldsymbol{\sigma}^{jk} = 2\varepsilon^{ijk}\hat{\boldsymbol{\sigma}}_i$. Let us find explicit form of $\hat{\boldsymbol{\sigma}}_i$. Starting from the definition of $\boldsymbol{\sigma}^{ij}$, we get in Dirac's representation

$$
\begin{aligned}
\boldsymbol{\sigma}^{ij} &= \frac{i}{2}\left[\gamma^i, \gamma^j\right] = i\gamma^i\gamma^j \\
&= i \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} = i \begin{pmatrix} -\sigma^i\sigma^j & 0 \\ 0 & -\sigma^i\sigma^j \end{pmatrix} \\
&= \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \varepsilon^{ijk} . \qquad (4.96)
\end{aligned}
$$

Consequently, $\hat{\sigma}_k$ reads

$$\begin{aligned}\hat{\sigma}_k &= \frac{1}{4}\varepsilon_{klm}\sigma^{lm} = \frac{1}{4}\varepsilon_{klm}\varepsilon^{lmp}\begin{pmatrix}\sigma_p & 0\\ 0 & \sigma_p\end{pmatrix}\\ &= \frac{1}{4}\varepsilon_{lmk}\varepsilon^{lmp}\begin{pmatrix}\sigma_p & 0\\ 0 & \sigma_p\end{pmatrix}\\ &= \frac{1}{4}(\delta^m_m\delta^p_k - \delta^p_m\delta^m_k)\begin{pmatrix}\sigma_p & 0\\ 0 & \sigma_p\end{pmatrix}\\ &= \frac{1}{2}\begin{pmatrix}\sigma_k & 0\\ 0 & \sigma_k\end{pmatrix}.\end{aligned} \tag{4.97}$$

Algebra of generators $\hat{\sigma}_k$

Taking commutator

$$[\hat{\sigma}_k, \hat{\sigma}_l] = \begin{pmatrix}\frac{1}{4}[\sigma_k, \sigma_l] & 0\\ 0 & \frac{1}{4}[\sigma_k, \sigma_l]\end{pmatrix} = \frac{i}{2}\varepsilon_{klm}\begin{pmatrix}\sigma_m & 0\\ 0 & \sigma_m\end{pmatrix} = i\varepsilon_{klm}\hat{\sigma}_m.$$

Hence $\hat{\sigma}_i$ are correct generators of rotations in Dirac's space of wave functions (so-called *bispinor space*).

By denoting

$$2\hat{\sigma}_k = \Sigma_k = \begin{pmatrix}\sigma_k & 0\\ 0 & \sigma_k\end{pmatrix}, \tag{4.98}$$

we obtain that

$$S(\boldsymbol{R}) = e^{-\frac{i}{2}\theta n^k \Sigma_k} = \begin{pmatrix}e^{-\frac{i}{2}\theta \boldsymbol{n}_k\sigma_k} & 0\\ 0 & e^{-\frac{i}{2}\theta \boldsymbol{n}_k\sigma_k}\end{pmatrix}. \tag{4.99}$$

Consider now action of $S(\boldsymbol{R})$ on $u(p)$:

$$\begin{aligned}S(\boldsymbol{R})u(p) &= \begin{pmatrix}e^{-\frac{i}{2}\theta \boldsymbol{n}_k\sigma_k} & 0\\ 0 & e^{-\frac{i}{2}\theta \boldsymbol{n}_k\sigma_k}\end{pmatrix}u(p)\\ &= \sqrt{E+m}\begin{pmatrix}e^{-\frac{i}{2}\theta \boldsymbol{n}_k\sigma_k} & 0\\ 0 & e^{-\frac{i}{2}\theta \boldsymbol{n}_k\sigma_k}\end{pmatrix}\begin{pmatrix}\chi\\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi\end{pmatrix}\\ &= \sqrt{E+m}\begin{pmatrix}e^{-\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}}\chi\\ \frac{e^{-\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}}\boldsymbol{\sigma}\cdot\boldsymbol{p}e^{\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}}}{E+m}e^{-\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}}\chi\end{pmatrix}.\end{aligned} \tag{4.100}$$

Define $\chi_{\boldsymbol{R}} = e^{-\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}}\chi$, where $\frac{\boldsymbol{\sigma}}{2}$ are generators of rotations in representation with spin 1/2. So, χ can be identified with a 2-component spinor.

2-component wave function that transforms under rotations as

$$\chi_{\boldsymbol{R}} = e^{-\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}}\chi,$$

is known in quantum mechanics as *Pauli spinor*. In mathematics this is better known as *Weyl spinor*.

We can now use the following identity (it can be easily check to the leading order in θ)

$$e^{-\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}}\boldsymbol{\sigma}\cdot\boldsymbol{p}e^{\frac{i}{2}\theta\boldsymbol{n}\cdot\boldsymbol{\sigma}} = \boldsymbol{\sigma}e^{-i\theta\boldsymbol{n}\cdot\boldsymbol{J}}\cdot\boldsymbol{p}. \tag{4.101}$$

(cf. with the similar relation $S^{-1}(\boldsymbol{L})\gamma^\mu S(\boldsymbol{L}) = \boldsymbol{L}^\mu{}_\nu\gamma^\nu = \gamma^\nu\boldsymbol{L}^\mu{}_\nu$). From this we finally get the rotated Dirac's wave function in the form

$$S(\boldsymbol{R})u(p) = \sqrt{E+m}\begin{pmatrix}\chi_{\boldsymbol{R}} \\ \frac{\sigma\cdot \boldsymbol{p}_{\boldsymbol{R}}}{E+m}\chi_{\boldsymbol{R}}\end{pmatrix} = u_{\boldsymbol{R}}(p_{\boldsymbol{R}})\,. \tag{4.102}$$

We thus see that $u_{\boldsymbol{R}}$ is constructed from the rotated spinor $\chi_{\boldsymbol{R}}$ and $p_{\boldsymbol{R}}$ is the rotated 4-momentum, i.e., $p_{\boldsymbol{R}} = (p_0, \boldsymbol{p}_{\boldsymbol{R}})$.

Here we change the former notation $\chi^{(\lambda)}, \lambda = 1, 2$, to the notation $\chi_\lambda, \lambda = \pm\frac{1}{2}$ that is more usual in situations when spin matters.

Let us now introduce Pauli spinors $\chi_{\frac{1}{2}} = \begin{pmatrix}1\\0\end{pmatrix}$ and $\chi_{-\frac{1}{2}} = \begin{pmatrix}0\\1\end{pmatrix}$ and denote them generally as χ_λ with $\lambda = \pm\frac{1}{2}$. Then

$$u_\lambda(p) = \sqrt{E+m}\begin{pmatrix}\chi_\lambda \\ \frac{\sigma\cdot \boldsymbol{p}}{E+m}\chi_\lambda\end{pmatrix}, \tag{4.103}$$

and $\lambda = \pm\frac{1}{2}$ (as seen in the box) has a meaning of the "third component of spin".

Action of Rotation Generators on Free Solutions of Dirac's Equation in the Rest Frame

Let us look what happens, when we apply generators of rotations (angular momentum operators) on free solutions of Dirac equation in the rest frame. We know that these solutions have general form

$$\psi^+_\lambda(x) = \sqrt{2m}\begin{pmatrix}\chi_\lambda\\0\end{pmatrix}\frac{e^{-ik_0t+i\boldsymbol{k}\cdot\boldsymbol{x}}}{\sqrt{2\pi}^3}\,,$$

$$\psi^-_\lambda(x) = \sqrt{2m}\begin{pmatrix}0\\\chi_\lambda\end{pmatrix}\frac{e^{ik_0t-i\boldsymbol{k}\cdot\boldsymbol{x}}}{\sqrt{2\pi}^3}\,,$$

with $\boldsymbol{k}\cdot\boldsymbol{x}$ being zero in the rest frame and $k_0 = m$. We also know that

$$\boldsymbol{J}_i \leftrightarrow \hat{\sigma} = \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix},$$

and in particular

$$\sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},$$

which leads to

$$\boldsymbol{J}_3\psi^+_{\frac{1}{2}} = \frac{1}{2}\psi^+_{\frac{1}{2}}\,, \quad \boldsymbol{J}_3\psi^+_{-\frac{1}{2}} = -\frac{1}{2}\psi^+_{-\frac{1}{2}}\,,$$

$$\boldsymbol{J}_3\psi^-_{\frac{1}{2}} = \frac{1}{2}\psi^-_{\frac{1}{2}}\,, \quad \boldsymbol{J}_3\psi^-_{-\frac{1}{2}} = -\frac{1}{2}\psi^-_{-\frac{1}{2}}\,.$$

These relations imply that Dirac's bispinors describe particle with spin 1/2.

4.3 Lorentz Boosts and Dirac Wave Function

Lorentz boost in x direction can be written as

$$\begin{aligned} t' &= \gamma(t - vx) \quad \Leftrightarrow \quad x_0' = \gamma(x_0 - \beta x)\,, \\ x' &= \gamma(x - vt) \quad \Leftrightarrow \quad x' = \gamma(x - \beta x_0)\,, \\ y' &= y\,, \\ z' &= z\,, \end{aligned} \tag{4.104}$$

where $v = \beta$ and $-1 < \beta < 1$ (note that we use the notation $c = 1$).

The Lorentz transformations are often also written in a way that resembles rotations in 3D using hyperbolic functions. This is possible, because β and γ satisfy identity

$$\gamma^2 - \gamma^2\beta^2 \;=\; \gamma^2(1-\beta^2) \;=\; 1\,, \tag{4.105}$$

which allows us to set $\gamma \equiv \cosh\zeta$ and $\gamma\beta \equiv \sinh\zeta$. With this, we can rewrite (4.104) equivalently as

$$\begin{aligned} x_0' &= x_0 \cosh\zeta \;-\; x \sinh\zeta\,, \\ x' &= x \cosh\zeta \;-\; x_0 \sinh\zeta\,, \\ y' &= y\,, \\ z' &= z\,. \end{aligned} \tag{4.106}$$

In this connection we note that

$$\begin{aligned} \cosh\zeta &= \gamma \;=\; \frac{1}{\sqrt{1-\beta^2}} \;=\; \frac{1}{\sqrt{1-\tanh^2\zeta}}\,, \\ \tanh\zeta &= \frac{\sinh\zeta}{\cosh\zeta} \;=\; \beta \quad\Rightarrow\quad \zeta \;=\; \tanh^{-1}\beta\,. \end{aligned} \tag{4.107}$$

Since $\beta \in (-1, 1)$, we get that $\zeta \in (-\infty, \infty)$. The new variable ζ is known as *rapidity*.

Rapidity is a standard parameter that quantifies relativistic velocities in particle physics. In 1D motion, rapidities are additive whereas velocities must be combined via Einstein's velocity-addition formula.

Similar equations can also be written for boost in z direction

$$\begin{aligned} x_0' &= x_0 \cosh\zeta \;-\; z \sinh\zeta\,, \\ x' &= x\,, \\ y' &= y\,, \\ z' &= z \cosh\zeta \;-\; x_0 \sinh\zeta\,, \end{aligned} \tag{4.108}$$

and similarly for y direction.

Aforestated boost transformations can be written in terms of boost matrices as

$$
\boldsymbol{L}_x = \begin{pmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
$$

$$
\boldsymbol{L}_z = \begin{pmatrix} \cosh\zeta & 0 & 0 & -\sinh\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\zeta & 0 & 0 & \cosh\zeta \end{pmatrix}, \tag{4.109}
$$

and similarly for $\boldsymbol{L}_y$.

Consider now a particle with mass m in its rest frame. In such a case it has a four-momentum $p^\mu = (m, 0, 0, 0)$. Let us now boost to the second frame, which moves with velocity v in the $-z$-direction of the first (i.e., $-v$ velocity in the positive z-axis direction). Then in the new frame the z momentum of the particle will appear increased. In particular, we can write this boost transformation from the rest frame in the form

$$
p' = \boldsymbol{L}_z p = \begin{pmatrix} \cosh\zeta & 0 & 0 & \sinh\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\zeta & 0 & 0 & \cosh\zeta \end{pmatrix} \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} E \\ 0 \\ 0 \\ q \end{pmatrix}, \tag{4.110}
$$

where $E = m\cosh\zeta = m\gamma$ and $q = m\sinh\zeta = mv\gamma$ (relativistic three-momentum). So, in this case, $\tanh\zeta = q/E = \beta$.

If ζ is infinitesimal (i.e., $|\zeta| \ll 1$), we can write the boost matrix in z direction as

$$
\boldsymbol{L}_z \approx \begin{pmatrix} 1 & 0 & 0 & \zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \zeta & 0 & 0 & 1 \end{pmatrix} = \boldsymbol{1} + \zeta \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{4.111}
$$

At the same time, we can use Eq. (3.52) and write

$$
\begin{aligned}
\boldsymbol{L}_z &\approx \boldsymbol{1} - \frac{i}{4}\omega_{\mu\nu}\boldsymbol{M}^{\mu\nu} = \boldsymbol{1} - \frac{i}{4}(\omega_{03}\boldsymbol{M}^{03} + \omega_{30}\boldsymbol{M}^{30}) \\
&= \boldsymbol{1} - \frac{i}{2}\omega_{03}\boldsymbol{M}^{03}.
\end{aligned} \tag{4.112}
$$

Here we used the fact that $\omega_{\mu\nu} = 0$ except for boost in 3rd direction where $\omega_{03} = -\omega_{30} \neq 0$. By using the relation $\boldsymbol{K}^i = \frac{1}{2}\boldsymbol{M}^{0i}$ we can identify ω_{03} with ζ and

Note that M^{03} is consistent with the defining formula (3.28).

$$
\boldsymbol{K}^3 = i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{M}^{03} = 2i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{4.113}
$$

As an independent check we can now go back to finite transformations, we finally get the rotated Dirac's wave function in the form

$$S(\boldsymbol{R})u(p) = \sqrt{E+m}\begin{pmatrix}\chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}_{\boldsymbol{R}}}{E+m}\chi_{\boldsymbol{R}}\end{pmatrix} = u_{\boldsymbol{R}}(p_{\boldsymbol{R}})\,. \tag{4.102}$$

So, indeed, we have

$$u(p) \xrightarrow{\boldsymbol{R}} u_{\boldsymbol{R}}(p) = S(\boldsymbol{R})u(\boldsymbol{R}^{-1}p)\,.$$

We thus see that $u_{\boldsymbol{R}}$ is constructed from the rotated spinor $\chi_{\boldsymbol{R}}$ and $p_{\boldsymbol{R}}$ is the rotated 4-momentum, i.e., $p_{\boldsymbol{R}} = (p_0, \boldsymbol{p}_{\boldsymbol{R}})$.

Let us now introduce Pauli spinors $\chi_{\frac{1}{2}} = \begin{pmatrix}1\\0\end{pmatrix}$ and $\chi_{-\frac{1}{2}} = \begin{pmatrix}0\\1\end{pmatrix}$ and denote them generally as χ_λ with $\lambda = \pm\frac{1}{2}$. Then

Here we change the former notation $\chi^{(\lambda)}, \lambda = 1, 2$, to the notation $\chi_\lambda, \lambda = \pm\frac{1}{2}$ that is more usual in situations when spin matters.

$$u_\lambda(p) = \sqrt{E+m}\begin{pmatrix}\chi_\lambda \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda\end{pmatrix}, \tag{4.103}$$

and $\lambda = \pm\frac{1}{2}$ (as seen in the box) has a meaning of the "third component of spin".

Action of Rotation Generators on Free Solutions of Dirac's Equation in the Rest Frame

Let us look what happens, when we apply generators of rotations (angular momentum operators) on free solutions of Dirac equation in the rest frame. We know that these solutions have general form

$$\psi^+_\lambda(x) = \sqrt{2m}\begin{pmatrix}\chi_\lambda\\0\end{pmatrix}\frac{e^{-ik_0t+i\boldsymbol{k}\cdot\boldsymbol{x}}}{\sqrt{2\pi}^3}\,,$$

$$\psi^-_\lambda(x) = \sqrt{2m}\begin{pmatrix}0\\\chi_\lambda\end{pmatrix}\frac{e^{ik_0t-i\boldsymbol{k}\cdot\boldsymbol{x}}}{\sqrt{2\pi}^3}\,,$$

with $\boldsymbol{k}\cdot\boldsymbol{x}$ being zero in the rest frame and $k_0 = m$. We also know that

$$\boldsymbol{J}_i \leftrightarrow \hat{\sigma} = \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix},$$

and in particular

$$\sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},$$

which leads to

$$\boldsymbol{J}_3\psi^+_{\frac{1}{2}} = \frac{1}{2}\psi^+_{\frac{1}{2}}\,, \quad \boldsymbol{J}_3\psi^+_{-\frac{1}{2}} = -\frac{1}{2}\psi^+_{-\frac{1}{2}}\,,$$

$$\boldsymbol{J}_3\psi^-_{\frac{1}{2}} = \frac{1}{2}\psi^-_{\frac{1}{2}}\,, \quad \boldsymbol{J}_3\psi^-_{-\frac{1}{2}} = -\frac{1}{2}\psi^-_{-\frac{1}{2}}\,.$$

These relations imply that Dirac's bispinors describe particle with spin 1/2.

4.3 Lorentz Boosts and Dirac Wave Function

Lorentz boost in x direction can be written as

$$\begin{aligned} t' &= \gamma(t - vx) \quad \Leftrightarrow \quad x'_0 = \gamma(x_0 - \beta x)\,, \\ x' &= \gamma(x - vt) \quad \Leftrightarrow \quad x' = \gamma(x - \beta x_0)\,, \\ y' &= y\,, \\ z' &= z\,, \end{aligned} \tag{4.104}$$

Here the system S' moves with respect to the system S with the velocity $v \equiv v_x$.

where $v = \beta$ and $-1 < \beta < 1$ (note that we use the notation $c = 1$).

The Lorentz transformations are often also written in a way that resembles rotations in 3D using hyperbolic functions. This is possible, because β and γ satisfy identity

$$\gamma^2 - \gamma^2\beta^2 = \gamma^2(1 - \beta^2) = 1\,, \tag{4.105}$$

which allows us to set $\gamma \equiv \cosh\zeta$ and $\gamma\beta \equiv \sinh\zeta$. With this, we can rewrite (4.104) equivalently as

$$\begin{aligned} x'_0 &= x_0 \cosh\zeta - x \sinh\zeta\,, \\ x' &= x \cosh\zeta - x_0 \sinh\zeta\,, \\ y' &= y\,, \\ z' &= z\,. \end{aligned} \tag{4.106}$$

In this connection we note that

$$\begin{aligned} \cosh\zeta &= \gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\tanh^2\zeta}}\,, \\ \tanh\zeta &= \frac{\sinh\zeta}{\cosh\zeta} = \beta \quad \Rightarrow \quad \zeta = \tanh^{-1}\beta\,. \end{aligned} \tag{4.107}$$

Since $\beta \in (-1, 1)$, we get that $\zeta \in (-\infty, \infty)$. The new variable ζ is known as *rapidity*.

Rapidity is a standard parameter that quantifies relativistic velocities in particle physics. In 1D motion, rapidities are additive whereas velocities must be combined via Einstein's velocity-addition formula.

Similar equations can also be written for boost in z direction

$$\begin{aligned} x'_0 &= x_0 \cosh\zeta - z \sinh\zeta\,, \\ x' &= x\,, \\ y' &= y\,, \\ z' &= z \cosh\zeta - x_0 \sinh\zeta\,, \end{aligned} \tag{4.108}$$

and similarly for y direction.

Aforestated boost transformations can be written in terms of boost matrices as

$$\boldsymbol{L}_x = \begin{pmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\boldsymbol{L}_z = \begin{pmatrix} \cosh\zeta & 0 & 0 & -\sinh\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\zeta & 0 & 0 & \cosh\zeta \end{pmatrix}, \tag{4.109}$$

and similarly for $\boldsymbol{L}_y$.

Consider now a particle with mass m in its rest frame. In such a case it has a four-momentum $p^\mu = (m, 0, 0, 0)$. Let us now boost to the second frame, which moves with velocity v in the $-z$-direction of the first (i.e., $-v$ velocity in the positive z-axis direction). Then in the new frame the z momentum of the particle will appear increased. In particular, we can write this boost transformation from the rest frame in the form

$$p' = \boldsymbol{L}_z p = \begin{pmatrix} \cosh\zeta & 0 & 0 & \sinh\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\zeta & 0 & 0 & \cosh\zeta \end{pmatrix} \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} E \\ 0 \\ 0 \\ q \end{pmatrix}, \tag{4.110}$$

where $E = m\cosh\zeta = m\gamma$ and $q = m\sinh\zeta = mv\gamma$ (relativistic three-momentum). So, in this case, $\tanh\zeta = q/E = \beta$.

If ζ is infinitesimal (i.e., $|\zeta| \ll 1$), we can write the boost matrix in z direction as

$$\boldsymbol{L}_z \approx \begin{pmatrix} 1 & 0 & 0 & \zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \zeta & 0 & 0 & 1 \end{pmatrix} = \mathbf{1} + \zeta \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{4.111}$$

At the same time, we can use Eq. (3.52) and write

$$\begin{aligned} \boldsymbol{L}_z &\approx \mathbf{1} - \frac{i}{4}\omega_{\mu\nu}\boldsymbol{M}^{\mu\nu} = \mathbf{1} - \frac{i}{4}(\omega_{03}\boldsymbol{M}^{03} + \omega_{30}\boldsymbol{M}^{30}) \\ &= \mathbf{1} - \frac{i}{2}\omega_{03}\boldsymbol{M}^{03}. \end{aligned} \tag{4.112}$$

Here we used the fact that $\omega_{\mu\nu} = 0$ except for boost in 3rd direction where $\omega_{03} = -\omega_{30} \neq 0$. By using the relation $\boldsymbol{K}_i = \frac{1}{2}\boldsymbol{M}^{0i}$ and the fact that

See the formula (3.28) for M^{0i} in fundamental representation.

$$\boldsymbol{K}_3 = i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{M}^{03} = 2i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{4.113}$$

we can identify ω_{03} with ζ. As an independent check we can now go back to finite transformations, i.e.

$$\begin{aligned} \boldsymbol{L}_z &= \exp\left(-\frac{i}{2}\omega_{03}\boldsymbol{M}^{03}\right) = \exp\left\{\zeta\begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&0\end{pmatrix}\right\} \\ &= \begin{pmatrix}\cosh\zeta&0&0&\sinh\zeta\\0&1&0&0\\0&0&1&0\\\sinh\zeta&0&0&\cosh\zeta\end{pmatrix}, \end{aligned} \tag{4.114}$$

where in the last step we used the fact that the matrix $\begin{pmatrix}0&1\\1&0\end{pmatrix}$ is involutive. This result indeed coincides with L_z given in (4.109), with $v \to -v$.

In the bispinor space the ensuing matrix responsible for boosts is given by

$$S(\boldsymbol{L}) = \exp\left\{-\frac{i}{2}\omega_{0i}\sigma^{0i}\right\}. \tag{4.115}$$

In particular, in the z direction we have

$$S(\boldsymbol{L}_z) = \exp\left\{-\frac{i}{2}\zeta\sigma^{03}\right\}, \tag{4.116}$$

where $\sigma^{03} = \frac{i}{2}\left[\gamma^0, \gamma^3\right] = i\gamma^0\gamma^3$, so

$$S(\boldsymbol{L}_z) = \exp\left\{-\frac{i}{2}\zeta i\gamma^0\gamma^3\right\} = \exp\left\{\frac{\zeta}{2}\gamma^0\gamma^3\right\}. \tag{4.117}$$

Note that the matrix $S(\boldsymbol{L}_z)$ is non-unitary. This is consequence of the fact that there are no finite dimensional unitary representations for non-compact groups (e.g. Lorentz group) of which boosts are examples.

Let us now compute the explicit form of $S(\boldsymbol{L}_z)$. Firstly, we observe that $(\gamma^0\gamma^3)^2 = \gamma^0\gamma^3\gamma^0\gamma^3 = -\gamma^0\gamma^0\gamma^3\gamma^3 = -\mathbf{1}(-\mathbf{1}) = \mathbf{1}$, which implies that $\gamma^0\gamma^3$ is an involutive matrix. By using the identity

$$e^{x\boldsymbol{N}} = (\cosh x)\,\mathbf{1} + (\sinh x)\,\boldsymbol{N}, \tag{4.118}$$

which is valid for any involutive matrix $\boldsymbol{N}$, we obtain

$$S(\boldsymbol{L_z}) = \cosh\frac{\zeta}{2} + \gamma^0\gamma^3\sinh\frac{\zeta}{2}. \tag{4.119}$$

Note that in bispinor representation $S(\boldsymbol{L})$, the arguments of hyperbolic sine and cosine are $\zeta/2$ whereas in fundamental representation $\boldsymbol{L}$, the arguments are just ζ.

Secondly, using *hyperbolic half-argument formulas*, one can write

$$\cosh\frac{\zeta}{2} = \sqrt{\frac{\cosh\zeta + 1}{2}} = \sqrt{\frac{E/m + 1}{2}} = \sqrt{\frac{E+m}{2m}}, \tag{4.120}$$

and

$$\begin{aligned} \sinh\frac{\zeta}{2} &= \sqrt{\frac{\cosh\zeta - 1}{2}} = \sqrt{\frac{E-m}{2m}} = \sqrt{\frac{E^2-m^2}{2m(E+m)}} \\ &= \sqrt{\frac{q^2}{2m(E+m)}}, \end{aligned} \tag{4.121}$$

(since $\cosh\zeta = \gamma$ and $E = \gamma m$ we used $\cosh\zeta = E/m$). Here m is rest mass and q relativistic three-momentum. From (4.120) and (4.121) we can read off that

$$S(\boldsymbol{L}_z) = \sqrt{\frac{E+m}{2m}}\left[\mathbb{1} + \frac{q}{E+m}\gamma^0\gamma^3\right]. \tag{4.122}$$

Thirdly, we observe that in Dirac's representation

$$\gamma^0\gamma^3 = (\sigma^3 \otimes \mathbb{1})(i\sigma^2 \otimes \sigma^3) = (\sigma^3 i\sigma^2 \otimes \sigma^3) = \begin{pmatrix} \mathbf{0} & \sigma^3 \\ \sigma^3 & \mathbf{0} \end{pmatrix}. \tag{4.123}$$

Consequently, we can write the bispinor representation of boost transformation in z-direction as

$$S(\boldsymbol{L}_z) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} \mathbb{1} & \frac{q}{E+m}\sigma^3 \\ \frac{q}{E+m}\sigma^3 & \mathbb{1} \end{pmatrix}. \tag{4.124}$$

This result might be generalized to boost in general 3-velocity $\boldsymbol{v} = \boldsymbol{q}/E$ direction. Ensuing transformation takes the form

$$S(\boldsymbol{L}) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} \mathbb{1} & \frac{\boldsymbol{q}}{E+m}\boldsymbol{\sigma} \\ \frac{\boldsymbol{q}}{E+m}\boldsymbol{\sigma} & \mathbb{1} \end{pmatrix}. \tag{4.125}$$

As an example, we can consider a particle in the rest frame, i.e., with four-momentum $\bar{p}^\mu = \begin{pmatrix} m \\ \mathbf{0} \end{pmatrix}$. In this case the Dirac spinor is

$$u_\lambda(\bar{p}) = \sqrt{2m}\begin{pmatrix} \chi_\lambda \\ 0 \end{pmatrix}. \tag{4.126}$$

When boosting system to a frame with velocity $-v$ in z-direction we get

$$p = \boldsymbol{L}_z\bar{p} = \begin{pmatrix} E \\ 0 \\ 0 \\ q \end{pmatrix}, \tag{4.127}$$

and

$$u_\lambda(p) = S(\boldsymbol{L}_z)u_\lambda(\bar{p}) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} \mathbf{1} & \frac{q\sigma^3}{E+m} \\ \frac{q\sigma^3}{E+m} & \mathbf{1} \end{pmatrix}\sqrt{2m}\begin{pmatrix} \chi_\lambda \\ 0 \end{pmatrix}. \tag{4.128}$$

Consequently

$$u_\lambda(P) = \sqrt{E+m}\begin{pmatrix} \chi_\lambda \\ \frac{q\sigma^3}{E+m}\chi_\lambda \end{pmatrix}, \tag{4.129}$$

which coincides with our former result (4.63).

4.4 Spin Sums and Projection Operators

Let us recall that negative energy solution were plane-waves of the form

$$v_\lambda(p)e^{ipx} = v_\lambda e^{i\omega_p t - i\boldsymbol{p}\cdot\boldsymbol{x}}\,, \tag{4.130}$$

with amplitude $v(p)$ satisfying

$$(\gamma^\mu p_\mu + m)v(p) = 0 \quad \Rightarrow \quad v(p) = \sqrt{E+m}\begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda \\ \chi_\lambda \end{pmatrix}. \tag{4.131}$$

We will now show that the following "ortho-normality" relations hold

$$\begin{aligned}
\overline{u}_\lambda(p)u_{\lambda'}(p) &= 2m\delta_{\lambda\lambda'}\,, \\
\overline{v}_\lambda(p)v_{\lambda'}(p) &= -2m\delta_{\lambda\lambda'}\,, \\
\overline{u}_\lambda(p)v_{\lambda'}(p) &= 0\,, \\
\overline{v}_\lambda(p)u_{\lambda'}(p) &= 0\,,
\end{aligned} \tag{4.132}$$

where $\overline{u}_\lambda(p) = u_\lambda^\dagger(p)\gamma^0$, $\overline{v}_\lambda(p) = v_\lambda^\dagger(p)\gamma^0$.

We already know that $\overline{\psi}(x)\psi(x)$ is a Lorentz scalar, it is not difficult to see that also $\overline{u}_\lambda(p)u_{\lambda'}(p)$, $\overline{v}_\lambda(p)v_{\lambda'}(p)$, and $\overline{u}_\lambda(p)v_{\lambda'}(p)$ are Lorentz scalars. In fact, e.g.

$$\begin{aligned}
\overline{u}_{L,\lambda}(p)v_{L,\lambda}(p) &= \overline{u}_\lambda(\boldsymbol{L}^{-1}p)S^{-1}(\boldsymbol{L})S(\boldsymbol{L})v_\lambda(\boldsymbol{L}^{-1}p) \\
&= \overline{u}_\lambda(\boldsymbol{L}^{-1}p)v_\lambda(\boldsymbol{L}^{-1}p)\,.
\end{aligned} \tag{4.133}$$

So, in order to prove (4.132) we can go to the rest frame where bisponors acquire the simple form (cf. Eq. (4.66))

$$\begin{aligned}
u_{1/2}(m,\boldsymbol{0}) &= \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad u_{-1/2}(m,\boldsymbol{0}) = \sqrt{2m}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \\
v_{1/2}(m,\boldsymbol{0}) &= \sqrt{2m}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad v_{-1/2}(m,\boldsymbol{0}) = \sqrt{2m}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}.
\end{aligned} \tag{4.134}$$

From these results (4.132) follow automatically.

Upshot of (4.132) is that the four bispinors $u_\lambda(p)$ and $v_\lambda(p)$ are linearly independent and so they form a complete basis in the space of Dirac spinors.

Positive Definite Norm

Normalization in (4.132) is Lorentz invariant, however, the *norm* is *not positive-definite*. In fact, we know that the positive-definite

4.4 Spin Sums and Projection Operators

Let us recall that negative energy solution were plane-waves of the form

$$v_\lambda(p)e^{ipx} = v_\lambda e^{i\omega_p t - i\boldsymbol{p}\cdot\boldsymbol{x}}\,, \tag{4.130}$$

with amplitude $v(p)$ satisfying

$$(\gamma^\mu p_\mu + m)v(p) = 0 \quad \Rightarrow \quad v(p) = \sqrt{E+m}\begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda \\ \chi_\lambda \end{pmatrix}. \tag{4.131}$$

We will now show that the following "ortho-normality" relations hold

$$\begin{aligned} \bar{u}_\lambda(p)u_{\lambda'}(p) &= 2m\delta_{\lambda\lambda'}\,, \\ \bar{v}_\lambda(p)v_{\lambda'}(p) &= -2m\delta_{\lambda\lambda'}\,, \\ \bar{u}_\lambda(p)v_{\lambda'}(p) &= 0\,, \\ \bar{v}_\lambda(p)u_{\lambda'}(p) &= 0\,, \end{aligned} \tag{4.132}$$

where $\bar{u}_\lambda(p) = u^\dagger_\lambda(p)\gamma^0$, $\bar{v}_\lambda(p) = v^\dagger_\lambda(p)\gamma^0$.

We already know that $\bar{\psi}(x)\psi(x)$ is a Lorentz scalar, it is not difficult to see that also $\bar{u}_\lambda(p)u_{\lambda'}(p)$, $\bar{v}_\lambda(p)v_{\lambda'}(p)$, and $\bar{u}_\lambda(p)v_{\lambda'}(p)$ are Lorentz scalars. In fact, e.g.

$$\begin{aligned} \bar{u}_{L,\lambda}(p)v_{L,\lambda}(p) &= \bar{u}_\lambda(\boldsymbol{L}^{-1}p)S^{-1}(\boldsymbol{L})S(\boldsymbol{L})v_\lambda(\boldsymbol{L}^{-1}p) \\ &= \bar{u}_\lambda(\boldsymbol{L}^{-1}p)v_\lambda(\boldsymbol{L}^{-1}p)\,. \end{aligned} \tag{4.133}$$

So, in order to prove (4.132) we can go to the rest frame where bisponors acquire the simple form (cf. Eq. (4.66))

$$\begin{aligned} u_{1/2}(m,\boldsymbol{0}) &= \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad u_{-1/2}(m,\boldsymbol{0}) = \sqrt{2m}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \\ v_{1/2}(m,\boldsymbol{0}) &= \sqrt{2m}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad v_{-1/2}(m,\boldsymbol{0}) = \sqrt{2m}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \end{aligned} \tag{4.134}$$

From these results (4.132) follow automatically.

Upshot of (4.132) is that the four bispinors $u_\lambda(p)$ and $v_\lambda(p)$ are linearly independent and so they form a complete basis in the space of Dirac spinors.

Positive Definite Norm

Normalization in (4.132) is Lorentz invariant, however, the *norm* is *not positive-definite*. In fact, we know that the positive-definite density per unit volume is

$$\rho(p) = J^0(p) = \bar{\psi}_p(x)\gamma^0\psi_p(x)\,.$$

This indicates that instead of (4.132) one could use scalar product with γ^0 insertion. This would indeed do, e.g.

$$\begin{aligned}\bar{\psi}^{(+)}_{\lambda,p}(x)\gamma^0\psi^{(+)}_{\lambda',p}(x) &= \bar{u}_\lambda(p)\gamma^0 u_{\lambda'}(p)\\ &= \frac{1}{2m}\left[\bar{u}_\lambda(p)\gamma^0\not{p}u_{\lambda'}(p) + \bar{u}_\lambda(p)\not{p}\gamma^0 u_{\lambda'}(p)\right]\\ &= \frac{1}{2m}\bar{u}_\lambda(p)\{\gamma^0,\not{p}\}u_{\lambda'}(p) = \frac{p^0}{m}\delta_{\lambda\lambda'}\,.\end{aligned}$$

Similarly, we can show that

$$\bar{v}_\lambda(p)\gamma^0 v_{\lambda'}(p) = \frac{p^0}{m}\delta_{\lambda\lambda'} \quad \text{and} \quad \bar{v}_\lambda(\tilde{p})\gamma^0 u_{\lambda'}(p) = 0\,,$$

with $\tilde{p}^\mu = (p^0, -\boldsymbol{p})$. In deriving these relations we used that

$$\begin{aligned}\bar{u}_\lambda(p)(\not{p} - m) = 0 &\quad\Leftrightarrow\quad \frac{1}{m}\bar{u}_\lambda(p)\not{p} = \bar{u}_\lambda(p)\,,\\ \bar{v}_\lambda(p)(\not{p} + m) = 0 &\quad\Leftrightarrow\quad -\frac{1}{m}\bar{v}_\lambda(p)\not{p} = \bar{v}_\lambda(p)\,.\end{aligned}$$

For purpose of this section, however, the "scalar product" defined via (4.132) will suffice because *a*) it is simpler and *b*) the semi-definiteness of the norm will not play any conceptual role in our reasonings.

To proceed, we observe that the following relations hold

$$\begin{aligned}(\not{p} - m)u_\lambda(p) &= 0\,,\\ (\not{p} + m)v_\lambda(p) &= 0\,,\\ (\not{p} - m)v_\lambda(p) &= (\not{p} + m - 2m)v_\lambda(p) = -2mv_\lambda(p)\,,\\ (\not{p} + m)u_\lambda(p) &= (\not{p} - m + 2m)u_\lambda(p) = 2mu_\lambda(p)\,. \qquad (4.135)\end{aligned}$$

Let us now consider the operators

Note that $\sum_\lambda u_\lambda(p)\overline{u}_\lambda(p)$ is in fact a short-hand notation for the tensor product $\sum_\lambda u_\lambda(p)\otimes\overline{u}_\lambda(p)$. Similarly also for $\tilde{\Lambda}^-(p)$.

$$\begin{aligned}\tilde{\Lambda}^+(p) &= \sum_\lambda u_\lambda(p)\overline{u}_\lambda(p)\,,\\ \tilde{\Lambda}^-(p) &= \sum_\lambda v_\lambda(p)\overline{v}_\lambda(p)\,. \qquad (4.136)\end{aligned}$$

When applied to positive and negative-energy bisponors they yield

$$\begin{aligned}\tilde{\Lambda}^+(p)u_\lambda(p) &= \sum_{\lambda'} u_{\lambda'}(p)\overline{u}_{\lambda'}(p)u_\lambda(p) = 2m\sum_{\lambda'} u_{\lambda'}(p)\delta_{\lambda\lambda'} = 2mu_\lambda(p)\,,\\ \tilde{\Lambda}^+(p)v_\lambda(p) &= \sum_{\lambda'} u_{\lambda'}(p)\underbrace{\overline{u}_{\lambda'}(p)v_\lambda(p)}_{0} = 0\,. \qquad (4.137)\end{aligned}$$

Similarly

$$\tilde{\Lambda}^-(p)u_\lambda(p) = 0\,, \quad \text{and} \quad \tilde{\Lambda}^-(p)v_\lambda(p) = -2m v_\lambda(p)\,. \tag{4.138}$$

By comparing these relations with operators $(\gamma^\mu p_\mu \pm m)$ we can make the following identification:

$$\begin{aligned} \tilde{\Lambda}^+(p) &= \sum_\lambda u_\lambda(p)\overline{u}_\lambda(p) = \gamma^\mu p_\mu + m\,, \\ \tilde{\Lambda}^-(p) &= \sum_\lambda v_\lambda(p)\overline{v}_\lambda(p) = \gamma^\mu p_\mu - m\,. \end{aligned} \tag{4.139}$$

Let us now define two new operators

$$\begin{aligned} \Lambda^+(p) &\equiv \frac{\tilde{\Lambda}^+(p)}{2m} = \frac{\gamma^\mu p_\mu + m}{2m}\,, \\ \Lambda^-(p) &\equiv -\frac{\tilde{\Lambda}^-(p)}{2m} = \frac{-(\gamma^\mu p_\mu - m)}{2m}\,. \end{aligned} \tag{4.140}$$

It can be easily recognized that these operators are projection operators as they fulfill the necessary conditions for projection operators, in particular:

- $$\begin{aligned} \left[\Lambda^\pm(p)\right]^2 &= \left[\frac{\pm(\gamma^\mu p_\mu \pm m)}{2m}\right]^2 = \frac{(\gamma^\mu p_\mu)^2 \pm 2\gamma^\mu p_\mu m + m^2}{4m^2} \\ &= \frac{\pm(\gamma^\mu p_\mu \pm m)}{2m} = \Lambda^\pm(p)\,, \end{aligned}$$
- $\Lambda^+(p) + \Lambda^-(p) = \mathbb{1}\,,$
- $\Lambda^\pm\Lambda^\mp = 0\,.$

Let us also note that

$$\begin{aligned} \mathrm{Tr}\left(\Lambda^\pm(p)\right) &= \frac{\pm\mathrm{Tr}(\gamma^\mu p_\mu)}{2m} + 2 \\ &= \frac{\pm p_\mu \mathrm{Tr}\,\gamma^\mu}{2m} + 2 = 2\,. \end{aligned} \tag{4.141}$$

Projection operators Λ^+ and Λ^- project over positive and negative energy states, respectively.

Exercises: Dirac equation — technical developments I

Dirac field bilinears

Dirac wave-functions have 4 complex components,

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \in \mathbb{C}^4\,.$$

Apart from the usual Hermitian conjugation $\psi^\dagger(x) = (\psi_1^*(x), \psi_2^*(x), \psi_3^*(x), \psi_4^*(x))$, we also define the Dirac conjugation $\bar{\psi}(x) = \psi^\dagger(x)\gamma^0$.

Exercise 4.1 Show that in the standard representation of γ-matrices, Eq. (3.13), (as well as in any unitarily equivalent representation) the Hermitian conjugation acts as follows:

$$
\begin{aligned}
&1) \quad (\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0\,,\\
&2) \quad (\sigma^{\mu\nu})^\dagger = \gamma^0\sigma^{\mu\nu}\gamma^0\,,\\
&3) \quad (\gamma^5)^\dagger = \gamma^5\,.
\end{aligned}
$$

Exercise 4.2 Show that the expression

$$\bar{\psi}(x)[\gamma^\mu,\gamma^\nu]\psi(x)\,,$$

transforms as a (skew-symmetric) tensor field under Lorentz transformations.

[**Hint:** Recall Eq. (4.12).]

Exercise 4.3 Show that the expression

$$\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x) \tag{4.142}$$

transforms as a pseudovector field under Lorentz transformations.

[**Hint:** Use the identity $\varepsilon_{\mu\nu\rho\sigma}L^\mu_{\ \alpha}L^\nu_{\ \beta}L^\rho_{\ \gamma}L^\sigma_{\ \delta} = (\det \boldsymbol{L})\,\varepsilon_{\alpha\beta\gamma\delta}$ for the determinant of a matrix $\boldsymbol{L}$.]

Exercise 4.4 Show that the set of matrices

$$\{\mathbf{1},\ \gamma^\mu,\ [\gamma^\mu,\gamma^\nu],\ \gamma^5\gamma^\mu,\ \gamma^5\}_{\mu,\nu=0,1,2,3}\,,$$

is a basis of the vector space $\mathbb{C}^{4,4}$ (4×4 complex matrices).

[**Hint:** Make use of the basic properties of Γ_i matrices.]

Dirac equation and its solutions

The Dirac equation, and its Dirac-conjugated equation read

$$(i\gamma^\mu\partial_\mu - m)\psi = 0\,, \qquad (\partial_\mu\bar{\psi})i\gamma^\mu + m\bar{\psi} = 0\,.$$

Exercise 4.5 Verify that the *axial current*

$$J_5^\mu(x) = \bar{\Psi}(x)\gamma^\mu\gamma^5\psi(x)\,,$$

satisfies the equation

$$\partial_\mu J_5^\mu = 2im\bar{\psi}\gamma^5\psi\,,$$

whenever $\psi(x)$ is a solution of the Dirac equation.

– variational principle

In field theory, the equations of motion for an n-component field $\phi_a(x)$ follow from the *stationary action principle*:

$$
\begin{aligned}
S[\phi_a(x)+\delta\phi_a(x)] - S[\phi_a(x)] &\approx \delta S[\phi_a(x)] = 0\,,\\
S[\phi_a(x)] &= \int d^4x\, L(\phi_a(x),\partial_\mu\phi_a(x))\,,
\end{aligned}
$$

where L is the *Lagrangian density*, and where variations of the field $\delta\phi_a(x)$ vanish on the boundary. The equations of motion are then the corresponding *Euler–Lagrange equations*

$$\partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu \phi_a)} \right) - \frac{\partial L}{\partial \phi_a} = 0 .$$

Exercise 4.6 Consider the Lagrangian density

$$L(\psi, \psi^*, \partial_\mu \psi, \partial_\mu \psi^*) = -\frac{\hbar^2}{2m}(\partial_j \psi^*)(\partial_j \psi) + \frac{i\hbar}{2}(\psi^* \partial_t \psi - \psi \partial_t \psi^*) - V(\boldsymbol{x}, t)\psi^* \psi .$$

Show that the stationary action principle for the action $S[\psi, \psi^*] = \int L\, d^4x$ yields the Schrödinger equation (and its complex conjugate) as its Euler–Lagrange equations.

[**Hint:** Regard ψ and ψ^* as two independent fields.]

Exercise 4.7 Consider the action

$$\begin{aligned} S[\psi, \bar{\psi}] &= \int d^4x\, \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \\ &= \int d^4x\, \bar{\psi}_\alpha(x)(i\gamma^\mu \partial_\mu - m\mathbf{1})_{\alpha\beta}\psi_\beta(x) , \end{aligned}$$

(where $\alpha, \beta = 1, 2, 3, 4$) and show that the stationary action principle yields the Dirac equation and its Dirac conjugate.

– solutions of the Dirac equation

The Dirac equation (4.4) is solved by harmonic waves

$$\begin{aligned} \psi_p^{(+)}(x) &= e^{-ip_\mu x^\mu} u(p) \quad \text{(positive energy)} , \\ \psi_p^{(-)}(x) &= e^{ip_\mu x^\mu} v(p) \quad \text{(negative energy)} , \end{aligned}$$

where $p_0 \equiv \omega_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2} > 0$, and the *polarization spinors* $u(p)$ and $v(p)$ are solutions of the algebraic equations

$$(\not{p} - m)u(p) = 0 , \qquad (\not{p} + m)v(p) = 0 .$$

Exercise 4.8 Working in the standard representation of γ-matrices, show that the equation

$$(\not{p} - m)u(p) = 0 , \qquad E \equiv p_0 \equiv \omega_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2} ,$$

has 2 linearly independent solutions, and they can be written as

$$u(p) = \begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m}\chi \end{pmatrix} , \qquad \chi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

Analogously, equation $(\not{p} + m)v(p) = 0$ has 2 linearly independent solutions

$$v(p) = \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E+m}\chi \\ \chi \end{pmatrix} , \qquad \chi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

[**Hint:** Note that $(\not{p} - m)(\not{p} + m)w = 0$ for any $w \in \mathbb{C}^4$.]

Exercise 4.9 Write massless ($m = 0$) Dirac equation explicitly in the standard representation of γ-matrices, and find similarity with the Maxwell equations of electrodynamics.

– Lorentz transformations of Dirac wave-functions

Recall the fundamental representation, and the spin representation of the Lorentz group:

$$\boldsymbol{L} = \exp\left(-\frac{i}{4}\omega_{\mu\nu}\boldsymbol{M}^{\mu\nu}\right), \qquad S(\boldsymbol{L}) = \exp\left(-\frac{i}{4}\omega_{\mu\nu}\boldsymbol{\sigma}^{\mu\nu}\right).$$

Similarly to the rotation and boost generators in the fundamental representation, the rotation and boost generators in the bi-spinor representation are, respectively,

$$2\hat{\boldsymbol{\sigma}}_i = \Sigma_i = \frac{1}{2}\varepsilon_{ijk}\boldsymbol{\sigma}^{jk}, \qquad \boldsymbol{\sigma}^{0i} = -\boldsymbol{\sigma}^{i0}.$$

A rotation $S(R)$ in the bi-spinor representation can be specified by 3 parameters $\theta_1, \theta_2, \theta_3$ ($\omega_{jk} = \varepsilon_{ijk}\theta^i$, $\omega_{0i} = 0$) as

$$S(\boldsymbol{R}) = e^{-\frac{i}{4}\omega_{jk}\boldsymbol{\sigma}^{jk}} = e^{-\frac{i}{2}\theta^i\Sigma_i}.$$

Exercise 4.10 Show that the identity

$$S(\boldsymbol{L})^{-1}\gamma^\mu S(\boldsymbol{L}) = \boldsymbol{L}^\mu{}_\nu\gamma^\nu,$$

when specialized to rotations $S(\boldsymbol{R})$, reduces to Eq. (4.101).

[**Hint:** Use the standard representation of γ-matrices, where $\Sigma^i = \mathbb{I}\otimes\sigma^i$.]

Exercise 4.11 Consider a boost in the x-direction by velocity β. This Lorentz transformation is described by the matrix

$$\boldsymbol{L}_x(\zeta) = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\gamma = \sqrt{1-\beta^2}$, and ζ is the *rapidity* of the boost. Cast $\boldsymbol{L}$ in an exponential form to find the corresponding generator.

Exercise 4.12 By composing two boosts in the x-direction, with rapidities ζ_1 and ζ_2, respectively, derive the relativistic formula for the addition of velocities.

[**Hint:** Use the identities for hyperbolic functions: $\sinh(\zeta_1+\zeta_2) = \sinh\zeta_1\cosh\zeta_2 + \sinh\zeta_2\cosh\zeta_1$, and $\cosh(\zeta_1+\zeta_2) = \cosh\zeta_1\cosh\zeta_2 + \sinh\zeta_1\sinh\zeta_2$.]

– matrix exponentials

The exponential of a matrix $\mathbb{A}$ is defined by the series

$$e^{\mathbb{A}} = \sum_{n=0}^{\infty}\frac{\mathbb{A}^n}{n!}.$$

The rule $e^{\mathbb{A}+\mathbb{B}} = e^{\mathbb{A}}e^{\mathbb{B}}$ holds true when $[\mathbb{A},\mathbb{B}] = 0$, but otherwise one has to use the Baker–Campbell–Hausdorff formula in one of the variants

$$e^{\mathbb{A}+\mathbb{B}} = e^{\mathbb{A}}e^{\mathbb{B}}e^{-\frac{1}{2}[\mathbb{A},\mathbb{B}]}\dots,$$

$$e^{\mathbb{A}}e^{\mathbb{B}} = e^{\mathbb{A}+\mathbb{B}+\frac{1}{2}[\mathbb{A},\mathbb{B}]+\cdots},$$

where '...' denotes terms with ever more nested commutators.

Exercise 4.13 Take

$$\mathbb{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

and calculate a) $e^{\mathbb{A}+\mathbb{B}}$, b) $e^{\mathbb{A}}e^{\mathbb{B}}$, c) $[\mathbb{A}, \mathbb{B}]$.

Suppose $\mathbb{A}(\alpha)$ is matrix-valued function of α. What is the derivative of $e^{\mathbb{A}(\alpha)}$? The general formula reads

$$\frac{d}{d\alpha}e^{\mathbb{A}(\alpha)} = \int_0^1 e^{\lambda\mathbb{A}(\alpha)}\frac{d\mathbb{A}(\alpha)}{d\alpha}e^{(1-\lambda)\mathbb{A}(\alpha)}d\lambda\,,$$

which reduces to a simple form $\frac{d}{d\alpha}e^{\mathbb{A}(\alpha)} = \frac{d\mathbb{A}(\alpha)}{d\alpha}e^{\mathbb{A}(\alpha)}$ when $[\mathbb{A}(\alpha), \mathbb{A}(\alpha')] = 0$ (e.g., when $\mathbb{A}(\alpha) = \alpha\mathbb{A}_0$).

Exercise 4.14 Take

$$\mathbb{A}(\alpha) = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix},$$

and calculate:

a) the explicit form of $e^{\mathbb{A}(\alpha)}$,
b) the derivative of the result of a),
c) the expression $\frac{d\mathbb{A}(\alpha)}{d\alpha}e^{\mathbb{A}(\alpha)}$,
d) the derivative of $e^{\mathbb{A}(\alpha)}$.

Dirac Equation — applications I 5

5.1 Electromagnetic Coupling of Electrons

Non-relativistic Charged Particle

For a free particle of mass m, the Hamiltonian is simply $H_0 = \boldsymbol{p}^2/2m$. If the particle has charge q, then in the presence of an external electromagnetic field the Hamiltonian changes to

The change from $\boldsymbol{p}$ to $\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}$ is known as *minimal substitution* and the ensuing interaction term is known as *minimal coupling*.

$$H = \frac{\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right)^2}{2m} + q\phi\,, \tag{5.1}$$

where $\boldsymbol{A}$ and ϕ are the vector and scalar potential, respectively.

Moreover, we also have to insist on a gauge condition which can be taken in the form $\nabla \cdot \boldsymbol{A} = 0$ (so-called Coulomb gauge). The magnetic field is connected with the vector potential via $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ while scalar potential with electric field via $\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t}$. So, in particular, situation with zero electric field and constant magnetic field $\boldsymbol{B} = (0, 0, B)$ we can set $\boldsymbol{A} = \frac{1}{2}B(-y, x, 0)$ and $\phi = 0$.

In this case we can write the Hamiltonian (5.1) to the first order in q as

$$H = \frac{1}{2m}\left[\boldsymbol{p}^2 - \frac{q}{c}(\boldsymbol{p}\boldsymbol{A} + \boldsymbol{A}\boldsymbol{p})\right] + O(q^2)\,. \tag{5.2}$$

Generally, the commutator of a function f with the derivation $\left[\frac{d}{dx}, f\right]$ is equal to $\frac{df}{dx}$ since (operating on a test function u)

$$\left[\frac{d}{dx}, f\right]u = \left(\frac{d}{dx}(fu)\right) - f\frac{du}{dx} = \\ = \left(\frac{df}{dx}\right)u + f\frac{du}{dx} - f\frac{du}{dx} = \left(\frac{df}{dx}\right)u.$$

In quantum mechanics, $\boldsymbol{p}\boldsymbol{A} \neq \boldsymbol{A}\boldsymbol{p}$, but we can relate these two expressions via $p^i A^i = A^i p^i + \left[p^i, A^i\right]$, where $\left[p^i, A^i\right] = -i\hbar\left[\nabla_i, A^i\right] = -i\hbar\nabla \cdot \boldsymbol{A} = 0$ (in Coulomb gauge). Then

$$H = \frac{\boldsymbol{p}^2}{2m} - \frac{q}{mc}\boldsymbol{A}\cdot\boldsymbol{p} + O(q^2)\,. \tag{5.3}$$

By evaluating explicitly $\boldsymbol{A}\cdot\boldsymbol{p}$ we get

$$\boldsymbol{A}\cdot\boldsymbol{p} = \frac{1}{2}B(-y, x, 0)(p_x, p_y, p_z) = \frac{B}{2}(xp_y - yp_x) = \frac{B}{2}L_z\,. \tag{5.4}$$

Here L_z is 3rd component of the (orbital) angular momenta. Due to rotational invariance of the experimental setup we can easily generalize this result to arbitrarily oriented $\boldsymbol{B}$. In particular, we can write $\boldsymbol{A}\cdot\boldsymbol{p} = \boldsymbol{L}\cdot\boldsymbol{B}/2$. Consequently, we get that the Hamiltonian (5.3) is composed of two parts $H = H_0 + H_{\text{EM}}$, where

$$H_{\text{EM}} = -\frac{q}{2mc}\boldsymbol{B}\cdot\boldsymbol{L}\,. \tag{5.5}$$

Particularly for electron it is conventional to write this in the form ($q = e$ and $m = m_e$)

$$H_{\text{EM}} = -\frac{e\hbar}{2m_e c}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} = -g_L\mu_{\text{B}}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} = \boldsymbol{\mu}\cdot\boldsymbol{B}\,, \tag{5.6}$$

Dirac Equation — Applications I

5

5.1 Electromagnetic Coupling of Electrons

Non-relativistic Charged Particle

For a free particle of mass m, the Hamiltonian is simply $H_0 = \boldsymbol{p}^2/2m$. If the particle has charge q, then in the presence of an external electromagnetic field the Hamiltonian changes to

$$H = \frac{\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right)^2}{2m} + q\phi , \tag{5.1}$$

where $\boldsymbol{A}$ and ϕ are the vector and scalar potential, respectively.

The change from $\boldsymbol{p}$ to $\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}$ is known as *minimal substitution* and the ensuing interaction term is known as *minimal coupling*.

Moreover, we also have to insist on a gauge condition which can be taken in the form $\nabla \cdot \boldsymbol{A} = 0$ (so-called Coulomb gauge). The magnetic field is connected with the vector potential via $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ while scalar potential with electric field via $\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t}$. So, in particular, in situation with zero electric field and constant magnetic field $\boldsymbol{B} = (0, 0, B)$ we can set $\boldsymbol{A} = \frac{1}{2}B(-y, x, 0)$ and $\phi = 0$.

In this case we can write the Hamiltonian (5.1) to the first order in q as

$$H = \frac{1}{2m}\left[\boldsymbol{p}^2 - \frac{q}{c}(\boldsymbol{pA} + \boldsymbol{Ap})\right] + O(q^2) . \tag{5.2}$$

In quantum mechanics, $\boldsymbol{pA} \neq \boldsymbol{Ap}$, but we can relate these two expressions via $p^i A^i = A^i p^i + \left[p^i, A^i\right]$, where $\left[p^i, A^i\right] = -i\hbar\left[\nabla_i, A^i\right] = -i\hbar\nabla\cdot\boldsymbol{A} = 0$ (in Coulomb gauge). Then

Generally, the commutator of a function f with the derivation $\left[\frac{d}{dx}, f\right]$ is equal to $\frac{df}{dx}$ since (operating on a test function u)

$$\left[\frac{d}{dx}, f\right]u = \left(\frac{d}{dx}(fu)\right) - f\frac{du}{dx} = \\ = \left(\frac{df}{dx}\right)u + f\frac{du}{dx} - f\frac{du}{dx} = \left(\frac{df}{dx}\right)u .$$

$$H = \frac{\boldsymbol{p}^2}{2m} - \frac{q}{mc}\boldsymbol{A}\cdot\boldsymbol{p} + O(q^2) . \tag{5.3}$$

By evaluating explicitly $\boldsymbol{A}\cdot\boldsymbol{p}$ we get

$$\boldsymbol{A}\cdot\boldsymbol{p} = \frac{1}{2}B(-y, x, 0)(p_x, p_y, p_z) = \frac{B}{2}(xp_y - yp_x) = \frac{B}{2}L_z . \tag{5.4}$$

Here L_z is 3rd component of the (orbital) angular momenta. Due to rotational invariance of the experimental setup we can easily generalize this result to arbitrarily oriented $\boldsymbol{B}$. In particular, we can write $\boldsymbol{A}\cdot\boldsymbol{p} = \boldsymbol{L}\cdot\boldsymbol{B}/2$. Consequently, we get that the Hamiltonian (5.3) is composed of two parts $H = H_0 + H_{\mathrm{I}} + H_{\mathrm{EM}}$, where

$$H_{\mathrm{I}} = -\frac{q}{2mc}\boldsymbol{B}\cdot\boldsymbol{L} . \tag{5.5}$$

Particularly for electron it is conventional to write this in the form ($q = e$ and $m = m_e$)

$$H_{\mathrm{I}} = -\frac{e\hbar}{2m_e c}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} = -g_L\mu_{\mathrm{B}}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} = \boldsymbol{\mu}\cdot\boldsymbol{B} , \tag{5.6}$$

where $\mu_B = e\hbar/(2m_e c)$ is the *Bohr magneton*. The factor g_L is known a the g-factor (here $g_L = 1$) and $\boldsymbol{\mu}$ the *orbital magnetic moment* of an electron. The term $g_L\mu_B$ is known as the *gyromagnetic ratio*.

Experimental atomic physics tells us that:

- electrons have spin 1/2
- gyromagnetic ratio for spin of electron (and more generally 1/2 particle) is twice the gyromagnetic ratio for orbital angular momentum (i.e. $g_s = 2g_L$).

The great achievement of Dirac was to show that his equation allows to consistently handle quantum theory of electron and to incorporate both aforementioned experimental observations. We will also see that Dirac's equation can do even more, namely it ensures that to each 1/2-particle must exists *antiparticle* and it correctly predicts *fine structure* in the spectrum of hydrogen atom.

Dirac's Charged Particle

Let us now derive the gyromagnetic ratio for spin 1/2 particle directly from Dirac's equation.

When considering a relativistic situation, we have to work with the four-vector gauge potential A_μ. The minimal coupling prescription is then given by

$$p_\mu \to p_\mu - qA_\mu \quad \Rightarrow \quad \partial_\mu \to \partial_\mu + iqA_\mu\,, \tag{5.7}$$

where $p_\mu = i\frac{\partial}{\partial x^\mu}$, respectively. It is easy to see that the space components of this prescription provide consistent minimal prescription known from non-relativistic physics, namely

$$p_i \to p_i - qA_i \;\Leftrightarrow\; p^i \to p^i - qA^i \;\Leftrightarrow\; -i\boldsymbol{\nabla} \to -i\boldsymbol{\nabla} - q\boldsymbol{A}\,. \tag{5.8}$$

With this notation the Dirac equation takes the form

Typical non-minimal coupling that could be added to (5.9) would be $\boldsymbol{\sigma}^{\mu\nu}\boldsymbol{F}_{\mu\nu}$. Non-minimal coupling terms typically lead to dipole and higher multipole terms.

$$(i\partial\!\!\!/ - qA\!\!\!/ - m)\psi(x) = 0\,. \tag{5.9}$$

This can be further rewritten in the Schrödinger-like form

$$\begin{aligned} i\frac{\partial\psi}{\partial t} &= [\boldsymbol{\alpha}(-i\boldsymbol{\nabla} - q\boldsymbol{A}) + \beta m + q\phi]\,\psi \\ &= [\boldsymbol{\alpha}\boldsymbol{p} + \beta m]\,\psi + [-q\boldsymbol{\alpha}\boldsymbol{A} + q\phi]\,\psi \\ &= (H_0 + H_{\text{int}})\,\psi\,. \end{aligned} \tag{5.10}$$

To extract more physics, let us concentrate on the non-relativistic limit. We write $\psi = \begin{pmatrix}\varphi \\ \chi\end{pmatrix}$ and use Dirac's representation. Then from

$$i\frac{\partial\psi}{\partial t} = \left[\begin{pmatrix}0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0\end{pmatrix}(-i\nabla - q\boldsymbol{A}) + \begin{pmatrix}q\phi + m & 0 \\ 0 & q\phi - m\end{pmatrix}\right]\psi, \tag{5.11}$$

we get two coupled equations

$$i\frac{\partial\varphi}{\partial t} = \boldsymbol{\sigma}\cdot\boldsymbol{\Pi}\chi + q\phi\varphi + m\varphi\,,$$
$$i\frac{\partial\chi}{\partial t} = \boldsymbol{\sigma}\cdot\boldsymbol{\Pi}\varphi + q\phi\chi - m\chi\,. \tag{5.12}$$

Here we have introduced the so-called *kinematic momentum* Π_μ as

$$p_\mu - qA_\mu = \Pi_\mu \quad\Rightarrow\quad -i\nabla - q\boldsymbol{A} = \boldsymbol{\Pi}\,. \tag{5.13}$$

Similarly as for a free particle we pass to the non-relativistic limit by factoring out from ψ the fast oscillating factor, i.e.

$$\psi = \begin{pmatrix}\varphi\\ \chi\end{pmatrix} = e^{-imt}\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix}. \tag{5.14}$$

With this Eq. (5.12) reduces to

$$i\frac{\partial\tilde{\varphi}}{\partial t} = \boldsymbol{\sigma}\cdot\boldsymbol{\Pi}\tilde{\chi} + q\phi\tilde{\varphi}\,, \tag{5.15}$$
$$i\frac{\partial\tilde{\chi}}{\partial t} = \boldsymbol{\sigma}\cdot\boldsymbol{\Pi}\tilde{\varphi} + q\phi\tilde{\chi} - 2m\tilde{\chi}\,. \tag{5.16}$$

Since in Eq. (5.16) the $2m\tilde{\chi}$ term dominates over $i\partial\tilde{\chi}/\partial t$ we can drop the $i\partial\tilde{\chi}/\partial t$ term and rewrite Eq. (5.16) as

$$\tilde{\chi} = \frac{\boldsymbol{\sigma}\boldsymbol{\Pi}}{2m}\tilde{\varphi} + \frac{q\phi}{2m}\tilde{\chi} \sim \frac{\boldsymbol{\sigma}\boldsymbol{\Pi}}{2m}\tilde{\varphi}\,, \tag{5.17}$$

where the last approximation reflects the fact that the interaction energy $q\phi$ is typically much smaller than the rest-mass energy mc^2, so $|q\phi/2mc^2| \ll 1$.

Inserting Eq. (5.17) into Eq. (5.15), we get

$$i\frac{\partial\tilde{\varphi}}{\partial t} = \left(\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})}{2m}\right)\tilde{\varphi} + q\phi\tilde{\varphi} = \left(\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})}{2m} + q\phi\right)\tilde{\varphi}\,. \tag{5.18}$$

This can further be reduced by using an analogue of the well known identity

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{a})(\boldsymbol{\sigma}\cdot\boldsymbol{b}) &= \sigma^i\sigma^j a^i b^j = (\delta^{ij} + i\varepsilon^{ijk}\sigma^k)a^i b^j\\ &= \boldsymbol{a}\cdot\boldsymbol{b} + i\sigma^k(\boldsymbol{a}\times\boldsymbol{b})^k\\ &= \boldsymbol{a}\cdot\boldsymbol{b} + i\boldsymbol{\sigma}\cdot(\boldsymbol{a}\times\boldsymbol{b})\,,\end{aligned} \tag{5.19}$$

which is true only if $\boldsymbol{a}$ and $\boldsymbol{b}$ are c-numbered vectors. For operators, this identity must be modified. In particular, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are generic (non-commuting) vector operators we should write

$$(\boldsymbol{\sigma}\cdot\boldsymbol{a})(\boldsymbol{\sigma}\cdot\boldsymbol{b}) = \frac{1}{4}\left\{\sigma^i,\sigma^j\right\}\left\{a^i,b^j\right\} + \frac{1}{4}\left[\sigma^i,\sigma^j\right]\left[a^i,b^j\right], \tag{5.20}$$

where we used the decomposition into a symmetric and an anti-symmetric parts. Specifically, for the scalar products of $\boldsymbol{\sigma}$ with the kinetic momentum, we get

$$
\begin{aligned}
(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\Pi}) &= \frac{1}{4}2\delta^{ij}\left\{\Pi^i,\Pi^j\right\} + \frac{1}{4}2i\varepsilon^{ijk}\sigma^k\left[\Pi^i,\Pi^j\right] \\
&= \boldsymbol{\Pi}^2 + \frac{1}{2}i\varepsilon^{ijk}\sigma^k\left[p^i - qA^i, p^j - qA^j\right] \\
&= \boldsymbol{\Pi}^2 - \frac{1}{2}i\varepsilon^{ijk}\sigma^k q\left(\left[p^i, A^j\right] + \left[A^i, p^j\right]\right) \\
&= \boldsymbol{\Pi}^2 - \frac{1}{2}q\sigma^k\varepsilon^{ijk}\left(\nabla_i A^j - \nabla_j A^i\right) \\
&= \boldsymbol{\Pi}^2 - q\sigma^k(\nabla\times\boldsymbol{A})^k = \boldsymbol{\Pi}^2 - q\boldsymbol{\sigma}\cdot\boldsymbol{B}\,. \qquad (5.21)
\end{aligned}
$$

Recall that $(\nabla\times\boldsymbol{A})^k = \varepsilon^{klm}\partial_l A^m$, i.e. $2(\nabla\times\boldsymbol{A})^k = \varepsilon^{klm}(\partial_l A^m - \partial_m A^l)$.

Thus, Eq. (5.18) finally reduces to

$$
i\frac{\partial\tilde{\varphi}}{\partial t} = \left[\frac{(\boldsymbol{p}-q\boldsymbol{A})^2}{2m} - \frac{q\boldsymbol{\sigma}\cdot\boldsymbol{B}}{2m} + q\phi\right]\tilde{\varphi}\,, \qquad (5.22)
$$

which is nothing but the Pauli equation of the non-relativistic quantum physics. Hence we gain certain confidence that we are on the right track.

We can again write Dirac's non-relativistic Hamiltonian in (5.22) as $H_0 + H_{\mathrm{I}} + H_{\mathrm{EM}}$. By restoring $\hbar$ and c we get

$$
\begin{aligned}
H_{\mathrm{I}} &= -\frac{e\hbar}{2mc}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} - \frac{q\hbar}{2mc}\frac{\boldsymbol{\sigma}}{\hbar}\cdot\boldsymbol{B} \\
&= -\frac{e\hbar}{2mc}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} - \frac{q\hbar}{mc}\frac{\hat{\boldsymbol{\sigma}}}{\hbar}\cdot\boldsymbol{B}\,, \qquad (5.23)
\end{aligned}
$$

where $\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma}/2$ is generator of rotations for spin-1/2 particles. Specifically for electrons we can write

$$
\begin{aligned}
H_{\mathrm{EI}} &= -g_L\mu_{\mathrm{B}}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} - g_s\mu_{\mathrm{B}}\frac{\hat{\boldsymbol{\sigma}}}{\hbar}\cdot\boldsymbol{B} \\
&= \boldsymbol{\mu}\cdot\boldsymbol{B}\,, \qquad (5.24)
\end{aligned}
$$

where

$$
\boldsymbol{\mu} = -\mu_{\mathrm{B}}\left(g_L\frac{\boldsymbol{L}}{\hbar} + g_s\frac{\hat{\boldsymbol{\sigma}}}{\hbar}\right), \qquad (5.25)
$$

is the *total magnetic moment*. Note that $g_s = 2g_L = 2$. The fact that $g_s = 2$ is a nontrivial prediction of Dirac theory derived within the non-relativistic context of the Dirac equation. The g-factor of electron has now been measured to something like 12 figures of accuracy and it is not precisely 2, it differs by a tiny amount. Understanding this small difference goes, however, beyond Dirac's theory as it requires a full-fledged Quantum Field Theory.

The current value of electron g-factor is give by $g_s/2 = 1{,}001159652180\ldots$.

Dirac Particle in Electromagnetic Field

Dirac particle with charge q (minimally) coupled to an electromagnetic field with four-dimension vector potential A_μ is described by the equation

$$(\gamma^\mu \Pi_\mu - m)\psi(x) = 0, \qquad \Pi_\mu = i\partial_\mu - qA_\mu .$$

Here Π_μ is kinematic momentum.

Exercise 5.1 Consider Dirac's particle in a uniform magnetic field and zero electric field. Assume that the field $\boldsymbol{B}$ is along the z axis. Show that the vector potential A^μ may be chosen so that $A^0 = A^x = A^y = 0$ and $A^z = Bx$ (this choice is known as *Landau gauge*). Assuming stationary states of the form

$$\psi(t, \boldsymbol{x}) = e^{-iEt}\begin{pmatrix}\varphi(\boldsymbol{x})\\ \chi(\boldsymbol{x})\end{pmatrix}.$$

Eliminate χ and show that φ fulfills

$$(E^2 - m^2)\varphi(\mathbf{x}) = [\hat{\boldsymbol{p}}^2 + q^2B^2x^2 - qB(\sigma^3 + 2x\,\hat{p}^y)]\varphi(\mathbf{x}).$$

with $\hat{p}^{x,y,z} = -i\partial_{x,y,x}$.

[**Hint:** Work in the standard representation of γ-matrices, and eliminate $\chi(x)$ from the pair of equations following from Eq. (5.9).]

Exercise 5.2 Since $\hat{p}^y$, $\hat{p}^z$ and σ^3 commute with the right-hand side of the equation in Exercise 5.1, one can consider the ansatz

$$\varphi(\mathbf{x}) = e^{i(p^y y + p^z z)} f(x).$$

Here p^y and p^z are c-numbered constants and f is an eigenvector σ^3. Can you justify the ansatz? Show that the resulting differential equation for f has the form

$$\left[-\frac{d^2}{dx^2} + (qBx - p^y)^2 - qB\sigma^3\right] f(x) = [E^2 - m^2 - (p^z)^2] f(x). \tag{5.26}$$

By introducing the auxiliary variables (assume $qB > 0$)

$$a = \frac{E^2 - m^2 - p_z^2}{qB}, \quad \xi = \sqrt{qB}\left(x - \frac{p^y}{qB}\right),$$

show that the equation for f reduces to

$$\left(-\frac{d^2}{d\xi^2} + \xi^2 - \sigma^3\right) f = af,$$

If f is corresponding eigenvector of σ^3 with eigenvalues $\alpha = \pm$, then

$$f = \begin{pmatrix} f_1 \\ 0 \end{pmatrix} \quad \text{for } \alpha = 1, \qquad f = \begin{pmatrix} 0 \\ f_{-1} \end{pmatrix} \quad \text{for } \alpha = -1.$$

From this deduce that the solution which vanishes at infinity has the form

$$f_\alpha = c e^{-\xi^2/2} H_n(\xi),$$

provided $a + \alpha = 2n + 1$ (n is integer, $n = 0, 1, 2, \ldots$, c is a normalization constant). H_n are Hermite polynomials. From this deduce the energy levels of a relativistic particle in constant magnetic field are (in full units)

$$E = (\pm)\sqrt{m^2c^4 + p_z^2c^2 + \hbar c^2 qB(2n + 1 - \alpha)}.$$

Notice that energy levels have both discrete ($n, \alpha = -1; n + 1, \alpha = -1$) and continuous (in p_y) degeneracy.

Exercise 5.3 How looks the non-relativistic expansion of the previous energy spectrum? Compare it with the usual Landau levels which are result of an analogous non-relativistic treatment.

Exercise 5.4 Show that the following identity holds

$$[(i\partial\!\!\!/ - qA\!\!\!/)^2 - m^2]\psi = \left[(i\partial - qA)^2 - \frac{q}{2}\boldsymbol{\sigma}_{\mu\nu}\boldsymbol{F}_{\mu\nu} - m^2\right]\psi .$$

Here $\boldsymbol{\sigma}_{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ and $F_{\mu\nu}$ is the usual electromagnetic stress tensor.

6 Dirac Equation — Technical Developments II

6.1 Representations of γ Matrices

We have seen that γ-matrices satisfy the defining relation

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} . \tag{6.1}$$

The question naturally arises, how uniquely are the γ-matrices determined by the Clifford product (6.1). This is answered by the so-called *Fundamental theorem of Clifford algebra* (W. Pauli 1936):

Theorem 6.1.1 (Fundamental theorem of Clifford algebra)
If two distinct sets of γ-matrices are given, that both satisfy the Clifford algebra relation

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} ,$$

then they are connected to each other by similarity transformation

$$\gamma'^\mu = \boldsymbol{S}^{-1}\gamma^\mu \boldsymbol{S} .$$

If, in addition, the γ-matrices are (anti-)Hermitian (as in Dirac's particular case where $\gamma^0 = \gamma^{0\dagger}$ and $\gamma^i = -\gamma^{i\dagger}$) then $\boldsymbol{S}$ itself is unitary, i.e.

$$\gamma'^\mu = \boldsymbol{U}^\dagger \gamma^\mu \boldsymbol{U} .$$

This transformation is unique up to a multiplicative factor, which in case of $\boldsymbol{U}$ must have absolute value 1.

Clearly when a set of γ^μ matrices satisfies Clifford algebra product (6.1) and is (anti)Hermitian, then also $\gamma'^\mu = \boldsymbol{U}^\dagger\gamma^\mu\boldsymbol{U}$ satisfy the Clifford product and are (anti)Hermitian. Hard part of the proof of Theorem 6.1.1 is to show existence and uniqueness of $\boldsymbol{U}$.

Though all physical consequences should be independent of a particular choice of γ-matrics representation, different sets might present different technical advantages.

Let us now review the most typical representations of γ-matrices.

Dirac's Representation

So far we have worked with the so-called Dirac's representation of γ-matrices. This is historically the oldest representation of γ-matrices. It was found by Dirac — hence the name *Dirac's representation*. In this representation γ matrices are given by

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} . \tag{6.2}$$

By using properties of tensor product, we can conveniently rewrite these matrices as

$$\gamma^0 = \sigma_3 \otimes \mathbb{1} , \quad \boldsymbol{\gamma} = i\sigma_2 \otimes \boldsymbol{\sigma} , \quad \gamma^5 = \sigma_1 \otimes \mathbb{1} . \tag{6.3}$$

Dirac Equation — Technical Developments II

6

6.1 Representations of γ Matrices

We have seen that γ-matrices satisfy the defining relation

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} . \tag{6.1}$$

The question naturally arises, how uniquely are the γ-matrices determined by the Clifford product (6.1). This is answered by the so-called *Fundamental theorem of Clifford algebra* (W. Pauli 1936):

Theorem 6.1.1 (Fundamental theorem of Clifford algebra)
If two distinct sets of γ-matrices are given, that both satisfy the Clifford algebra relation

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} ,$$

then they are connected to each other by similarity transformation

$$\gamma'^\mu = \boldsymbol{S}^{-1}\gamma^\mu \boldsymbol{S} .$$

If, in addition, the γ-matrices are (anti-)Hermitian (as in Dirac's particular case where $\gamma^0 = \gamma^{0\dagger}$ and $\gamma^i = -\gamma^{i\dagger}$) then $\boldsymbol{S}$ itself is unitary, i.e.

$$\gamma'^\mu = \boldsymbol{U}^\dagger \gamma^\mu \boldsymbol{U} .$$

This transformation is unique up to a multiplicative factor, which in case of $\boldsymbol{U}$ must have absolute value 1.

Clearly when a set of γ^μ matrices satisfies Clifford algebra product (6.1) and is (anti)Hermitian, then also $\gamma'^\mu = \boldsymbol{U}^\dagger\gamma^\mu\boldsymbol{U}$ satisfy the Clifford product and are (anti)Hermitian. Hard part of the proof of Theorem 6.1.1 is to show existence and uniqueness of $\boldsymbol{U}$.

Though all physical consequences should be independent of a particular choice of γ-matrics representation, different sets might present different technical advantages.

Let us now review the most typical representations of γ-matrices.

Dirac's Representation

So far we have worked with the so-called Dirac's representation of γ-matrices. This is historically the oldest representation of γ-matrices. It was found by Dirac — hence the name *Dirac's representation*. In this representation γ matrices are given as

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} . \tag{6.2}$$

By using properties of tensor product, we can conveniently rewrite these matrices in the form

$$\gamma^0 = \sigma_3 \otimes \mathbb{1} , \quad \boldsymbol{\gamma} = i\sigma_2 \otimes \boldsymbol{\sigma} , \quad \gamma^5 = \sigma_1 \otimes \mathbb{1} . \tag{6.3}$$

This representation is particularly convenient for taking the non-relativistic limit, e.g., for a free particle we have already seen that in Dirac's representation

$$\begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix} \stackrel{NR}{\rightarrow} \begin{pmatrix} \chi_{NR} \\ 0 \end{pmatrix}. \tag{6.4}$$

An analogous behavior was observed for Dirac's wave function of an electron in external electromagnetic field.

Let us also recall that the Lorentz group elements associated with rotations and boosts have in Dirac's representation the explicit forms [cf. Eq. (4.99) and Eq. (4.125)]

$$S_D(\boldsymbol{R}) = \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{n}_k\boldsymbol{\sigma}_k} & 0 \\ 0 & e^{-\frac{i}{2}\boldsymbol{n}_k\boldsymbol{\sigma}_k} \end{pmatrix}, \tag{6.5}$$

$$S_D(\boldsymbol{B}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \mathbb{1} & \frac{\boldsymbol{p}\cdot\boldsymbol{\sigma}}{E+m} \\ \frac{\boldsymbol{p}\cdot\boldsymbol{\sigma}}{E+m} & \mathbb{1} \end{pmatrix}. \tag{6.6}$$

Chiral (Weyl) Representation

Another important representation is *chiral or Weyl representation* that is given by γ-matrices

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{6.7}$$

Again, we can rewrite this representation in terms of tensor products as

$$\gamma^0 = \sigma_1 \otimes \mathbb{1}, \quad \boldsymbol{\gamma} = i\sigma_2 \otimes \boldsymbol{\sigma}, \quad \gamma^5 = -\sigma_3 \otimes \mathbb{1}. \tag{6.8}$$

Chiral representation is important for description of massless spin 1/2 particles as in this representation the Dirac massless equation decouples into two autonomous equations, one for upper and one for the lower component of the Dirac spinor.

Chiral representation is also instrumental in discussions related to bispinor representation of Lorentz group. To this end we recall that for rotations we can write

$$S(\boldsymbol{R}) = e^{-\frac{i}{4}\omega_{ij}\sigma^{ij}}, \quad i,j = 1,2,3, \tag{6.9}$$

with $\boldsymbol{\sigma}^{ij} = \frac{i}{2}[\gamma^i, \gamma^j] = i\gamma^i\gamma^j$. By defining

$$\boldsymbol{\sigma}^{ij} = 2\varepsilon^{ijk}\hat{\boldsymbol{\sigma}}^k = 2\varepsilon^{ijk} \begin{pmatrix} \frac{\sigma_k}{2} & 0 \\ 0 & \frac{\sigma_k}{2} \end{pmatrix}, \tag{6.10}$$

we get

$$S_W(\boldsymbol{R}) = e^{-i\boldsymbol{\theta}\cdot\hat{\boldsymbol{\sigma}}} = \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma}} & 0 \\ 0 & e^{\frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma}} \end{pmatrix}. \tag{6.11}$$

This has clearly the same structure both in chiral and Dirac's representation because γ^i, $i = 1,2,3$ are the same.

On the other hand, general boost in bispinor representation is given by the prescription

$$S(\boldsymbol{B}) = e^{-\frac{i}{2}\omega_{0i}\sigma^{0i}} = e^{\frac{1}{2}\zeta_i\gamma^0\gamma^i} . \tag{6.12}$$

By using the fact that in the chiral representation

$$\sigma^{0i} = i\begin{pmatrix} -\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} , \tag{6.13}$$

we get that

$$S_W(\boldsymbol{B}) = e^{\frac{\zeta\cdot\gamma^0\gamma}{2}} = \begin{pmatrix} e^{-\frac{\zeta\cdot\sigma}{2}} & 0 \\ 0 & e^{\frac{\zeta\cdot\sigma}{2}} \end{pmatrix} . \tag{6.14}$$

Since the group elements $S_W(\boldsymbol{R})$ and $S_W(\boldsymbol{B})$ have a block diagonal form it seems that the bispinor representation is *reducible* to two independent spinor representations. So apparently we do not need to work with bispinors, but it would be enough to work with spinors only. This, indeed, is true for massless particles. For massive particles the issue is more complicated and is related to the concept of *parity*.

In order to understand this better we should discuss discrete transformations of Lorentz group. This will be done in the following section.

6.2 Discrete Transformations

Space Reflection (Parity) Transformation

Parity transformation is acting on the position 4-vector as follows:

$$x^\mu = (t, \boldsymbol{x}) \xrightarrow{P} x^\mu_P = (t, -\boldsymbol{x}) . \tag{6.15}$$

From this we can see, that this transformation can be written in the form

$$x^\mu_P = \boldsymbol{P}^\mu{}_\nu x^\nu , \tag{6.16}$$

where the matrix $\boldsymbol{P}$ reads

$$\boldsymbol{P}^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \eta^{\mu\nu} = \eta_{\mu\nu} . \tag{6.17}$$

This satisfies the defining property of Lorentz group, namely

$$\boldsymbol{P}^\mu{}_\alpha \boldsymbol{P}^\nu{}_\beta \eta_{\mu\nu} = \eta_{\alpha\beta} , \tag{6.18}$$

and hence $\boldsymbol{P}$ is an element of Lorentz group. We can also see that $\det \boldsymbol{P} = -1$, so it corresponds to an improper orthochronous Lorentz transformation.

To find the bispinor representation of the parity transformation, we recall that the covariance of Dirac equation under Lorentz group requires that (cf. Eq. (3.51))

$$S(\boldsymbol{P})\gamma^\mu S^{-1}(\boldsymbol{P}) = \boldsymbol{P}_\nu{}^\mu\gamma^\nu \quad \Leftrightarrow \quad S^{-1}(\boldsymbol{P})\gamma^\mu S(\boldsymbol{P}) = \boldsymbol{P}^\mu{}_\nu\gamma^\nu\,, \tag{6.19}$$

should be satisfies. In this case, we cannot solve the equation in terms of infinitesimal transformations. Fortunately, it can be solved directly. In this case note that

Note that $\delta_{\mu 0}$ and γ^μ are not Einstein summed.

$$S(\boldsymbol{P})\,\gamma^\mu S^{-1}(\boldsymbol{P}) = -(-1)^{\delta_{\mu 0}}\gamma^\mu\,. \tag{6.20}$$

Using the fact, that $\{\gamma^0, \gamma^\mu\} = 2\eta^{0\mu}$, we obtain that

$$\begin{aligned}\gamma^0\gamma^\mu &= -\gamma^\mu\gamma^0 + 2\eta^{0\mu} \\ &= \begin{cases} -\mathbb{1} + 2\cdot\mathbb{1} = \mathbb{1},\ \mu; & \text{for } \mu = 0 \\ -\gamma^i\gamma^0\,; & \text{for } \mu \neq 0,\ \mu = i\,. \end{cases}\end{aligned} \tag{6.21}$$

Take $S(\boldsymbol{P}) = \gamma^0$, then (6.20) reduces to the

$$\gamma^0\gamma^\mu(\gamma^0)^{-1} = \gamma^0\gamma^\mu\gamma^0 = -(-1)^{\delta_{\mu 0}}\gamma^\mu\,. \tag{6.22}$$

Most generally, we can chose $S(\boldsymbol{P})$ to be

$$S(\boldsymbol{P}) = e^{i\phi}\gamma^0\,. \tag{6.23}$$

So, finally the parity transformed bispinor takes the form

$$\psi_P(x) = S(\boldsymbol{P})\psi(\boldsymbol{P}^{-1}x) = e^{i\phi}\gamma^0\psi(\boldsymbol{P}^{-1}x) = e^{i\phi}\gamma^0\psi(x_P)\,. \tag{6.24}$$

If one requires that after two parity transformations one should return to the original state, i.e., $S(\boldsymbol{P})\,\psi_P(x_P) = \psi(x)$ then

$$e^{i2\phi}(\gamma^0)^2\psi(x) = e^{i2\phi}\psi(x) = \psi(x)\,. \tag{6.25}$$

This implies that $\phi = 0$ or $\phi = \pi$ (modulo 2π) or, in other words, that

$$S(\boldsymbol{P}) = \pm\gamma^0 = \eta_P\gamma^0\,, \tag{6.26}$$

where η_P is the so-called *intrinsic parity*, i.e, yet another quantum number of a particle. Result (6.26) is, however, not entire correct. In fact, let us consider a bispinor representation of rotations, i.e.

$$\psi_{\boldsymbol{R}}(x) = S(\boldsymbol{R})\psi(\boldsymbol{R}^{-1}x) = \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma}} & 0 \\ 0 & e^{-\frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma}} \end{pmatrix}\psi(\boldsymbol{R}^{-1}x)\,. \tag{6.27}$$

For 2π rotation around z-axis we get:

$$\begin{aligned} S(\boldsymbol{R}, \theta = 2\pi, \text{around } z\text{-axis})\psi(x) &= \begin{pmatrix} e^{-i\pi\sigma_3} & 0 \\ 0 & e^{-i\pi\sigma_3} \end{pmatrix}\psi(x) \\ &= \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}\psi(x) \neq \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}\psi(x)\,. \end{aligned} \tag{6.28}$$

Hence, one should rotate by 4π around z axis to get the original bispinor. Since 4π rotation is an analogue of 4 reflections we should require that after 4 (and not 2) reflections the Dirac particle will get back to its original state $\psi(x)$. With this we finally get

$$S(\boldsymbol{P})\psi(x) = \pm i\gamma^0\psi(x) = i\eta_P\gamma^0\psi(x)\,. \tag{6.29}$$

Since in the Weyl representation

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \tag{6.30}$$

we see that under parity transformation the upper and lower spinors are interchanged in the bispinor. Hence, if we wish to work with a parity invariant 1/2-spin fermion, we must describe it with a bispinor wave function and Dirac's equation is the equation that stipulates the corresponding dynamics.

As an example of application of a parity transformation we consider now the positive-energy Dirac wave function for a free particle with momentum p. Then

$$u(p,\lambda)e^{-ipx} \longrightarrow \pm i\gamma^0 u(p,\lambda)e^{-ipx_P} = \pm iu(p_P,\lambda)e^{-ip_Px}\,. \tag{6.31}$$

Rewriting this in more detail and recalling that in the Dirac representation $\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}$ and $u(p,\lambda) \propto \begin{pmatrix} \chi_\lambda \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda \end{pmatrix}$:

$$\begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \chi_\lambda \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda \end{pmatrix} = \begin{pmatrix} \chi_\lambda \\ \frac{-\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda \end{pmatrix} = \begin{pmatrix} \chi_\lambda \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}_P}{E+m}\chi_\lambda \end{pmatrix}. \tag{6.32}$$

Here we have introduced a paritity transformed 4-momenta

$$p_P^\mu = (p_0, -\boldsymbol{p}) = (p_0, \boldsymbol{p}_P)\,. \tag{6.33}$$

Note, in particular, that the spatial part of the momentum has been parity transformed in (6.31), but the spin state has been unaltered. This is precisely what we would expect from parity transformation.

Similarly, we can now proceed with negative-energy solutions. In particular

$$v(p,\lambda)e^{ipx} = \pm i\gamma^0 v(p,\lambda)e^{ipx_p} = \mp iv(p_p,\lambda)e^{ip_px}\,. \tag{6.34}$$

So, the positive and negative-energy solutions have relative opposite intrinsic parities. After the reinterpretation of negative energy solutions this will imply that intrinsic parities for particle and antiparticle are reversed.

Parity of a Scalar Particle

For a complex wave function $\phi(x)$ of a relativistic scalar (Klein–Gordon) particle one can follow the usual prescription for Lorentz group transformation of scalar functions, i.e.

$$\phi(x) \longrightarrow \phi_L(x) = \phi(\boldsymbol{L}^{-1}x)\,, \tag{6.35}$$

Reason why η_P appears at all is that prescription (6.35) is valid only for *restricted Lorentz group*, i.e., the set of all Lorentz transformations that can be connected to the identity element by a continuous curve lying in the group. Passage to some of the remaining 3 component of Lorentz group is accompanied by an appearance of intrinsic quantum numbers.

which in the parity case are phrased as

$$\phi(x) \;\to\; \phi_P(x) \;=\; \eta_P \phi(x_P)\,.$$

Consistency of this prescription can be checked by looking at a state with a definite momentum $\phi(p,x) = e^{-ipx}$. In this case the prescription gives

$$\phi_P(p,x) \;=\; \{e^{-ipx_P}\} \;=\; e^{-ip_P x} \;=\; \phi(p_P,x)\,.$$

Time Reversal

Time reversal transformation acts on a position 4-vector so that

$$x^\mu \;=\; (t,\boldsymbol{x}) \;\overset{T}{\longrightarrow}\; x^\mu_T \;=\; (-t,\boldsymbol{x})\,. \tag{6.36}$$

This transformation can be described via transformation matrix

$$\boldsymbol{T}^\mu{}_\nu \;=\; \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \;=\; -\eta^{\mu\nu} \;=\; -\eta_{\mu\nu}\,, \tag{6.37}$$

which satisfies the defining relation of Lorentz group

$$\boldsymbol{T}^\mu{}_\alpha \boldsymbol{T}^\nu{}_\beta \eta_{\mu\nu} \;=\; \eta_{\alpha\beta}\,. \tag{6.38}$$

Again, we can see that $\boldsymbol{T}^\mu{}_\alpha$ is an element of a Lorentz group. Since $\det \boldsymbol{T} = -1$ and $\boldsymbol{T}^0{}_0 = -1$, this discrete group element corresponds to an improper non-orthochronous Lorentz transformation.

Let us also mention that such important vector quantities as velocity, momentum or angular momentum transform with respect to time reversals as follows

$$\boldsymbol{v} \overset{T}{\to} \boldsymbol{v_T} \;=\; -\boldsymbol{v}\,, \quad \boldsymbol{p} \overset{T}{\to} \boldsymbol{p_T} \;=\; -\boldsymbol{p}\,, \quad \boldsymbol{J} \overset{T}{\to} \boldsymbol{J_T} \;=\; -\boldsymbol{J}\,. \tag{6.39}$$

To find a bispinor representation of time reversal we cannot follow the same route as we did in the case of parity. This is because time reversal must be implemented via anti-unitary transformation (i.e., complex conjugation of the wave function is required) and Eq. (3.51) was not derived under such an assumption.

In non-relativistic quantum mechanics we know that the complete effect of a linear operator can be determined by specifying its action on a basis set of the vector space of physical states and then extend its application by exploiting the linearity of the maps. Similarly, the complete effect of an antilinear map can be determined by specifying its effect on a basis and extending the results using its *antilinearity*. Take, for instance, the momentum basis $|\boldsymbol{p}\rangle$. Then $\hat{\boldsymbol{T}}\,|\boldsymbol{p}\rangle = |-\boldsymbol{p}\rangle$. A generic state would then look like

$$|\psi\rangle \;=\; \sum_{\boldsymbol{p}} \widetilde{\psi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle\,. \tag{6.40}$$

From this we can see, that effect of time reversal is then

$$\begin{aligned}\hat{\boldsymbol{T}}\,|\psi\rangle &= \hat{\boldsymbol{T}}\sum_{\boldsymbol{p}}\widetilde{\psi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle = \sum_{\boldsymbol{p}}\widetilde{\psi}^*(\boldsymbol{p})\hat{\boldsymbol{T}}\,|\boldsymbol{p}\rangle \\ &= \sum_{\boldsymbol{p}}\widetilde{\psi}^*(\boldsymbol{p})\,|{-\boldsymbol{p}}\rangle = \sum_{\boldsymbol{p}}\widetilde{\psi}^*(-\boldsymbol{p})\,|\boldsymbol{p}\rangle \\ &= \sum_{\boldsymbol{p}}\widetilde{\psi}_T(\boldsymbol{p})\,|\boldsymbol{p}\rangle = |\psi_T\rangle\,. \end{aligned} \tag{6.41}$$

Thus we have that $\widetilde{\psi}_T(\boldsymbol{p}) = \widetilde{\psi}^*(-\boldsymbol{p})$ for scalar wave functions. If the state $|\phi\rangle = \sum_{\boldsymbol{p}}\widetilde{\phi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle$ is defined similarly, then the scalar product

$$\begin{aligned}\langle\phi|\psi\rangle &= \sum_{\boldsymbol{p}}\widetilde{\phi}^*(\boldsymbol{p})\widetilde{\psi}(\boldsymbol{p}) = \sum_{\boldsymbol{p}}\widetilde{\phi}^*(-\boldsymbol{p})\widetilde{\psi}(-\boldsymbol{p}) \\ &= \left[\sum_{\boldsymbol{p}}\widetilde{\phi}(-\boldsymbol{p})\widetilde{\psi}^*(-\boldsymbol{p})\right]^* = \left[\sum_{\boldsymbol{p}}\widetilde{\phi}_T^*(\boldsymbol{p})\widetilde{\psi}_T(\boldsymbol{p})\right]^* \\ &= \langle\phi_T|\psi_T\rangle^*\,. \end{aligned} \tag{6.42}$$

In order to get x-representation of our wave function we can apply Fourier transformation thus obtaining

$$\begin{aligned}\psi_T(\boldsymbol{x}) &= \int\widetilde{\psi}_T(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} = \int\widetilde{\psi}^*(-\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} \\ &= \int\widetilde{\psi}^*(\boldsymbol{p})e^{i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} = \left[\int\widetilde{\psi}(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\right]^* = \psi^*(\boldsymbol{x})\,. \end{aligned} \tag{6.43}$$

Note that from (6.43) also follows that $\psi_T(\boldsymbol{x},t) = \left[e^{-i\boldsymbol{H}t}\psi^*(\boldsymbol{x},0)\right] = \left[e^{i\boldsymbol{H}t}\psi(\boldsymbol{x},0)\right]^* = \psi^*(\boldsymbol{x},-t) = \psi^*(x_T)\,.$

In line with non-relativistic quantum mechanics, the effect of time reversal on the Dirac wave function can be written in the form

$$\psi(x) \xrightarrow{T} \psi_T(x) = \boldsymbol{B}\psi^*(x_T)\,, \tag{6.44}$$

where the matrix $\boldsymbol{B}$ acts on bispinorial indices.

To find $\boldsymbol{B}$ we complex conjugate Dirac equation and observe that the following chain of equivalencies holds

$$\begin{aligned} &\left(i\gamma^\mu\frac{\partial}{\partial x^\mu} - m\right)\psi(x) = 0 \\ \Leftrightarrow\quad &\left(i\gamma^\mu\frac{\partial}{\partial x_T^\mu} - m\right)\psi(x_T) = 0 \\ \Leftrightarrow\quad &\left(i\gamma^{0*}\partial_t - i\gamma^{i*}\partial_i - m\right)\psi^*(x_T) = 0\,. \end{aligned} \tag{6.45}$$

If we assume that there exists a matrix $\boldsymbol{B}$ such that

$$\boldsymbol{B}\left(\gamma^{0*},-\gamma^{i*}\right)\boldsymbol{B}^{-1} = \left(\gamma^0,\gamma^i\right)\,, \tag{6.46}$$

Since matrices $\left(\gamma^{0*},-\gamma^{i*}\right)$ satisfy Clifford algebra, the similarity transformation (6.46) is guaranteed by the Fundamental theorem of Clifford algebra.

then with the definition of $\psi_T(x)$ in (6.44) we get

$$0 = \boldsymbol{B}\left(i\gamma^{0*}\partial_t - i\gamma^{i*}\partial_i - m\right)\psi^*(x_T) = \left(i\gamma^\mu\partial_\mu - m\right)\psi_T(x)\,. \tag{6.47}$$

From this we can see, that effect of time reversal is then

$$\begin{aligned}\hat{\boldsymbol{T}}\,|\psi\rangle &= \hat{\boldsymbol{T}}\sum_{\boldsymbol{p}}\widetilde{\psi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle = \sum_{\boldsymbol{p}}\widetilde{\psi}^*(\boldsymbol{p})\hat{\boldsymbol{T}}\,|\boldsymbol{p}\rangle \\ &= \sum_{\boldsymbol{p}}\widetilde{\psi}^*(\boldsymbol{p})\,|-\boldsymbol{p}\rangle = \sum_{\boldsymbol{p}}\widetilde{\psi}^*(-\boldsymbol{p})\,|\boldsymbol{p}\rangle \\ &= \sum_{\boldsymbol{p}}\widetilde{\psi}_T(\boldsymbol{p})\,|\boldsymbol{p}\rangle = |\psi_T\rangle\,. \end{aligned} \tag{6.41}$$

Thus we have that $\widetilde{\psi}_T(\boldsymbol{p}) = \widetilde{\psi}^*(-\boldsymbol{p})$ for scalar wave functions. If the state $|\phi\rangle = \sum_{\boldsymbol{p}}\widetilde{\phi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle$ is defined similarly, then the scalar product

$$\begin{aligned}\langle\phi|\psi\rangle &= \sum_{\boldsymbol{p}}\widetilde{\phi}^*(\boldsymbol{p})\widetilde{\psi}(\boldsymbol{p}) = \sum_{\boldsymbol{p}}\widetilde{\phi}^*(-\boldsymbol{p})\widetilde{\psi}(-\boldsymbol{p}) \\ &= \left[\sum_{\boldsymbol{p}}\widetilde{\phi}(-\boldsymbol{p})\widetilde{\psi}^*(-\boldsymbol{p})\right]^* = \left[\sum_{\boldsymbol{p}}\widetilde{\phi}_T^*(\boldsymbol{p})\widetilde{\psi}_T(\boldsymbol{p})\right]^* \\ &= \langle\phi_T|\psi_T\rangle^*\,. \end{aligned} \tag{6.42}$$

In order to get x-representation of our wave function we can apply Fourier transformation thus obtaining

$$\begin{aligned}\psi_T(\boldsymbol{x}) &= \int\widetilde{\psi}_T(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} = \int\widetilde{\psi}^*(-\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} \\ &= \int\widetilde{\psi}^*(\boldsymbol{p})e^{i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} = \left[\int\widetilde{\psi}(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\right]^* = \psi^*(\boldsymbol{x})\,. \end{aligned} \tag{6.43}$$

Note that from (6.43) also follows that $\psi_T(\boldsymbol{x},t) = \left[e^{-i\boldsymbol{H}t}\psi^*(\boldsymbol{x},0)\right] = \left[e^{i\boldsymbol{H}t}\psi(\boldsymbol{x},0)\right]^* = \psi^*(\boldsymbol{x},-t) = \psi^*(x_T)$.

In line with non-relativistic quantum mechanics, the effect of time reversal on the Dirac wave function can be written in the form

$$\psi(x) \xrightarrow{T} \psi_T(x) = \boldsymbol{B}\psi^*(x_T)\,, \tag{6.44}$$

where the matrix $\boldsymbol{B}$ acts on bispinorial indices.

To find $\boldsymbol{B}$ we complex conjugate Dirac equation and observe that the following chain of equivalencies holds

$$\begin{aligned} &\left(i\gamma^\mu\frac{\partial}{\partial x^\mu} - m\right)\psi(x) = 0 \\ \Leftrightarrow\quad &\left(i\gamma^\mu\frac{\partial}{\partial x_T^\mu} - m\right)\psi(x_T) = 0 \\ \Leftrightarrow\quad &\left(i\gamma^{0*}\partial_t - i\gamma^{i*}\partial_i - m\right)\psi^*(x_T) = 0\,. \end{aligned} \tag{6.45}$$

If we assume that there exists a matrix $\boldsymbol{B}$ such that

$$\boldsymbol{B}\left(\gamma^{0*}, -\gamma^{i*}\right)\boldsymbol{B}^{-1} = \left(\gamma^0, \gamma^i\right)\,, \tag{6.46}$$

Since matrices $\left(\gamma^{0*}, -\gamma^{i*}\right)$ satisfy Clifford algebra, the similarity transformation (6.46) is guaranteed by the Fundamental theorem of Clifford algebra.

then with the definition of $\psi_T(x)$ in (6.44) we get

$$0 = \boldsymbol{B}\left(i\gamma^{0*}\partial_t - i\gamma^{i*}\partial_i - m\right)\psi^*(x_T) = \left(i\gamma^\mu\partial_\mu - m\right)\psi_T(x)\,. \tag{6.47}$$

This clearly shows that $\psi_T(x)$ satisfies Dirac's equation if $\psi(x)$ does. The Fundamental theorem of Clifford algebra ensures that $\boldsymbol{B}$ is unitary and (aside from a phase factor) unique.

To retrieve $\boldsymbol{B}$ we use the fact that

$$\begin{aligned} \gamma^{0\dagger} &= \gamma^0 \quad \Rightarrow \quad \gamma^{0*} = (\gamma^0)^t \,, \\ \gamma^{i\dagger} &= -\gamma^i \quad \Rightarrow \quad \gamma^{i*} = -(\gamma^i)^t \,, \end{aligned} \tag{6.48}$$

and

$$\boldsymbol{C}(\gamma^\mu)^t \boldsymbol{C}^{-1} = -\gamma^\mu \,. \tag{6.49}$$

Here $\boldsymbol{C} = i\gamma^0\gamma^2$ is the so-called *charge conjugation matrix* (we will derive it in the following section). With this we can write $\boldsymbol{B}$ in the form

$$\begin{aligned} \boldsymbol{B} &= \eta_T \gamma_5 \boldsymbol{C} = i\eta_T \gamma^0\gamma^1\gamma^2\gamma^3 i\gamma^0\gamma^2 = \eta_T \gamma^1\gamma^3 \\ &= \eta_T \begin{pmatrix} 0 & \sigma^1 \\ -\sigma^1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^3 \\ -\sigma^3 & 0 \end{pmatrix} = \eta_T \begin{pmatrix} i\sigma^2 & 0 \\ 0 & i\sigma^2 \end{pmatrix} . \end{aligned} \tag{6.50}$$

Here η_T is a complex number of unit amplitude, that is allowed by the Fundamental theorem of Clifford algebra. It is easy to see that $\boldsymbol{B}$ proposed in (6.50) satisfies (6.46). Indeed

$$\begin{aligned} \boldsymbol{B}(\gamma^0)^* \boldsymbol{B}^{-1} &= \gamma_5 \boldsymbol{C}(\gamma^0)^t \boldsymbol{C}^{-1}\gamma_5 = \gamma^0 \,, \\ \boldsymbol{B}(\gamma^i)^* \boldsymbol{B}^{-1} &= -\gamma_5 \boldsymbol{C}(\gamma^i)^t \boldsymbol{C}^{-1}\gamma_5 = -\gamma^i \,. \end{aligned} \tag{6.51}$$

Note that since γ^1 and γ^3 are identical both in Dirac's and chiral representation, the form (6.50) remains the same in both representations. Consequently, time reversal (in contrast to parity) does not swap upper and lower components of Dirac's bispinor.

To find a constraint on the phase factor η_T we might require that $(\psi_T)_T = \psi$. This in turn gives

$$\psi(x) = [\psi_T(x)]_T = \eta_{TT}\gamma^1\gamma^3 \left[\eta_T \gamma^1\gamma^3 \psi^*(x)\right]^* = -\eta_{TT}\eta_T^* \psi(x) \,. \tag{6.52}$$

Hence $\eta_{TT}\eta_T^* = -1$. So, relative sign between η_{TT} and η_T^* must be -1. Since in experiments it is always a relative intrinsic parity between two wave functions that is measured (and not the phase factor itself), one conventionally chooses $\eta_{TT} = i\eta_{TT}^R$ and $\eta_T = i\eta_T^R$ with $\eta_{TT}^R = \pm 1$ and $\eta_T^R = \mp 1$. With this convention we can finally write

$$\psi_T(x) = i\eta_T^R \gamma^1\gamma^3 \psi^*(x_T) \,, \tag{6.53}$$

where η_T^R is *intrinsic time reversal*, i.e., another quantum number of the particle (akin to charge or spin). Instead of η_T^R one simply writes only η_T. Aside from η_T one can also check that

$$\boldsymbol{B}^t = -\boldsymbol{B}, \quad \boldsymbol{B}^* = -\boldsymbol{B}^{-1} \,. \tag{6.54}$$

As an example we now compute a time reversal of the bispinor $u(p, \lambda)$.

This can be done as follows. We first write

$$
\begin{aligned}
\psi_T^{(+)}(x,p) &= \boldsymbol{B}\left(u(p,\lambda)e^{-ipx_T}\right)^* \\
&= i\eta_T\sqrt{E+m}\begin{pmatrix} i\sigma^2 & 0 \\ 0 & i\sigma^2 \end{pmatrix}\begin{pmatrix} \chi_\lambda^* \\ \frac{\boldsymbol{p}\cdot\boldsymbol{\sigma}^*}{E_T+m}\chi_\lambda^* \end{pmatrix} e^{-ip_T x} \\
&= i\eta_T\sqrt{E_T+m}\begin{pmatrix} i\sigma^2\chi_\lambda^* \\ \frac{-\boldsymbol{p}\cdot(i\sigma^2)\boldsymbol{\sigma}^*(i\sigma^2)}{E_T+m} i\sigma^2\chi_\lambda^* \end{pmatrix} e^{-ip_T x}\,. \qquad (6.55)
\end{aligned}
$$

Now we can use the fact that ($\sigma^3/2$ is a spin projection operator to z-direction)

$$
\begin{aligned}
\frac{\sigma^3}{2} i\sigma^2\chi_\lambda^* &= -\left(\frac{\sigma^3}{2} i(\sigma^2)^*\chi_\lambda\right)^* = \left(\frac{\sigma^3}{2} i\sigma^2\chi_\lambda\right)^* \\
&= -\left(i\sigma^2\frac{\sigma^3}{2}\chi_\lambda\right)^* = -\lambda\left(i\sigma^2\chi_\lambda\right)^* = -\lambda i\sigma^2\chi_\lambda^*\,. \qquad (6.56)
\end{aligned}
$$

Eq. (6.56) implies that $i\sigma^2\chi_\lambda^* \propto \chi_{-\lambda}$. By employing the identity $(i\sigma^2)\boldsymbol{\sigma}^*(i\sigma^2) = \boldsymbol{\sigma}$ we can rewrite (6.55) in the form

In the standard representation χ_λ are real and it is easy to see that $i\sigma^2\chi_\lambda = \chi_{-\lambda}(-1)^{1/2+\lambda}$.

$$
\begin{aligned}
u_T(p,\lambda) &= i\eta_T(-1)^{1/2+\lambda}\sqrt{E_T+m}\begin{pmatrix} \chi_{-\lambda} \\ \frac{\boldsymbol{p}_T\cdot\boldsymbol{\sigma}}{E_T+m}\chi_{-\lambda} \end{pmatrix} \\
&= i\eta_T(-1)^{1/2+\lambda}u(p_T,-\lambda)\,. \qquad (6.57)
\end{aligned}
$$

Note that if one is interested only on a relative phase factor between different spin components, then (6.57) can be reduced to the relation

$$
u_T(p,\lambda) = i(-1)^{1/2+\lambda}u(p_T,-\lambda) \equiv i\eta_\lambda u(p_T,-\lambda)\,. \qquad (6.58)
$$

In this case the condition $\eta_T\eta_T^* = -1$ reduces to $\eta_\lambda\eta_{-\lambda} = -1$.

Time reversal of a Scalar Particle

For a complex wave function $\phi(x)$ of a relativistic scalar (Klein-Gordon) particle

$$
\phi(x) \rightarrow \eta_T\phi_T(x)\,,
$$

with $\phi_T(x) = \phi(x_T)^*$. Consistency of this prescription can be confirmed by looking at a state with a definite momentum $\phi(p,x) = e^{-ipx}$. In this case the prescription gives

$$
\phi_T(p,x) = \{e^{-ipx_T}\}^* = e^{-ip_T x} = \phi(p_T,x)\,.
$$

Charge Conjugation

Last of discrete symmetries we will discuss is a *charge conjugation* $\boldsymbol{C}$, which simply accounts for a change of electric charge to its negative value, i.e.

$$
q \overset{\boldsymbol{C}}{\rightarrow} -q\,. \qquad (6.59)
$$

Note that charge conjugation is not a discrete symmetry of Lorentz group, though, there is a deep connection with the Lorentz symmetry via the so-called CPT theorem.

The issue of charge conjugation is best discussed when electromagnetic field is coupled to a charged particle via minimal coupling, i.e.

$$\left[\left(i\partial\!\!\!/ - qA\!\!\!/\right) - m\right]\psi(x) = 0\,. \tag{6.60}$$

Charge conjugated wave function $\psi_c(x)$ must satisfy

$$\left[\left(i\partial\!\!\!/ + qA\!\!\!/\right) - m\right]\psi_c(x) = 0\,. \tag{6.61}$$

Note that from (6.60) directly follows that

$$\psi^\dagger(x)\left[-i\gamma^{\mu,\dagger}\overleftarrow{\partial}_\mu - q\gamma^{\mu,\dagger}A_\mu - m\right] = 0\,. \tag{6.62}$$

By multiplying this equation from right by γ^0 and using relation $\gamma^0\gamma^\dagger\gamma^0 = \gamma$, we get

$$\begin{aligned} 0 &= \psi^\dagger(x)\gamma^0\left[-i\gamma^\mu\overleftarrow{\partial}_\mu - q\gamma^\mu A_\mu - m\right] \\ &= \overline{\psi}(x)\left[-i\gamma^\mu\overleftarrow{\partial}_\mu - q\gamma^\mu A_\mu - m\right], \\ \Leftrightarrow \quad & \left[-i(\gamma^\mu)^t\partial_\mu - q(\gamma^\mu)^t A_\mu - m\right]\overline{\psi}^t(x) = 0\,. \end{aligned} \tag{6.63}$$

We might thus take the charge-conjugated Dirac's wave function in the form

$$\psi_c(x) = \boldsymbol{C}\overline{\psi}^t(x)\,. \tag{6.64}$$

The matrix $\boldsymbol{C}$ must be chosen so that ψ_c satisfies Dirac equation with opposite charge. From all above we get that

$$\left(-i\boldsymbol{C}(\gamma^\mu)^t\boldsymbol{C}^{-1}\partial_\mu - q\boldsymbol{C}(\gamma^\mu)^t\boldsymbol{C}^{-1}A_\mu - m\right)\psi_c(x) = 0\,. \tag{6.65}$$

Assuming that $\boldsymbol{C}$ satisfies

Existence of such matrix $\boldsymbol{C}$ is guaranteed by the Fundamental theorem of Clifford algebra.

$$\boldsymbol{C}(\gamma^\mu)^t\boldsymbol{C}^{-1} = -\gamma^\mu\,, \tag{6.66}$$

we get from (6.63) that

$$\left(i\gamma^\mu\partial_\mu + q\gamma^\mu\partial_\mu - m\right)\psi_c(x) = \left(i\partial\!\!\!/ + qA\!\!\!/ - m\right)\psi_c(x) = 0\,, \tag{6.67}$$

provided $\psi(x)$ satisfies (6.60) (or equivalently (6.63)).

As for the matrix $\boldsymbol{C}$, it can be checked that

In some convention $\boldsymbol{C}$ is taken as $i\gamma^2\gamma^0$, which differs by sign.

$$\boldsymbol{C} = i\gamma^0\gamma^2 = \begin{pmatrix} 0 & i\sigma_2 \\ i\sigma_2 & 0 \end{pmatrix}. \tag{6.68}$$

$\boldsymbol{C}$ Matrix in Weyl/Chiral Representation

It can be checked that in Weyl/chiral representation the $\boldsymbol{C}$ matrix is given by

$$C = i\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} -i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix}.$$

Note that

$$\boldsymbol{C} = -\boldsymbol{C}^{-1} = -\boldsymbol{C}^{\dagger} = -\boldsymbol{C}^{t}. \tag{6.69}$$

$\psi_c(x)$ describes particle with the same mass and the same spin direction, but with opposite charge and energy. *Change conjugation is antilinear transformation.*

Let us now compute $\psi_c(x)$ for $\psi^{(-)}(x)$ describing a spin-down negative energy electron at rest in absence of external field. Begin with (we omit the normalization factor $\sqrt{2m}$)

$$\psi^{(-)}(x) \equiv \psi(x) = e^{imt}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{6.70}$$

and take $\boldsymbol{C} = i\gamma^0\gamma^2$. Then $\psi_c(x)$ can be written as

$$\begin{aligned} \psi_c(x) &= \eta_c \boldsymbol{C}\overline{\psi}^{t} = i\eta_c\gamma^0\gamma^2(\psi^{\dagger}(x)\gamma^0)^t \\ &= \eta_c \boldsymbol{C}(\gamma^0)^t\psi^*(x) = \eta_c(-\gamma^0)\boldsymbol{C}\psi^*(x) \\ &= \eta_c(-\gamma^0)i\gamma^0\gamma^2 e^{-imt}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = -\eta_c e^{-imt}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{6.71}$$

Therefore, charge conjugation of a negative-energy spin-down electron is equivalent to positive-energy spin-up solution. We thuse see that the spinor $v(p,\lambda)$ is associated with the anti-particle. This connection will be discussed in some detail in the following section when Dirac's hole theory will be considered.

Full Lorentz Group — Brief Summary

- Lorentz transformations $x^\mu \to x'^\mu = \boldsymbol{L}^{\mu}_{\cdot\,\nu}\, x^\nu$ preserve the invariance of the space-time interval $x^\mu x_\mu = x'^\mu x'_\mu$.
- This constraints the matrices $\boldsymbol{L}^{\mu}_{\cdot\,\nu}$ to obey

$$\boldsymbol{L}^{\mu}_{\cdot\,\alpha}\boldsymbol{L}^{\nu}_{\cdot\,\beta}\eta_{\mu\nu} = \eta_{\alpha\beta} \quad\Leftrightarrow\quad \eta = \boldsymbol{L}^T\eta\boldsymbol{L}.$$

  where $\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$.
- Condition $\eta = \boldsymbol{L}^T\eta\boldsymbol{L}$ allows classification of transformations depending on whether

$$\det\boldsymbol{L} = \pm 1 \quad \text{and} \quad \boldsymbol{L}^{0}_{\cdot\,0} = \pm\left[1 + \sum_i (L^{i}_{\cdot\,i})^2\right]^{1/2}$$

- Consequently, the full Lorentz group splits into 4 pairwise disjoint and non compact connected sets:

$$\mathcal{L}^{\uparrow}_{+}: \quad \det \boldsymbol{L} = 1\,, \quad \boldsymbol{L}^{0}_{\cdot\,0} \geq 1\,,$$

$$\mathcal{L}^{\uparrow}_{-}: \quad \det \boldsymbol{L} = -1\,, \quad \boldsymbol{L}^{0}_{\cdot\,0} \geq 1\,,$$

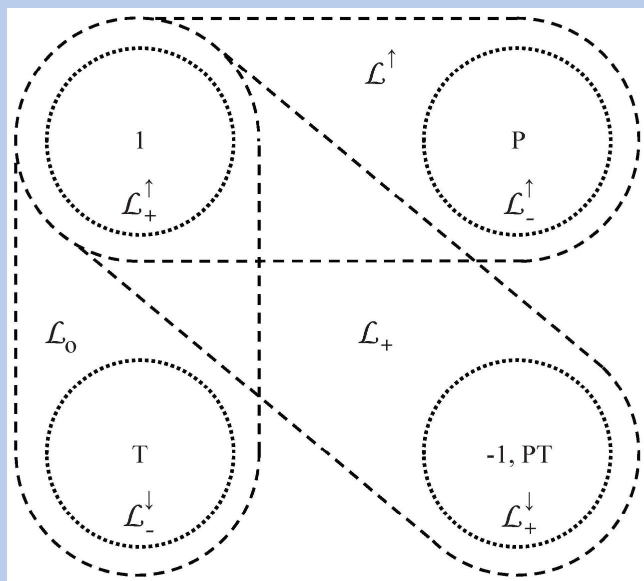
$$\mathcal{L}^{\downarrow}_{+}: \quad \det \boldsymbol{L} = 1\,, \quad \boldsymbol{L}^{0}_{\cdot\,0} \leq 1\,,$$

$$\mathcal{L}^{\downarrow}_{-}: \quad \det \boldsymbol{L} = -1\,, \quad \boldsymbol{L}^{0}_{\cdot\,0} \leq 1\,.$$

The transformation matrices $\boldsymbol{L}$ in $\mathcal{L}^{\uparrow}_{+}$ form a soubgroup — *the proper orthochronous Lorentz group* (or *restricted Lorentz group*). All other transformations in the *full Lorentz group* can be obtained from $\boldsymbol{L} \subset \mathcal{L}^{\uparrow}_{+}$ by using two discrete transformations

- Parity : $\boldsymbol{P}^{\mu}_{\cdot\,\nu} = \eta_{\mu\nu}$
- Time Reversal: $\boldsymbol{T}^{\mu}_{\cdot\,\nu} = -\eta_{\mu\nu}$.

Clearly if $\boldsymbol{L} \subset \mathcal{L}^{\uparrow}_{+}$ then $\boldsymbol{PL} \subset \mathcal{L}^{\uparrow}_{-}$, $\boldsymbol{TL} \subset \mathcal{L}^{\downarrow}_{-}$ and $\boldsymbol{PTL} \subset \mathcal{L}^{\downarrow}_{+}$.



Remarkably, *nature* is invariant under the proper orthochronous Lorenzt group $\mathcal{L}^{\uparrow}_{+}$ but *not* under the full Lorentz group.

- Parity is *violated* in the weak interactions.
- Time reversal is *violated* in K-meson (kaon), B-meson and D-meson decays. It is also theoretically present in theory of strong interactions (quantum chromodynamics — QCD).

Exercise: Dirac Equation — Technical Developments II

Helicity

The *helicity* operator h, defined

$$\boldsymbol{h} = \frac{1}{2}\frac{\boldsymbol{\Sigma}\cdot\mathbf{p}}{|\mathbf{p}|} = \frac{\hat{\boldsymbol{\sigma}}\cdot\mathbf{p}}{|\mathbf{p}|},$$

measures the spin projection into the direction of particle's motion.

Exercise 6.1 Show that $(2\boldsymbol{h})^2 = \mathbb{1}$.

Exercise 6.2 Show that the helicity operator $\boldsymbol{h}$ commutes with the Dirac Hamiltonian

$$H_D = -\gamma^0\gamma^j p_j + m\gamma^0 .$$

Exercise 6.3 The helicity projection operators are defined as

$$P^{(+)} = \frac{1}{2}(\mathbb{1} + 2\boldsymbol{h}), \qquad P^{(-)} = \frac{1}{2}(\mathbb{1} - 2\boldsymbol{h}).$$

a) Show that $P^{(+)}$ and $P^{(-)}$ are orthogonal projection operators.

b) Show that for any wave-function ψ

$$\boldsymbol{h}\,(P^{(+)}\psi) = \frac{1}{2}P^{(+)}\psi, \qquad \boldsymbol{h}\,(P^{(-)}\psi) = -\frac{1}{2}P^{(-)}\psi,$$

hence deduce that $P^{(+)}$ and $P^{(-)}$ project ψ onto states with helicity $1/2$ and $-1/2$, respectively.

Chirality

Define the operators — so-called *chiral projection operators*

$$P_R = \frac{\mathbb{1} + \gamma^5}{2}, \qquad P_L = \frac{\mathbb{1} - \gamma^5}{2},$$

so that $P_R + P_L = \mathbb{1}$.

Exercise 6.4 Show that P_R and P_L are orthogonal projectors, i.e.,

$$1)\quad P_R P_L = P_L P_R = \mathbf{0},$$

$$2)\quad P_R^2 = P_R\,,\; P_L^2 = P_L\,.$$

[**Hint:** $(\gamma^5)^2 = \mathbb{1}$.]

Exercise 6.5 Show that a Dirac wave-function ψ can be decomposed with chiral projection operators as

$$\psi = \psi_R + \psi_L\,, \qquad \psi_R = \frac{\mathbb{1} + \gamma^5}{2}\psi\,, \qquad \psi_L = \frac{\mathbb{1} - \gamma^5}{2}\psi\,,$$

where ψ_R and ψ_L are eigenstates of the *chirality* operator γ^5:

$$\gamma^5\psi_R = \psi_R\,, \qquad \gamma^5\psi_L = -\psi_L\,.$$

Exercise 6.6 Show that (for all 4 cases) $[P^{(\pm)}, P_{R,L}] = 0$.

Exercise 6.7 Consider the Dirac spinor $u(p)$ in the massless case ($m = 0$). Calculate the helicity of chiral-projected states $u_R = P_R u(p)$, and $u_L = P_L u(p)$.

[**Hint:** Show that $\Sigma^i = \gamma^0\gamma^i\gamma^5$, and use the massless Dirac equation $\gamma^\mu p_\mu u(p) = 0$.]

[**Result:** $\boldsymbol{h}\, u_R = \frac{1}{2} u_R, \boldsymbol{h}\, u_L = -\frac{1}{2} u_L$.]

Discrete Transformations of Dirac Fields

Parity (or space reflection) is the Lorentz transformation

$$x^\mu \mapsto x^\mu_P = (t, -\mathbf{x}), \qquad \boldsymbol{L}_P = \mathrm{diag}(1,-1,-1,-1), \qquad \det(\boldsymbol{L}_P) = -1 .$$

The Dirac wave-function transforms under parity as

$$\psi_P(t,\mathbf{x}) = i\eta_P \gamma^0 \psi(t,-\mathbf{x}) .$$

Exercise 6.8 Find a spin representation $S(\boldsymbol{L}_P)$ of the parity transformation.
[**Hint:** Solve the equation $S(\boldsymbol{L}_P)^{-1}\gamma^\mu S(\boldsymbol{L}_P) = (\boldsymbol{L}_P)^\mu{}_\nu \gamma^\nu$.]

Time reversal (or time inversion) is the Lorentz transformation

$$x^\mu \mapsto x^\mu_T = (-t, \mathbf{x}), \qquad \boldsymbol{L}_T = \mathrm{diag}(-1,1,1,1), \qquad \det(\boldsymbol{L}_T) = -1 .$$

The Dirac wave-function transforms under time reversal as

$$\psi_T(t,\mathbf{x}) = i\eta_T \gamma^1\gamma^3 \psi^*(-t,\mathbf{x}) .$$

Exercise 6.9 Dirac wave-function is propagating in x^3-direction with helicity $\frac{1}{2}$, i.e.

$$\psi(t,\mathbf{x}) = e^{-ip_\mu x^\mu} u_+(p), \qquad u_+(p) = \sqrt{E+m}\begin{pmatrix} \chi_+ \\ \frac{\sigma\cdot\mathbf{p}}{E+m}\chi_+ \end{pmatrix},$$

$$\text{where} \qquad \chi_+ \equiv \chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{p} = (0,0,p) .$$

What is the corresponding wave-function after time inversion?

Charge conjugation is the transformation, which reverses sign of the electric charge (in the Dirac equation minimally coupled to an electromagnetic field). The Dirac wave-function transforms under charge conjugation as

$$\psi_C(x) = i\eta_C \gamma^0\gamma^2 \bar{\psi}^T(x) = -i\eta_C \gamma^2 \psi^*(x) .$$

Exercise 6.10 What is the result of the CPT transformation $\psi \mapsto \psi_{CPT} = [(\psi_C)_P]_T$?

Exercise 6.11 Show what is the effect of charge conjugation on the right-handed chiral wave function ψ_R.

Dirac Equation — Applications II

7

7.1 Dirac's Hole Theory and Positron

Although we have found that the Dirac theory accommodates negative energy solutions whose existence should not be ignored, we have as yet not examined the physical significance of those solutions.

According to quantum theory of radiation, an excited atomic state can lose its energy discontinuously by spontaneously emitting a photon, even in absence of any external field — this is why all atomic states (save for the ground state) have *finite live time* and due to $\Delta t \Delta E \geq \hbar/2$ also *finite energy width*.

In the Dirac theory, however, the so-called ground state of an atom is not really the lowest energy state since there exists a continuum of negative-energy states from $-mc^2$ to $-\infty$. This remains to be true even in cases when a Coulomb potential is included. Indeed, the minimal substitution implies at the classical level that

$$E \ = \ \pm\sqrt{(\boldsymbol{p} - e\boldsymbol{A})^2c^2 + m^2c^4} \ + \ e\phi c \, , \tag{7.1}$$

In zero magnetic field $\boldsymbol{A} = 0$ and the maximum energy for the negative energy spectrum

$$E^<_{\text{max}} = -mc^2 \ + \ e\phi_{\text{max}}c \ \leq \ -mc^2 \, . \tag{7.2}$$

In addition $E^<$ is clearly unbounded from below as $e\phi \leq 0$.

We know that an excited atomic state makes a radiative transition to the *ground state*. Similarly, we might expect that atomic electron in the ground state with energy $mc^2 - E_{BE}$ (here E_{BE} is a binding energy) can emit spontaneously a photon of a sufficient energy that will allow the electron to bridge the energy gap to negative-energy states. Furthermore, once it reaches a negative energy states, it will keep on lowering its energy indefinitely. This scenario leads to the so-called *radiation catastrophe*, i.e., atom would radiate as "crazy" without ever attaining a stable state. Since we know that the ground state of the atom is stable, one must somehow prevent a catastrophic transition to states of $E^<$.

In 1930 Dirac proposed that all states within $E^<$ are completely filled under normal conditions. Since the subsequent Dirac's argument can be easily understood by considering only a free particle, we confine our following discussion to spin-1/2 free particle. To this end we assume that *real electrons* are described solely by positive-energy states. These are the states with $E = \sqrt{\boldsymbol{p}^2 + m^2}$. All states of negative energy are occupied by electrons — one electron in each state of negative energy with given $\boldsymbol{p}$ and spin projection λ. In this way a real electron of positive energy is prevented from falling into energetically lower and lower states by radiation emission. Hence, the radiation catastrophe is averted by Pauli's exclusion principle, which prevents these transitions.

In absence of any field (electromagnetic, etc.) the *vacuum* represents the negative energy ($E^<$) continuum (so-called *Dirac sea*), whose states are completely occupied with electrons. Occasionally one of the negative-energy electrons in the Dirac sea can absorb a photon of energy $\hbar\omega \geq 2mc^2$ and transit into positive energy states. As a result, a *hole* is created in Dirac sea. One might naturally ask, what is the meaning of such a hole in the occupied "sea" of negative states.

Even though the spectrum of energies is identical with Klein-Gordon particle, a non-existence of Pauli exclusion principle for scalar particles means that no Dirac's sea can be formed.

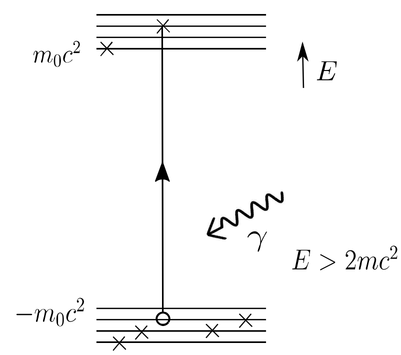


Figure 7.1: Formation of a hole within Dirac's sea.

The observable energy of the Dirac sea with a single hole in it is

$$E_{obs} = E_{vac} - (-|E_e|) = E_{vac} + |E_e|. \tag{7.3}$$

So, E_{vac} has increased, hence we expect that the absence of a negative-energy electron appears as the presence of a positive-energy particle — a *hole*.

Similarly, when a hole is created in Dirac's sea, the total charge of the Dirac's sea becomes

$$Q_{obs} = Q_{vac} - e = Q_{vac} - (-|e|) = Q_{vac} + |e|. \tag{7.4}$$

Thus, a hole in the sea of negative-energy states looks like a positive-energy particle of charge $|e|$.

Once we accept that (a) negative-energy are completely filled under normal conditions, (b) negative-energy electron can absorb a photon of energy $\hbar\omega > 2mc^2$ (just like a positive-energy electron can) to become a positive-energy electron, we are unambiguously led to the existence of a particle of a charge $|e|$ with a positive energy. This particle is called *positron*. The absorption of two-photon quanta by a negative-energy electron can be formally written as

Positron was experimentally discovered by Carl David Anderson on 2 August 1932, by observing cosmic rays in a cloud chamber (the Nobel Prize for Physics in 1936). Name *positron* first appeared in his paper.

$$e^-_{E<0} + 2\gamma \rightarrow e^-_{E>0}, \tag{7.5}$$

We may also consider a closely related process when a positive-energy emits photon and falls to the negative-energy sea, i.e.

$$e^-_{E>0} + e^+_{E>0} \rightarrow 2\gamma, \tag{7.6}$$

which is allowed only when a hole is present in Dirac's sea. Process (7.6) basically describes that both electron and hole/positron disappear and two photon quanta are generated. This process, called e^-e^+ annihilation is often observed in solids.

Note that a single γ quantum (SQA) annihilation is forbidden, for real processes, for a free e^-e^+ pair, due to the impossibility of balancing both momentum and energy conservation simultaneously. The most probable is the creation of two or more photon quanta. Nonetheless, in the presence of a third body, e.g., a nucleus in atom, SQA is allowed since the third body can recoil and allow simultaneous momentum and energy conservation.

We can gain further information on hole/positron by looking on its momentum $\boldsymbol{p}_{hole}$. Again, absence of *momentum* $\boldsymbol{p}^e$ in the Dirac sea appears as a presence of $-\boldsymbol{p}^e$ momentum of a hole. Indeed

$$\boldsymbol{p}_{obs} = \boldsymbol{p}_{vac} - \boldsymbol{p}^e \,, \tag{7.7}$$

or equivalently the momentum of hole/positron is

$$\boldsymbol{p}_{hole} = \boldsymbol{p}_{obs} - \boldsymbol{p}_{vac} = -\boldsymbol{p}^e \,. \tag{7.8}$$

Similarly, one can argue that the absence of spin up for $E < 0$ electron is manifested as a presence of spin down of $E > 0$ positron. These conclusions can be summarise in the following table below. In the table

	Q	E	$\boldsymbol{p}$	$\hat{\sigma}$	$\boldsymbol{h}$
$E < 0$ electron state	$-\lvert e\rvert$	$-\lvert E\rvert$	$\boldsymbol{p}$	$\frac{\hbar}{2}\boldsymbol{\Sigma}$	$\frac{1}{2}\boldsymbol{\Sigma}\cdot\frac{\boldsymbol{p}}{\lvert\boldsymbol{p}\rvert}$
positron state	$+\lvert e\rvert$	$+\lvert E\rvert$	$-\boldsymbol{p}$	$-\frac{\hbar}{2}\boldsymbol{\Sigma}$	$\frac{1}{2}\boldsymbol{\Sigma}\cdot\frac{\boldsymbol{p}}{\lvert\boldsymbol{p}\rvert}$

Table 7.1: Comparison of properties of negative-energy electron and ensuing positive-energy hole/positron

we have introduces the notion of *helicity*

$$\boldsymbol{h} = \frac{1}{2}\boldsymbol{\Sigma}\cdot\frac{\boldsymbol{p}}{|\boldsymbol{p}|}\,, \tag{7.9}$$

i.e., projection of the spin along the direction of particle's 3-momentum.

"Helicity"

Sometimes the choice of the Pauli spinors

$$\chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

with the particle spin quantized along the z-axis is not the most convenient basis in the spinor space. Instead of the z-axis, one can choose any quantization direction. Particularly special is the situation when the direction is chosen to be along particle's 3-momentum $\boldsymbol{p}$.

It should be noted that the spin operator $\hat{\sigma} = \frac{1}{2}\boldsymbol{\Sigma}$ does not commute with free Dirac's Hamiltonian H_D. In fact,

$$[H_D, \hat{\sigma}] = i\boldsymbol{\alpha}\times\boldsymbol{p} \neq 0\,,$$

where $\boldsymbol{\alpha}$ is Dirac's $\boldsymbol{\alpha}$-matrix (see, Eq. (3.7)).

So far, the index λ in Dirac's bispinors referred to spin projection along z-axis in the rest frame. Since the z-axis spin projection is not conserved under time evolution, it is better to label the bispinor with another intrinsic quantum number that preserves a spin projection under time evolution. *Helicity h* provides such a quantum number. Indeed, since $\boldsymbol{p}$ commutes with Dirac's free Hamiltonian we have

$$[H_D, \boldsymbol{h}] = i(\boldsymbol{\alpha} \times \boldsymbol{p}) \cdot \frac{\boldsymbol{p}}{|\boldsymbol{p}|} = 0 .$$

Here we used the identity $(\boldsymbol{A} \times \boldsymbol{B}) \cdot \boldsymbol{C} = (\boldsymbol{B} \times \boldsymbol{C}) \cdot \boldsymbol{A}$.

It can be also easily checked that

$$\boldsymbol{h}^2 = \frac{(\hat{\sigma} \cdot \boldsymbol{p})(\hat{\sigma} \cdot \boldsymbol{p})}{|\boldsymbol{p}|^2} = \frac{\{\hat{\sigma}_i, \hat{\sigma}_j\} p_i p_j}{2|\boldsymbol{p}|^2} = \frac{1}{4} .$$

So, by Hamilton-Cayley theorem the eigenvalues of $\boldsymbol{h}$ are $\pm 1/2$. Particles with helicity $+1/2$ are conventionally referred as *right-handed*, and those with $-1/2$ are *left-handed*.

Note that (free) right-handed and left-handed particle states retain their identity under the boosts only if they are massless, because only in massless case it is impossible to change particle's helicity by bringing it to rest frame and reversing its direction of motion.

In passing we might notice that the charge conjugate negative-energy solutions presented in Section 6.2 have all traits that are possessed by hole/positrons (cf. Table 7.1). In other word, positrons can be identified with charge conjugate negative-energy electrons.

7.2 Antiparticles

On the basis of our previous discussion, it would seem that only fermions (specifically spin-1/2 fermions) can have antiparticles. It is, however, a striking feature of relativistic quantum theory that it naturally ensures existence of antiparticle to any kind of particle (be it boson or fermion). Though, Quantum Field Theory provides more natural framework for discussing this issue, the essence can be already understood in the context of relativistic quantum mechanics.

To understand what is involved, let us observe that Lorentz transformations do not necessarily leave invariant the order of events. For instance, suppose that event at $\boldsymbol{x}_2$ occurs later that at $\boldsymbol{x}_1$, i.e. $\boldsymbol{x}_0^2 > \boldsymbol{x}_0^1$. A second observer who sees the first observer moving with velocity $\boldsymbol{v}$ will see the events separated by time difference

$$x_2'^0 - x_1'^0 = \boldsymbol{L}^0{}_\mu(\boldsymbol{v})(x_2^\mu - x_1^\mu) , \tag{7.10}$$

or more explicitly

For boosts in generic velocity direction one has

$$\boldsymbol{L}^i{}_j = \delta^i_j + \boldsymbol{v}_i \boldsymbol{v}_j \frac{(\gamma - 1)}{\boldsymbol{v}^2} ,$$
$$\boldsymbol{L}^0{}_i = -\gamma \boldsymbol{v}_i ,$$
$$\boldsymbol{L}^0{}_0 = \gamma , \quad \gamma = (1 - \boldsymbol{v}^2)^{-1/2} .$$

$$x_2'^0 - x_1'^0 = \gamma(x_2^0 - x_1^0) - \gamma \boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1) . \tag{7.11}$$

This is *negative* if

$$\boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1) > (x_2^0 - x_1^0) , \tag{7.12}$$

which provides a seeming causality paradox. In fact, suppose that 1st observer sees a radioactive decay $A \to B + C$ at $\boldsymbol{x}_1$, followed by absorption of particle B, e.g. $B + D \to E$ at $\boldsymbol{x}_2$.

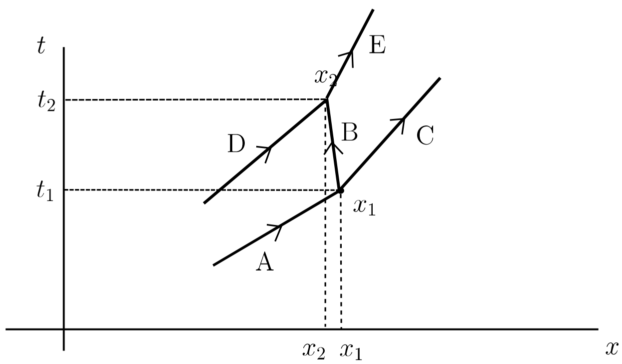


Figure 7.2: A radioactive decay $A \to B + C$ at $\boldsymbol{x}_1$, followed by absorption of particle B, e.g. $B + D \to E$ at $\boldsymbol{x}_2$.

Will then the 2nd observer see B particle absorbed at $\boldsymbol{x}_2$ before it is emitted at $\boldsymbol{x}_1$?

This paradox disappears if we note that the speed $|\boldsymbol{v}| \leq 1$, so that we can write

$$\begin{aligned} & (x_2^0 - x_1^0) \; < \; \boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1) \; = \; |\boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1)| \; \leq \; |\boldsymbol{v}||(\boldsymbol{x}_2 - \boldsymbol{x}_1)| \\ \Rightarrow \quad & (x_2^0 - x_1^0) \; < \; |(\boldsymbol{x}_2 - \boldsymbol{x}_1)| \, . \end{aligned} \tag{7.13}$$

This is clearly impossible because to travel from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ would require the average velocity greater than 1 (that is $|\boldsymbol{v}| > c$), since

$$\frac{|\boldsymbol{x}_2 - \boldsymbol{x}_1|}{(x_2^0 - x_1^0)} \; > \; 1 \, . \tag{7.14}$$

Note that temporal order can only be affected if the events x_1 and x_2 are *space-like separated*, i.e., when:

$$(x_1 - x_2)^\mu (x_1 - x_2)_\mu \; < \; 0 \, .$$

Space-like, time-like and light-like separations are Lorentz invariant concepts.

Consequently, temporal order raises no problem in classical physics, but it plays an important role in quantum theories.

In fact, in quantum theory the uncertainty principle tells us that when we specify that a particle is at position $\boldsymbol{x}_1$ at time t_1, we cannot also define its velocity precisely. Consequently, there is a certain chance of particle getting from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ even if x_1 and x_2 are space-like separated, i.e. $|\boldsymbol{x}_1 - \boldsymbol{x}_2| > |\boldsymbol{x}_1^0 - \boldsymbol{x}_2^0|$.

To be more precise, from quantum mechanical commutation relations one can derive that under Lorentz transformations [we set $(t_1, \boldsymbol{x}_1) = (0, 0)$ and $(t_2, \boldsymbol{x}_2) = (t, x, 0, 0)$]

See, e.g., Zhi-Yong Wang and Cai-Dong Xiong, Physics Letters B 659 (2008) 707–711.

$$(x_1' - x_2')^2 \; = \; ct'^2 \; - \; \boldsymbol{x}'^2 \; = \; c^2 t^2 \; - \; \boldsymbol{x}^2 \; + \; \hbar^2 c^2 \hat{\boldsymbol{H}}^{-2}/4 \, , \tag{7.15}$$

where $\hat{\boldsymbol{H}}^2 = \boldsymbol{p}^2 c^2 + m^2 c^4$ and both $\boldsymbol{p}$ and $\boldsymbol{x}$ are quantum mechanical operators with the usual canonical commutation relations. So, *quantum-mechanical Lorentz transformations* (in contrast to classical ones) *do not* generally *preserve* the notion of *time-like, space-like* or *light-like separation* under Lorenzt transformation. Indeed, let us start with *time-like (or* *light-like)* interval $c^2t'^2 - \boldsymbol{x}'^2$. Since $\hat{\boldsymbol{H}}^2 \geq m^2c^4$ (in the sense of eigenvalues or quantum-mechanical averages), we get for such a time-like (or light-like) interval that

$$\begin{aligned} 0 \leq c^2t'^2 - \boldsymbol{x}'^2 &= c^2t^2 - \boldsymbol{x}^2 + \hbar^2c^2\frac{\hat{\boldsymbol{H}}^{-2}}{4} \\ &\leq c^2t^2 - \boldsymbol{x}^2 + \frac{\hbar^2}{4m^2c^2} \\ &= c^2t^2 - \boldsymbol{x}^2 + \left(\frac{\lambda\!\!\!^{-}}{2}\right)^2 . \end{aligned} \tag{7.16}$$

Here $\lambda\!\!\!^{-} = \frac{\hbar}{mc}$ is the (reduced) Compton wavelength of the particle. Consequently, the particle in the "unprimed" frame can propagate over *space-like* interval provided that

$$0 > c^2t^2 - \boldsymbol{x}^2 \geq -\left(\frac{\lambda\!\!\!^{-}}{2}\right)^2 . \tag{7.17}$$

So, we see that the notion of time-like separation is not conserved in quantum mechanics.

Inequality (7.17) is particular version of the so-called *Weinberg formula*, which has a general form

$$0 > c^2(t_2 - t_1)^2 - (\boldsymbol{x}_1 - \boldsymbol{x}_2)^2 \geq -\left(\frac{\lambda\!\!\!^{-}}{2}\right)^2 . \tag{7.18}$$

Such space-time intervals are very "narrow" even for elementary particle masses, e.g. if m is the mass of electron, then $\lambda\!\!\!^{-} = 3.9 \times 10^{-13}$m. We are thus faced again with a paradox; if one observer sees a particle emitted at $\boldsymbol{x}_1$ and absorbed at $\boldsymbol{x}_2$, and if the Weinberg formula is satisfied, then a second observer may see the particle absorbed at $\boldsymbol{x}_2$ at a time t_2 before it is emitted at $\boldsymbol{x}_1$ at time t_1.

Though in our preceding argument it was the "primed" frame where events were time-like and "unprimed" frame where events could be space-like, the illustrative figures have the role of frames reversed so that we could make easier connection with our earlier classical causal paradox.

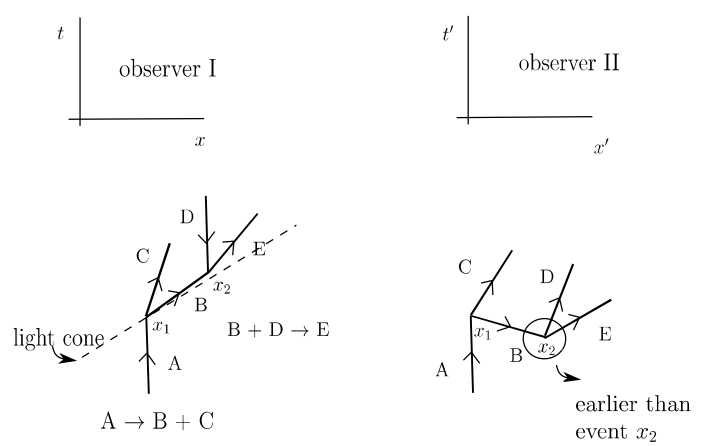


Figure 7.3: Quantum causality paradox. Time-like intervals are not generally preserved under Lorentz transformations. The violation is quantified by Weinberg formula.

There is only one known way out of this paradox. The second observer must see a particle emitted at $\boldsymbol{x}_2$ and absorbed at $\boldsymbol{x}_1$. But in general the particle seen by the second observer must necessarily be different from the first one.

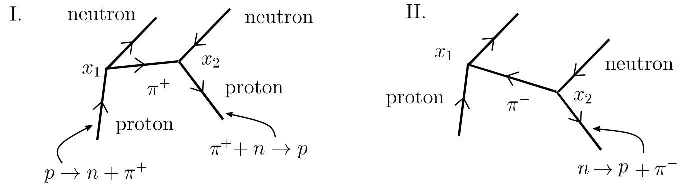


Figure 7.4: Antiparticles solve the quantum causality paradox.

For instance, if the first observer sees a proton turn into a neutron and positive π-meson at $\boldsymbol{x}_1$ and then sees π^+ and some other neutron turn into proton at $\boldsymbol{x}_2$, then the second observer must see the neutron at $\boldsymbol{x}_2$ turn into proton and a particle of a *negative* charge, which is then absorbed by a proton at $\boldsymbol{x}_1$ that turns into a neutron. Since the rest mass is a Lorentz invariant, the mass of the negative charged particle seen by the second observer will be equal to that of π^+. There is indeed experimentally observed such a particle and it is called negative π-meson (or shortly pion π^-).

This analysis allows us to make the following general statement: **For every type of charged particle there is an oppositely charged particle of equal mass.** Note, that this conclusion is not obtainable in non-relativistic quantum mechanics, nor in relativistic classical mechanics.

"Feynman–Stueckelberg Interpretation of Antiparticles"

Uncertainty relations allow that a particle can "tunnel" from time-like to space-like regions. This situation is depicted at Fig. 7.4. By comparing both figures in Fig. 7.4 we can make the following identifications:

- in frame II: π^- brings to the vertex x_1 a *positive energy*
  $\leftrightarrow$ in frame II: π^+ leaves the vertex x_1 with a *negative energy*

- in frame II: π^- brings to the vertex x_1 certain value of *spin projection*
  $\leftrightarrow$ in frame II: π^+ leaves the vertex x_1 with the *opposite value of spin projection*

- in frame II: π^- brings to the vertex x_1 certain value of *helicity*
  $\leftrightarrow$ in frame II: π^+ leaves the vertex x_1 with the *same value of helicity* (both spin and momenta are reversed).

Similar statements hold true also in connection with the vertex x_2.

Above parallelisms leads to the following *Feynman–Stueckelberg interpretation of antiparticles:*

Antiparticle can be viewed as a particle with *negative energy*, *charge* and *spin* moving *backward* in time.

7.3 Central field problem: exact solution

One of the key successes of Dirac's theory was the correct prediction of a *fine structure* in the energy levels of hydrogen. In this section we will discuss the issue of a relativistic electron in a central potential with a particular emphasize on a Coulomb potential.

We start with an eigenvalue problem for Dirac's Hamiltonian describing a particle in a central scalar potential. This reads

$$H_D\psi \;=\; [\boldsymbol{\alpha}\cdot\boldsymbol{p} \;+\; \beta m \;+\; V(r)]\psi \;=\; E\psi\,, \tag{7.19}$$

where in the Dirac representation

$$\boldsymbol{\alpha} \;=\; \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \quad \beta \;=\; \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{7.20}$$

An important observation in this context is that the *total angular momentum*

$$\boldsymbol{J} \;=\; \boldsymbol{r}\times\boldsymbol{p} \;+\; \hat{\boldsymbol{\sigma}}\,, \tag{7.21}$$

(and not orbital angular momentum $\boldsymbol{L} = \boldsymbol{r}\times\boldsymbol{p}$ or spin $\hat{\boldsymbol{\sigma}}$ separately) is conserved.

Note on conserved quantities in Quantum Mechanics

Conserved quantity (say $J_{\mu\nu}$) must transform state vector in the same spacetime point namely

$$\psi'(x) \;=\; e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\psi(x)\,. \tag{7.22}$$

So, in particular $J_{\mu\nu}$ can be easily read off from the linear $\omega^{\mu\nu}$ term in the $|\omega^{\mu\nu}| << 1$ expansion, where

$$\psi'(x) \;\sim\; \left(\mathbf{1} \;-\; \frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}\right)\psi(x)\,.$$

Reason why such $J_{\mu\nu}$ should be conserved follows from the fact that

$$H\psi = E\psi \quad\Leftrightarrow\quad e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}He^{\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\psi \;=\; Ee^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\psi\,.$$

Now, if $J_{\mu,\nu}$ is conserved then $[J_{\mu\nu}; H] = 0$. This is, however, equivalent to the statement that

$$e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}He^{\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}} \;=\; H\,,$$

which then implies that also

$$\psi'(x) \;=\; e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\psi(x)\,,$$

is an eigenstate of the Hamiltonian with the same energy E as the state $\psi(x)$. Since this statement is true at the level of energy eigenstates, it must be true also for any state vector since these these can be written as a linear combination of energy eigenstates.

We know that to the *linear* order in $\omega^{\mu\nu}$

$$\begin{aligned}
\psi'(x) &= S(L)\psi(L^{-1}x) \underbrace{\longmapsto}_{|\omega|<<1} \left(\mathbb{1} - \frac{i}{4}\boldsymbol{\sigma}_{\mu\nu}\omega^{\mu\nu}\right)\psi(L^{-1}x) \\
&= \left(\mathbb{1} - \frac{i}{4}\boldsymbol{\sigma}_{\mu\nu}\omega^{\mu\nu}\right)\psi(x^\rho - \omega^\rho{}_\nu x^\nu) \\
&= \left(\mathbb{1} - \frac{i}{4}\boldsymbol{\sigma}_{\mu\nu}\omega^{\mu\nu} - \underbrace{\omega^\rho{}_\nu x^\nu}_{\omega^{\rho\nu}x_\nu}\frac{\partial}{\partial x^\rho}\right)\psi(x) \\
&= \left[\mathbb{1} - \frac{i}{4}\boldsymbol{\sigma}_{\mu\nu}\omega^{\mu\nu} - \frac{1}{2}\omega^{\mu\nu}(x_\nu\partial_\mu - x_\mu\partial_\nu)\right]\psi(x) \\
&= \left\{\mathbb{1} - \frac{i}{2}\omega^{\mu\nu}\left[\frac{1}{2}\boldsymbol{\sigma}_{\mu\nu} + i(x_\mu\partial_\nu - x_\nu\partial_\mu)\right]\right\}\psi(x)\,.
\end{aligned}$$

So,

$$J_{\mu\nu} = \frac{1}{2}\boldsymbol{\sigma}_{\mu\nu} + i(x_\mu\partial_\nu - x_\nu\partial_\mu)\,,$$

is our candidate conserved quantity. Since the components J_{0i} are associated with boosts, they cannot be conserved, i.e., they cannot commute with Hamiltonian H_D, since boosts inevitably change the value of energy. Thus, the only possible candidate conserved quantity can be J_{ij} $(i, j = 1, 2, 3)$, which after contraction with $\frac{1}{4}\varepsilon^{ijk}$ yields the total angular momentum (7.21).

In particular, neither $\frac{1}{2}\sigma_{\mu\nu}$ nor $i(x_\mu\partial_\nu - x_\nu\partial_\mu)$ are conserved separately. Note that non of them separately transforms the state vector according to prescription (7.22).

The first thing to note is that $\{H_D; \boldsymbol{J}^2; J_z\}$ can be simultaneously diagonalized.

Proof:

Both in Weyl and Dirac representations of the matrix $\hat{\sigma}$ is block diagonal with identical blocks and hence the angular momentum operator $\boldsymbol{J}$ acts in the same way on the upper and lower bispinor components

$$\boldsymbol{J}\psi = \begin{pmatrix} \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix}\begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix}, \tag{7.23}$$

To compute the commutator $[H_D, \boldsymbol{J}]$ we turn for definiteness to Dirac's representation. In this case

$$\begin{aligned}
[H_D, \boldsymbol{J}] &= \left[\begin{pmatrix} [V(r)+m]\mathbb{1}_{2\times 2} & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & [V(r)-m]\mathbb{1}_{2\times 2} \end{pmatrix}, \begin{pmatrix} \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix}\right] \\
&= \underbrace{[\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m, \hat{\boldsymbol{\sigma}}]}_{A} + \underbrace{[\boldsymbol{\alpha}\cdot\boldsymbol{p} + V(r), \boldsymbol{r}\times\boldsymbol{p}]}_{B}\,. \qquad (7.24)
\end{aligned}$$

The commutator "A" reduces to

$$\begin{aligned}
&\left[\begin{pmatrix} m & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -m \end{pmatrix}, \begin{pmatrix} \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix}\right] \\
&\quad = \begin{pmatrix} \frac{m}{2}\boldsymbol{\sigma} & \frac{1}{2}(\boldsymbol{\sigma}\cdot\boldsymbol{p})\cdot\boldsymbol{\sigma} \\ \frac{1}{2}(\boldsymbol{\sigma}\cdot\boldsymbol{p})\cdot\boldsymbol{\sigma} & -\frac{m}{2}\boldsymbol{\sigma} \end{pmatrix} - \begin{pmatrix} \frac{m}{2}\boldsymbol{\sigma} & \frac{1}{2}\boldsymbol{\sigma}\cdot(\boldsymbol{\sigma}\cdot\boldsymbol{p}) \\ \frac{1}{2}\boldsymbol{\sigma}\cdot(\boldsymbol{\sigma}\cdot\boldsymbol{p}) & -\frac{m}{2}\boldsymbol{\sigma} \end{pmatrix} \\
&\quad = \begin{pmatrix} 0 & \frac{1}{2}[\boldsymbol{\sigma}\cdot\boldsymbol{p};\boldsymbol{\sigma}] \\ \frac{1}{2}[\boldsymbol{\sigma}\cdot\boldsymbol{p};\boldsymbol{\sigma}] & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2}[\sigma_i;\boldsymbol{\sigma}]p_i \\ \frac{1}{2}[\sigma_i;\boldsymbol{\sigma}]p_i & 0 \end{pmatrix} \\
&\quad = i\begin{pmatrix} 0 & \boldsymbol{\sigma}\times\boldsymbol{p} \\ \boldsymbol{\sigma}\times\boldsymbol{p} & 0 \end{pmatrix} = i\boldsymbol{\alpha}\times\boldsymbol{p}\,. \qquad (7.25)
\end{aligned}$$

Similarly, the commutator "B" can be written as

$$[\boldsymbol{\alpha}\cdot\boldsymbol{p} + V(r), \boldsymbol{r}\times\boldsymbol{p}] = \underbrace{[V(r), \boldsymbol{r}\times\boldsymbol{p}]}_{a} + \underbrace{[\boldsymbol{\alpha}\cdot\boldsymbol{p}, \boldsymbol{r}\times\boldsymbol{p}]}_{b}\,, \qquad (7.26)$$

where the commutator "a" reads

$$\begin{aligned}
[V(r), \boldsymbol{r}\times\boldsymbol{p}]_k &= \epsilon^{ijk}[V(r), r^i p^j] = \epsilon^{ijk} r^i [V(r), p^j] \\
&= \epsilon^{ijk} r^i \left(i\frac{\partial}{\partial x^j}V(r)\right) = \epsilon^{ijk}\left(-\frac{i r^i r^j}{r}\frac{dV}{dr}\right) = 0\,, \qquad (7.27)
\end{aligned}$$

and "b" is

$$\begin{aligned}
[\boldsymbol{\alpha}\cdot\boldsymbol{p}, \boldsymbol{r}\times\boldsymbol{p}]_k &= \alpha^i\epsilon^{lmk}[p^i, r^l p^m] = \alpha^i\epsilon^{lmk}[p^i, r^l]p^m \\
&= -i\alpha^i\epsilon^{imk}p^m = -i(\boldsymbol{\alpha}\times\boldsymbol{p})_k\,. \qquad (7.28)
\end{aligned}$$

By putting together all results obtained we get that indeed $[H_D, \boldsymbol{J}] = 0$. This directly implies that also $[H_D, \boldsymbol{J}^2] = 0$ and $[H_D, J_z] = 0$. Since the algebra for the angular momentum ensures that $[\boldsymbol{J}^2, J_i] = 0$ for all $i = 1, 2, 3$ but $[J_i, J_k] \neq 0$ for all $i \neq j$, we see that the triple of operators H_D, $\boldsymbol{J}^2$ and J_z pairwise commute and hence $\{H_D, \boldsymbol{J}^2, J_z\}$ can be simultaneously diagonalized. This concludes the proof

A consequence of our previous analysis is that the Dirac spinor angular momentum eigenstates satisfy

$$\boldsymbol{J}^2\psi = j(j+1)\psi; \quad J_3\psi \equiv J_z\psi = m\psi \quad \text{with} \quad m \in (-J, ..., J)\,, \qquad (7.29)$$

and (due to a diagonal nature of $\boldsymbol{J}$) they must be composed of two-component Pauli spinors with the same angular momentum eigenvalues, i.e., ψ from (7.29) can be written as

$$\psi = \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix} = \begin{pmatrix} \varphi_{j,m} \\ \chi_{j,m} \end{pmatrix}. \qquad (7.30)$$

Note that we have already observed this type of behavior in the case of a free Dirac particle with spin in place of total angular momentum.

There is yet another operator that commutes with H_D and $\boldsymbol{J}$. Intuitively, we expect that we must be able to specify whether the electron spin is parallel or antiparallel to the total angular momentum. In non-relativistic Quantum Mechanics these two possibilities would be distinguished by the eigenvalues of the operator

$$\boldsymbol{\sigma} \cdot \boldsymbol{J} \;=\; \boldsymbol{\sigma}\left(\boldsymbol{L} + \frac{\boldsymbol{\sigma}}{2}\right) . \tag{7.31}$$

For a relativistic electron we might try a 4×4 generalization of (7.31), namely $\boldsymbol{\Sigma} \cdot \boldsymbol{J}$ or $\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{J}$. It can be, however, checked that this would not work.

Recall that $\Sigma = 2\hat{\sigma}$

One might thus try $\beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}$, which has the same non-relativistic limit as $\boldsymbol{\Sigma} \cdot \boldsymbol{J}$. In this case

$$[H_D, \beta\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{J}] \;=\; \frac{1}{4}[H_D; \beta] \quad \text{or} \quad [H_D, \beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}] \;=\; \frac{1}{2}[H_D, \beta] . \tag{7.32}$$

The proof of this statement is quite straightforward. First we might observe that

$$\begin{aligned} [H_D, \beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}] &= [H_D, \beta](\boldsymbol{\Sigma} \cdot \boldsymbol{J}) \;+\; \beta[H_D, \boldsymbol{\Sigma} \cdot \boldsymbol{J}] \\ &= -2\beta(\boldsymbol{\alpha} \cdot \boldsymbol{p})(\boldsymbol{\Sigma} \cdot \boldsymbol{J}) \;+\; 2i\beta(\boldsymbol{\alpha} \times \boldsymbol{p}) \cdot \boldsymbol{J} . \end{aligned} \tag{7.33}$$

Now, in Dirac's representation we have

$$\begin{aligned} (\boldsymbol{\alpha} \cdot \boldsymbol{p})(\boldsymbol{\Sigma} \cdot \boldsymbol{J}) &= (\sigma^1 \otimes \boldsymbol{\sigma} \cdot \boldsymbol{p})(\mathbb{1} \otimes \boldsymbol{\sigma} \cdot \boldsymbol{J}) \\ &= [\sigma^1 \otimes (\boldsymbol{\sigma} \cdot \boldsymbol{p})(\boldsymbol{\sigma} \cdot \boldsymbol{J})] \;=\; \sigma^1 \otimes \sigma^i \sigma^j p^i J^j \\ &= \sigma^1 \otimes (\delta^{ij} + i\epsilon^{ijk}\sigma^k) p^i J^j \\ &= \underbrace{(\sigma_1 \otimes \boldsymbol{p} \cdot \boldsymbol{J})}_{\gamma^5 \boldsymbol{p} \cdot \boldsymbol{J}} + i \underbrace{[\sigma_1 \otimes \boldsymbol{\sigma} \cdot (\boldsymbol{p} \times \boldsymbol{J})]}_{i\boldsymbol{\alpha} \cdot (\boldsymbol{p} \times \boldsymbol{J})} , \end{aligned} \tag{7.34}$$

This finally implies that

$$\begin{aligned} [H_D, \beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}] &= -2\beta\gamma^5 \boldsymbol{p} \cdot \boldsymbol{J} \;-\; 2\beta[i\boldsymbol{\alpha} \cdot (\boldsymbol{p} \times \boldsymbol{J})] \;+\; 2i\beta(\boldsymbol{\alpha} \times \boldsymbol{p}) \cdot \boldsymbol{J} \\ &= -2\beta\gamma^5 \boldsymbol{p} \cdot \boldsymbol{J} \;=\; -2\beta\gamma^5 \left[\boldsymbol{p} \cdot \left(\boldsymbol{L} + \frac{\boldsymbol{\Sigma}}{2}\right)\right] \;=\; -2\beta\gamma^5 \boldsymbol{p} \cdot \frac{\boldsymbol{\Sigma}}{2} \\ &= -2\beta\left(\sigma_1 \otimes \frac{\boldsymbol{\sigma}}{2}\right) \cdot \boldsymbol{p} \;=\; -\beta\boldsymbol{\alpha} \cdot \boldsymbol{p} \;=\; \frac{1}{2}[H_D, \beta] . \end{aligned} \tag{7.35}$$

This result [see also Eq. (7.32)] indicates that we can define a new operator K:

$$\begin{aligned} K &= \beta\boldsymbol{\Sigma} \cdot \boldsymbol{J} - \frac{\beta}{2} \;=\; \beta\left[\boldsymbol{\Sigma}\left(\boldsymbol{L} + \frac{\boldsymbol{\Sigma}}{2}\right) - \frac{1}{2}\right] \;=\; \beta\left(\boldsymbol{\Sigma}\boldsymbol{L} + \frac{\boldsymbol{\Sigma}^2}{2} - \frac{1}{2}\right) \\ &= \beta\left(\boldsymbol{\Sigma}\boldsymbol{L} + \frac{3}{2} - \frac{1}{2}\right) \;=\; \beta(\boldsymbol{\Sigma}\boldsymbol{L} + 1) , \end{aligned} \tag{7.36}$$

which commutes with H_D, i.e., $[H_D; K] = 0$. Furthermore, since $\boldsymbol{J}$ commutes with β and $\boldsymbol{\Sigma} \cdot \boldsymbol{L}$ then also $[\boldsymbol{J}, K] = 0$ and hence $[\boldsymbol{J}^2, K] = 0$. Consequently, for Dirac's particle in a central potential we can construct a simultaneous eigenfunctions of H_D; K; $\boldsymbol{J}^2$ and J_z. The corresponding eigenvalues will be further denoted as E, $-\kappa$, $j(j+1)$ and m, respectively.

It should be noted that κ and j are not totally independent (similarly as m and j are not). To see how they are related let us consider K^2, this gives

Here we use the identity: $\sigma^i\sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k$.

$$\begin{aligned}
K^2 &= \beta(\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1)\beta(\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1) \;=\; \beta^2(\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1)^2 \\
&= (\boldsymbol{\Sigma}\cdot\boldsymbol{L})(\boldsymbol{\Sigma}\cdot\boldsymbol{L}) + 2\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1 \\
&= \begin{pmatrix}\sigma^i & 0\\ 0 & \sigma^i\end{pmatrix}\begin{pmatrix}\sigma^j & 0\\ 0 & \sigma^j\end{pmatrix} L_i L_j + 2\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1 \\
&= \boldsymbol{L}^2 + i\Sigma^k\epsilon^{ijk}L_iL_j + 2\boldsymbol{\Sigma}\boldsymbol{L}+1 \\
&= \boldsymbol{L}^2 - \boldsymbol{\Sigma}\boldsymbol{L} + 2\boldsymbol{\Sigma}\boldsymbol{L}+1 \;=\; \boldsymbol{L}^2 + \boldsymbol{\Sigma}\boldsymbol{L}+1\,.
\end{aligned} \tag{7.37}$$

At the same time we have

$$\boldsymbol{J}^2 = \left(\boldsymbol{L}+\frac{\boldsymbol{\Sigma}}{2}\right)^2 = \boldsymbol{L}^2 + \frac{2}{2}\boldsymbol{\Sigma}\cdot\boldsymbol{L} + \frac{\Sigma^2}{4} = \boldsymbol{L}^2 + \boldsymbol{\Sigma}\cdot\boldsymbol{L} + \frac{3}{4}\,, \tag{7.38}$$

so, we can write $K^2 = \boldsymbol{J}^2 + \frac{1}{4}$. Thus, eigenvalues $\boldsymbol{J}^2$ and K^2 are related to each other by the relation

$$\kappa^2 = j(j+1) + \frac{1}{4} = \left(j+\frac{1}{2}\right)^2, \tag{7.39}$$

or in other words, $\kappa = \pm\left(j+\frac{1}{2}\right)$. Since $j = 1/2, 3/2, 5/2, \dots$ we see that κ is a *non-zero integer* which can be both positive and negative.

From (7.36) follows that K has an explicit form

$$K = \begin{pmatrix}\boldsymbol{\sigma}\boldsymbol{L}+1 & 0\\ 0 & -\boldsymbol{\sigma}\boldsymbol{L}-1\end{pmatrix} = \begin{pmatrix}\boldsymbol{\sigma}\boldsymbol{J}-\frac{1}{2} & 0\\ 0 & -\boldsymbol{\sigma}\boldsymbol{J}+\frac{1}{2}\end{pmatrix}. \tag{7.40}$$

Note:

Pictorially speaking, the sign of κ determines whether the spin is *antiparallel* ($\kappa > 0$) or *parallel* ($\kappa < 0$) to $\boldsymbol{J}$ in the non-relativistic limit.

So, if Dirac's wave function ψ (assumed to be an energy eigenfunction) is a simultaneous eigenfunction of K, $\boldsymbol{J}^2$ and J_z then it must satisfy

$$\underbrace{(\boldsymbol{\sigma}\boldsymbol{L}+1)\psi_+ = -\kappa\psi_+\,,\quad (\boldsymbol{\sigma}\boldsymbol{L}+1)\psi_- = \kappa\psi_-\,,}_{K\psi = -\kappa\psi} \tag{7.41}$$

and

$$\begin{aligned}
\boldsymbol{J}^2\psi_\pm &= (\boldsymbol{L}+\boldsymbol{\sigma}/2)^2\psi_\pm = j(j+1)\psi_\pm\,, \\
J_z\psi_\pm &= (L_3+\sigma_3/2)^2\psi_\pm = m\psi_\pm\,.
\end{aligned} \tag{7.42}$$

Note that the operator $\boldsymbol{L}^2 = \boldsymbol{J}^2 - \boldsymbol{\Sigma}\cdot\boldsymbol{L} - \frac{3}{4}$ when acted on ψ_+ and ψ_-

gives

$$
\begin{aligned}
\boldsymbol{L}^2\psi_+ &= j(j+1)\psi_+ + \kappa\psi_+ + \frac{1}{4}\psi_+ \\
\Rightarrow\quad l_+(l_++1) &= j(j+1) + \kappa + \frac{1}{4}\,, \qquad (7.43)\\
\boldsymbol{L}^2\psi_- &= j(j+1)\psi_- - \kappa\psi_- + \frac{1}{4}\psi_- \\
\Rightarrow\quad l_-(l_-+1) &= j(j+1) - \kappa + \frac{1}{4}\,. \qquad (7.44)
\end{aligned}
$$

So, any two-component eigenfunction of $(\boldsymbol{\sigma L}+1)$ and $\boldsymbol{J}^2$ is also an eigenfunction of $\boldsymbol{L}^2$. Thus, although the four-component bi-spinor $\psi = \binom{\psi_+}{\psi_-}$ is not an eigenfunction of $\boldsymbol{L}^2$ (since H_D does not commute with $\boldsymbol{L}^2$) ψ_+ and ψ_- *separately* are eigenfunctions of $\boldsymbol{L}^2$ whose eigenvalues are denoted as $l_+(l_++1)$ and $l_-(l_-+1)$, respectively. From (7.43)-(7.44) we can read off the relation between $l_\pm$ and $\pm\kappa$, namely

$$
\begin{aligned}
\psi_+ :\quad -\kappa &= j(j+1) - l_+(l_++1) + \frac{1}{4}\,,\\
\psi_- :\quad \kappa &= j(j+1) - l_-(l_-+1) + \frac{1}{4}\,. \qquad (7.45)
\end{aligned}
$$

Using, in addition, the fact that $\kappa = \pm\left(j+\frac{1}{2}\right)$ we can determine l_+ and l_- for a given κ. The result is depicted in the table Tab. 7.2.

	l_+	l_-
$\kappa = j+\frac{1}{2}$	$j+\frac{1}{2}$	$j-\frac{1}{2}$
$\kappa = -\left(j+\frac{1}{2}\right)$	$j-\frac{1}{2}$	$j+\frac{1}{2}$

Table 7.2: Relation among $\pm\kappa$ and l_+ and l_-.

To see how this table is constructed, let us look, e.g., at the first upper entry in the first column, i.e., $j+\frac{1}{2}$. This results from the first equation in (7.45) by writing

$$
\begin{aligned}
-\left(j+\tfrac{1}{2}\right) &= j(j+1) - l_+(l_++1) + \tfrac{1}{4}\\
&= j^2 + j - l_+^2 - l_+ + \tfrac{1}{4}\\
\Rightarrow\quad l_+^2 + l_+ &= j^2 + 2j + \tfrac{3}{4} = \left(j+\tfrac{1}{2}\right)^2 + \left(j+\tfrac{1}{2}\right). \qquad (7.46)
\end{aligned}
$$

This confirms the stated entry in the table. Similarly we would derive another 3 entries in the table Tab. 7.2.

Thus, for given j we see that $l_\pm$ can assume two possible values corresponding to two possible values of κ, so instead of κ one can use l_+ and l_-. This is particularly convenient since the spectrosopic notation involves orbital angular momentum rather than κ and our results can be better compared with existent spectroscopic data.

Table 7.4

n	$n_c = n - \lvert\kappa\rvert \geqslant 0$	$\kappa = \pm(j + 1/2)$	spectroscopic notation
1	0	-1 $j = 1/2;\ l_+ = 0$	$1s_{1/2}$
2	1	-1 $j = 1/2;\ l_+ = 0$	$2s_{1/2}$
2	1	1 $j = 1/2;\ l_+ = 1$	$2p_{1/2}$
2	0	-2 $j = 3/2;\ l_+ = 1$	$2p_{3/2}$
3	2	-1 $j = 1/2;\ l_+ = 0$	$3s_{1/2}$
3	2	1 $j = 1/2;\ l_+ = 1$	$3p_{1/2}$
3	1	-2 $j = 3/2;\ l_+ = 1$	$3p_{3/2}$
3	1	2 $j = 3/2;\ l_+ = 2$	$3d_{3/2}$
3	0	-3 $j = 5/2;\ l_+ = 2$	$3d_{5/2}$

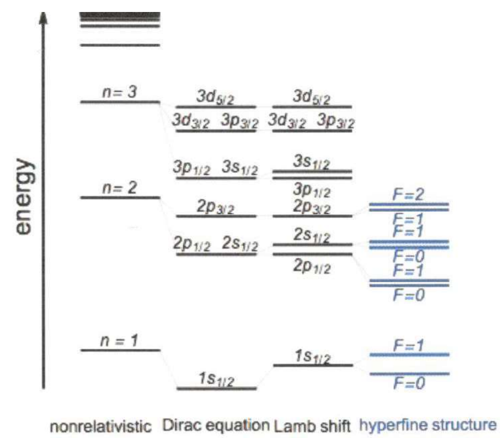


Figure 7.5: Schematic energy level diagram for H atom with relativistic corrections included.

For $n = 2$, $Z = 1$

$$E(2p_{3/2} - 2p_{1/2}) \simeq -\frac{1}{2}\frac{\alpha^4 mc^2}{8}\left[\frac{1}{2} - 1\right] = \frac{\alpha^4 mc^2}{32} = 4,53 \cdot 10^{-5}\ \text{eV}$$

7.4 Relativistic higher-spin wave equations

Apart from Klein-Gordon wave equation (for spin - 0 particle) and Dirac's wave equation (for spin – 1/2 particle), there exists a number of relativistic higher-spin wave equations.

Examples of higher-spin wave equations:

- "Maxwell equation" (for *massless* spin - 1 particle)

  $$\partial_\mu \partial^\mu A^\nu \;=\; e\bar{\psi}\gamma^\nu\psi \quad \text{(valid in Lorentz gauge where } \partial_\mu A^\mu = 0\text{)}$$

  This wave equation describes a photon in interaction with electrically charges spin 1/2 particle. For completeness, this equation should be complemented also with corresponding Dirac's equation with electromagnetic potential included (e.g., via minimal substitution). Note that the role of wave function for a photon is played by the gauge potential (i.e., wave function carries a vector index as an index of internal symmetry, which in turn means that it transforms in the vector representation of Lorentz group).

Single photon equation without any source term is described by the conventional system of Maxwell equations. In this connection it is interesting to observe that Maxwell equations can be equivalently written as

$$i\frac{\partial \boldsymbol{E}}{\partial t} = \frac{1}{i}\boldsymbol{S}\cdot\boldsymbol{\nabla}(i\boldsymbol{B}),$$
$$i\frac{\partial i\boldsymbol{B}}{\partial t} = \frac{1}{i}\boldsymbol{S}\cdot\boldsymbol{\nabla}(\boldsymbol{E}),$$

where $(S_i)_{jk} = -i\varepsilon_{ijk}$ represents angular momentum generator in *adjoint representation* (recall that $[S_i, S_j] = i\varepsilon_{ijk}S_k$).

Recall that the intensity of electromagnetic radiation (i.e., square of its amplitude — be it $\boldsymbol{E}$ or $\boldsymbol{B}$) is due to Einstein's explanation of the *photoelectric effect* proportional to the (average) density of photons in the radiation. In particular, if the intensity of monochromatic electromagnetic field is sufficiently low so that it can support only *one* quantum of energy — photon, then the corresponding field can be interpreted as being proportional to *probability density amplitude*.

This system of equations should be compared to the system of equations for massless spin 1/2 particle (so called *Weyl equations*)

$$i\frac{\partial \varphi}{\partial t} = \frac{1}{i}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}(\chi),$$
$$i\frac{\partial \chi}{\partial t} = \frac{1}{i}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}(\varphi).$$

Here Dirac's bispinor

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}.$$

So, we see that matrices S_i play for photon (i.e., particle with spin $s = 1$ and rest mass $m_0 = 0$) the same role as Pauli matrices for spin 1/2 massless particles (so called Weyl fermions). Similarly, $(\boldsymbol{E}, i\boldsymbol{B})$ is analogous to (φ, χ).

Note that

$$\sum_{i=1}^{3} S_i^2 = s(s+1)\mathbb{1}_{3\times 3} = 2\cdot\mathbb{1}_{3\times 3},$$

which implies that S indeed describes a particle with spin $s = 1$.

- *Proca equation* (for *massive* spin - 1 particle)

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) + mA^\nu = 0.$$

This wave equation can be, in principle, used to describe such particles as, e.g., $W^\pm$ or Z^0 bosons.

- *Rarita–Schwinger equation* (for massive spin - 3/2 particle)

$$\left(\epsilon^{\mu\nu\rho\sigma}\gamma^5\gamma_\nu\partial_\rho + m\sigma^{\mu\sigma}\right)\psi_\sigma = 0,$$

where $\sigma^{\mu\nu} = (i/2)[\gamma^\mu, \gamma^\nu]$ is a Lorentz group generator in bispinor representation. This wave equation can be, in principle, used to describe *gravitino* that is predicted in supersymmetric theories (SUSY).

- *Bargmann-Wigner equation* (for massive arbitrary spin free particle) - quite complicated system of equations

Relativistic wave equations have a number of conceptual difficulties.

- No simple way to include multi-particle interactions.
- Strictly single-particle description does not allows (is not applicable for) unstable particles or resonances.
- Various paradoxes: "Zitterbewegung", Klein's paradox, problematic probabilistic interpretations (we already seen for Klein-Gordon particle).
- Single-particle picture is not tenable beyond energies that allow for pair (or multi-particle) production.
- No elementary particles has been observed beyond spin 1 particles, though there are various theoretical reasons supporting existence of higher spin particles, e.g., SUSY predicts 3/2-spin particle — *gravitino*, and quantum gravity predicts spin 2 particle — *graviton*.

These arguments provide good reasons for abandoning relativistic wave equations — higher-spin wave equations in particular. On the other hand, lower-spin wave equations (such as those with spin 0, 1/2 and 1) are often used as a starting point in setting up corresponding Quantum Field Theories via the so-called *second quantization procedure* (see following chapter).

7.5 Exercises: Dirac particle in central potential

Dirac particle with charge q in central electric field with potential $V(r) = q\phi(r)$ (and zero magnetic field) is described by the equation

$$i\partial_t\Psi(t,\mathbf{x}) = H_D\Psi(t,\mathbf{x}), \quad H_D = \boldsymbol{\alpha}\cdot\hat{\mathbf{p}} + m\beta + V(r). \tag{7.66}$$

where, in the standard representation

$$\alpha^i \equiv \gamma^0\gamma^i = \begin{pmatrix} \mathbb{O} & \sigma^i \\ \sigma^i & \mathbb{O} \end{pmatrix} = \sigma^1 \otimes \sigma^i \quad , \quad \beta \equiv \gamma^0 = \begin{pmatrix} \mathbb{I} & \mathbb{O} \\ \mathbb{O} & -\mathbb{I} \end{pmatrix} = \sigma^3 \otimes \mathbb{I}. \tag{7.67}$$

Exercise 7.1 Show that the operator of total (i.e., orbital plus spin) angular momentum

$$J^i = L^i + \frac{1}{2}\Sigma^i \quad , \quad L^i = \varepsilon_{ijk}x^j\hat{p}^k = \varepsilon_{ijk}x^j(-i\partial_k) \quad , \quad \Sigma^i = \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} = \mathbb{I}\otimes\sigma^i \tag{7.68}$$

commutes with the Dirac Hamiltonian H_D.

Problem 7.1 Show that

$$[H_D, \beta\mathbf{\Sigma}\cdot\mathbf{J}] = \frac{1}{2}[H_D, \beta]. \tag{7.69}$$

- *Bargmann-Wigner equation* (for massive arbitrary spin free particle) - quite complicated system of equations

Relativistic wave equations have a number of conceptual difficulties.

- No simple way to include multi-particle interactions.
- Strictly single-particle description does not allows (is not applicable for) unstable particles or resonances.
- Various paradoxes: "Zitterbewegung", Klein's paradox, problematic probabilistic interpretations (we already seen for Klein-Gordon particle).
- Single-particle picture is not tenable beyond energies that allow for pair (or multi-particle) production.
- No elementary particles has been observed beyond spin 1 particles, though there are various theoretical reasons supporting existence of higher spin particles, e.g., SUSY predicts 3/2-spin particle — *gravitino*, and quantum gravity predicts spin 2 particle — *graviton*.

These arguments provide good reasons for abandoning relativistic wave equations — higher-spin wave equations in particular. On the other hand, lower-spin wave equations (such as those with spin 0, 1/2 and 1) are often used as a starting point in setting up corresponding Quantum Field Theories via the so-called *second quantization procedure* (see following chapter).

Exercises: Dirac particle in central potential

Dirac particle with charge q in central electric field with potential $V(r) = q\phi(r)$ (and zero magnetic field) is described by the equation

$$i\partial_t \Psi(t,\mathbf{x}) = H_D \Psi(t,\mathbf{x}), \qquad H_D = \boldsymbol{\alpha}\cdot\hat{\mathbf{p}} + m\beta + V(r).$$

where, in the standard representation

$$\alpha^i \equiv \gamma^0\gamma^i = \begin{pmatrix} \mathbf{0} & \sigma^i \\ \sigma^i & \mathbf{0} \end{pmatrix} = \sigma^1 \otimes \sigma^i, \qquad \beta \equiv \gamma^0 = \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{pmatrix} = \sigma^3 \otimes \mathbb{1}.$$

Exercise 7.1 Show that the operator of total (i.e., orbital plus spin) angular momentum

$$J^i = L^i + \frac{1}{2}\Sigma^i, \qquad L^i = \varepsilon_{ijk} x^j \hat{p}^k = \varepsilon_{ijk} x^j(-i\partial_k), \qquad \Sigma^i = \begin{pmatrix} \sigma^i & \mathbf{0} \\ \mathbf{0} & \sigma^i \end{pmatrix} = \mathbb{1}\otimes\sigma^i,$$

commutes with the Dirac Hamiltonian H_D.

Exercise 7.2 Prove that

$$[H_D, \beta\boldsymbol{\Sigma}\cdot\mathbf{J}] = \frac{1}{2}[H_D,\beta],$$

hence, show that the operator

$$K \equiv \beta\boldsymbol{\Sigma}\cdot\mathbf{J} - \frac{1}{2}\beta = \beta(\boldsymbol{\Sigma}\cdot\mathbf{L} + 1),$$

commutes with the Hamiltonian H_D.

Exercise 7.3 Show that the operator K commutes with $\mathbf{J}$.

Exercise 7.4 Show that $K^2 = \mathbf{J}^2 + \frac{1}{4}$.

In summary, the operators

$$H_D\,,\quad K\,,\quad \mathbf{J}^2\,,\quad J_3\,,$$

all mutually commute. In fact, they constitute the *complete set of observables* for Dirac particle in central potential and consequently their simultaneous eigenfunctions uniquely label the states.

Part II

INTRODUCTION TO QUANTUM FIELD THEORY

7.5 Why Quantum Field Theory

Let us now put forward a couple of reasons that explain a conceptual inevitability of quantum field theory.

I. — The combination of quantum mechanics and special relativity implies that particle number is not conserved. Relativity necessarily brings in the possibility of conversion of mass into energy and vice versa, i.e., the creation and annihilation of particles. For instance, β decay of the neutron via $n \to p + e^- + \bar{\nu}_e$ or positron-electron annihilation $e^+e^- \to 2\gamma$. The latter case shows that there are situations when the number of particles of given species is not conserved, even though the number of particles of all types taken together is conserved.

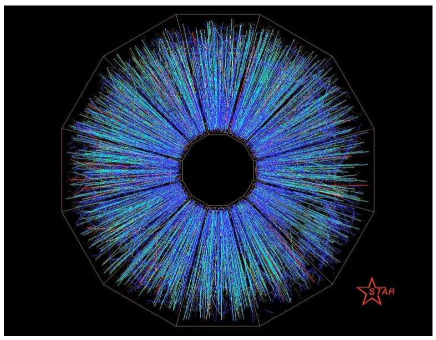


Figure 7.6: Here we show an extreme demonstration of particle creation. This comes from the Relativistic Heavy Ion Collider (RHIC) at Brookhaven, Long Island. This machine crashes gold nuclei together, each containing 197 nucleons. The resulting explosion contains up to 10,000 particles, captured here in all their beauty by the STAR detector.

It should be stressed that the creation of particles is impossible to avoid whenever one tries to locate a particle of mass m within its Compton wavelength. Indeed, from Heisenberg's uncertainty relation we find that (consider motion along x direction)

$$\begin{aligned}
\sigma_E^2\sigma_x^2 &\geq \frac{1}{4}\left|\langle[\hat{H}, x]\rangle_\psi\right|^2 = \frac{1}{4}\left|\left\langle[\sqrt{p_x^2c^2 + m^2c^4}, x]\right\rangle_\psi\right|^2 \\
&= \frac{\hbar^2}{4}\left|\left\langle p_x c^2/\sqrt{p_x^2c^2 + m^2c^4}\right\rangle_\psi\right|^2 \\
&= \frac{\hbar^2}{4}\left|c + O\left(\left\langle\frac{m^2c^4}{p_x^2c^2}\right\rangle_\psi\right)\right|^2 \sim \frac{\hbar^2c^2}{4}\,.
\end{aligned} \tag{7.65}$$

This implies that $\sigma_E\sigma_x = \Delta E\Delta x \sim \hbar c/2$. If we assume that $\Delta x \sim \lambda\!\!\!^{-}_C = \hbar/mc$, then we have $\Delta E \sim mc^2$. Therefore, in a relativistic theory, the fluctuations of the energy are enough to allow the creation of particles out of the vacuum. In the case of spin $\frac{1}{2}$ particle, the Dirac sea picture shows clearly how, when the energy fluctuations are of order mc^2, electrons from the Dirac sea can be excited to positive energy states, thus creating electron-positron pairs. So, at distances shorter than particle's Compton wavelength there is a high probability that we will see particles swarming around the original particle we put in.

Note about Compton wavelength

Particle's (reduced) Compton wavelength $\lambda\!\!\!^{-}_C = \hbar/mc$ is always smaller than the corresponding de Broglie wavelength $\lambda_{dB} = \hbar/|\boldsymbol{p}|$. In fact, we can say that:

- λ_{dB} is the distance at which the wave nature of a particle is apparent
- $\lambda\!\!\!^{-}_C$ is the distance at which the concept of a single particle breaks

In order to discuss such processes, the usual formalism of many-body quantum mechanics with wave functions of fixed number of particles, has to be augmented by including the possibility of creation and annihilation of particles via interaction.

II. — Ordinary (non-relativistic) point-particle QM can deal with the quantum description of a many-body system in terms of many body wave functions. This is important, e.g., in atomic, molecular or condense matter physics. Similar generalization for relativistic particles would be desirable. Problem with this generalization, however, starts already at classical level. There does not exist any generalization to a relativistically invariant interacting many-body theory — not even for 2 interacting particles. This is known as *Leutwyler's no-interaction theorem*

H. Leutwyler, Il Nuovo Cimento **37**, 556 (1965).

Any *finite number* of point particles cannot interact in such a way that the principles of special relativity are respected, i.e. that the system states correspond to some representation of the Lorentz (or more generally Poincare) group/algebra. Accordingly, classical relativistic point particles are necessarily free, as a consequence of Poincare invariance.

Note

The only exception are two particles in one spatial dimension confined to each other by a linearly rising potential.

In contrast to point particles, strings (i.e., 1 dimensional objects) can interact relativistically in higher dimensions, without violating Leutwyler's non-interaction theorem.

Hence, it is not surprising that particle physics is based on Quantum Field Theory (i.e., infinite number of degrees of freedom) rather than on relativistic point particle quantum mechanics.

III. — We know of classical field that is fundamental in physics — the electromagnetic field. Analyses of Bohr and Rosenfeld showed that there are difficulties in having a quantum description of various charged particle phenomena (such as those that occur in atomic physics) while retaining a classical description of the electromagnetic field. One has to quantize the electromagnetic field (e.g., to get Lamb shift correctly); this is independent of any many-particle interpretation that might emerge from quantization.

IV. — Because all particles of the same type are the same. What we mean by this is that, for instance, two electrons are identical in every way, regardless of where they came from and what they have been through. The same is true of every other fundamental particle. Let us illustrate this through a rather prosaic story. Suppose we capture a proton from a cosmic ray which we identify as coming from a supernova lying 8 billion light years away. We compare this proton with one freshly created in a particle accelerator here on Earth. And the two are exactly the same! How is this possible? Why are not there errors in proton production? How can two objects, manufactured so far apart in space and time, be identical in all respects? One explanation that might be offered is that there's a sea of proton "stuff" filling the universe and when we make a proton we somehow dip our hand into this stuff and from it mould a proton. Then it's not surprising that protons produced in different parts of the universe are identical: they're made of the same stuff. It turns out that this is roughly what happens. The "stuff" is the proton field or, if one looks closely enough, the quark field.

Note on field quantization I

There are two complementary approaches that are typically employed in field quantization.

a) One can postulate fields as the basic dynamical variables and show that the result can be interpreted in many-body terms.

b) One can start with point-particles as the basic objects of interest and derive (or construct) field operator as an efficient way of organizing the many-particle states.

Here we will work with the first approach as this gets us into the subject quickly. The second approach is often a starting point in non-relativistic field theory that is typically employed in condensed matter physics. Approach **a)** is known as *Quantum Field Theory* (QFT) or *Theory of Quantized Fields*, while **b)** is known as *Second Quantization*.

Note on field quantization II

There is a number of different types of quantization schemes, each with its own merits and drawbacks.

- *Canonical quantization* — will also be used as our starting point. It emulates the conventional quantization procedure used in Quantum Mechanics. In particular, time is singled out as a special coordinate and manifest Lorentz invariance is renounced. The advantage of canonical quantization is that it quantizes only physical modes, which ensures that unitarity is manifest. In simple cases such as scalar and fermion fields or quantum electrodynamics (QED) this method is relatively easy to apply. In more complicated (though experimentally important) cases it is impractical.

- *Gupta-Bleuler or Covariant quantization* — maintains full Lorentz symmetry (contrary to canonical quantization), which is clearly a great advantage. The disadvantage of this approach is that "ghosts" or unphysical states of negative norm are allowed to propagate in the theory, and are eliminated only when one applies constraints to the state vectors. This approach was historically most successful in QED. It is rather limited in its scope and rarely used beyond QED.

- *Functional integral* — will be utilized in the second part of this lecture. It is a simple, intuitive method with a close connection with classical physics (it employs, e.g., Lagrangian density, action functional or Hamilton variational principle). It provides an excellent tool for various semiclassical approximations. In addition, the functional integral is formulated in manifestly Lorentz covariant fashion. Corresponding "ghosts" are killed by another type of "ghosts" (the so called "good ghosts" or Faddeev-Popov ghosts). It is also an ideal bookkeeping tool allowing for a systematic perturbation expansion in terms of the so-called *Feynman diagrams*. The disadvantage of this approach is that functional integration is mathematically delicate operation that may not even exist in Minkowski space.

- *Becchi-Rouet-Stora-Tyutin (BRST) & Batalin-Vilkovisky quantization schemes* — are fully covariant quantization methods. They are used in complicated systems, such as non-abelian Yang-Mills theories, string theory or quantum gravity. They can be (and as a rule are) expressed in terms of functional integrals.

- *Stochastic quantization* — is conceptually very different from previous quantization schemes. The main idea is to view Euclidean field theory (i.e., field theory where time variable is analytically continued to imaginary values) as an equilibrium limit of a statistical system coupled to a thermal reservoir. This system evolves in a fictitious time direction t (5th parameter) until it reaches equilibrium limit as $t \to \infty$. The coupling to a heat reservoir is simulated by means of a stochastic noise. Stochastic quantization is particularly suitable for numerical applications.

8 Some useful background from quantum mechanics and classical field theory

Let us first recall the familiar path to the quantization of a classical dynamical system in particle mechanics. For the purpose of illustration we consider a 1-D motion of a particle in a conservative potential. Let q be the (generalized) coordinate of the particle, $\dot{q} = dq/dt$ the velocity, and $L(q, \dot{q})$ the Lagrangian. According to Hamilton's principle, the dynamics of the particle is determined by the condition

$$\delta S[q] \ = \ \delta \int_{t_1}^{t_2} dt \, L(q, \dot{q}) \ = \ 0 \, , \tag{8.1}$$

which provides an actual physical trajectory $q(t)$ from (q_1, t_1) to (q_2, t_2). Eq. (8.1) states that the action functional is stationary around classical trajectory, i.e. small variations from classical path, $q(t) \to q(t) + \delta q(t)$, leave the action unchanged to the first order in the variation.

This is analogous to the situation when $df(x) = 0$, which implies that $f(x)$ is stationary around the extremal point x_0 of $f(x)$, i.e., $f(x_0) = f(x_0 + dx)$ to the first order in dx.

Hamilton's principle gives us the well known Euler–Lagrange equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}} \ - \ \frac{\partial L}{\partial q} \ = \ 0 \, . \tag{8.2}$$

In order to carry out the formal quantization based on this equation, we rewrite it in the Hamiltonian form, by defining the momentum p conjugate to q as

$$p \ \equiv \ \frac{\partial L}{\partial \dot{q}} \, , \tag{8.3}$$

and introduce the Hamiltonian via the Legendre transformation

$$H(p, q) \ = \ p\dot{q} \ - \ L(q, \dot{q}) \, . \tag{8.4}$$

Note in passing that H does not dependent on $\dot{q}$ since

$$\begin{aligned} \mathrm{d}H \ &= \ (dp)\dot{q} \ + \ p d\dot{q} \ - \ \frac{\partial L}{\partial q} dq \ - \ \frac{\partial L}{\partial \dot{q}} d\dot{q} \\ &= \ (dp)\dot{q} \ - \ \frac{\partial L}{\partial q} \mathrm{d}q \, . \end{aligned} \tag{8.5}$$

In terms of H, the Euler-Lagrange equation can be equivalently rewritten as

$$\begin{aligned} \{q, H\}_{PB} \ &= \ \frac{\partial H}{\partial q} \ = \ \dot{q} \, , \\ \{p, H\}_{PB} \ &= \ -\frac{\partial H}{\partial q} \ = \ \dot{p} \, , \end{aligned} \tag{8.6}$$

where $\{\cdot, \cdot\}_{PB}$ denotes a Poisson bracket. To quantize (8.6) we let q become a Hermitan operator in a Hilbert space and replace p by $-i\partial/\partial q$, so that the conjugate momentum and coordinate satisfy a commutation relation

$$[q, p] \ = \ i\hbar \, , \tag{8.7}$$

which corresponds to the classical Poisson bracket $\{q, p\}_{PB} = 1$. The dynamics of our particle is contained in the Schrödinger equation

$$H(p, q)\psi(t) \;=\; i\frac{\partial \psi(t)}{\partial t}\,, \tag{8.8}$$

where $\psi(t)$ is a wave function (or state vector) in the Hilbert space. In this formulation all time dependence is carried by ψ while p and q are not time dependent. This approach is known as the *Schrödinger picture*.

Alternatively, we may transfer whole time dependence to the operators $q(t)$ and $p(t)$ while ψ will be time independent. This is known as *Heisenberg picture*. Both pictures are equivalent as they can be connected via unitarity transformation. In fact, from Schrödinger equation it follows that

There also exist intermediate pictures, such as *Dirac picture* or *thermo-field dynamics* that will be discussed later on.

$$\psi_S(t) \;=\; e^{-iH_S t}\psi_S(0) \;=\; e^{-iHt}\psi_H\,, \tag{8.9}$$

where the value $\psi_S(t = 0)$ is set to coincide with ψ_H. Similarly, operators in both pictures are connected according to prescription

$$O_H(t) \;=\; e^{iH_S t} O_S e^{-iH_S t}\,. \tag{8.10}$$

It is clear that the unitary transformation (8.10) is constructed so that matrix elements of all observables are identical at all times provided they coincide at some reference time t_0 (here $t_0 = 0$). This is precisely what we require in quantum theory, where dynamical problem typically consist in finding, at a later time t, matrix elements of operators, which represent physical observables, provided we know the matrix elements at some initial time. In Schrödinger picture this is done by solving Schrödinger equation.

In Heisenberg picture, one solves the equation of motion for the Heisenberg operator $O_H(t)$

When O_H has also explicit time dependence then this generalizes to

$$\frac{\mathrm{d}O_H(t)}{\mathrm{d}t} \;=\; i\,[H, O_H(t)] \;+\; \frac{\partial O_H(t)}{\partial t}.$$

$$\frac{\mathrm{d}O_H(t)}{\mathrm{d}t} \;=\; i\,[H_S, O_H(t)]\;. \tag{8.11}$$

which directly follows from (8.10). This can also be alternatively viewed as a consequence of Dirac's quantization condition.

Note

- As long as we deal with *energy eigenfunctions* and *eigenvalues* in *non-relativistic* theory, there is a little practical difference between Schrödineger and Heisenberg picture, as in the absence of external time-varying forces we have

$$H_H(t) \;=\; H_S \;\equiv\; H\,,$$

  and hence $\frac{\mathrm{d}H}{\mathrm{d}t} = 0$. As for energy eigenfunctions, the Schrödinger wave function is $\psi_n(q, t) = e^{-i\omega_n t}u_n(q)$ while the corresponding Heisenberg wave function is simply $u_n(q)$. Spectrum is (due to a unitary similarity between both pictures) identical.

- In *relativistic field theory*, the Heisenberg picture is more convenient, since the explicit representation of the state vector ψ is considerably more complicated than in the non-relativistic case (such a ψ is a solution of the so-called functional Schrödinger equation), and the dynamics of operators is easier to describe and solve (even if only perturbatively) than the dynamics of ψ.

- *Lorentz invariance* can be more readily implemented in the Heisenberg picture, which puts time together with space coordinates in the field operators. So, one can, for instance, formulate Lorentz covariant field equations. Note also, that time and space are both treated on equal footing, in particular both are c-numbers.

- In Quantum Mechanics are both pictures unitarily equivalent. This unitary equivalence is guaranteed by the so-called *Stone-von Neumann uniqueness theorem*, which states that all irreducible representations of the canonical commutation relation are for a finite number of degrees of freedom unitarily equivalent to that of Schrödinger.

  In QFT is the unitary equivalence violated. This violation can be, in turn related to the concept of *renormalization*.

In the Heisenberg picture it follows that the CCR retain the form

$$[\hat{q}(t), \hat{p}(t)] \ = \ i \, , \tag{8.12}$$

Sub-index $\boldsymbol{H}$ and hat over operators will be mostly suppressed further on.

which is again dictated by Dirac's quantization condition. For an arbitrary t the operators can be represented as

$$\hat{p}(t) \ = \ -i\frac{\partial}{\partial q(t)}, \quad \hat{q}(t) \ = \ q(t) \, , \tag{8.13}$$

in q-representation, and

$$\hat{p}(t) \ = \ p(t), \quad \hat{q}(t) \ = \ i\frac{\partial}{\partial p(t)} \, , \tag{8.14}$$

in p-representation.

From (8.11) we can directly write the equations of motion for the canonical variables (the Heisenberg dynamical equations) in the form

$$\begin{aligned} \frac{\mathrm{d}p(t)}{\mathrm{d}t} \ &= \ i\,[H, p(t)] \, , \\ \frac{\mathrm{d}q(t)}{\mathrm{d}t} \ &= \ i\,[H, q(t)] \, . \end{aligned} \tag{8.15}$$

To completely determine the dynamical problem in Quantum Mechanics, we must still specify the matrix elements of p and q at the initial time.

Let us illustrate the aforemntioned quantization methodology on a simple problem. To this end we consider Lagrangian of the form

$$L = \frac{1}{2}m\dot{q}^2 - \frac{\omega^2}{2}mq^2 . \tag{8.16}$$

The corresponding action functional is given by

$$S = \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{1}{2}m\dot{q}^2 - \frac{\omega^2}{2}mq^2\right] = \frac{m}{2}\int_{t_1}^{t_2} \mathrm{d}t \left[\dot{q}^2 - \omega^2 q^2\right] . \tag{8.17}$$

Requirement that $\delta S = 0$ yields the equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 , \tag{8.18}$$

which reduces to the

$$\ddot{q} + \omega^2 q = 0 . \tag{8.19}$$

This is the equation of motion in a *configuration space*.

Corresponding equations in the *phase space* are obtained by defining the conjugate momentum p (consider further $m = 1$)

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} = \dot{q}. \tag{8.20}$$

Ensuing Hamiltonian is defined via Legendre transformation

$$\begin{aligned} H(p,q) &= p\dot{q} - L(q,\dot{q}) \\ &= p^2 - \frac{1}{2}p^2 + \frac{\omega^2}{2}q^2 = \frac{1}{2}\left(p^2 + \omega^2 q^2\right) . \end{aligned} \tag{8.21}$$

This is nothing but the Hamiltonian of linear harmonic oscillator. Hamilton equations of motion read

$$\begin{aligned} \frac{\mathrm{d}q}{\mathrm{d}t} &= \{q, H\}_{PB} = \frac{\partial H}{\partial p} = p , \\ \frac{\mathrm{d}p}{\mathrm{d}t} &= \{p, H\}_{PB} = -\frac{\partial H}{\partial q} = -\omega^2 q , \end{aligned} \tag{8.22}$$

which are clearly equivalent to the Euler-Lagrange equation (8.19).

We can now pass to Quantum Mechanics by making change

$$\{\cdot,\cdot\}_{PB} \rightarrow -\frac{i}{\hbar}[\cdot,\cdot] , \tag{8.23}$$

or, in more explicit terms

This can also be written as: $\{f,g\}_{PB} = d \rightarrow [O_f, O_g] = i\hbar O_d$.

$$O_{\{f,g\}_{PB}} = -\frac{i}{\hbar}\left[O_f, O_g\right] . \tag{8.24}$$

Here, Q_f denotes an operator representation of a classical dynamical function f. This allows to write

$$\begin{aligned} \dot{q} &= \{q, H\}_{PB} = p \rightarrow \left[\hat{q}, \hat{H}\right] = i\hbar\hat{p} , \\ \dot{p} &= \{p, H\}_{PB} = -\omega^2 q \rightarrow \left[\hat{p}, \hat{H}\right] = -i\hbar\omega^2\hat{q} , \end{aligned} \tag{8.25}$$

which is equivalent to the operator (configuration-space) equation (again without using hats)

$$\ddot{q} = -\omega^2 q\,. \tag{8.26}$$

To diagonalize the system of equations (8.25) one conventionally defines ladder operators (setting $\hbar = 1$)

$$\begin{aligned} a(t) &= \frac{\omega q(t) + ip(t)}{\sqrt{2\omega}}\,, \\ a^\dagger(t) &= \frac{\omega q(t) - ip(t)}{\sqrt{2\omega}}\,. \end{aligned} \tag{8.27}$$

These can also be equivalently introduced as

$$\begin{aligned} q(t) &= \frac{1}{\sqrt{2\omega}}\left(a^\dagger(t) + a(t)\right)\,, \\ p(t) &= i\sqrt{\frac{\omega}{2}}\left(a^\dagger(t) - a(t)\right)\,. \end{aligned} \tag{8.28}$$

It is not difficult to see that the the above ladder operators indeed diagonalize the system of equations (8.25), in fact

$$\begin{aligned} \dot{a}^\dagger(t) &= -i\left[a^\dagger(t), H\right] = -i\left[\frac{\omega q(t) - ip(t)}{\sqrt{2\omega}}, H\right] \\ &= \frac{\omega p(t) + i\omega^2 q(t)}{\sqrt{2\omega}} = i\omega a^\dagger(t)\,. \end{aligned} \tag{8.29}$$

By Hermitian conjugation we obtain analogous equation for $a(t)$. Corresponding solutions can be written in the form

$$\begin{aligned} a^\dagger(t) &= a^\dagger(0)e^{i\omega t}\,, \\ a(t) &= a(0)e^{-i\omega t}\,. \end{aligned} \tag{8.30}$$

With these we can rewrite $q(t)$ as

$$q(t) = \frac{1}{\sqrt{2\omega}}\left(a^\dagger(0)e^{i\omega t} + a(0)e^{-i\omega t}\right)\,, \tag{8.31}$$

and similar formula could be written also for $p(t)$. In terms of $a(t)$ and $a^\dagger(t)$ the Hamiltonian reads

$$\begin{aligned} H &= \frac{1}{2}\omega\left(a^\dagger(t)a(t) + a(t)a^\dagger(t)\right) \\ &= \frac{1}{2}\omega\left(a^\dagger(0)a(0) + a(0)a^\dagger(0)\right) \\ &= \omega\left(a^\dagger(0)a(0) + \frac{1}{2}\right)\,. \end{aligned} \tag{8.32}$$

There exists yet another advantage by introducing the ladder operators. In particular, from functional analysis is known that if there exist two operators A and $A^\dagger$ such that $\left[A, A^\dagger\right] = \lambda \in \mathbb{R}^+$ is satisfied, then the eigenvalues of $A^\dagger A$ operator are $0, \lambda, 2\lambda, 3\lambda, \ldots$.

In finite dimensional Hilbert spaces there is no couple of operators A and B such that $[A, B] = 1$.

$a^\dagger(t)$ the Hamiltonian reads

$$\begin{aligned} H &= \frac{1}{2}\omega\left(a^\dagger(t)a(t) + a(t)a^\dagger(t)\right) \\ &= \frac{1}{2}\omega\left(a^\dagger(0)a(0) + a(0)a^\dagger(0)\right) \\ &= \omega\left(a^\dagger(0)a(0) + \frac{1}{2}\right). \end{aligned} \tag{9.32}$$

There exists yet another advantage by introducing the ladder operators. In particular, from functional analysis is known that if there exist two operators A and $A^\dagger$ such that $\left[A, A^\dagger\right] = \lambda \in \mathbb{R}^+$ is satisfied, then the eigenvalues of $A^\dagger A$ operator are $0, \lambda, 2\lambda, 3\lambda, \ldots$.

In finite dimensional Hilbert spaces there is no couple of operators A and B such that $[A, B] = 1$.

Since $[a(0), a^\dagger(0)] = 1$, we see that the spectrum of H is $\omega_n = \omega(n + 1/2)$ with $\omega = 0, 1, 2, \ldots$. By denoting the eigenstates related to ω_n as $|n\rangle$ we can write

$$H|n\rangle = \omega_n |n\rangle = \omega\left(n + \frac{1}{2}\right)|n\rangle, \quad n = 0, 1, \ldots. \tag{9.33}$$

Important property of the spectrum is that it is equidistant, namely $\omega_n - \omega_{n-1} = \omega$.

The *lowest* energy state $|0\rangle$ (also known as *vacuum state* or *ground state*) satisfies

$$H|0\rangle = \omega\left(a^\dagger(0)a(0) + \frac{1}{2}\right)|0\rangle = \frac{1}{2}\omega|0\rangle. \tag{9.34}$$

Because $a^\dagger(0)a(0)|0\rangle = 0$ we have that

$$\langle 0|a^\dagger(0)a(0)|0\rangle = \|a(0)|0\rangle\|^2 = 0, \tag{9.35}$$

which implies that the vacuum state is annihilated by the $a(0)$ operator.

Suppose now that

$$H|n\rangle = \omega_n |n\rangle, \tag{9.36}$$

then by applying $a^\dagger(0)$ to $|n\rangle$ we get

$$\begin{aligned} H[a^\dagger(0)|n\rangle] &= \left(\left[H, a^\dagger(0)\right] + a^\dagger(0)H\right)|n\rangle \\ &= \left(\omega a^\dagger(0) + a^\dagger(0)\omega_n\right)|n\rangle \\ &= (\omega + \omega_n)[a^\dagger(0)|n\rangle]. \end{aligned} \tag{9.37}$$

This implies that $a^\dagger(0)|n\rangle \propto |n+1\rangle$. Similarly, by applying $a(0)$ to $|n\rangle$ we obtain

$$H[a(0)|n\rangle] = (\omega_n - \omega)[a(0)|n\rangle]. \tag{9.38}$$

This implies that $a(0)|n\rangle \propto |n-1\rangle$, which terminates at $|0\rangle$ state.

In particular, by applying creation operator on vacuum state we get

$$a^\dagger(0)|0\rangle \propto |1\rangle, \tag{9.39}$$

and generally

$$(a^{\dagger}(0))^n |0\rangle \propto |n\rangle \,. \tag{9.40}$$

After normalization of states to the $\langle n|m\rangle = \delta_{nm}$, we get

$$|n\rangle = \frac{(a^{\dagger})^n}{\sqrt{n!}} |0\rangle \,. \tag{9.41}$$

which can be easily proved by induction.

Cute note

For energy eigenstates $|n\rangle$ we can write

$$\begin{aligned} 2\langle n|H|n\rangle &= 2\omega_n = \langle n|p^2|n\rangle + \omega^2 \langle n|q^2|n\rangle \\ &= ||p\,|n\rangle\,||^2 + ||\omega q\,|n\rangle\,||^2 \geq 2||p\,|n\rangle\,||\,||\omega q\,|n\rangle\,||\,. \end{aligned}$$

Where last inequality follows from *triangle inequality*.

We now use *virial theorem* which implies that $\langle n|p|n\rangle = \langle n|q|n\rangle = 0$. The latter is a direct consequence of the fact that for any operator A and any energy eigenstate $|n\rangle$ we have

$$\begin{aligned} \langle n|\,[A, H]\,|n\rangle &= \langle n|AH|n\rangle - \langle n|HA|n\rangle \\ &= \omega_n \langle n|A|n\rangle - \omega_n \langle n|A|n\rangle = 0\,. \end{aligned}$$

Thus if $A = q$

$$\langle n|\,[q, H]\,|n\rangle = \langle n|ip|n\rangle = 0\,,$$

or when $A = p$

$$\langle n|\,[p, H]\,|n\rangle = -i\omega \langle n|q|n\rangle = 0\,.$$

Consequently

$$2\langle n|H|n\rangle \geq 2||p\,|n\rangle\,||\,||\omega q\,|n\rangle\,|| \geq 2\omega\Delta p\Delta q \geq \hbar\omega\,.$$

So, uncertainty relation prohibits $\langle n|H|n\rangle = \omega_n$ to be smaller that $\hbar\omega/2$. Eigenvalue $\omega_n = \hbar\omega/2$ represents the so-called zero mode of H (in this specific case we call it also *zero-mode fluctuation* or *ground state fluctuation*).

It is now straightforward to generalize this procedure to N degrees of freedom. We introduce N Hermitian operators $q_i(t)$, $i = 1, \ldots, N$ in the Heisenberg picture and N conjugate momenta $p_i(t)$. The dynamics is again given by the $2N$ classical equations of motion

$$\begin{aligned} \{p_i, H\}_{PB} &= -\frac{\partial H}{\partial q_i} = \dot{p}_i\,, \\ \{q_i, H\}_{PB} &= \frac{\partial H}{\partial p_i} = \dot{q}_i, \quad i = 1, \ldots, N\,. \end{aligned} \tag{9.42}$$

Again, our goal is to find the matrix elements of p_i and q_i (or their functions), at arbitrary times provided that at initial time, say $t = 0$

they satisfy the restrictions

$$\begin{aligned} \left[p_i(0), q_j(0)\right] &= -i\delta_{ij}\,, \\ \left[p_i(0), p_j(0)\right] &= 0\,, \\ \left[q_i(0), q_j(0)\right] &= 0\,. \end{aligned} \tag{9.43}$$

By employing Dirac's quantization condition we can from (9.42) write Heisenberg equations of motion in the form

$$\begin{aligned} \dot{p}_i(t) &= i\,[H, p_i(t)]\,, \\ \dot{q}_i(t) &= i\,[H, q_i(t)]\,. \end{aligned} \tag{9.44}$$

9.2 Classical fields

A *field* is a quantity defined at every point of space $\boldsymbol{x}$ and time t. While classical particle mechanics deals with a finite number of generalized coordinates $q_i(t)$, indexed by a discrete label "i", in field theory we are interested in the dynamics of fields $\phi_a(\boldsymbol{x}, t)$, where both "$a$" and "$\boldsymbol{x}$" are considered as labels. We are thus dealing with a system with an infinite (uncountably infinite) number of degrees of freedom — at least one for each point $\boldsymbol{x}$ in space.

Notice, that the concept of position has been relegated from a dynamical variable in particle mechanics (so called wave mechanics or first quantization) to a mere label in field theory.

► **Example — Electromagnetic field**
$\boldsymbol{E}(\boldsymbol{x}, t)$ and $\boldsymbol{B}(\boldsymbol{x}, t)$, both of these fields are 3 spatial (or euclidean) vectors — *3 dimensional vector fields*. In covariant treatment of electromagnetism one introduces instead of $\boldsymbol{E}$ and $\boldsymbol{B}$ electromagnetic potential A^μ $\boldsymbol{E}, \boldsymbol{B} \to A^\mu(\boldsymbol{x}, t) = (\phi, \boldsymbol{A})$ $(\mu = 0, \ldots, 3)$, where A^μ is a vector in spacetime — *4 dimensional vector field*. Connection between $\boldsymbol{E}, \boldsymbol{B}$ and A^μ is done by the Maxwell relation

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \quad \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}\,. \tag{9.45}$$

These defining relations directly imply sourceless Maxwell equations (so called Bianchi identities)

It should be stressed that (9.46) are not field equations, since there are no sources; rather, they impose constraints on the electric and magnetic fields.

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0\,, \quad \boldsymbol{\nabla} \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}\,. \tag{9.46}$$

To proceed we start again with *Lagrangian*. This is not only convenient starting point for a covariant treatment but it also allows to formulate constructively a field-theoretical systems solely on the basis of required/expected symmetries. Besides, for simple systems the Lagrangian formalism also provides a straightforward passage to Hamiltonian (i.e., canonical) formalism via Legendre transformation.

In the Lagrangian formalism, the dynamics is governed by a Lagrangian, which is a function of $\phi_i(\boldsymbol{x}, t)$, $\dot{\phi}(\boldsymbol{x}, t)$ and $\boldsymbol{\nabla}\phi_i(\boldsymbol{x}, t)$. We change our Lagrangian (L) to Lagrangian density ($\mathcal{L}$), i.e.

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \int \mathrm{d}^3\boldsymbol{x}\, \mathcal{L}(\phi_i(\boldsymbol{x}, t), \partial_\mu \phi_i(\boldsymbol{x}, t)). \tag{9.47}$$

In principle, we could consider also higher derivative terms (or even non-local interactions), but in all systems, studied in this course, the Lagrangian will be of the form given above. The action is

$$\begin{aligned} S &= \int_{t_1}^{t_2} dt\ !L = \int_{t_1}^{t_2} dt \int d^3\boldsymbol{x}\, \mathcal{L}(\phi_i, \partial_\mu \phi_i) \\ &= \int d^4x\, \mathcal{L}(\phi_i, \partial_\mu \phi_i)\,. \end{aligned} \tag{9.48}$$

More on Lagrangian in field theory

In particle mechanics L depends on q_i and $\dot{q}_i$, but not $\ddot{q}_i$. In field theory we similarly restrict the Lagrangian $\mathcal{L}$ to ϕ_i and $\dot{\phi}_i$. In principle, there is nothing to stop $\mathcal{L}$ from depending on $\boldsymbol{\nabla}\phi$, $\boldsymbol{\nabla}^2\phi$, $\boldsymbol{\nabla}^3\phi, \ldots$ In cases when we require Lorentz invariance, we will consider only dependence of $\mathcal{L}$ on $\boldsymbol{\nabla}\phi$ (this is not needed in non-relativistic context). We will also not consider $\mathcal{L}$ explicitly dependent on x^μ.

Since we strive for a self-contained theory we do not consider Lagrange densities with external fields (i.e. fields that do not have their dynamics described within the Lagrangian itself).

Corresponding equations of motion are obtained via the principle of stationary action (Hamilton's principle)

$$\begin{aligned} \delta S &= \int \mathrm{d}^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_i} \delta\phi_i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \delta(\partial_\mu \phi_i) \right] \\ &= \int \mathrm{d}^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_i} \delta\phi_i - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) \delta\phi_i + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \delta\phi_i \right) \right]. \end{aligned} \tag{9.49}$$

We can now neglect the surface term due to condition that $\delta\phi_i = 0$ on the surface. By the *fundamental lemma of calculus of variations* this leads to the Euler-Lagrange equations of motion for fields ϕ_i

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_i} = 0\,. \tag{9.50}$$

Let us now list a couple of illustrative field systems.

- ▶ **Example I — Klein-Gordon field**
  Consider Lagrangian for a real scalar field $\phi(t, \boldsymbol{x})$ in the form

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2 \\ &= \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 - \frac{1}{2}m^2\phi^2\,. \end{aligned} \tag{9.51}$$

For equation of motion we need

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2 \phi\,, \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi = (\dot{\phi}, -\nabla \phi)\,. \tag{9.52}$$

By using the Euler-Lagrange equation we arrive at the equation of the motion

$$\ddot{\phi} - \nabla^2 \phi + m^2 \phi = 0 \quad \Leftrightarrow \quad \Box \phi + m^2 \phi = 0\,. \tag{9.53}$$

This is nothing but the Klein-Gordon equation. In this time, however for a classical field and not for the wave function.

► **Example II — First order field Lagrangian**
As a second example we consider Lagrangian with a complex field that is linear in time derivative (rather than quadratic), namely

$$\mathcal{L} = \frac{i}{2}\left(\psi^* \dot{\psi} - \dot{\psi}^* \psi\right) - \nabla \psi^* \nabla \psi - m \psi^* \psi\,. \tag{9.54}$$

For equation of motion we need

$$\frac{\partial \mathcal{L}}{\partial \psi^*} = \frac{i}{2}\dot{\psi} - m\psi\,, \quad \frac{\partial \mathcal{L}}{\partial \dot{\psi}^*} = -\frac{i}{2}\psi\,, \quad \frac{\partial \mathcal{L}}{\partial \nabla \psi^*} = -\nabla \psi\,. \tag{9.55}$$

The corresponding Euler-Lagrange equation thus implies equation of motion

$$i\frac{\partial \psi}{\partial t} = -\nabla^2 \psi + m\psi\,. \tag{9.56}$$

Note that this equation looks like Schrödinger equation, but it is not. Its interpretation is very different and besides the field ψ is a classical field with no probabilistic interpretation à la wave function.

► **Example III —Maxwell equations**
Consider a Lagrangian of the form

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A^\mu)^2\,. \tag{9.57}$$

The minus sign in front of the gradient term ensures that the kinetic term for A_i is positive. Note, also that $\mathcal{L}$ has no kinetic term $(\dot{A}_0)^2$ for A_0, so that A_0 can not correspond to a physical degree of freedom. For equation of motion we need to compute

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\partial^\mu A^\nu + \partial^\nu(\partial_\rho A^\rho)\,, \tag{9.58}$$

which yields the dynamical equation in the form

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)}\right) = -\partial_\mu\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) \equiv \partial_\mu F^{\mu\nu} = 0\,. \tag{9.59}$$

Here we have defined the field strength tensor $F^{\mu\nu} = (\partial^\mu A^\nu - \partial^\nu A^\mu$.

By identifying now A^μ with electromagnetic potential and using the relations for electric and magnetic fields

$$\boldsymbol{E} = -\boldsymbol{\nabla} A^0 - \frac{\partial \boldsymbol{A}}{\partial t}, \quad \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}, \tag{9.60}$$

we can check that (9.59) is equivalent to the first set of Maxwell equations, i.e.

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = 0, \quad \frac{\partial \boldsymbol{E}}{\partial t} = \boldsymbol{\nabla} \times \boldsymbol{B}. \tag{9.61}$$

Indeed, for instance

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{E} &= \partial_k E^k = \left\{E^k = -\partial_k A^0 - \partial_0 A^k = \partial^k A^0 - \partial^0 A^k\right\} \\ &= \partial_k \partial^k A^0 - \partial_k \partial^0 A^k = \partial_k \left(\partial^k A^0 - \partial^0 A^i\right) = 0. \end{aligned} \tag{9.62}$$

The second series of Maxwell equations, i.e.,

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0, \quad \frac{\partial \boldsymbol{B}}{\partial t} = -\boldsymbol{\nabla} \times \boldsymbol{E}, \tag{9.63}$$

do not correspond to field equations. These merely impose constraints on the electric and magnetic fields. In fact they are equivalent to *Bianchi identity*

$$\partial_\gamma F_{\alpha\beta} + \partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} = 0, \tag{9.64}$$

which is a simple identity implied by the structure of $F_{\mu\nu}$.

It can also be checked that the Lagrangian (9.57) is (upto 4-divergence) identical to the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \tag{9.65}$$

Exercises: Some useful background from quantum mechanics and classical field theory

Variational calculus

Recall that for a function of n variables

$$f : \mathbb{R}^n \to \mathbb{R} \quad , \quad (u_1, \dots, u_n) \mapsto f(u_1, \dots, u_n)\,,$$

the partial derivative, with respect to the k-th variable, is defined

$$\begin{aligned} \frac{\partial f}{\partial u_k} &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left(f(u_1, \dots, u_{k-1}, u_k + \varepsilon, u_{k+1}, \dots, u_n) - f(u_1, \dots, u_n)\right) \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big(f(u_i + \varepsilon \delta_{ik}) - f(u_i)\Big)\,. \end{aligned}$$

A *functional*

$$F : \mathcal{M} \to \mathbb{R} \quad , \quad \phi \mapsto F[\phi] \quad , \quad \phi : x \mapsto \phi(x)\,,$$

attributes a number to each function $\phi \in \mathcal{M}$ (for us, typically, $\phi(x)$ will be fields defined on the Minkowski spacetime). In analogy with the partial derivative, the *functional* (or *variational*) derivative of a functional F, with respect to variations at a spacetime point x_0, is formally defined as

$$\frac{\delta F}{\delta \phi(x_0)} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big(F[\phi(\cdot) + \varepsilon \delta(\cdot - x_0)] - F[\phi(\cdot)]\Big)\,.$$

Then, for a *variation* of the functional F we may write

$$\delta F \equiv F[\phi + \delta\phi] - F[\phi] = \int \frac{\delta F}{\delta \phi(x)} \delta\phi(x)\, dx\,.$$

This should be compared with the differential of a function f, i.e.

$$df = \sum_{k=1} \frac{\partial f}{\partial u_k} du_k\,.$$

While, df is a linear function in coordinate increments, δF is a linear functional in field (or function) increments.

Exercise 9.1 Using the definition (9.2), calculate the functional derivative $\frac{\delta F}{\delta \phi(x_0)}$ for the functional:

$$1) \quad F[\phi] = \int f(x)\phi(x)\, d^4x\,.$$

$$2) \quad F[\phi] = \phi(x_1)\,.$$

$$3) \quad F[\phi] = \exp\left(\int f(x)\phi(x)\, d^4x\right)\,.$$

In point 3), calculate the functional derivative also using a generalization of the composite function differentiation theorem, and the result of 1).

We note the following properties of functional derivatives:

$$1) \quad \frac{\delta}{\delta \phi(x_0)}\Big(\alpha(x) F[\phi] + G[\phi]\Big) = \alpha(x) \frac{\delta F}{\delta \phi(x_0)} + \frac{\delta G}{\delta \phi(x_0)}\,.$$

$$2) \quad \frac{\delta}{\delta \phi(x_0)}\Big(F[\phi] G[\phi]\Big) = \frac{\delta F}{\delta \phi(x_0)} G[\phi] + F[\phi] \frac{\delta G}{\delta \phi(x_0)}\,.$$

$$3) \quad \frac{\delta}{\delta \phi(x_0)} f(G[\phi]) = f'(G[\phi]) \frac{\delta G}{\delta \phi(x_0)} \quad , \quad f : \mathbb{R} \to \mathbb{R}\,.$$

Exercise 9.2 Derive the field-theoretic Euler–Lagrange equations for the action

$$S[\phi] = \int \mathcal{L}(\phi(x), \partial_\mu \phi(x), x)\, d^4x \,.$$

[**Hint:** Calculate $\frac{\delta S}{\delta \phi(x_0)}$, and set equal to zero.]

Exercise 9.3 Calculate the functional derivative $\frac{\delta S}{\delta w(x)}$ for the entropy functional

$$S[w] \;=\; -k \int w(x) \ln w(x)\, dx \,.$$

Exercise 9.4 Derive (one half of) the vacuum Maxwell equations from the action

$$S[A_\mu] \;=\; -\frac{1}{4} \int F_{\mu\nu} F^{\mu\nu}\, d^4x \,, \qquad F_{\mu\nu} \;=\; \partial_\mu A_\nu - \partial_\nu A_\mu \,,$$

where $F_{\mu\nu}$ is the *Faraday tensor* (i.e. the electromagnetic tensor), and A_μ is the four-potential of the electromagnetic field. Show that the second half of the Maxwell equations is a (trivial) consequence of the definition of $F_{\mu\nu}$.

[**Hint:** Use the Euler–Lagrange equations $\partial_\mu \big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)}\big) - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0$.]

Lagrangian and Hamiltonian formalism of classical field theory

Exercise 9.5 Consider a one-dimensional infinite chain of masses m, connected via springs with spring constant κ. The Lagrangian of this system is

$$L(q_n, \dot{q}_n) \;=\; \sum_n \frac{m}{2} \dot{q}_n^2 \;-\; \sum_n \frac{\kappa}{2} (q_{n+1} - q_n)^2 \,, \qquad n \in \mathbb{Z} \,,$$

where q_n denotes the displacement of the n-th mass point from its equilibrium position.

Write for this system the Euler–Lagrange equations of motion. Moreover, derive the Hamiltonian, and write the Hamilton's canonical equations. (Make use of the Poisson brackets.)

Exercise 9.6 We introduce the displacement field $\phi(na, t) = q_n(t)$, where a is the distance between two neighbouring equilibrium positions. In the (continuum) limit $a \to 0$, with the density $\rho = m/a$ and the tension $T = \kappa a$ kept fixed, the Lagrangian in the previous exercise takes the form

$$L[\phi(x), \dot{\phi}(x)] \;=\; \int \mathcal{L}(\phi(x), \dot{\phi}(x), \partial_x \phi(x))\, dx \,,$$

$$\mathcal{L}(\phi, \partial_t \phi, \partial_x \phi) \;=\; \frac{\rho}{2} (\partial_t \phi)^2 \;-\; \frac{T}{2} (\partial_x \phi)^2 \,,$$

where $\mathcal{L}$ is the Lagrangian density corresponding to the "continuum" Lagrangian L.

Show that the (variational) Euler–Lagrange equations for a continuum Lagrangian $L[\phi, \dot{\phi}]$ reduce to field-theoretic Euler–Lagrange equations for the corresponding Lagrangian density $\mathcal{L}(\phi, \partial_t \phi, \partial_x \phi)$. Find these equations for the system (a string) described above.

From now on, let us consider (up to) three spatial dimensions. In the continuum limit, the canonical momentum field is defined

$$\pi(\boldsymbol{x}) \;=\; \frac{\delta L}{\delta \dot{\phi}(\boldsymbol{x})} \;=\; \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \,,$$

and the Hamiltonian reads

$$H[\phi(\boldsymbol{x}),\pi(\boldsymbol{x})] = \int \pi(\boldsymbol{x})\dot{\phi}(\boldsymbol{x})\,d^3x - L[\phi(\boldsymbol{x}),\dot{\phi}(\boldsymbol{x})]$$
$$= \int \mathcal{H}(\phi(\boldsymbol{x}),\pi(\boldsymbol{x}),\partial_j\phi(\boldsymbol{x}))\,d^3x\,,$$

with

$$\mathcal{H}(\phi,\pi,\partial_j\phi) = \pi\,\partial_t\phi - L(\phi,\partial_t\phi,\partial_j\phi)\,,$$

the corresponding Hamiltonian density.

Exercise 9.7 Calculate the canonical momentum field π, the Hamiltonian density $\mathcal{H}$, and the continuum Hamiltonian H for the string described in the previous exercise.

The field-theoretic Poisson bracket of two functionals $F[\phi(\boldsymbol{x}),\pi(\boldsymbol{x})]$ and $G[\phi(\boldsymbol{x}),\pi(\boldsymbol{x})]$ is defined as

$$\{F,G\} = \int d^3x\left(\frac{\delta F}{\delta\phi(\boldsymbol{x})}\frac{\delta G}{\delta\pi(\boldsymbol{x})} - \frac{\delta G}{\delta\phi(\boldsymbol{x})}\frac{\delta F}{\delta\pi(\boldsymbol{x})}\right).$$

Exercise 9.8 Calculate the Poisson brackets between functionals

$$F_{\boldsymbol{y}}[\phi,\pi] = \phi(\boldsymbol{y}) \quad \text{and} \quad G_{\boldsymbol{x}}[\phi,\pi] = \pi(\boldsymbol{x})\,.$$

Exercise 9.9 Consider the Lagrangian density of a one-component real Klein–Gordon field $\phi(x)$,

$$\mathcal{L}(\phi,\partial_\mu\phi) = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2\,, \qquad x = (x^0,\boldsymbol{x})\,,$$

and derive the corresponding Euler–Lagrange equation. Show that the same equation is obtained by passing to the Hamiltonian formalism, and combining the ensuing (field-theoretic) Hamilton's canonical equations.

[**Hint:** Use the Poisson-bracket formulation

$$\dot{\phi}(t,\boldsymbol{y}) = \{\phi(t,\boldsymbol{y}),H\}\,, \qquad \dot{\pi}(t,\boldsymbol{y}) = \{\pi(t,\boldsymbol{y}),H\}\,,$$

of the canonical equations.]

Classical energy-momentum tensor

Suppose the Lagrangian density $\mathcal{L}$ does not explicitly depend on the spacetime point x. For the field multiplet $\Phi(x) = \{\phi_r(x)\}_{r=1}^n$ we can write

$$\frac{\partial}{\partial x^\nu}\mathcal{L}(\phi_r(x),\partial_\mu\phi_r(x)) = \partial_\nu\phi_r\frac{\partial\mathcal{L}}{\partial\phi_r} + \partial_\mu(\partial_\nu\phi_r)\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\,,$$
$$\frac{\partial\mathcal{L}}{\partial(\phi_r)} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\right),$$

where the second equation is the Euler–Lagrange equations of motion. Subtracting the left-hand side from the right, we find the continuity equations

$$\partial_\mu T^\mu{}_\nu = 0\,, \qquad T^\mu{}_\nu = \partial_\nu\phi_r\frac{\partial L}{\partial(\partial_\mu\phi_r)} - \delta^\mu_\nu\,\mathcal{L}\,,$$

where $T_{\mu\nu}$ is the (canonical) energy-momentum tensor. Note that $T^0{}_0 = \mathcal{H}$ is the Hamiltonian density of the field.

Writing $\partial_\mu T^\mu{}_\nu = \partial_t T^0{}_\nu + \partial_i T^i{}_\nu$, and integrating (for a fixed time t) over the space

$\mathbb{R}^3$, we find

$$\partial_t \int d^3x\, T^0{}_\nu = -\int d^3x\, \partial_i T^i{}_\nu = -\int d^2\Sigma_i T^i{}_\nu = 0\,,$$

that is, the *total* four-momentum of the field

$$P_\nu = \int d^3x\, T^0{}_\nu\,,$$

is constant in time, $P^\nu = (H, \boldsymbol{P})$.

Exercise 9.10 Show that the energy-momentum tensor of a one-component Klein–Gordon field reads

$$T_{\mu\nu} = \partial_\mu\phi\,\partial_\nu\phi - g_{\mu\nu}\frac{1}{2}(\partial_\rho\phi\,\partial^\rho\phi - m^2\phi^2)\,. \tag{9.66}$$

In particular, find the energy density T^0_0, and the momentum density T^0_k.

Exercise 9.11 Find the canonical energy-momentum tensor $T^{(\mathrm{can})}_{\mu\nu}$ of the electromagnetic field with Lagrangian density (9.65). Note that it is not symmetric in $\mu \leftrightarrow \nu$, and show that it can be augmented by the term $(\partial_\nu A_\rho)F_\mu{}^\rho$ without affecting the continuity equations, thus arriving at

$$T^{(\mathrm{sym})}_{\mu\nu} = F_{\mu\rho}F^\rho_\nu + \frac{1}{4}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\,,$$

the symmetric energy-momentum tensor of the electromagnetic field.
[Hint: Use the electromagnetic Lagrangian of Eq. (9.65).]

Generally covariant form of an action $S[\phi_r] = \int \mathcal{L}(\phi, \partial_\mu\phi) d^4x$ is

$$S[\phi_r, g_{\mu\nu}] = \int \mathcal{L}(\phi_r, \partial_\mu\phi_r, g_{\mu\nu})\sqrt{-g}\, d^4x\,, \qquad g = \det(g_{\mu\nu})\,.$$

The *Hilbert energy-momentum tensor* is defined as

$$T^{(H)}_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g^{\mu\nu}}\,,$$

Exercise 9.12 Show that

$$T^{(H)}_{\mu\nu} = 2\frac{\partial\mathcal{L}}{\partial g^{\mu\nu}} - g_{\mu\nu}\mathcal{L}\,.$$

[Hint: Analyse partial derivatives $\frac{\partial}{\partial g^{\mu\nu}}$ of the determinant $\det(g_{\mu\nu}) = [\det(g^{\mu\nu})]^{-1}$.]

Exercise 9.13 Calculate the Hilbert energy-momentum tensor for:
a) the Klein–Gordon field
b) the electromagnetic field.

Normal modes

Exercise 9.14 Consider the Lagrangian of an infinite chain from Exercise 9.5, and solve the equations of motion by the method of modes. Show that the general solution reads

$$q_n(t) = \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} dk\, \left[a_k e^{-i(\omega_k t - kan)} + a_k^* e^{i(\omega_k t - kan)}\right]\,,$$

$$\omega_k = \sqrt{\frac{2\kappa}{m}}\sqrt{1-\cos(ka)}\,,$$

where $a_k \in \mathbb{C}$ are constant amplitudes. Perform the continuum limit $a \to 0$.

Quantum field theory preliminaries

Exercise 9.15 Show that (for $a > 0$)

$$\int_{-\infty}^{+\infty} \delta(x^2 - a^2)\phi(x)\,dx \;=\; \frac{\phi(a)}{2a} + \frac{\phi(-a)}{2a}.$$

[**Hint:** Rewrite as two integrals $\int_0^{+\infty}$, and substitute $x^2 = y$.]

Exercise 9.16 For a generic function $g(x)$, if x_i denote all the points of g at which $g(x_i) = 0$ (and provided $g'(x_i) \neq 0$), then

$$\delta(g(x)) = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|}.$$

Show that this formula correctly reproduces the results:
a) $\delta(ax) \;=\; \frac{1}{|a|}\delta(x)$
b) $\delta(x^2 - a^2) \;=\; \frac{1}{2a}\delta(x - a) \;+\; \frac{1}{2a}\delta(x + a)$.

We also note that for a multidimensional δ-function, and a constant matrix $\mathbb{A}$

$$\delta(\mathbb{A}(\boldsymbol{x} - \boldsymbol{x}_0)) \;=\; \frac{1}{|\det \mathbb{A}|}\delta(\boldsymbol{x} - \boldsymbol{x}_0)\,.$$

Exercise 9.17 Show that the expression

$$2\omega_{\boldsymbol{p}}\,\delta^{(3)}(\boldsymbol{p} - \boldsymbol{q})\,, \quad \omega_{\boldsymbol{p}} \equiv \sqrt{\boldsymbol{p}^2 + m^2}\,,$$

is Lorentz invariant.

[**Hint:** Consider boosts along the x^3-axis.]

Exercise 9.18 Show that for ladder operators $\hat{a}$ and $\hat{a}^\dagger$, $[\hat{a}, \hat{a}^\dagger] = 1$, and a 'vacuum' state $|0\rangle$, for which $\langle 0\rangle\, 0 = 1$, and $\hat{a}\,|0\rangle = 0$, the following identity holds:

$$\langle 0|\, \hat{a}^n (\hat{a}^\dagger)^m \,|0\rangle \;=\; n!\,\delta_{nm}\,, \quad (\forall n, m \in \mathbb{N}_0)\,.$$

[**Hint:** Show (and use) the identity $e^{-\alpha \hat{a}^\dagger}\,\hat{a}\,e^{\alpha \hat{a}^\dagger} = \hat{a} + \alpha$, $\forall \alpha \in \mathbb{R}$.]

Quantization of Scalar Field – I | 10

We will be particularly interested in relativistic field theories. We have seen that for relativistic (scalar) field theories, equations of motion, i.e. Euler-Lagrange equations read

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_i} = 0 . \tag{10.1}$$

Should this equation be covariant under Lorentz transformations, $\mathcal{L}$ must transform as scalar density (of weight 1), i.e.

$$\mathcal{L}(x) \equiv \mathcal{L}(\phi_i(x), \partial_\mu \phi_i(x)) \xrightarrow{L} \mathcal{L}_{\boldsymbol{L}}(x) = |\det \boldsymbol{L}| \, \mathcal{L}(\boldsymbol{L}^{-1} x) . \tag{10.2}$$

In general, a scalar density of weight w would transform as

$$\mathcal{L}(x) \xrightarrow{L} \mathcal{L}_{\boldsymbol{L}}(x) = |\det \boldsymbol{L}|^w \, \mathcal{L}(\boldsymbol{L}^{-1} x).$$

Let us construct the simplest free (real-field) scalar theory with maximally second time derivative in equation of motion.

The simplest $\mathcal{L}$ that is a scalar density and has bounded from below potential energy is

$$\mathcal{L} = \frac{1}{2} \eta^{\mu\nu} \partial_\mu \phi(x) \partial_\nu \phi(x) - \frac{1}{2} m^2 \phi^2(x) . \tag{10.3}$$

Proof of the fact that $\mathcal{L}$ is a scalar density is quite simple. First we realize that

$$\begin{aligned} \phi(x) &\xrightarrow{\boldsymbol{L}} \phi_{\boldsymbol{L}} = \phi(\boldsymbol{L}^{-1} x) \\ \partial_\mu \phi(x) &\xrightarrow{\boldsymbol{L}} (\boldsymbol{L}^{-1})^{\nu}_{\cdot\,\mu} \partial_\nu \phi(\boldsymbol{L}^{-1} x) . \end{aligned} \tag{10.4}$$

Thus

$$\begin{aligned} &\partial_\mu \phi(x) \partial_\nu \phi(x) \eta^{\mu\nu} \\ &\quad \xrightarrow{\boldsymbol{L}} (\boldsymbol{L}^{-1})^{\alpha}_{\cdot\,\mu} \partial_\alpha \phi(\boldsymbol{L}^{-1} x) (\boldsymbol{L}^{-1})^{\beta}_{\cdot\,\nu} \partial_\beta \phi(\boldsymbol{L}^{-1} x) \boldsymbol{L}^{\mu}{}_{\gamma} \eta^{\gamma\delta} \boldsymbol{L}^{\nu}{}_{\delta} \\ &\quad = \partial_\gamma \phi(\boldsymbol{L}^{-1} x) \partial_\delta \phi(\boldsymbol{L}^{-1} x) \eta^{\gamma\delta} . \end{aligned} \tag{10.5}$$

In addition, the term $\phi^2(x)$ transforms simply as $\phi^2(x) \xrightarrow{\boldsymbol{L}} \phi^2(\boldsymbol{L}^{-1} x)$. Since $|\det \boldsymbol{L}| = 1$, the Lagrangian (10.2) is a scalar density. Consequently, the action functional is a Lorentz scalar.

As before, the link between the Lagrangian formalism and canonical quantum theory is established via Hamiltonian formalism. To this end, we start by defining *the conjugate momentum* $\pi(x)$ to field $\phi(x)$ as

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} = \frac{\delta S[\phi]}{\delta \dot{\phi}(x)} . \tag{10.6}$$

This should not be confused with total (conserved) momentum, which will be defined shortly.

The *Hamiltonian density* is then given by

$$\mathcal{H}(\pi(x), \phi(x)) \equiv \mathcal{H}(x) = \pi(x)\dot{\phi}(x) - \mathcal{L}(x), \tag{10.7}$$

where, as in classical mechanics, we eliminate $\dot{\phi}(x)$ in favour of $\pi(x)$ everywhere in $\mathcal{H}(x)$. The *Hamiltonian* is then the simply

$$H = \int d^3\boldsymbol{x}\,\mathcal{H}(x). \tag{10.8}$$

Quantization starts by identifying commutators via Poisson brackets. In particular, we know that

$$-\frac{i}{\hbar}\left[O_f, O_g\right] = O_{\{f,g\}_{PB}}. \tag{10.9}$$

Poisson brackets in field theory

For two functions f and g that depend on phase-space and time variables their Poisson bracket $\{f, g\}_{PB}$ is another function that depends on phase space and time.

Given two functions $f(p_i, q_i, t)$ and $g(p_i, q_i, t)$ with generalized momenta $p_1, \ldots, p_N$, generalized positions $q_2, \ldots, q_N$ and time t, the Poisson bracket takes the form

$$\{f, g\}_{PB} = \sum_{i=1}^{N}\left[\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right].$$

Here it is implicit that time arguments are both in functions f, g and phase-space variables identical. As a particular case one has

$$\{q_k, p_l\}_{PB} = \delta_{kl}.$$

Passage to continuous degrees of freedom — i.e. *fields*, is obtained by formally replacing

$$q_i(t) \to \phi_a(\boldsymbol{x}, t), \quad p_i(t) \to \pi_a(\boldsymbol{x}, t),$$
$$\frac{\partial}{\partial q_i(t)} \to \frac{\delta}{\delta\phi_a(\boldsymbol{x}, t)}, \quad \frac{\partial}{\partial p_i(t)} \to \frac{\delta}{\delta\pi_a(\boldsymbol{x}, t)},$$
$$\sum_{i=1}^{N} \to \sum_a \int d^3\boldsymbol{x}.$$

With this the Poisson brackets between two functionals A and B defined on the field phase space read

$$\{A(\pi, \phi), B(\pi, \phi)\}_{PB} = \sum_a \int d^3\boldsymbol{x}\left[\frac{\delta A}{\delta\phi_a(\boldsymbol{x})}\frac{\delta B}{\delta\pi_a(\boldsymbol{x})} - \frac{\delta A}{\delta\pi_a(\boldsymbol{x})}\frac{\delta B}{\delta\phi_a(\boldsymbol{x})}\right],$$

(time variable is the same on both sides and is customarily omitted). In particular, for canonical field variables we obtain

$$\{\phi_a(\boldsymbol{x}), \pi_b(\boldsymbol{y})\}_{PB} = \delta_{ab}\delta(\boldsymbol{y} - \boldsymbol{x}).$$

So, that in the Shrödinger picture the canonical commutation relation becomes $[\phi(\boldsymbol{x}), \pi(\boldsymbol{x}')] = i\delta(\boldsymbol{x} - \boldsymbol{x}')$.

Starting with Schrödinger picture is merely historical coincidence. One could use Dirac's quantization prescription to set up Heisenberg picture directly without going via Schrödinger picture.

In the second step, we construct Hamiltonian

$$\begin{aligned} H &= \int d^3\boldsymbol{x} \left[\pi(x)\dot{\phi}(x) - \mathcal{L}(x)\right] \\ &= \int d^3\boldsymbol{x}\, \pi(x)\dot{\phi}(x) - L\,. \end{aligned} \tag{10.10}$$

Since in our case $\pi(\boldsymbol{x}) = \dot{\phi}(\boldsymbol{x})$, we get

$$H = \frac{1}{2}\int d^3\boldsymbol{x} \left[\pi(\boldsymbol{x})^2 + (\boldsymbol{\nabla}\phi(\boldsymbol{x}))^2 + m^2\phi(\boldsymbol{x})^2\right]. \tag{10.11}$$

At this stage we can pass to the Heisenberg picture

$$\phi(\boldsymbol{x}) \rightarrow \phi(\boldsymbol{x}, t) \equiv \phi(x) = e^{iHt}\phi(\boldsymbol{x})e^{-iHt}. \tag{10.12}$$

(we assume that Heisenberg and Schrödinger picture coincide at the reference time $t_0 = 0$). Similarly for $\pi(\boldsymbol{x})$

$$\pi(\boldsymbol{x}) \rightarrow \pi(x) = e^{iHt}\pi(\boldsymbol{x})e^{-iHt}. \tag{10.13}$$

In Heisenberg picture this provides, among others, the equal-time commutation relations

$$\begin{aligned} [\phi(\boldsymbol{x}, t), \pi(\boldsymbol{x}', t)] &= i\delta(\boldsymbol{x} - \boldsymbol{x}')\,, \\ [\phi(\boldsymbol{x}, t), \phi(\boldsymbol{x}', t)] &= [\pi(\boldsymbol{x}, t), \pi(\boldsymbol{x}', t)] = 0\,. \end{aligned} \tag{10.14}$$

Equations of motion for Heisenberg picture fields are then dictated by the Heisenberg-Hamilton relations. The first of these relations is

$$\dot{\phi}(x) = i\,[H, \phi(x)] = i\int d^3\boldsymbol{x}'\, [\mathcal{H}(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]\,. \tag{10.15}$$

To this end we need to evaluate $[\mathcal{H}(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]$. One has

$$\begin{aligned} [\mathcal{H}(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)] &= \frac{1}{2}\left[\pi^2(\boldsymbol{x}', t) + (\boldsymbol{\nabla}'\phi(\boldsymbol{x}', t))^2 + m^2\phi^2(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)\right] \\ &= \frac{1}{2}\left[\pi^2(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)\right] \\ &= \frac{1}{2}\pi(\boldsymbol{x}', t)\,[\pi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)] \\ &\quad + \frac{1}{2}\,[\pi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]\,\pi(\boldsymbol{x}', t) \\ &= -i\pi(\boldsymbol{x}', t)\delta(\boldsymbol{x} - \boldsymbol{x}')\,. \end{aligned} \tag{10.16}$$

Thus

$$[H, \phi(\boldsymbol{x}, t)] = -i\int d^3\boldsymbol{x}'\pi(\boldsymbol{x}', t)\delta(\boldsymbol{x} - \boldsymbol{x}') = -i\pi(\boldsymbol{x}, t)\,, \tag{10.17}$$

which leads to

$$\dot{\phi}(\boldsymbol{x},t) \;=\; -i(i\pi(\boldsymbol{x},t)) \;=\; \pi(\boldsymbol{x},t)\,. \tag{10.18}$$

Similarly, the second relation reads

$$\dot{\pi}(\boldsymbol{x},t) \;=\; i\,[H,\pi(\boldsymbol{x},t)] \;=\; i\int \mathrm{d}^3\boldsymbol{x}'\,[\mathcal{H}(\boldsymbol{x}',t),\pi(\boldsymbol{x},t)]\;. \tag{10.19}$$

To evaluate this we need

$$\begin{aligned}
[\mathcal{H}(\boldsymbol{x}',t),\pi(\boldsymbol{x},t)] \;&=\; \frac{1}{2}\left[(\boldsymbol{\nabla}'\phi(\boldsymbol{x}',t))^2\,,\pi(\boldsymbol{x},t)\right] \\
&\quad + \frac{1}{2}m^2\left[\phi^2(\boldsymbol{x}',t),\pi(\boldsymbol{x},t)\right] \\
&=\; i\boldsymbol{\nabla}'\phi(\boldsymbol{x}',t)\boldsymbol{\nabla}'\delta(\boldsymbol{x}-\boldsymbol{x}') \\
&\quad + im^2\phi(\boldsymbol{x}',t)\delta(\boldsymbol{x}-\boldsymbol{x}')\,,
\end{aligned} \tag{10.20}$$

and hence

$$\begin{aligned}
[H,\pi(\boldsymbol{x},t)] \;&=\; i\int \mathrm{d}^3\boldsymbol{x}'\left[\boldsymbol{\nabla}'\phi(\boldsymbol{x}',t)\boldsymbol{\nabla}'\delta(\boldsymbol{x}-\boldsymbol{x}')+m^2\phi(\boldsymbol{x}',t)\delta(\boldsymbol{x}-\boldsymbol{x}')\right] \\
&=\; i\left[-\boldsymbol{\nabla}^2\phi(\boldsymbol{x},t)+m^2\phi(\boldsymbol{x},t)\right]\;.
\end{aligned} \tag{10.21}$$

This leads to the second equation of motion of the form

$$\dot{\pi}(\boldsymbol{x},t) \;=\; \boldsymbol{\nabla}^2\phi(\boldsymbol{x},t)-m^2\phi(\boldsymbol{x},t)\,. \tag{10.22}$$

Recall that $\dot{\phi}(\boldsymbol{x},t)=\pi(\boldsymbol{x},t)$, then

$$\ddot{\phi}(\boldsymbol{x},t) \;=\; \dot{\pi}(\boldsymbol{x},t) \;=\; \boldsymbol{\nabla}^2\phi(\boldsymbol{x},t)\;-\;m^2\phi(\boldsymbol{x},t)\,. \tag{10.23}$$

This is an equation of motion for the Heisenberg field, which can be cast into more familiar form

$$\ddot{\phi}\;-\;\boldsymbol{\nabla}^2\phi \;=\; -m^2\phi \quad\Leftrightarrow\quad \partial_\mu\partial^\mu\phi\;+\;m^2\phi\;=\;0\,. \tag{10.24}$$

Which can be succinctly rewritten as

$$\left(\Box\;+\;m^2\right)\phi \;=\; 0\,. \tag{10.25}$$

Momentum operator and energy momentum tensor

Our Hamiltonian reads

$$\begin{aligned}
H \;&=\; \int \mathrm{d}^3\boldsymbol{x}\,\left[\pi(x)\dot{\phi}(x)-\mathcal{L}(x)\right] \\
&=\; \int \mathrm{d}^3\boldsymbol{x}\,\dot{\phi}^2(\boldsymbol{x},t)-L(\boldsymbol{x},t)\,.
\end{aligned} \tag{10.26}$$

With this the Heisenberg field obeys the evolution equation $\phi(\boldsymbol{x}, t) = e^{iHt}\phi(\boldsymbol{x}, 0)e^{-iHt}$, which can be equivalently rewritten as

$$\phi(\boldsymbol{x}, t - \tau) \;=\; e^{-iH\tau}\phi(\boldsymbol{x}, t)e^{iH\tau}\,. \tag{10.27}$$

We can now ask question:"How does operator $\boldsymbol{P}$, that is responsible for spatial translations, look like?" If we translate the physical system by a spatial displacement $\boldsymbol{a}$, then $\phi(\boldsymbol{x}, t) \to \phi(\boldsymbol{x} - \boldsymbol{a}, t)$. Idea is that the momentum operator $\boldsymbol{P}$ should be the generator of these translations. In other words, we require that

$$e^{i\boldsymbol{P}\cdot\boldsymbol{a}}\phi(\boldsymbol{x}, t)e^{-i\boldsymbol{P}\cdot\boldsymbol{a}} \;=\; \phi(\boldsymbol{x} - \boldsymbol{a}, t)\,. \tag{10.28}$$

The sign in the generator is dictated by the presumed covariant form $e^{-iP_\mu a^\mu}\phi(\boldsymbol{x}, t)e^{iP_\mu a^\mu} \;=\; \phi(x - a)$.

Let $\boldsymbol{a}$ be infinitesimal, then we can write (10.28) as

$$\phi(\boldsymbol{x}, t) \;+\; i\,[\boldsymbol{P}\cdot\boldsymbol{a}, \phi(\boldsymbol{x}, t)] \;+\; O(\boldsymbol{a}^2)\,. \tag{10.29}$$

On the other hand, according to Taylor's expansion

$$\phi(\boldsymbol{x} - \boldsymbol{a}, t) \;=\; \phi(\boldsymbol{x}, t) \;-\; \boldsymbol{a}\cdot\boldsymbol{\nabla}\phi(\boldsymbol{x}, t) \;+\; O(\boldsymbol{a}^2)\,. \tag{10.30}$$

This implies that

$$i\,[\boldsymbol{P}\cdot\boldsymbol{a}, \phi(\boldsymbol{x}, t)] \;=\; -\boldsymbol{a}\cdot\boldsymbol{\nabla}\phi(\boldsymbol{x}, t)\,. \tag{10.31}$$

Since $\boldsymbol{a}$ is arbitrary, we have

$$i\left[\boldsymbol{P}^k, \phi(\boldsymbol{x}, t)\right] \;=\; -\boldsymbol{\nabla}_k\phi(\boldsymbol{x}, t)\,. \tag{10.32}$$

To construct $\boldsymbol{P}^k$, we observe that

$$\begin{aligned}\left[\pi(\boldsymbol{x}', t)\boldsymbol{\nabla}'_k\phi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)\right] \;&=\; [\pi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]\,\boldsymbol{\nabla}'_k\phi(\boldsymbol{x}', t)\\ &=\; -i\delta(\boldsymbol{x} - \boldsymbol{x}')\boldsymbol{\nabla}'_k\phi(\boldsymbol{x}', t)\,,\end{aligned} \tag{10.33}$$

and so we can choose

$$\boldsymbol{P}^k \;=\; -\int d^3\boldsymbol{x}'\pi(\boldsymbol{x}', t)\boldsymbol{\nabla}_k\phi(\boldsymbol{x}', t)\,. \tag{10.34}$$

Indeed, in this case

$$\begin{aligned}\left[\boldsymbol{P}^k, \phi(\boldsymbol{x}, t)\right] \;&=\; -\int d^3\boldsymbol{x}'[\pi(\boldsymbol{x}, t)\boldsymbol{\nabla}_k\phi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]\\ &= (-1)(-i)\int d^3\boldsymbol{x}'\delta(\boldsymbol{x} - \boldsymbol{x}')\boldsymbol{\nabla}_k\phi(\boldsymbol{x}', t) \;=\; i\boldsymbol{\nabla}_k\phi(\boldsymbol{x}, t)\,.\end{aligned} \tag{10.35}$$

This is precisely what we have required for $\boldsymbol{P}^k$. Relation (10.34) thus fixes $\boldsymbol{P}^k$ modulo additive c-number (this point will be discussed shortly).

By using the fact that $\boldsymbol{\nabla}_k \;=\; \partial_k \;=\; -\partial^k, (k = 1, 2, 3)$ we can write

$$\boldsymbol{P}^k \;=\; \int d^3\boldsymbol{x}\,\pi(\boldsymbol{x}, t)\partial^k\phi(\boldsymbol{x}, t)\,. \tag{10.36}$$

In addition, we can check that $\boldsymbol{P}^k$ is independent of t and so it represents a conserved quantity. To see this let us consider the commutator

$[H, \boldsymbol{P}]$. Recalling that

$$H = \int d^3\boldsymbol{x} \left[\frac{1}{2}\pi^2(\boldsymbol{x},t) + \frac{1}{2}(\boldsymbol{\nabla}\phi(\boldsymbol{x},t))^2 + \frac{1}{2}m^2\phi^2(\boldsymbol{x},t) \right]. \qquad (10.37)$$

we can directly write

$$\begin{aligned}
[H, \boldsymbol{P}] &= -\int d^3\boldsymbol{x} d^3\boldsymbol{x}' \left[\pi(\boldsymbol{x},t)\partial^k\phi(\boldsymbol{x},t), \left(\frac{1}{2}\pi^2(\boldsymbol{x}',t) + \right.\right. \\
&\quad \left.\left. \frac{1}{2}(\boldsymbol{\nabla}\phi(\boldsymbol{x}',t))^2 + \frac{1}{2}m^2\phi^2(\boldsymbol{x}',t) \right) \right] \\
&= \int d^3\boldsymbol{x}' \, [H, \pi(\boldsymbol{x}',t)] \, \partial'^k\phi(\boldsymbol{x}',t) + \pi(\boldsymbol{x}',t) \left[H, \partial'^k\phi(\boldsymbol{x},t) \right] \\
&= \int d^3\boldsymbol{x} \left(-i\dot{\pi}(\boldsymbol{x},t)\partial^k\phi(\boldsymbol{x},t) \right) + \pi(\boldsymbol{x},t)\partial^k \underbrace{[H, \phi(\boldsymbol{x},t)]}_{-i\pi(\boldsymbol{x},t)} \\
&= -i\int d^3\boldsymbol{x}' \left[\left(\boldsymbol{\nabla}'\boldsymbol{\nabla}'\phi(\boldsymbol{x}',t) - m^2\phi(\boldsymbol{x}',t) \right) \partial^k\phi(\boldsymbol{x}',t) \right. \\
&\quad \left. + \, \pi(\boldsymbol{x}',t)\partial'^k\pi(\boldsymbol{x}',t) \right] \\
&= -i\int d^3\boldsymbol{x} \, \boldsymbol{\nabla}\boldsymbol{\nabla}\phi(\boldsymbol{x},t)\boldsymbol{\nabla}_k\phi(\boldsymbol{x},t) \; + \; \text{surface terms} \\
&= -i\int d^3\boldsymbol{x} \underbrace{\left\{ -\frac{1}{2}\boldsymbol{\nabla}_k \left[\boldsymbol{\nabla}_l\phi(\boldsymbol{x},t)\boldsymbol{\nabla}_l\phi(\boldsymbol{x},t) \right] \right\}}_{\text{surface term}} \\
&= 0. \qquad (10.38)
\end{aligned}$$

Thus, for a field that at each fixed t and $|\boldsymbol{x}| \to \infty$ goes quickly to zero, the commutator of $[H, \boldsymbol{P}] = 0$.

Let us now set $P^0 = H$, then we can combine the above results for H and $\boldsymbol{P}$ to a single 4-vector

$$P^\mu = (P^0, \boldsymbol{P}) = \int d^3\boldsymbol{x} \, T^{0\mu}(x), \qquad (10.39)$$

where the tensor

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}(\phi(x), \dot{\phi}(x)). \qquad (10.40)$$

It is easy to see that such $T^{\mu\nu}$ provides a correct 4-momentum. In fact, from (10.40) we have that

$$T^{00} = \partial^0\phi\partial^0\phi - \mathcal{L} = \partial_t\phi\partial_t\phi - \mathcal{L} = \pi\dot{\phi} - \mathcal{L} = \mathcal{H}, \qquad (10.41)$$

and

$$T^{0k} = \partial^0\phi\partial^k\phi = -\dot{\phi}\boldsymbol{\nabla}_k\phi = -\pi\boldsymbol{\nabla}_k\phi. \qquad (10.42)$$

This consistently implies both $H = P^0$ and $\boldsymbol{P}^k$. In addition, from the equation of motion it follows that

$$\partial_\mu T^{\mu\nu} = 0. \qquad (10.43)$$

$[H, \boldsymbol{P}]$. Recalling that

$$H = \int \mathrm{d}^3\boldsymbol{x} \left[\frac{1}{2}\pi^2(\boldsymbol{x},t) + \frac{1}{2}(\boldsymbol{\nabla}\phi(\boldsymbol{x},t))^2 + \frac{1}{2}m^2\phi^2(\boldsymbol{x},t)\right]. \tag{3.37}$$

we can directly write

$$\begin{aligned}
[H,\boldsymbol{P}] &= -\int d^3\boldsymbol{x}\mathrm{d}^3\boldsymbol{x}' \Big[\pi(\boldsymbol{x},t)\partial^k\phi(\boldsymbol{x},t), \Big(\frac{1}{2}\pi^2(\boldsymbol{x}',t)+ \\
&\quad \frac{1}{2}(\boldsymbol{\nabla}\phi(\boldsymbol{x}',t))^2 + \frac{1}{2}m^2\phi^2(\boldsymbol{x}',t)\Big)\Big] \\
&= \int d^3\boldsymbol{x}'\,[H,\pi(\boldsymbol{x}',t)]\,\partial'^k\phi(\boldsymbol{x}',t) + \pi(\boldsymbol{x}',t)\left[H,\partial'^k\phi(\boldsymbol{x},t)\right] \\
&= \int \mathrm{d}^3\boldsymbol{x}\left(-i\dot{\pi}(\boldsymbol{x},t)\partial^k\phi(\boldsymbol{x},t)\right) + \pi(\boldsymbol{x},t)\partial^k \underbrace{[H,\phi(\boldsymbol{x},t)]}_{-i\pi(\boldsymbol{x},t)} \\
&= -i\int d^3\boldsymbol{x}'\left[\left(\boldsymbol{\nabla}'\boldsymbol{\nabla}'\phi(\boldsymbol{x}',t) - m^2\phi(\boldsymbol{x}',t)\right)\partial^k\phi(\boldsymbol{x}',t)\right. \\
&\quad \left. + \pi(\boldsymbol{x}',t)\partial'^k\pi(\boldsymbol{x}',t)\right] \\
&= -i\int d^3\boldsymbol{x}\,\boldsymbol{\nabla}\boldsymbol{\nabla}\phi(\boldsymbol{x},t)\boldsymbol{\nabla}_k\phi(\boldsymbol{x},t) + \text{surface terms} \\
&= -i\int d^3\boldsymbol{x}\underbrace{\left\{-\frac{1}{2}\boldsymbol{\nabla}_k\left[\boldsymbol{\nabla}_l\phi(\boldsymbol{x},t)\boldsymbol{\nabla}_l\phi(\boldsymbol{x},t)\right]\right\}}_{\text{surface term}} \\
&= 0.
\end{aligned} \tag{3.38}$$

Thus, for a field that at each fixed t and $|\boldsymbol{x}| \to \infty$ goes quickly to zero, the commutator of $[H, \boldsymbol{P}] = 0$.

Let us now set $P^0 = H$, then we can combine the above results for H and $\boldsymbol{P}$ to a single 4-vector

$$P^\mu = (P^0, \boldsymbol{P}) = \int \mathrm{d}^3\boldsymbol{x}\, T^{0\mu}(x), \tag{3.39}$$

where the tensor

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}(\phi(x), \dot{\phi}(x)). \tag{3.40}$$

It is easy to see that such $T^{\mu\nu}$ provides a correct 4-momentum. In fact, from (3.40) we have that

$$T^{00} = \partial^0\phi\partial^0\phi - \mathcal{L} = \partial_t\phi\partial_t\phi - \mathcal{L} = \pi\dot{\phi} - \mathcal{L} = \mathcal{H}, \tag{3.41}$$

and

$$T^{0k} = \partial^0\phi\partial^k\phi = -\dot{\phi}\boldsymbol{\nabla}_k\phi = -\pi\boldsymbol{\nabla}_k\phi. \tag{3.42}$$

This consistently implies both $H = P^0$ and $\boldsymbol{P}^k$. In addition, from the equation of motion it follows that

$$\partial_\mu T^{\mu\nu} = 0. \tag{3.43}$$

Indeed, by using the explicit form for $\mathcal{L}$ given by (3.3) we have

$$\begin{aligned}
\partial_\mu T^{\mu\nu} &= \partial_\mu \left[\partial^\mu \phi \partial^\nu \phi \, - \, \eta^{\mu\nu} \left(\frac{1}{2}(\partial_\alpha \phi)^2 - \frac{1}{2} m^2 \phi^2 \right) \right] \\
&= \underbrace{(\partial_\mu \partial^\mu \phi)}_{-m^2\phi} \partial^\nu \phi \, + \, \partial^\mu \phi \partial_\mu \partial^\nu \phi \, - \, (\partial^\nu \partial_\alpha \phi) \partial^\alpha \phi + m^2 \phi \partial^\nu \phi \\
&= -m^2 \phi \partial^\nu \phi \, + \, m^2 \phi \partial^\nu \phi \, = \, 0 \, , \qquad (3.44)
\end{aligned}$$

Let us finally observe that (3.39) can be rewritten in explicitly covariant manner, so that ensuing P^μ will be a genuine 4-vector. In particular, we can write (3.39) as

$$P^\mu \, = \, (P^0, \boldsymbol{P}) \, = \, \int dV n_\mu T^{\mu\nu}(x) \, . \qquad (3.45)$$

Here, the measure dV is over the space-like slice (of the $4D$ spacetime) that is orthogonal to the unit time-like vector n^μ. We claim that P^μ transforms covariantly under a change of n^μ, i.e., P^μ is a genuine Lorentzian 4-vector. To see this, we write

$$\begin{aligned}
\int d^3\boldsymbol{x}\, T^{0\mu}(x) &= \int \mathrm{d}^4 x\, \delta(x_0) T^{0\mu}(x) = \int \mathrm{d}^4 x \left[\frac{\partial}{\partial x^0} \theta(x_0) \right] T^{0\mu}(x) \\
&= \int d^4 x \frac{\partial}{\partial x^\nu} \theta(n_0^\alpha x_\alpha) T^{\nu\mu}(x) \, . \qquad (3.46)
\end{aligned}$$

Here, $n_0^\alpha = (1, 0, 0, 0)$ is a time-like unit vector that is manifestly orthogonal to the space-like slice over which we integrate. Now, we relabel x^μ in (3.46) to x'^μ and take the Lorentz transformation $x^\mu \xrightarrow{\boldsymbol{L}} x'^\mu = \boldsymbol{L}^\mu_{\cdot\,\nu} x^\nu$. With this we can write

$$\begin{aligned}
&\int \mathrm{d}^4 x' \frac{\partial}{\partial x'^\nu} \theta(n_0^\alpha x'_\alpha) T^{\nu\mu}(x') \\
&\quad = \int d^4 x \, |\det \boldsymbol{L}| \boldsymbol{L}_\nu^{\cdot\,\sigma} \frac{\partial}{\partial x^\sigma} \theta(n_{0,\alpha} \boldsymbol{L}^\alpha_{\cdot\,\beta} x^\beta) \boldsymbol{L}^\nu_{\cdot\,\gamma} \boldsymbol{L}^\mu_{\cdot\,\delta} T^{\gamma\delta}(\underbrace{\boldsymbol{L}^{-1} x'}_{x}) \\
&\quad = \int d^4 x \, \eta_\gamma^\sigma \frac{\partial(n_{0,\alpha} \boldsymbol{L}^\alpha_{\cdot\,\beta} x^\beta)}{\partial x^\sigma} \delta(n_{0,\alpha} \boldsymbol{L}^\alpha_{\cdot\,\beta} x^\beta) \boldsymbol{L}^\mu_{\cdot\,\delta} T^{\gamma\delta}(x) \\
&\quad = \boldsymbol{L}^\mu_{\cdot\,\delta} \int d^4 x \, n'_\sigma \delta(n' x) T^{\sigma\delta}(x) \\
&\quad = \boldsymbol{L}^\mu_{\cdot\,\delta} \int dV n'_\sigma T^{\sigma\delta}(x) \, , \qquad (3.47)
\end{aligned}$$

Here n_β denotes time-like 4-vector orthogonal to space-like slice over which we integrate. Note that space-like property of vectors is preserved under proper Lorentz transformations.

with $n'_\nu = \boldsymbol{L}^\mu_{\cdot\,\nu} n_{0,\mu}$ (which equivalently means that $n'^\nu = (\boldsymbol{L}^{-1})^\nu_{\cdot\,\mu} n_0^\mu$). So, that finally

$$P^\mu(n_0) \, = \, P^\mu(Ln') \, = \, L^\mu_{\cdot\,\alpha} \int dV n'_\beta T^{\beta\alpha}(x) \, = \, L^\mu_{\cdot\,\alpha} P^\alpha(n') \, . \qquad (3.48)$$

Thus, the 4-vector P^μ defined by (3.45) transforms as a true relativistic 4-vector under a change of the space-like slice over which we integrate. The order 2 (and type $(2, 0)$) tensor $T^{\mu\nu}(x)$ is called *energy-momentum* *tensor* and in this particular case it satisfies the symmetry condition $T^{\mu\nu} = T^{\nu\mu}$.

Particle interpretation

Let us now Fourier decompose $\phi(x)$ as

$$\phi(x) = \int \mathrm{d}^4 p e^{-ipx} \tilde{\phi}(p). \tag{3.49}$$

With this we get

$$\begin{aligned}(\Box + m^2)\phi(x) = 0 \quad &\Rightarrow \quad \int \mathrm{d}^4 p e^{-ipx}(p^2 - m^2)\tilde{\phi}(p) = 0 \\ &\Rightarrow \quad (p^2 - m^2)\tilde{\phi}(p) = 0. \end{aligned} \tag{3.50}$$

Solution of this equation has the generic form

$$\begin{aligned}\tilde{\phi}(p) &= f(p)\delta(p^2 - m^2) \\ &= f(p)\frac{\delta(p_0 + \sqrt{\boldsymbol{p}^2 + m^2})}{2\underbrace{\sqrt{\boldsymbol{p}^2 + m^2}}_{\omega_{\boldsymbol{p}}}} + f(p)\frac{\delta(p_0 - \sqrt{\boldsymbol{p}^2 + m^2})}{2\sqrt{\boldsymbol{p}^2 + m^2}}. \end{aligned} \tag{3.51}$$

Here $f(p)$ is an arbitrary function that is not zero at $p^2 = m^2$. Eq. (3.50) implies that that $\phi(x)$ can be written in the form

$$\begin{aligned}\phi(x) &= \int \frac{\mathrm{d}^4 p}{2\omega_{\boldsymbol{p}}} e^{-ipx}[f(p)\delta(p_0 + \omega_{\boldsymbol{p}}) + f(p)\delta(p_0 - \omega_{\boldsymbol{p}})] \\ &= \int \frac{\mathrm{d}^3 \boldsymbol{p}}{2\omega_{\boldsymbol{p}}} \left[e^{-i\omega_{\boldsymbol{p}} t_0 + i\boldsymbol{p}\boldsymbol{x}} f(\omega_{\boldsymbol{p}}, \boldsymbol{p}) + e^{i\omega_{\boldsymbol{p}} t_0 + i\boldsymbol{p}\boldsymbol{x}} f(-\omega_{\boldsymbol{p}}, \boldsymbol{p})\right] \\ &= \int \frac{\mathrm{d}^3 \boldsymbol{p}}{2\omega_{\boldsymbol{p}}} \left[e^{-ipx} f(p) + e^{ipx} g(p)\right]. \end{aligned} \tag{3.52}$$

Here we have set $g(p) = f(-\omega_{\boldsymbol{p}}, -\boldsymbol{p})$. So we can finally rewrite it as

$$\phi(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3 2\omega_{\boldsymbol{p}}} \left(a(p)e^{-ipx} + b(p)e^{ipx}\right), \tag{3.53}$$

where

$$a(p) = (2\pi)^3 f(p), \quad b(p) = (2\pi)^3 g(p). \tag{3.54}$$

By requiring that ϕ is Hermitian (in order to have Hermitian Hamiltonian), we have that $b(p) = a^\dagger(p)$, which gives

$$\phi(x) = \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}} \left[a(p)e^{-ipx} + a^\dagger(p)e^{ipx}\right]. \tag{3.55}$$

Similarly for conjugate field momentum $\pi(x) = \dot{\phi}(x)$, we obtain

$$\begin{aligned}\pi(x) = \dot{\phi}(x) &= \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}} \left[(-i\omega_{\boldsymbol{p}})a(p)e^{-ipx} + i\omega_{\boldsymbol{p}} a^\dagger(p)e^{ipx}\right] \\ &= -\frac{i}{2}\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\left[a(p)e^{-ipx} - a^\dagger(p)e^{ipx}\right]. \qquad (3.56)\end{aligned}$$

The measure in the integral (3.55) is manifestly Lorentz invariant. Indeed

$$\frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}} = \frac{d^4 p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\theta(p^0).$$

At this stage we recall that the canonical commutation relations read

$$\begin{aligned}[\phi(\boldsymbol{x},t), \pi(\boldsymbol{y},t)] &= \left[\phi(\boldsymbol{x},t), \dot{\phi}(\boldsymbol{y},t)\right] = i\delta(\boldsymbol{x}-\boldsymbol{y}) \\ [\phi(\boldsymbol{x},t), \phi(\boldsymbol{y},t)] &= [\pi(\boldsymbol{x},t), \pi(\boldsymbol{y},t)] = 0. \qquad (3.57)\end{aligned}$$

From these one can compute commutation relations for $a(p)$ and $a^\dagger(p)$. It is not difficult to see that the following holds

$$\begin{aligned}\left[a(p), a^\dagger(p')\right] &= (2\pi)^3 2\omega_p \delta(\boldsymbol{p}-\boldsymbol{p}') \\ [a(p), a(p')] &= \left[a^\dagger(p), a^\dagger(p')\right] = 0. \qquad (3.58)\end{aligned}$$

The consistency check can be done as follows. First, we define the notation

$$\begin{aligned}\sum_{\boldsymbol{p}} &= \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_p}, \\ \delta_{\boldsymbol{p}\boldsymbol{p}'} &= (2\pi)^3 2\omega_p \delta(\boldsymbol{p}-\boldsymbol{p}'). \qquad (3.59)\end{aligned}$$

With these, we have, for instance

$$\begin{aligned}\sum_{\boldsymbol{p}'} \delta_{\boldsymbol{p}\boldsymbol{p}'} f(\boldsymbol{p}') &= f(\boldsymbol{p}), \\ \left[a(p), a^\dagger(p')\right] &= \delta_{\boldsymbol{p}\boldsymbol{p}'}. \qquad (3.60)\end{aligned}$$

In addition, the field operators acquire succinct forms

$$\phi(x) = \sum_{\boldsymbol{p}} \left(a(p)e^{-ipx} + a^\dagger(p)e^{ipx}\right) \qquad (3.61)$$

and

$$\pi(x) = \dot{\phi}(x) = \sum_{\boldsymbol{p}} (-i\omega_p)\left(a(p)e^{-ipx} - a^\dagger(p)e^{ipx}\right). \qquad (3.62)$$

Now we set $x \equiv (t, \boldsymbol{x})$ and $y \equiv (t, \boldsymbol{y})$ and assume validity of (3.58). With this convention we can write

$$\begin{aligned}
[\phi(x), \pi(y)] &= \sum_{\boldsymbol{p}\boldsymbol{p}'} (-i\omega_{\boldsymbol{p}'}) \left[a(p)e^{-ipx}, (-a^\dagger(p'))e^{ip'y}\right] \\
&\quad + (-i\omega_{\boldsymbol{p}'}) \left[a^\dagger(p)e^{ipx}, a(p')e^{-ip'y}\right] \\
&= \sum_{\boldsymbol{p}} \left[(i\omega_p)e^{i\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{y})} + (i\omega_p)e^{-i\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{y})}\right] \\
&= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{i2\omega_p}{2\omega_p} e^{i\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{y})} \\
&= i\delta(\boldsymbol{x} - \boldsymbol{y})\,.
\end{aligned} \tag{3.63}$$

Similarly, it can be checked that

$$[\phi(x), \phi(y)] = [\pi(x), \pi(y)] = 0\,. \tag{3.64}$$

One can alternatively start from ϕ, π and commutation relations (3.57) and prove (3.58). This is a bit lengthier but a straightforward exercise.

Energy — Hamiltonian

Hamiltonian can be written in the form

$$H = \int \mathrm{d}^3\boldsymbol{x}\, \underbrace{\frac{1}{2}\pi^2(x)}_{H_1} + \underbrace{\frac{1}{2}(\nabla\phi(x))^2}_{H_2} + \underbrace{m^2\phi^2}_{H_3}\,. \tag{3.65}$$

Let us now compute the respective terms explicitly in terms of mode operators a and $a^\dagger$:

It is normally tedious to do this computation in its full length. But we can take advantage that H is time independent and drop all time-dependent terms because these must cancel anyway.

$$\begin{aligned}
H_1 &= \int \mathrm{d}^3\boldsymbol{x}\, \frac{1}{2}\pi^2 \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'} (-i\omega_p)(-i\omega_{p'}) \int \mathrm{d}^3\boldsymbol{x} \left[a(p)a(p')e^{-i(p+p')x}\right. \\
&\quad \left. - a(p)a^\dagger(p')e^{-i(p-p')x} + a^\dagger(p)a^\dagger(p')e^{i(p+p')x} - a^\dagger(p)a(p')e^{i(p-p')x}\right] \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'} (-\omega_p\omega_{p'}) \left[-a(p)a^\dagger(p')(2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')\right. \\
&\quad \left. - a^\dagger(p)a(p')(2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')\right] \\
&= \frac{1}{4}\sum_{\boldsymbol{p}} \omega_p \left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right]\,.
\end{aligned} \tag{3.66}$$

We stress that this result represents only *time independent* contribution to H_1. Similarly we can compute time independent contributions to H_2

and H_3, namely

$$
\begin{aligned}
H_2 &= \frac{1}{2}\int \mathrm{d}^3\boldsymbol{x}\,(\boldsymbol{\nabla}\phi)^2 \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}(i\boldsymbol{p})(-i\boldsymbol{p}')\int \mathrm{d}^3\boldsymbol{x}[a(p)a^\dagger(p')e^{-i(p-p')x} + a^\dagger(p)a(p')e^{i(p-p')x}] \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}\boldsymbol{p}\boldsymbol{p}'(2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')[a(p)a^\dagger(p') + a^\dagger(p)a(p')] \\
&= \frac{1}{2}\sum_{\boldsymbol{p}}\frac{\boldsymbol{p}^2}{2\omega_p}\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right] , \qquad (3.67)
\end{aligned}
$$

$$
H_3 = \frac{1}{2}\sum_{p}\frac{m^2}{2\omega_p}\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right] . \qquad (3.68)
$$

which finally gives

$$
H = H_1 + H_2 + H_3 = \frac{1}{2}\sum_{\boldsymbol{p}}\omega_p\left(a(p)a^\dagger(p) + a^\dagger(p)a(p)\right) . \qquad (3.69)
$$

We could rewrite this as

$$
\sum_{\boldsymbol{p}}\omega_{\boldsymbol{p}}\left(a^\dagger(p)a(p) + \frac{1}{2}\delta_{\boldsymbol{p}\boldsymbol{p}}\right),
$$

which mimics the linear harmonic oscillator, though now we have sum of infinitely many of them and the contribution from $\frac{1}{2}\omega_{\boldsymbol{p}}\delta_{\boldsymbol{p}\boldsymbol{p}}$ diverges.

Spatial momentum

Spatial momentum can be written as

$$
\boldsymbol{P} = -\int \mathrm{d}^3\boldsymbol{x}\,\pi(x)\boldsymbol{\nabla}\phi(x) . \qquad (3.70)
$$

Analogous calculation as for H gives

$$
\boldsymbol{P} = \frac{1}{2}\sum_{\boldsymbol{p}}\boldsymbol{p}\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right] , \qquad (3.71)
$$

which does not suffer with vacuum divergences. Together with H, we can form four-vector

$$
P^\mu = \frac{1}{2}\sum_{\boldsymbol{p}}p^\mu\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right] . \qquad (3.72)
$$

Particle interpretation

Similarly as in linear harmonic oscillator, the physical states of our quantum system are excitations over a ground state, which we identify with the vacuum state $|0\rangle$. Here $|0\rangle$ satisfies $a(p)\,|0\rangle = 0$ for all $\boldsymbol{p}$.

Ground state and excited states

The fact that the *ground state* must be annihilated by $a(p)$ can be easily understood via similar reasoning as done for linear harmonic oscillator.

and H_3, namely

$$
\begin{aligned}
H_2 &= \frac{1}{2}\int d^3\boldsymbol{x}\,(\boldsymbol{\nabla}\phi)^2 \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}(i\boldsymbol{p})(-i\boldsymbol{p}')\int d^3\boldsymbol{x}\,[a(\boldsymbol{p})a^\dagger(\boldsymbol{p}')e^{-i(p-p')x} + a^\dagger(\boldsymbol{p})a(\boldsymbol{p}')e^{i(p-p')x}] \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}\boldsymbol{p}\boldsymbol{p}'(2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')[a(\boldsymbol{p})a^\dagger(\boldsymbol{p}') + a^\dagger(\boldsymbol{p})a(\boldsymbol{p}')] \\
&= \frac{1}{2}\sum_{\boldsymbol{p}}\frac{\boldsymbol{p}^2}{2\omega_{\boldsymbol{p}}}\left[a(\boldsymbol{p})a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p})a(\boldsymbol{p})\right], \qquad (3.67)
\end{aligned}
$$

$$
H_3 = \frac{1}{2}\sum_{p}\frac{m^2}{2\omega_{\boldsymbol{p}}}\left[a(\boldsymbol{p})a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p})a(\boldsymbol{p})\right]. \qquad (3.68)
$$

which finally gives

$$
H = H_1 + H_2 + H_3 = \frac{1}{2}\sum_{\boldsymbol{p}}\omega_{\boldsymbol{p}}\left(a(\boldsymbol{p})a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p})a(\boldsymbol{p})\right). \qquad (3.69)
$$

We could rewrite this as

$$
\sum_{\boldsymbol{p}}\omega_{\boldsymbol{p}}\left(a^\dagger(\boldsymbol{p})a(\boldsymbol{p}) + \frac{1}{2}\delta_{\boldsymbol{p}\boldsymbol{p}}\right),
$$

which mimics the linear harmonic oscillator, though now we have sum of infinitely many of them and the contribution from $\frac{1}{2}\omega_{\boldsymbol{p}}\delta_{\boldsymbol{p}\boldsymbol{p}}$ diverges.

Spatial momentum

Spatial momentum can be written as

$$
\boldsymbol{P} = -\int d^3\boldsymbol{x}\,\pi(x)\boldsymbol{\nabla}\phi(x). \qquad (3.70)
$$

Analogous calculation as for H gives

$$
\boldsymbol{P} = \frac{1}{2}\sum_{\boldsymbol{p}}\boldsymbol{p}\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right], \qquad (3.71)
$$

which does not suffer with vacuum divergences. Together with H, we can form four-vector

$$
P^\mu = \frac{1}{2}\sum_{\boldsymbol{p}}p^\mu\left[a(\boldsymbol{p})a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p})a(\boldsymbol{p})\right]. \qquad (3.72)
$$

Particle interpretation

Similarly as in linear harmonic oscillator, the physical states of our quantum system are excitations over a ground state, which we identify with the vacuum state $|0\rangle$. Here $|0\rangle$ satisfies $a(\boldsymbol{p})\,|0\rangle = 0$ for all $\boldsymbol{p}$.

Ground state and excited states

The fact that the *ground state* must be annihilated by $a(p)$ can be easily understood via similar reasoning as done for linear harmonic oscillator.

We first denote the ground state as $|0\rangle$, then

$$H|0\rangle = \omega_0|0\rangle\,,$$

where ω_0 is the *minimal eigenvalue* of H, which by positive definiteness of H cannot be negative.

We further note that

$$[H, a(p)] = \sum_{\boldsymbol{p}'} \omega_{\boldsymbol{p}'}[a^\dagger(\boldsymbol{p}')a(p'), a(\boldsymbol{p})] = -\omega_{\boldsymbol{p}} a(\boldsymbol{p})\,,$$

which implies that

$$Ha(\boldsymbol{p})|0\rangle = \left[a(\boldsymbol{p})H - \omega_{\boldsymbol{p}} a(\boldsymbol{p})\right]|0\rangle = (\omega_0 - \omega_{\boldsymbol{p}})a(\boldsymbol{p})|0\rangle\,.$$

So, $a(p)|0\rangle$ is also eigenstate of H with the eigenvalue $\omega_0 - \omega_{\boldsymbol{p}}$. This, however, is in contradiction with the fact that ω_0 is the smallest eigenvalue. The only way how to resolve this contradiction is to assume that $a(\boldsymbol{p})|0\rangle = 0$, for all $\boldsymbol{p}$.

In order to see how to construct *excited states* in our theory, we use the relation

$$[H, a^\dagger(\boldsymbol{p})] = \omega_{\boldsymbol{p}} a^\dagger(\boldsymbol{p})\,,$$

which implies that

$$Ha^\dagger(\boldsymbol{p})|0\rangle = (\omega_0 + \omega_{\boldsymbol{p}})a^\dagger(\boldsymbol{p})|0\rangle\,,$$

and more generally

$$H[a^\dagger(\boldsymbol{p})]^n|0\rangle = (\omega_0 + n\omega_{\boldsymbol{p}})[a^\dagger(\boldsymbol{p})]^n|0\rangle\,,$$

$$H\left\{[a^\dagger(\boldsymbol{p}_1)]^{n_1}[a^\dagger(\boldsymbol{p}_2)]^{n_2}\ldots\right\}|0\rangle$$
$$= (\omega_0 + n_1\omega_{\boldsymbol{p}_1} + n_2\omega_{\boldsymbol{p}_2} + \ldots)\left\{[a^\dagger(\boldsymbol{p}_1)]^{n_1}[a^\dagger(\boldsymbol{p}_2)]^{n_2}\ldots\right\}|0\rangle\,.$$

This shows that states $\prod_i [a^\dagger(\boldsymbol{p}_i)]^{n_i}|0\rangle$ correspond to excited states with the energy eigenvalue $(\omega_0 + \sum_i n_i\omega_{\boldsymbol{p}_i})$. In addition, there are no other eigenstates of H than those created via application of $a^\dagger$ on the vacuum state. If this would not be the case, we could, in contrast assume that there is an exited state $|\alpha_k\rangle$ such that

$$H|\alpha_k\rangle = (\omega_0 + n_1\omega_{\boldsymbol{p}_1} + n_2\omega_{\boldsymbol{p}_2} + \ldots + \alpha_k\omega_{\boldsymbol{p}_k} + \ldots)|\alpha_k\rangle\,,$$

where α_k is not a positive integer. By employing the identity

$$Ha(\boldsymbol{p})|0\rangle = a(\boldsymbol{p})H - \omega_{\boldsymbol{p}} a(\boldsymbol{p})\,,$$

we would obtain that

$$Ha^n(\boldsymbol{p}_k)|\alpha_k\rangle = \left[\omega_0 + n_1\omega_{\boldsymbol{p}_1} + \ldots + (\alpha_k - n)\omega_{\boldsymbol{p}_k} + \ldots\right] a^n(\boldsymbol{p}_k)|\alpha_k\rangle\,.$$

This equation implies that also the state $a^n(\boldsymbol{p}_k)|\alpha_k\rangle$ is an eigenstate of the Hamiltonian. On the other hand, since this is true for any positive integer n, we can always choose a sufficiently large n that will ensure that the ensuing eigenvalue of $a^n(p_k)|\alpha_k\rangle$ is negative. This cannot be, however, right since H is positive definite operator. The only way out of this paradox is that α_k is a positive integer.

Typical states are:

- $|p\rangle = a^\dagger(\boldsymbol{p})\,|0\rangle$ — *single-particle state*, it creates excitation with energy $\omega_{\boldsymbol{p}}$ and momenta $\boldsymbol{p}$ (as will be seen shortly).

  In linear harmonic oscillator, $|1\rangle = a^\dagger\,|0\rangle$ creates excited state with energy ω — here no momentum is included, thus one cannot really speak about particle.

- $|p_1, p_2\rangle = a^\dagger(\boldsymbol{p}_1)a^\dagger(\boldsymbol{p}_2)\,|0\rangle$ — *two-particle state*, it creates two excitations, one with energy $\omega_{\boldsymbol{p}_1}$ and momenta $\boldsymbol{p}_1$, second with energy $\omega_{\boldsymbol{p}_2}$ and momenta $\boldsymbol{p}_2$ (again the momentum parts will be seen shortly).

  In linear harmonic oscillator,

  $$|2\rangle = \frac{a^\dagger a^\dagger}{\sqrt{2}}\,|0\rangle\ , \qquad (3.73)$$

  creates 2 excited states above $|0\rangle$ with energy $\omega + \omega = 2\omega$.

- *Three and higher particle states* can be constructed in the same way.

To bolster our particle interpretation we can construct a *particle number operator*

$$N \;=\; \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_p} a^\dagger(\boldsymbol{p})a(\boldsymbol{p}) \;=\; \sum_{\boldsymbol{p}} a^\dagger(\boldsymbol{p})a(\boldsymbol{p})\,, \qquad (3.74)$$

Clearly an analogue of the *energy-level operator* in linear harmonic oscillator. There $N = a^\dagger a$.

which satisfies

$$\begin{aligned}[N, a^\dagger(p)] \;&=\; \sum_{\boldsymbol{p}'}[a^\dagger(\boldsymbol{p}')a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] \\ &=\; \sum_{\boldsymbol{p}'} a^\dagger(\boldsymbol{p}')\underbrace{[a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p}')]}_{\delta_{\boldsymbol{p}\boldsymbol{p}'}} \;=\; a^\dagger(\boldsymbol{p})\,. \end{aligned} \qquad (3.75)$$

So, that

$$N a^\dagger(\boldsymbol{p}) \;=\; a^\dagger N + a^\dagger(\boldsymbol{p}) \;=\; a^\dagger(\boldsymbol{p})(N+1)\,. \qquad (3.76)$$

Operator N thus counts particles (or excitations) in a given state. Indeed, for instance

$$N\,|0\rangle \;=\; \sum_{\boldsymbol{p}} a^\dagger(\boldsymbol{p})a(\boldsymbol{p})\,|0\rangle \;=\; 0\,, \qquad (3.77)$$

so for vacuum, $N = 0$. Similarly,

$$N\,|p\rangle \;=\; Na^{\dagger}(\boldsymbol{p})\,|0\rangle \;=\; a^{\dagger}(\boldsymbol{p})(N+1)\,|0\rangle \;=\; a^{\dagger}(\boldsymbol{p})\,|0\rangle \;=\; |p\rangle\;, \tag{3.78}$$

$|p\rangle$ is an eigenstate of N with $N = 1$. Analogously

$$\begin{aligned} N\,|p_1, p_2\rangle \;&=\; Na^{\dagger}(\boldsymbol{p}_1)a^{\dagger}(\boldsymbol{p}_2)\,|0\rangle \;=\; a^{\dagger}(\boldsymbol{p}_1)(N+1)a^{\dagger}(\boldsymbol{p}_2)\,|0\rangle \\ &=\; a^{\dagger}(\boldsymbol{p}_1)a^{\dagger}(\boldsymbol{p}_2)(N+2)\,|0\rangle \;=\; 2\,|p_1, p_2\rangle\;, \end{aligned} \tag{3.79}$$

$|p_1 p_2\rangle$ is a 2-particle eigenstate of N and similarly for higher excited states. By the same token we could proceed to higher particle states.

Note on vacuum energy

In the Hamiltonian (3.69) we have seen that

$$\frac{1}{2}\left(a(\boldsymbol{p})a^{\dagger}(\boldsymbol{p}) + a^{\dagger}(\boldsymbol{p})a(\boldsymbol{p})\right) \;=\; a^{\dagger}(\boldsymbol{p})a(\boldsymbol{p}) \;+\; \frac{1}{2}\delta_{\boldsymbol{pp}}\,.$$

Since $\delta_{\boldsymbol{pp}} = (2\pi^3)2\omega_{\boldsymbol{p}}\delta(\boldsymbol{0})$, we can observe that the (momentum) density of the ground state energy diverges. To understand the structure of this divergence let us first realize that

$$(2\pi)^3\delta(\boldsymbol{0}) \;=\; \int \mathrm{d}^3\boldsymbol{x}\, e^{\pm i\boldsymbol{px}}|_{\boldsymbol{p}=0} \;=\; \int \mathrm{d}^3\boldsymbol{x} \;=\; V\,,$$

this diverges is in the infinite volume limit. Then,

$$\frac{1}{2}\delta_{\boldsymbol{pp}} \quad=\quad \text{energy density}\cdot V\,,$$

which implies that the ground state energy reads

$$\begin{aligned} E_0 \;&=\; \frac{1}{2}\sum_{\boldsymbol{p}}\omega_{\boldsymbol{p}}\delta_{\boldsymbol{pp}'} \\ &=\; \frac{1}{2}\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}}\omega_{\boldsymbol{p}}(2\pi)^3 2\omega_{\boldsymbol{p}}\delta(\boldsymbol{0}) \\ &=\; \frac{1}{2}\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\omega_{\boldsymbol{p}}V\,. \end{aligned}$$

Here $\omega_{\boldsymbol{p}}V$ is the density of energy with given $\boldsymbol{p}$ in the entire space. This is divergent because for each frequency (momentum) there are infinitely many particles whose number is proportional to the volume of the space. Because the divergent factor V is related to small momenta, the corresponding divergence is called *infrared* or *IR divergence*.

Vacuum energy is relevant only for the gravitational field. If we neglect gravity then the presence of vacuum energy cannot be detected in usual experiments that are based only on transformations between excited states.

Naively one could assume that the (spatial) density of the ground state energy could be finite. Unfortunately even this quantity is divergent:

$$\lim_{V\to\infty}\frac{E_0}{V} \;=\; \frac{1}{2}\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\omega_{\boldsymbol{p}} \;=\; \infty\,.$$

So, the vacuum energy density diverges at $|\boldsymbol{p}| \to \infty$. This so-called *UV divergence* arises because the "oscillators" with large momentum have very large zero-point energy $1/2\omega_{\boldsymbol{p}} \approx |\boldsymbol{p}|/2$.

Let us now compare the ground state issue with the one in the linear harmonic oscillator (LHO):

- LHO
  - ω is $\boldsymbol{p}$ independent — no UV divergence
  - there is only 1 or finite number of oscillators in volume V — no IR divergence
- QFT
  - frequency $\omega_{\boldsymbol{p}}$ depends on $\boldsymbol{p}$ which changes smoothly over $\mathbb{R}^3$. For large $|\boldsymbol{p}|$ we have that $\omega_p \approx |\boldsymbol{p}|$, which implies UV divergence
  - in a given volume V, there is a large number ($\propto V$) of oscillators with fixed frequencies — IR divergence

Typically one wishes to ensure that vacuum has 0 energy and momentum. So, we subtract the ground state energy and define the Hamiltonian to be original Hamiltonian minus ground state energy. We wish to set $P^\mu = \sum_{\boldsymbol{p}} p^\mu a^\dagger(\boldsymbol{p})a(\boldsymbol{p})$. Now $P^\mu |0\rangle = 0$.

This change does not destroy the significance of the previous calculations with p^μ, since only commutators were involved. Adjustment by a c-number does not affect commutators.

The subtraction of the vacuum energy does not, however, remove the vacuum fluctuations of the quantum field. This can be evaluated from the so-called *correlation function*

$$\langle 0| \phi(\boldsymbol{x},t)\phi(\boldsymbol{y},t) |0\rangle \ .$$

Similarly as subtraction of vacuum energy in LHO does not remove fluctuation in $\boldsymbol{p}$ or $\boldsymbol{x}$.

$$\begin{aligned} P^\mu |p\rangle &= \sum_{\boldsymbol{p}'} p^{\mu\prime} a^\dagger(\boldsymbol{p}')a(\boldsymbol{p}')a^\dagger(\boldsymbol{p}) |0\rangle = \sum_{\boldsymbol{p}'} p^{\mu\prime} \left[a^\dagger(\boldsymbol{p}')a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})\right] |0\rangle \\ &= \sum_{\boldsymbol{p}'} p^{\mu\prime} a^\dagger(\boldsymbol{p}') \left[a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})\right] |0\rangle = \sum_{\boldsymbol{p}'} p^{\mu\prime} a^\dagger(\boldsymbol{p}')\delta_{\boldsymbol{p}'\boldsymbol{p}} |0\rangle \\ &= p^\mu a^\dagger(\boldsymbol{p}) |0\rangle = p^\mu |p\rangle \, . \end{aligned} \tag{3.80}$$

This means that $|p\rangle$ is an eigenstate of P^μ with eigenvalue p^μ.

For two particle states:

$$\begin{aligned} P^\mu |p_1, p_2\rangle &= \sum_{\boldsymbol{p}'} p^{\mu\prime} (a^\dagger(p')a(p')a^\dagger(p_1)a^\dagger(p_2) |0\rangle \\ &= \sum_{\boldsymbol{p}'} p^{\mu\prime} \left[a^\dagger(\boldsymbol{p}')a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p}_1)a^\dagger(\boldsymbol{p}_2)\right] |0\rangle \\ &= \sum_{\boldsymbol{p}'} p^{\mu\prime} a^\dagger(\boldsymbol{p}') \left[a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p}_1)a^\dagger(\boldsymbol{p}_2)\right] |0\rangle \\ &= \sum_{\boldsymbol{P}'} p^{\mu\prime} a^\dagger(\boldsymbol{p}') \left[a^\dagger(\boldsymbol{p}_2)\delta_{\boldsymbol{p}'\boldsymbol{p}_1} + a^\dagger(\boldsymbol{p}_1)\delta_{\boldsymbol{p}'\boldsymbol{p}_2}\right] |0\rangle \\ &= (p_1^\mu + p_2^\mu) a^\dagger(\boldsymbol{p}_1)a^\dagger(\boldsymbol{p}_2) |0\rangle \\ &= (p_1^\mu + p_2^\mu) |p_1, p_2\rangle \, . \end{aligned} \tag{3.81}$$

This means that $|p_1, p_2\rangle$ is an eigenstate of P^μ with $(p_1^\mu + p_2^\mu)$ as eigenvalue. Similalry, for higher particle states. This once more justifies the usage of the particle picture description.

Occupation number representation — 1st bite

In case we have n_{p_1} particles with the momenta p_1, n_{p_2} particles with the momenta p_2, etc., it is customary to write the corresponding $\sum_i n_{p_i}$-particle state as

$$|n_{p_1}, n_{p_2}, \cdots\rangle \;\equiv\; |\{n_p\}\rangle \;=\; \frac{(a^{\dagger}_{\boldsymbol{p}_1})^{n_{p_1}}}{\sqrt{n_{p_1}!}} \frac{(a^{\dagger}_{\boldsymbol{p}_2})^{n_{p_2}}}{\sqrt{n_{p_2}!}} \cdots |0\rangle \; .$$

Compare with LHO n-th excited state:

$$|n\rangle \;=\; \frac{(a^{\dagger})^n}{\sqrt{n!}} \, |0\rangle \; .$$

This way of representing momentum eigenstates of free fields is known as the occupation number representation. For instance, in the occupation number representation we would have for a 2-particle state

$$|p_i, p_j\rangle \;=\; |\underbrace{0}_{p_1}\; \underbrace{0}_{p_2}\; \ldots\; \underbrace{1}_{p_i}\; \ldots 0 \ldots\; \underbrace{1}_{p_j}\; 0 \ldots\rangle \; . \tag{3.82}$$

- occupation number representation takes into account automatically symmetric exchange of particles (bosonic indistinguishability) as no labeling of particles is present in the state description,
- occupation number basis is an orthonormal basis on the Hilbert space $\mathcal{H}_N$ for each fixed $N = \sum_i n_{p_i}$,
- normalization to 1 is intuitive $\langle n_{p_1} \ldots n_{p_k} \ldots | n'_{p_1} \ldots n'_{p_k} \ldots\rangle = 1$, only when $n' = n$ for all p_i.

The Hilbert space can be combined in the so called *Fock space*

$$\mathcal{H} \;=\; \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2 \ldots \;=\; \bigoplus_{N=0}^{\infty} \mathcal{H}_N \;\equiv\; \sum_{N=0}^{\infty}{}^{\oplus}\, \mathcal{H}_N \; .$$

Here, the Hilbert space $\mathcal{H}_0$ only contains one element, the vacuum $|0\rangle = |0, 0, 0, \ldots 0 \ldots\rangle$.

Note that because $[N, H] = 0$, we can place ourselves in the N-particle sector and stay there. This is, of course, the characteristic property of a free theory. Since, interactions generally *create* and *destroy* particles, they allow to move between different particle sectors of the Fock space.

Exercises: Some useful background from quantum mechanics and classical field theory

Quantum-field-theoretical formulation of many-body quantum mechanics

Let us consider a system of n indistinguishable bosonic particles described by the non-relativistic Schrödinger equation

$$i\hbar\partial_t \psi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, t) \;=\; H\psi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, t) \, ,$$

$$H \;=\; \sum_{p=1}^{n} \left[-\frac{\hbar^2}{2m} \Delta_{\boldsymbol{x}_p} + V(\boldsymbol{x}_p) \right] ,$$

with external (one-body) potential $V(\boldsymbol{x})$. This is know as *n-particle Schródinger equa-*

tion.

The wave-function is assumed to be symmetrized

$$\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,t) \;=\; {}^S\langle \boldsymbol{x}_1,\ldots\boldsymbol{x}_n|\psi(t)\rangle\,,$$

$$|\boldsymbol{x}_1,\ldots\boldsymbol{x}_n\rangle^S \;=\; \frac{1}{n!}\sum_{\pi\in S_n}|\boldsymbol{x}_{\pi(1)},\ldots,\boldsymbol{x}_{\pi(n)}\rangle\,.$$

The field-theoretic description of this system starts with the introduction of (abstract) creation and annihilation operators $\hat{a}^\dagger_{\boldsymbol{x}}$ and $\hat{a}_{\boldsymbol{x}}$ ($\forall\boldsymbol{x}\in\mathbb{R}^3$), and the (abstract) vacuum state $|0\rangle$, with the properties

$$[\hat{a}_{\boldsymbol{x}},\hat{a}_{\boldsymbol{y}}] \;=\; [\hat{a}^\dagger_{\boldsymbol{x}},\hat{a}^\dagger_{\boldsymbol{y}}] \;=\; 0\,,$$

$$[\hat{a}_{\boldsymbol{x}},\hat{a}^\dagger_{\boldsymbol{y}}] \;=\; \delta(\boldsymbol{x}-\boldsymbol{y})\,,\quad \hat{a}_{\boldsymbol{x}}\,|0\rangle \;=\; 0\,,\quad \langle 0|0\rangle \;=\; 1\,.$$

Exercise 3.1 Argue that one may identify

$$|\boldsymbol{x}_1,\ldots\boldsymbol{x}_n\rangle^S \;\equiv\; \frac{1}{\sqrt{n!}}\hat{a}^\dagger_{\boldsymbol{x}_1}\ldots\hat{a}^\dagger_{\boldsymbol{x}_n}\,|0\rangle\,,$$

by showing that

$$\frac{1}{n!}\,\langle 0|\,\hat{a}_{\boldsymbol{x}_n}\ldots\hat{a}_{\boldsymbol{x}_1}\hat{a}^\dagger_{\boldsymbol{y}_1}\ldots\hat{a}^\dagger_{\boldsymbol{y}_n}\,|0\rangle \;=\; {}^S\langle \boldsymbol{x}_1,\ldots\boldsymbol{x}_n|\boldsymbol{y}_1,\ldots\boldsymbol{y}_n\rangle^S\,.$$

(For simplicity consider only cases $n = 1$ and 2.)

If we define the *second-quantized Hamiltonian*

$$\hat{H} \;=\; \int d^3x\,\hat{a}^\dagger_{\boldsymbol{x}}\left[-\frac{\hbar^2}{2m}\Delta_{\boldsymbol{x}} \;+\; V(\boldsymbol{x})\right]\hat{a}_{\boldsymbol{x}}\,,$$

then the *second-quantized Schrödinger equation*

$$i\hbar\partial_t\,|\psi(t)\rangle \;=\; \hat{H}\,|\psi(t)\rangle\,,$$

encapsulates the n-particle Schrödinger equations for all n.

Exercise 3.2 Show that the second-quantized Schrödinger equation reduces to the n-particle Schrödinger equation when multiplied from the left by

$$\frac{1}{\sqrt{n!}}\,\langle 0|\,\hat{a}_{\boldsymbol{x}_n}\ldots\hat{a}_{\boldsymbol{x}_1}\,.$$

Time-dependent quantum field $\hat{\phi}(\boldsymbol{x},t)$ is defined as the Heisenberg picture of the operators $\hat{a}_{\boldsymbol{x}}$:

$$\hat{\phi}(\boldsymbol{x},t) \;=\; e^{\frac{i}{\hbar}t\hat{H}}\,\hat{a}_{\boldsymbol{x}}\,e^{-\frac{i}{\hbar}t\hat{H}}\,,\qquad \hat{\phi}^\dagger(\boldsymbol{x},t) \;=\; e^{\frac{i}{\hbar}t\hat{H}}\,\hat{a}^\dagger_{\boldsymbol{x}}\,e^{-\frac{i}{\hbar}t\hat{H}}\,.$$

Exercise 3.3 What are the equal-time commutation relations of the quantum fields $\hat{\phi}$ and $\hat{\phi}^\dagger$, and what dynamical equation do these fields satisfy?

– Two-body interaction

We now include a two-body interaction between the particles in the form of an interaction Hamiltonian

$$H_{int} \;=\; \frac{1}{2}\sum_{p\neq q}V_2(\boldsymbol{x}_p-\boldsymbol{x}_q)\,.$$

The second-quantized form of the interaction Hamiltonian is

$$\hat{H}_{int} = \frac{1}{2}\int d^3x\, d^3y\, \hat{a}^\dagger_{\boldsymbol{x}} \hat{a}^\dagger_{\boldsymbol{y}} V_2(\boldsymbol{x}-\boldsymbol{y}) \hat{a}_{\boldsymbol{y}} \hat{a}_{\boldsymbol{x}} \,.$$

Show that

$$\frac{1}{\sqrt{n!}} \langle 0|\, \hat{a}_{\boldsymbol{x}_n} \ldots \hat{a}_{\boldsymbol{x}_1} \hat{H}_{int}\, |\psi(t)\rangle = H_{int}\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,t)\,,$$

and hence that the second-quantized Schrödinger equation

$$i\hbar\partial_t\, |\psi(t)\rangle = (\hat{H} + \hat{H}_{int})\, |\psi(t)\rangle \,,$$

describes a non-relativistic quantum-mechanical system of an arbitrary (but fixed) number of pairwise-interacting indistinguishable bosons.

– Fermionic systems

Wave-functions of fermions are anti-symmetric:

$$\begin{aligned}
\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,t) &= {}^A\langle \boldsymbol{x}_1,\ldots \boldsymbol{x}_n\rangle\, \psi(t)\, |\boldsymbol{x}_1,\ldots \boldsymbol{x}_n\rangle^A \,, \\
&= \frac{1}{n!}\sum_{\pi\in\mathcal{S}_n} \mathrm{sgn}(\pi)\, |\boldsymbol{x}_{\pi(1)},\ldots,\boldsymbol{x}_{\pi(n)}\rangle \,.
\end{aligned}$$

They can be represented with a help of fermionic creation and annihilation operators $\hat{b}^\dagger_{\boldsymbol{x}}$ and $\hat{b}_{\boldsymbol{x}}$, which obey the *anti-commutation* rules

$$\{\hat{b}_{\boldsymbol{x}}, \hat{b}_{\boldsymbol{y}}\} = \{\hat{b}^\dagger_{\boldsymbol{x}}, \hat{b}^\dagger_{\boldsymbol{y}}\} = 0\,,$$

$$\{\hat{b}_{\boldsymbol{x}}, \hat{b}^\dagger_{\boldsymbol{y}}\} = \delta(\boldsymbol{x}-\boldsymbol{y})\,, \quad \hat{b}_{\boldsymbol{x}}\, |0\rangle = 0\,, \quad \langle 0|0\rangle = 1\,.$$

One then defines

$$|\boldsymbol{x}_1,\ldots \boldsymbol{x}_n\rangle^A \equiv \frac{1}{\sqrt{n!}} \hat{b}^\dagger_{\boldsymbol{x}_1} \ldots \hat{b}^\dagger_{\boldsymbol{x}_n}\, |0\rangle \,,$$

Note that $(\hat{b}^\dagger_{\boldsymbol{x}})^2 = 0$ agrees with $|\ldots,\boldsymbol{x},\ldots,\boldsymbol{x},\ldots\rangle^A = 0$ which is nothing but the *Pauli exclusion principle*.

The second-quantized Hamiltonian (for non-interacting fermionic systems) is constructed analogously to the bosonic case as:

$$\begin{aligned}
\hat{H} &= \int d^3x\, \hat{b}^\dagger_{\boldsymbol{x}} \left[-\frac{\hbar^2}{2m}\Delta_{\boldsymbol{x}} + V(\boldsymbol{x})\right] \hat{b}_{\boldsymbol{x}} \\
&= \int d^3x\, d^3y\, \hat{b}^\dagger_{\boldsymbol{x}} \hat{b}_{\boldsymbol{y}} \left[-\frac{\hbar^2}{2m}\Delta_{\boldsymbol{x}} + V(\boldsymbol{x})\right] \delta(\boldsymbol{x}-\boldsymbol{y})\,.
\end{aligned}$$

Exercise 3.4 Show that the second-quantized Schrödinger equation

$$i\hbar\partial_t\, |\psi(t)\rangle = \hat{H}\, |\psi(t)\rangle \,,$$

leads to an n-particle (fermionic) Schrödinger equation

$$i\hbar\partial_t\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,t) = H\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,t)\,, \qquad H = \sum_{p=1}^{n}\left[-\frac{\hbar^2}{2m}\Delta_{\boldsymbol{x}_p} + V(\boldsymbol{x}_p)\right] .$$

[**Hint:** Evaluate (and use) the commutator $[\hat{b}_{\boldsymbol{x}_p}, \hat{b}^\dagger_{\boldsymbol{x}} \hat{b}_{\boldsymbol{y}}]$.]