**KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI, GHANA**

# Investigation of Superluminal Motion of Free Spin-half Particles in Spacetime

by

Gazoya, Emmanuel Komi Dokodjo (BSc. MPhil Mathematics)

A thesis submitted to the Department of Mathematics, College of Science
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

AUGUST, 2016

# Certification

I hereby declare that this submission is my own work towards the PhD and that, to the best of my knowledge, it contains no material previously published by another person, nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

Gazoya, Emmanuel Komi Dokodjo, ID: 20361010
(Student)
Sign: ..........................................

Date: ..........................................

Certified by:
Dr. F.T. Oduro
(Supervisor)
Sign: ..........................................

Date: ..........................................

Dr. E. Prempeh
(Co-Supervisor)
Sign: ..........................................

Date: ..........................................

Certified by:
Dr. R.K. Avuglan
(Head of Department)
Sign: ..........................................

Date: ..........................................

*“Dans le domaine de la science, le hasard ne favorise que les esprits qui ont ´et´e pr´epar´es.”*

Louis Pasteur,

*"We must know, we shall know."*

Hilbert,

*"Mavorme morzorla."*

Emmanuel Gazoya

*"Le voyageur de l'infinie."*

Emmanuel Gazoya

*Dedicated to Vincentia Akua, my late mom; John 'the baptist', my late dad; Boye Benoit and Joe Gado, my late brothers; Michel, Gabriel, and Peter, my brothers...*

# *Acknowledgements*

To the Eternal Charioteer; the Leading Star of being; the Sustaining Infinite; the true Light which enlightens every man that comes into the world; He who was in the world and the world was made by Him, but the world knew Him not; the Light that is eternally shining in darkness, and the darkness comprehends it not; He who came unto his own, and his own receive Him not; He, by whom, through whom, and to whom are all things; to the only true God, whose Almighty Word chaos and darkness heard and took their flight, I give the glory!

I express here my sincere and heartfelt gratitude and appreciation to Dr. F.T. Oduro and Dr E. Prempeh, for their supervision in ceaseless dedication, time, sacrifice, advice and suggestions that have molded this work into what it is. I would like to thank Prof. S.K. Amponsah, Head, Dept. of Mathematics-KNUST for support and friendliness, as well as Prof. Banini-Manager, ARC/NNRI (GAEC), Dr. Nuviadenu, and Mr. Srinivasan . Let my family, father, mother, brothers and sister receive a big ”thank you” for spiritual and financial support. Finally, my thanks go to KNUST for hosting the Doctor of Philosophy program in Mathemat-
ics....

# Abbreviations

KG Klein-Gordon

ISF Invariant Spin Field

IRT Invariant Reduction Theorem

CST Cross Section Theorem

DT Decomposition Theorem

NFT Normal Form Theorem

iff if and only if

# Table of Contents


**Certification** iii

**Acknowledgements** vi

**Abbreviations** vii

**List of Tables** xii

**List of Figures** xiii

**Abstract** 1

**1 Introduction 2**

**Introduction** 2

1.1 Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2

1.2 Problem Statement . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.3 Objective of Study . . . . . . . . . . . . . . . . . . . . . . . . . . 5

1.4 Methodology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

1.5 Significance of Study . . . . . . . . . . . . . . . . . . . . . . . . . 6

1.6 Organization of Thesis . . . . . . . . . . . . . . . . . . . . . . . . 7

**2 Spin Transformations and the Rotation Group 8**

2.1 The Three-dimensional Proper Rotation Group . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

2.2 The Euler Angles . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

2.3 The Spinor Map in three Dimensions . . . . . . . . . . . . . . . . . 11

**3 Relativistic Wave Equations 30**

3.1 Special Relativity and Spacetime . . . . . . . . . . . . . . . . . . . 31

3.2 Contravariant and Covariant Vectors . . . . . . . . . . . . . . . . . 32

3.3 Lorentz Transformations . . . . . . . . . . . . . . . . . . . . . . . 40

3.3.1 Improper and Proper Rotations . . . . . . . . . . . . . . . . 46

3.3.2 Group Structure of Lorentz Transformations . . . . . . . . . 48

3.4 Klein-Gordon Equation . . . . . . . . . . . . . . . . . . . . . . . . 49

3.5 The Nonrelativistic Limit ........................ 54

3.6 A Wave Equation for Spin-$\frac{1}{2}$ Particles:
The Dirac Equation ........................... 55

3.7 Interaction with an Electromagnetic Field .............. 65

3.8 Covariant Notation ........................... 66

3.9 Covariance of the Dirac Equation .................. 69

3.10 Bilinear Covariants ........................... 72

**4 Classical Field Theory 78**

4.1 Introduction ............................. 78

4.2 The Hamilton Formalism ........................ 84

4.3 Conservation Laws in Classical Field
Theories ................................ 89

4.3.1 The Symmetrized Energy-Momentum Tensor ......... 102

**5 Free Motion of a Dirac Particle 107**

5.1 Introduction ............................. 107

5.2 Single-Particle Interpretation of the Plane
(Free) Dirac Waves .......................... 127

5.2.1 Lagrange Density and Energy-Momentum Tensor of the Free
Dirac Equation ......................... 134

5.2.2 Energies of the Solutions to the Free Dirac Equation in the
Canonical Formalism ..................... 137

5.2.3 Calculation of the Dirac Hamiltonian and the Velocity Operator ............................. 140

5.2.4 Nonrelativistic Limit of the Dirac Equation ......... 142

5.3 Superluminal Motion of Free Spin-1/2 Particles ........... 147

5.3.1 Evaluation of the Dirac spinor ................. 151

5.3.2 Determination of the expectation of trajectory of the free
spin-1/2 field ......................... 154

5.3.3 Calculation of the Expectation Value of the Overall Relative Linear Velocity Component . . . . . . . . . . . . . . . . . . 155
5.4 Superluminal Motion of Free Spinning Fermions in the Fiber Bundle Formalism . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 160
5.4.1 Homotopy in the Rotation and Lorentz Groups . . . . . . . 163
5.4.2 Particle-spin Motion . . . . . . . . . . . . . . . . . . . . . . 165
5.4.3 Field Motion and Invariant Fields . . . . . . . . . . . . . . . 165
5.4.4 The Set $\Sigma_x[f]$ and Its Invariance . . . . . . . . . . . . . . . 166
5.4.5 The Decomposition Theorem (DT) . . . . . . . . . . . . . . 168
5.4.6 The Invariant Reduction Theorem (IRT) and the Cross Section Theorem (CST) . . . . . . . . . . . . . . . . . . . . . . . 168
5.4.7 The Normal Form Theorem (NFT) . . . . . . . . . . . . . . 169
5.5 Bundle Interpretation of our Special Case of Free Spin-1/2 Particles in Spacetime . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 170

**6 Conclusions and Recommendations** **174**
6.1 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 174
6.2 Recommendations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 175

**References** **176**

# List of Tables

3.1 Bilinear Dirac Covariants . . . . . . . . . . . . . . . . . . . . . . . . 74

# List of Figures

3.1 Light cone: Variations of orbital and spin angular momenta .... 31

5.1 Spectrum of the energies of free Dirac equation. . . . . . . . . . . . 123

5.2 Electrons with positive and negative helicity: spin-up and spin-down126

5.3 Gap between the filled negative-energy states and the positive-energy states. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 129

5.4 Two reference systems: Lorentz boost in the $z$−direction. . . . . . . 155

5.5 Commutation diagram . . . . . . . . . . . . . . . . . . . . . . . . . 162

5.6 Invariant helical motion of free spin-half fields. . . . . . . . . . . . . 171

# Abstract

The possibility of free spin-1/2 particles (also called Dirac particles) superluminal motion in spacetime, is investigated. The universal cover of the entire Lorentz group L consists of $SL(2, \mathrm{C})$ and the spinor map so that to obtain a relativistically invariant description of the state of an electron, one looks to the representations of $SL(2,\mathrm{C})$, that is, to the 2-valued representation of L, known as spinors. We restrict our approach to realistic one-particle systems along with the "positive energy" and utilize the free Dirac waves propagating in the $z$−direction. The Dirac wave function $\psi$ $(x,t)$ is considered as a "classical field" and the corresponding wave equation is derived from a symmetrized Lagrange function. It is observed that variation in spin angular momentum (in the light cone) leads to causality violation, whereas variation in orbital angular momentum does not. Consequently, it is shown that the expectation value of the generalized translational relative velocity component of a free spin-half field indeed exceeds the speed of light. This result is also feasible in a fiber bundle formalism.

# Chapter 1

# Introduction



## 1.1 Background

The universal existential problem of particles capable of superluminal motion has been of great interest at the heart of many controversies and for decades occupied a lot of authors' attention in the field of mathematical physics. Though special relativity is not in conflict with faster-than-light phenomena, many regard as paradoxical the fact that the speed of relativistic particles always averages to the speed of light. Gregory Breit in 1928 was first to recognize this (Breit, 1932). At present (from 2008 to date), scientists are yet to finally decide, in respect of the conflicting conclusions of CERN experiments, on superluminal neutrino result.

When Goudsmit and Uhlenbeck proposed the hypothesis of the spin of the electron, they had in mind a mechanical picture of the electron as a small rigid body rotating about its axis. Such a picture had earlier been considered by Kronig and discarded on the advice of Pauli, Kramers, and Heisenberg, who deemed it a fatal flaw of this picture that the speed of rotation - calculated from the magnitude of the



spin and a plausible estimate of the radius of the electron - was in excess of the speed of light. However, the great success of the spin hypothesis in explaining the

Zeeman effect and the doublet structure of spectral lines quick1y led to its acceptance (Van der Waerden, 1960). Since the naive mechanical picture of spin proved untenable, physicists were left with the concept of spin minus its physical basis. Pauli pontificated that spin is ”an essentially quantum-mechanical property, ...a classically not describable two-valuedness” (Pauli, 1993) and he insisted that the lack of a concrete picture was a satisfactory state of affairs.

Thus physicists gradually came to regard the spin as an abstruse quantum property of the electron, a property not amenable to physical explanation.

Judging from statements found in modern textbooks on atomic physics and quantum theory, one would think our understanding of spin (or the lack thereof) has not made any significant progress since the early years of quantum mechanic (Ohanian, 1986). The spin is usually said to be a nonorbital, ”internal,” ”intrinsic,” or ”inherent” angular momentum (the words are often used interchangeably, although they should not be), and it is often treated as an irreducible entity that cannot be explained further. Sometimes the (unsubstantiated) suggestion is made that the spin is due to an (unspecified) internal structure of the electron” (Dirac, 1981).

## 1.2 Problem Statement

The prevailing acquiescence to this unsatisfactory situation becomes all the more puzzling when one realizes that the means for filling this gap have been at hand since 1939, when Belinfante (Belinfante, 1940) used the symmetrized energymomentum tensor to show that the momentum density gives rise to both orbital angular momentum and spin angular momentum. And sometimes the consolation is offered that the spin arises in a natural way from Dirac's equation or from the analysis of the representations of the Lorentz group. It is true that the Dirac equation contains a wealth of information about spin: the equation tells us the spinor wavefunctions are indeed endowed with a spin angular momentum of ~/2, it supplies the mathematical description of the kinematics of a free electron or other particle of spin one-half, and in conjunction with the principle of minimal coupling, it supplies the equations governing the dynamics of a charged particle immersed in an electromagnetic field, equations which directly yield the correct value of the gyromagnetic ratio for the electron. It is also true that the analysis of the representations of the Lorentz group is very informative: the analysis tells us that the quantum-mechanical wavefunctions must be certain types of tensors or spinors characterized by a value of the mass and (if the mass is not negative) an integer or half-integer value of the spin. But in all of this the spin merely plays (in most of modern textbooks) the role of an extra, nonorbital angular momentum. Clearly, this poor concept of the theory of spin can only lead to paradoxes.

Thus the mathematical formalism of the Dirac equation and of group theory demands the existence of the spin to achieve the conservation of angular momentum and to construct the generators of the rotation group, but fails to give us any understanding of the physical mechanism that leads to causality violation by variation in spin angular momentum (in the light-cone) arising therefrom, as we shall see in a later section.

## 1.3 Objective of Study

In this thesis, we go beyond the usual subluminal and luminal aspects of free fermions (as shown in other literature, where the unsymmetrized Dirac Lagrange density is used) to show these particles are superluminal, employing the symmetrized Lagrangian.

## 1.4 Methodology

It is now a well established fact that there exist, at the subatomic level, physical systems whose state is altered by a rotation of the system through $2\pi$ about some axis, but is returned to its original value by a rotation through $4\pi$. Any of the elementary particles in nature classified as fermions (electrons, protons, neutrons, neutrinos, etc.) possess what is called “half-integer spin” and, as a consequence, their quantum mechanical descriptions (wave functions) behave in precisely this way. We will study the behaviour of quantities such as the wave function of a free spin-1/2 particle, which depend not only on the object's configuration, but also on its motion in localized regions of spacetime. To obtain a high resolution probe, high energies and momenta are requisite. So, the covariant generalization of the Schr¨*o*dinger equation becomes a necessity.

We study the free motion of a Dirac particle in spacetime through the spin-1/2 field theory, and examine the restrictions in which a single-particle interpretation of the Dirac equation and its solutions, holds. Then the single-particle aspect is pursued further by investigating modified one-particle operators and the Ehrenfest's theorems, which state: *In nonrelativistic quantum mechanics there is always a correspondence between a relation of operators and that of classical objects (measurable values)*. At this point, having reached the limit of results review sufficient for our investigation, we utilize the symmetrized Dirac Lagrange density (believed to be fundamental in uncovering hidden information in the dynamics of the spin1/2 field) to derive the expectation values of the relative linear velocity component of a superluminal free fermion.

## 1.5 Significance of Study

No one can deny that the demonstration of particles that can move beyond the speed of light underpins the net advance of science. It is legitimate to anticipate that the energy generated by superluminal free electron and neutrino will be far higher than that of the photon. When experimentation finally masters the control of spin, this new form of higher energy can be turned to good in the context of intergalactic propulsion.

## 1.6 Organization of Thesis

Chapter 2 finds the one-parameter subgroup of the three-dimensional pure rotations generated by spin transformations of the group $SL(2,\mathbb{C})$ by the method of the spinor map. There, we construct the spinor map and show that spin transformations result in $2k\pi$ rotations, with $k$ = 1,2,3,...

In the third chapter, the principles of special relativity and Lorentz transformations, and their application to the construction of special relativistic wave equations for fermions and bosons, the Dirac and Klein-Gordon equations, are reviewed.

The study of one-body problem necessarily passes by the review of classical field theory, where one must reckon with the important statement of Noether's theorem by which *each continuous symmetry transformation leads to a conservation law,* with its applications of Invariance under Translation, Lorentz Invariance, and Internal Symmetries. This review is carried out in Chapter 4.

In Chapter 5, we study the free motion of a fermion in the single-particle interpretation of the plane (free) Dirac waves. Then the symmetrized Dirac Lagrange density is employed to calculate the expectation value of the overall linear velocity component of the spin-half field. Finally, it is shown that this result fits in the context of a fiber bundle formalism.

# Chapter 2

# Spin Transformations and the Rotation Group



The phenomenon of spin and the process of rotation may seem to be identical to common belief. However, when it comes to the exercise of mathematically precisely determining the difference between these two mechanisms, the problem becomes a little challenging. To the best of our knowledge, there does not exist a result which establishes an exact relationship between spin transformations and spatial rotations. In this chapter, we find the one-parameter subgroup of the threedimensional pure rotations generated by spin transformations, using the method of the spinor map, and we show that spin results in $2k\pi$ rotations, with $k$ = 1,2,3,...

## 2.1 The Three-dimensional Proper Rotation Group

**Definition 2.1.** A continuous linear transformation $g : \mathrm{R}^3 \to \mathrm{R}^3$ of the threedimensional Euclidean space which leaves the length $(x^1)^2 + (x^2)^2 + (x^3)^2$ of the coordinate vector $(x^1,x^2,x^3)$ invariant, is called a *three-dimensional rotation*.



The aggregate of all such linear transformations provides a continuous group, which is formed from the set of all *real orthogonal three-dimensional matrices* and is called the *three-dimensional rotation group.* The determinant of every orthogonal matrix is equal to +1, in which case the transformation describes a *pure* (or *proper*) *rotation*, or to −1, in which case it describes a *rotation-reflection*. The set of all pure rotations forms a group, which is a subgroup of the threedimensional rotation group, and is called the *proper rotation group*. Our interest here is in the *three-dimensional proper rotation group*. This group is denoted by *SO* (3).

## 2.2 The Euler Angles

Let $g$ be an element of the group *SO* (3), i.e., a three-dimensional orthogonal matrix with determinant unity. One can then express each such element in terms of a set of three parameters. An example of such parameters is that of Euler angles, which are defined as the three successive angles of rotation describing the transformation from a given Cartesian coordinate system to another by means of three successive rotations performed in a specific sequence (Carmeli, 1977).

The sequence will be started by rotating the original system of axes $\mathbf{x} = (x,y,z) = (x^1,x^2,x^3)$ by an angle $\varphi_1$ clockwise about the $z$ axis. The new coordinate system will be labelled $\xi = (\xi,\eta,\zeta)$. One thus has $\xi = g(\varphi_1)\mathbf{x}$, where

$$g(\varphi_1) = \begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 & 0 \\ \sin\varphi_1 & \cos\varphi_1 & 0 \\ & 0 & \end{pmatrix}. \tag{2.1}$$

$$0 \qquad 1$$

In the second stage the intermediate axes $\xi$ are rotated about the $\xi$ axis clockwise by an angle $\theta$ to another intermediate set which is denoted $\boldsymbol{\xi}' = (\xi', \eta', \zeta')$, and one has $\xi' = g(\theta)\xi$, where

$$g(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}. \tag{2.2}$$

The $\xi'$ axis is called the *line of nodes*. Finally the $\xi'$ axes are rotated clockwise by an angle $\varphi_2$ about the $\zeta'$ axis to produce the desired $\mathbf{x}$ system of axes, $\mathbf{x} = g(\varphi_2)\xi'$, where

$$g(\varphi_2) = \begin{pmatrix} \cos\varphi_2 & -\sin\varphi_2 & 0 \\ \sin\varphi_2 & \cos\varphi_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

The matrix of the complete transformation $\mathbf{x}' = \Pi\mathbf{x} = g(\varphi_2)g(\theta)g(\varphi_1)\mathbf{x}$ is given, therefore, by

$$\Pi = \begin{pmatrix} \cos\varphi_2\cos\varphi_1 - \cos\theta\sin\varphi_1\sin\varphi_2 & -\cos\varphi_2\sin\varphi_1 - \cos\theta\cos\varphi_1\sin\varphi_2 & -\sin\varphi_2\sin\theta \\ \sin\varphi_2\cos\varphi_1 + \cos\theta\sin\varphi_1\cos\varphi_2 & -\sin\varphi_2\sin\varphi_1 + \cos\theta\cos\varphi_1\cos\varphi_2 & -\cos\varphi_2\sin\theta \\ \sin\theta\sin\varphi_1 & \sin\theta\cos\varphi_1 & \cos\theta \end{pmatrix}. \qquad (2.4)$$

The angles $\varphi_1$, $\theta$, $\varphi_2$ are independent parameters, fully determining the rotation $g$.

They are called *Euler Angles*. By their definition, one has $0 \leq \varphi_1 \leq 2\pi$, $0 \leq \theta \leq \pi$, and

$0 \leq \varphi_2 \leq 2\pi$.

## 2.3 The Spinor Map in three Dimensions

We begin by establishing some notations.

**Definition 2.2.** $\mathbb{C}^{2\times 2}$ denotes the set of all $2 \times 2$ matrices

$$A = [a_{ij}] = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad (2.5)$$

with complex entries.

**Definition 2.3.** The special linear group of order two $SL(2,\mathbb{C})$ consists of all $2 \times 2$ complex matrices of determinant unity.

**Definition 2.4.** The group $SO\,(1,3)$ also called the special Lorentz group comprehends all proper rotations in 4-dimensional Lorentzian space $\mathbb{R}^{1,3}$ (one-time,

3-space) represented by $4 \times 4$ matrices of the form

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & [R] & \\ 0 & & & \end{pmatrix}$$

where $R \in SO(3)$, and is a subgroup of the Lorentz group.

In analogy to Section 1.7 of (Naber, 1992) where it is shown there exists a certain homomorphism (called the "spinor map") from the group $SL(2,\mathbb{C})$ onto the Lorentz group, we investigate the relationship between spin and spatial rotation, in three-dimensional real space. To do this, it suffices us, by the definition of the special Lorentz group, to restrict this homomorphism to three dimensions and employ it to determine the connection between elements of $SL(2,\mathbb{C})$ and the group $SO(3)$. This uncovers the remarkable link by which spin transformations result in $2k\pi$ rotations.

Using an overbar to designate complex conjugation, the conjugate transpose $A^{CT}$ of $A$ is defined by

$$A^{CT} = \begin{pmatrix} \bar{a}_{11} & \bar{a}_{21} \\ \bar{a}_{12} & \bar{a}_{22} \end{pmatrix}. \tag{2.6}$$

A matrix $H$ in $\mathbb{C}^{2\times 2}$ is said to be Hermitian if $H^{CT} = H$ and we denote by $\mathbb{H}_2$ the set of all Hermitian matrices in $\mathbb{C}^{2\times 2}$.

**Proposition 2.5.** *In three dimensions, any Hermitian matrix $H$ in $\mathbb{C}^{2\times 2}$ is uniquely expressible in the form*

$$H = \begin{pmatrix} x^3 & x^1 + ix^2 \\ x^1 - ix^2 & -x^3 \end{pmatrix} \tag{2.7}$$

*where $x^1$, $x^2$, $x^3$ are real.*

*Proof.* Let $H = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ be Hermitian in three dimensions. By the matrix equation $H = H^{CT}$, we have that $\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & -\bar{a} \end{pmatrix}$, and so $a = \bar{a}$,

$$\begin{pmatrix} c & -a \end{pmatrix} \begin{pmatrix} \bar{b} & -\bar{a} \end{pmatrix}$$

$b = \bar{c}$, and $c = \bar{b}$. It is obvious that $a$ is real, and $b$ and $c$ are complex conjugates.

Hence the matrix $H$ in (2.7) satisfies the definition of Hermitian matrix. Next, observe that since the terms of the entries of $H$ are linear, it follows that the expression of $H$ in the form (2.7) is unique. □

In addition, we have

$$\begin{aligned} H &= \begin{pmatrix} x^3 & 0 \\ 0 & -x^3 \end{pmatrix} + \begin{pmatrix} 0 & x^1 + ix^2 \\ x^1 - ix^2 & 0 \end{pmatrix} \\ &= x_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} + x_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3, \end{aligned} \tag{2.8}$$

where $\sigma_i$, $i = 1,2,3$, are the *Paulispinmatrices*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.9}$$

and the first term in the first line of the expansion of $H$ in (2.8) [given by $\begin{pmatrix} x_3 & 0 \\ 0 & -x_3 \end{pmatrix} = x^3\sigma_3$] represents sp

Thus, from Proposition 2.5, given any vector $\mathbf{x} = (x^1, x^2, x^3)$ in three dimensions, one can associate the Hermitian matrix

$$\begin{pmatrix} x_3 & x_1 + ix_2 \end{pmatrix}$$

$$H = \begin{pmatrix} & \\ x_1 - ix_2 & -x_3 \end{pmatrix}.$$

Elements of $SL(2,C)$ are often called *spin transformations*. Each matrix $A$ in $SL(2,C)$ gives rise to a mapping $M_A : H_2 \to H_2$ defined by

$$M_A(H) = AHA^{CT}$$

for every $H$ in $H_2$ [$M_A(H)$ is in $H_2$ since $\left(AHA^{CT}\right)^{CT} = \left(A^{CT}\right)^{CT} \cdot (AH)^{CT} = AH^{CT}A^{CT}$].

Moreover, $\det M_A(H) = \det\left(AHA^{CT}\right) = (\det A)(\det H)\left(\det A^{CT}\right) = \det H.$

But $M_A(H)$ can be uniquely written in the form

$$M_A(H) = \begin{pmatrix} x_{03} & x_{01} + ix_{02} \\ x_{01} - ix_{02} & -x_{03} \end{pmatrix} \tag{2.10}$$

some real numbers $x_0^a$, $a = 1,2,3$. Computing the determinants in (2.7) and for (2.10) therefore gives

$$-[(x^1)^2+(x^2)^2+(x^3)^2] = -[\left(x'^1\right)^2+\left(x'^2\right)^2+\left(x'^3\right)^2] = -(length\,of\,x)^2. \tag{2.11}$$

Therefore, the mapping $[x^a] \to [x'^a]$ defined by

$$x_{03} \quad x_{01} + ix_{02} \qquad x_3 \quad x_1 + ix_2 \qquad {}_{CF}$$

$$= A \qquad A, \tag{2.12}$$

$$x_{01} - ix_{02} \qquad -x_{03} \qquad x_1 - ix_2 \qquad -x_3$$

which is clearly linear, preserves the quadratic form $\eta x^a x^b$. It is known, in four dimensions (1-time, 3-space), that the matrix of this map is a general, homogeneous

Lorentz transformation, which includes the subgroup of rotational transformations $SO\ (3)$. For our purpose, we intend to construct (in three dimensions) this matrix explicitly from the entries of

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

Letting $h_1 = x^3$, $h_2 = x^1 + ix^2$, and $h_3 = x^1 - ix^2$, we have

$$\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & i & 0 \\ 1 & -i & 0 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

which we will write more compactly as

$$[h_i] = G\left[x^i\right]$$

and similarly for $[h'_i]$. The inverse of $G$ is obtained as

$$G^{-1} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{-i}{2} & \frac{i}{2} \\ 1 & 0 & 0 \end{pmatrix}.$$

Next, notice that

$$\begin{aligned}
AHA^{CT} &= \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} h_1 & h_2 \\ h_3 & -h_1 \end{pmatrix} \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix} \\
&= \begin{pmatrix} \alpha h_1 + \beta h_3 & \alpha h_2 - \beta h_1 \\ \gamma h_1 + \delta h_3 & \gamma h_2 - \delta h_1 \end{pmatrix} \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix} \\
&= \begin{pmatrix} \left(\alpha\bar{\alpha} - \beta\bar{\beta}\right) h_1 + \alpha\bar{\beta} h_2 + \beta\bar{\alpha} h_3 & \left(\alpha\bar{\gamma} - \beta\bar{\delta}\right) h_1 + \alpha\bar{\delta} h_2 + \beta\bar{\gamma} h_3 \\ \left(\gamma\bar{\alpha} - \delta\bar{\beta}\right) h_1 + \gamma\bar{\beta} h_2 + \delta\bar{\alpha} h_3 & \left(\gamma\bar{\gamma} - \delta\bar{\delta}\right) h_1 + \gamma\bar{\delta} h_2 + \delta\bar{\gamma} h_3 \end{pmatrix} \\
&= \begin{pmatrix} h_1' & h_2' \\ h_3' & -h_1' \end{pmatrix} \\
&= M_A\left(H\right),
\end{aligned}$$

which is equivalent to

$$h_1 \qquad \alpha\bar{\alpha} - \beta\bar{\beta} \quad \alpha\bar{\beta} \quad \alpha\bar{\delta} \quad \gamma\bar{\beta} \qquad \beta\bar{\alpha} \quad h_1 \tag{2.13}$$

$_0 = \alpha\bar{\gamma} - \beta\bar{\delta}$ $\quad \beta\bar{\gamma}$ $h_2$;

$h_2$

$_0$ $\gamma\bar{\alpha} - \delta\bar{\beta}$ $h_3$ $\quad \delta\bar{\alpha}$ $\quad h_3$

observe that since $AHA^{CT}$ is a three dimensional Hermitian matrix, its first and fourth entries differ only in a sign. So we have that $(\gamma\bar{\gamma} - \delta\bar{\delta})\, h_1 + \gamma\bar{\delta}h_2 + \delta\bar{\gamma}h_3 = -\left[(\alpha\bar{\alpha} - \beta\bar{\beta})\, h_1 + \alpha\bar{\beta}h_2 + \beta\bar{\alpha}h_3\right]$. That is why we ignore the row $(\gamma\bar{\gamma} - \delta\bar{\delta},\, \gamma\bar{\delta}, \delta\bar{\gamma})$ in the expression of the matrix in (2.13). We write this equation more concisely as

$$\left[h_i^{'}\right] = R_A\left[h_i\right]. \tag{2.14}$$

Consequently, the map $[x^a] \to [x^{'a}]$ defined by (2.12) is given by

$$[x^a] \xrightarrow[G]{} [h_i] \xrightarrow[R_A]{} \left[h_i^{'}\right] \xrightarrow[G^{-1}]{} \left[x^{'a}\right] \tag{2.15}$$

and the transformation $T_A$ determined via (2.12) or (2.15) by $A$ is

$$T_A = G^{-1}R_AG. \tag{2.16}$$

Calculating the matrix product $G^{-1}R_AG$ explicitly, we obtain its entries as

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{-i}{2} & \frac{i}{2} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha\bar{\alpha} - \beta\bar{\beta} & \alpha\bar{\beta} & \beta\bar{\alpha} \\ \alpha\bar{\gamma} - \beta\bar{\delta} & \alpha\bar{\delta} & \beta\bar{\gamma} \\ \gamma\bar{\alpha} - \delta\bar{\beta} & \gamma\bar{\beta} & \delta\bar{\alpha} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & i & 0 \\ 1 & -i & 0 \end{pmatrix},$$

that is

$$T_{11} = \frac{1}{2}\left(\alpha\bar{\delta} + \gamma\bar{\beta} + \beta\bar{\gamma} + \delta\bar{\alpha}\right) \qquad T_{12} = \frac{i}{2}\left(\alpha\bar{\delta} + \gamma\bar{\beta} - \beta\bar{\gamma} - \delta\bar{\alpha}\right)$$

$$T_{13} = \frac{1}{2}\left(\alpha\bar{\gamma} - \beta\bar{\delta} + \gamma\bar{\alpha} - \delta\bar{\beta}\right) \qquad T_{21} = \frac{i}{2}\left(-\alpha\bar{\delta} + \gamma\bar{\beta} - \beta\bar{\gamma} + \delta\bar{\alpha}\right)$$
$$T_{22} = \frac{1}{2}\left(\alpha\bar{\delta} - \gamma\bar{\beta} - \beta\bar{\gamma} + \delta\bar{\alpha}\right) \qquad T_{23} = \frac{i}{2}\left(-\alpha\bar{\gamma} + \beta\bar{\delta} + \gamma\bar{\alpha} - \delta\bar{\beta}\right)$$

$$T_{31} = \alpha\bar{\beta} + \beta\bar{\alpha} \qquad T_{32} = i\left(\alpha\bar{\beta} - \beta\bar{\alpha}\right) \qquad T_{33} = \alpha\bar{\alpha} - \beta\bar{\beta}. \tag{2.17}$$

By direct calculation one shows that

$$\det T_A = \det\left(G^{-1}R_AG\right) = \left(\det G^{-1}\right)\left(\det R_A\right)\left(\det G\right), \text{ i.e.,}$$

$$\det T_A = \det R_A = \alpha\bar{\alpha} + \beta\bar{\beta} > 0.$$

With $\det T_A$ being positive, we deduce that $T_A$ is a transformation which preserves orientation. Observe that if $A$ and $B$ are both in $SL(2,\mathbb{C})$, then

$$T_AT_B = \left(G^{-1}R_AG\right)\left(G^{-1}R_BG\right) = G^{-1}\left(R_AR_B\right)G. \tag{2.18}$$

Now, since

$$M_{AB}(H) = (AB)H\,(AB)^{CT}$$

$$= ABHB^{CT}A^{CT} = A\left(BHB^{CT}\right)A^{CT} = M_A\left(BHB^{CT}\right)$$

$$= M_A\left(M_B\left(H\right)\right) = M_A \circ M_B\left(H\right),$$

it follows that $M_{AB}=M_A \circ M_B$ and so $R_{AB} = R_A R_B$. Thus (2.18) gives

$T_A T_B=G^{-1}R_{AB}G$ and so



$$T_A T_B = T_{AB}. \tag{2.19}$$

Thus, the map $T_A$ preserves matrix multiplication, that is, it is a group homomorphism of $SL(2,C)$ to $SO(3)$. It is not one-to-one since it is clear from (2.17) that both $A$ and $-A$ have the same image in $SO(3)$. We claim that it is precisely two-to-one, i,e., if $A$ and $B$ are in $SL(2,C)$ and $T_A = T_B$, then $A = \pm B$. To see this note that $AB^{-1}$ is in $SL(2,C)$ and, since the spinor map is a homomorphism, $T_{AB^{-1}} = T_A T_{B^{-1}} = T_A (T_B)^{-1} = T_A (T_A)^{-1}=$ identity matrix.

At this point, it is important to notice that all the entries $T_{ij}$, $i, j = 1, 2,3$, of the matrix $T_A$ are real. To see this, we rewrite each $T_{ij}$ in (2.17) as follows

$$T_{11} = \frac{1}{2}\left[\left(\alpha\bar{\delta} + \overline{\alpha\bar{\delta}}\right) + \left(\beta\bar{\gamma} + \overline{\beta\bar{\gamma}}\right)\right] = \frac{1}{2} \times real\,part = real$$

$$T_{12} = \frac{i}{2}\left[\left(\alpha\bar{\delta} - \overline{\alpha\bar{\delta}}\right) + \left(\gamma\bar{\beta} - \overline{\gamma\bar{\beta}}\right)\right] = \frac{i}{2} \times imaginary\,part = real,$$

similarly for the remaining entries $T_{13}$, $T_{21}$, $T_{22}$, $T_{23}$, etc.

We deduce that $T_A$ is a real matrix, and hence a matrix of a transformation of the real three-dimensional space.

An element $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ of $SL(2,\mathbb{C})$ is said to be unitary if $A^{-1} = A^{CT}$, that is, if

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix} = \begin{pmatrix} \alpha\bar{\alpha} + \beta\bar{\beta} & \alpha\bar{\gamma} + \beta\bar{\delta} \\ \alpha\bar{\gamma} + \beta\bar{\delta} & \gamma\bar{\gamma} + \delta\bar{\delta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.20}$$

The set of all such matrices is denoted $SU(2)$ and is a subgroup of $SL(2,\mathbb{C})$, i.e., $SU(2)$ is also closed under the formation of products and inverses.

Notice that if $A$ is in $SU(2)$, then, by (2.20) $\det T_A = \alpha\bar{\alpha} + \beta\bar{\beta} = 1$. Moreover, by the construction of the three-dimensional Hermitian matrix $AHA^{CT}$ (from above) and by (2.20) we obtain the following three sets of equations (I), (II), and (III):

(I):

$$\alpha\bar{\alpha} - \beta\bar{\beta} = \delta\bar{\delta} - \gamma\bar{\gamma}$$

$$\alpha\bar{\beta} + \gamma\bar{\delta} = 0$$

$$\bar{\alpha}\beta + \bar{\gamma}\delta = 0,$$

(II):

$$\alpha\bar{\gamma} - \beta\bar{\delta} = \bar{\alpha}\gamma - \bar{\beta}\delta$$

$$\alpha\bar{\delta} + \beta\bar{\gamma} = \bar{\alpha}\delta + \bar{\beta}\gamma$$

$$\alpha\bar{\delta} - \beta\bar{\gamma} = \bar{\alpha}\delta - \bar{\beta}\gamma,$$

(III):

$$\alpha\bar{\alpha} + \beta\bar{\beta} = \gamma\bar{\gamma} + \delta\bar{\delta} = 1$$

$$\alpha\bar{\gamma} + \beta\bar{\delta} = \bar{\alpha}\gamma + \bar{\beta}\delta = 0.$$

From the last two equations in (II), we obtain $\alpha\bar{\delta} = \bar{\alpha}\delta = realnumber$, since $\alpha\bar{\delta}$ and $\bar{\alpha}\delta$ are conjugate to each other; similarly, $\beta\bar{\gamma} = \bar{\beta}\gamma = real$. Also, since $\alpha,\beta,\gamma,\delta$ and their conjugates are scalars, it follows that $\alpha\bar{\beta} = \bar{\alpha}\beta =$ real, $\gamma\bar{\delta} = \bar{\gamma}\delta =$real, $\alpha\bar{\gamma} = \bar{\alpha}\gamma$ =real, and $\beta\bar{\delta} = \bar{\beta}\delta$ =real. By these equations, the $3 \times 3$ real matrix $T_A$ simplifies to

$$T_A = \begin{pmatrix} \alpha\bar{\delta} + \beta\bar{\gamma} & 0 & \alpha\bar{\gamma} - \beta\bar{\delta} \\ 0 & \alpha\bar{\delta} - \beta\bar{\gamma} & 0 \\ 2\alpha\bar{\beta} & 0 & \alpha\bar{\alpha} - \beta\bar{\beta} \end{pmatrix}. \tag{2.21}$$

In order to analyze the elements of this matrix, we now introduce some definitions and properties of orthogonal matrices.

**Definition 2.6.** For an $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

the vectors

$$\begin{aligned} \mathbf{r}_1 &= (a_{11}, a_{12}, \ldots, a_{1n}) \\ \mathbf{r}_2 &= (a_{21}, a_{22}, \ldots, a_{2n}) \\ \vdots \quad & \quad \vdots \\ \mathbf{r}_m &= (a_{m1}, a_{m2}, \ldots, a_{mn}) \end{aligned}$$

formed from the rows of $A$ are called the *row vectors* of $A$, and the vectors

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \end{pmatrix}, \quad \ldots, \quad \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \end{pmatrix}$$

$a_{m1}$ $a_{m2}$ $a_{mn}$

formed from the columns of $A$ are called the *column vectors* of $A$. The subspace of $\mathbb{R}^n$ spanned by the row vectors is called the *row space* of $A$, and the subspace of $\mathbb{R}^m$ spanned by the column vectors is called the *column space* of $A$.

Matrices whose inverses can be obtained by transposition are sufficiently important that there is some terminology associated with them.

**Definition 2.7.** A square matrix with the property

$$A^{-1} = A^T$$

is said to be an *orthogonal matrix*, where the superscript $T$ stands for "transpose". To paraphrase this definition, a square matrix $A$ is orthogonal if and only if

$$AA^T = A^T A = I,$$

where $I$ is identity matrix. The following result makes it easy to determine when an $n \times n$ matrix $A$ is orthogonal.

**Theorem 2.8.** *The following are equivalent:*

*(a) $A$ is orthogonal.*

*(b) The row vectors of $A$ form an orthonormal set in $\mathbb{R}^n$ with the Euclidean inner*

*product.*

*(c) The column vectors of A form an orthonormal set in* $\mathbb{R}^n$ *with the Euclidean inner product.*

*Proof.* First, we show that if the square matrix $A$ is an orthogonal matrix, then its transpose $A^T$ is also orthogonal. Let $A$ be orthogonal, then

$$AA^T = A^T A = I.$$

But we have $AA^T = \left(A^T\right)^T A^T$ and $A^T A = A^T \left(A^T\right)^T$, which shows that

$$\left(A^T\right)^T A^T = A^T \left(A^T\right)^T = I.$$

Hence $A^T$ is also orthogonal. Next, explicitly, let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad so \quad A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Observe that the row vectors of $A$, which are given by

$$\mathbf{r}_1 = (a_{11}, a_{12}, \ldots, a_{1n}),\ \mathbf{r}_2 = (a_{21}, a_{22}, \ldots, a_{2n}),\ \ldots,\ \mathbf{r}_n = (a_{n1}, a_{n2}, \ldots, a_{nn})$$

turn to be the column vectors of $A^T$, and the column vectors of $A$, given by

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad \ldots, \quad \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}$$

become the row vectors of $A^T$. From the matrix equation $AA^T = A^TA = I$, we have

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Computing this in the terminology of the Euclidean inner product, we obtain

$$\begin{pmatrix} (\mathbf{r}_1)^2 & (\mathbf{r}_1\cdot\mathbf{r}_2) & \cdots & (\mathbf{r}_1\cdot\mathbf{r}_n) \\ (\mathbf{r}_2\cdot\mathbf{r}_1) & (\mathbf{r}_2)^2 & \cdots & (\mathbf{r}_2\cdot\mathbf{r}_n) \\ (\mathbf{r}_3\cdot\mathbf{r}_1) & (\mathbf{r}_3\cdot\mathbf{r}_2) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & (\mathbf{r}_n)^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

that is

$$\begin{aligned} \|\mathbf{r}_i\| &= 1, \quad i = 1,2,\ldots,n \\ \mathbf{r}_i\cdot\mathbf{r}_j &= 0, \quad i \neq j, \end{aligned}$$

which implies the row vectors, as well as the column vectors, of $A$ form an orthonormal set in $\mathbb{R}^n$ with the Euclidean inner product, as required. So, we have that (a)⇒(b) and (c). This completes the proof of the theorem. □

Now the row vectors of the matrix $T_A$ are

$$\boldsymbol{v}_1 = \left(\alpha\bar{\delta}+\beta\bar{\gamma},\, 0,\, \alpha\bar{\gamma}-\beta\bar{\delta}\right),\ \boldsymbol{v}_2 = \left(0,\, \alpha\bar{\delta}-\gamma\bar{\beta},\, 0\right),\ \boldsymbol{v}_3 = \left(2\alpha\bar{\beta},\, 0,\, \alpha\bar{\alpha}-\beta\bar{\beta}\right). \tag{2.22}$$

A necessary and sufficient condition for $T_A$ to be an orthogonal matrix is to have its row (or column) vectors to form an orthonormal set in $\mathbb{R}^3$ with the Euclidean inner product. A necessary condition for this to happen is to make $\alpha\bar{\delta} - \gamma\bar{\beta} = \pm 1$ so that $v_2$ is unimodular. Let $\alpha\bar{\delta} - \gamma\bar{\beta} = 1$, i.e.,

$$T_A = \begin{pmatrix} \alpha\bar{\delta} + \beta\bar{\gamma} & 0 & \alpha\bar{\gamma} - \beta\bar{\delta} \\ 0 & 1 & 0 \\ 2\alpha\bar{\beta} & 0 & \alpha\bar{\alpha} - \beta\bar{\beta} \end{pmatrix}, \qquad (2.23)$$

so that the middle entry of our matrix $T_A$ be $T_{22} = 1$, always. Then one can check and find that

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = \left(\alpha\bar{\delta} + \beta\bar{\gamma}\right) \cdot \left(2\alpha\bar{\beta}\right) + \left(\alpha\bar{\gamma} - \beta\bar{\delta}\right) \cdot \left(\alpha\bar{\alpha} - \beta\bar{\beta}\right) = \alpha\bar{\gamma} + \beta\bar{\delta} = 0,$$

by (III). Readily, we have $v_1 \cdot v_2 = 0$ and $v_2 \cdot v_3 = 0$. Therefore the vectors $v_1$, $v_2$, and $v_3$ form an orthogonal set. One can also check and find that

$$||\boldsymbol{v}_1||^2 = \left(\alpha\bar{\delta} + \beta\bar{\gamma}\right)^2 + \left(\alpha\bar{\gamma} - \beta\bar{\delta}\right)^2 = \alpha\bar{\delta} - \beta\bar{\gamma} = 1, \qquad (2.24)$$

$$||\boldsymbol{v}_3||^2 = \left(2\alpha\bar{\beta}\right)^2 + \left(\alpha\bar{\alpha} - \beta\bar{\beta}\right)^2 = \left(\alpha\bar{\alpha} + \beta\bar{\beta}\right)^2 = 1. \qquad (2.25)$$

So the three vectors $v_1$, $v_2$, and $v_3$ are unimodular. We deduce that the row vectors of $T_A$ form an orthonormal set with the Euclidean inner product, and so $T_A$ is an orthogonal matrix, under the restriction $\alpha\bar{\delta} - \gamma\bar{\beta} = 1$. Moreover $\det T_A = 1$, so that

$T_A$ is proper. It follows that the transformation represented by the matrix $T_A$ is a rotation of the 3-dimensional real space. Observe that the case $\alpha\bar{\delta} - \gamma\bar{\beta} = -1$ corresponds to $T_A$ being a rotation-reflection.

Thus the map $A \rightarrow T_A$ carries the subgroup $SU(2)$ of the group $SL(2,\mathbb{C})$ into the proper (or pure) rotation group $SO$ (3) This map is called the *spinor map* in three dimensions (Naber, 1992).

Now, consider rotation around and along the coordinate axes and let $\alpha\bar{\delta} + \beta\bar{\gamma} = 1$ in addition to the condition $\alpha\bar{\delta} - \beta\bar{\gamma} = 1$. It follows that $\beta\bar{\gamma} = 0$, i.e., $\beta = \gamma = 0$, from the set of equations (I), (II) and (III). This leads to the one-parameter restricted subgroup of spin transformations

$$A(\alpha) = \begin{pmatrix} \alpha & 0 \\ 0 & \frac{1}{\alpha} \end{pmatrix}$$

where $\alpha \neq 0$; equivalently

$$A(\psi) = \begin{pmatrix} \exp(i\psi/2) & 0 \\ 0 & \exp(-i\psi/2) \end{pmatrix}.$$

Writing out the action $X \longmapsto AXA^{CT}$ and calculating terms, we find that our spinor map takes this into

$$a_1(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{pmatrix}, \tag{2.26}$$

$$a_1(\psi) = \begin{pmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{pmatrix}, \tag{2.27}$$

and corresponding to rotations around and along the coordinate

$$a_3(\psi) = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.28}$$

axes $Ox^1$, $Ox^2$, and $Ox^3$, respectively, using the results of our earlier discussions on Euler angles.

Clearly, for any of the infinitesimal matrices $a_1(\psi)$, $a_2(\psi)$ and $a_3(\psi)$ to satisfy (2.21) it should be, for example, of the form

$$a_1 = \begin{pmatrix} \cos 2k\pi & -\sin 2k\pi & 0 \\ \sin 2k\pi & \cos 2k\pi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (2.29)$$

where $k = 1,2,3,\ldots$, for the case of rotation about the $Ox^3$ axis, and similarly for the other axes.

Thus, the transformations generated by the unitary subgroup $SU(2)$ in threedimensional real space are precisely rotations of angles which are multiples of $2\pi$. This result corroborates the physical evidence that a spinning mechanism generates a rotation of an angle of at least $2\pi$.

# Chapter 3

# Relativistic Wave Equations

The phenomena that we will be investigating take place over localized regions of spacetime. To probe spacetime on high resolution requires high energies and momenta. Therefore, we require covariant generalizations of the Schr¨*o*dinger equation

(Guildry, 1991). In this chapter we review the principles of special relativity and Lorentz transformations, and apply these to the construction of special relativistic wave equations for fermions and bosons, the Dirac and Klein-Gordon equations. These will be interpreted as single-particle wave equations. Only then will we have a mathematical apparatus potent enough to handle the theories that are of interest to us.

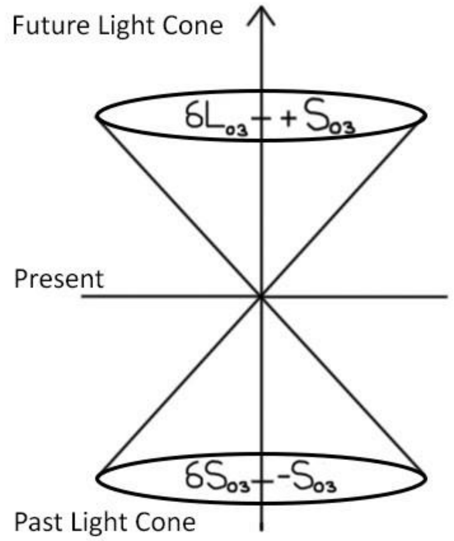


Figure 3.1: Light cone: Variations of orbital and spin angular momenta

## 3.1 Special Relativity and Spacetime

The 4-dimensional spacetime ($ct,x,y,z$) in which we find ourselves is called a *Minkowski space*, and a point in such a space is termed an *event*. The equation

$$s^2 = c^2t^2 - x^2 - y^2 - z^2 = 0 \tag{3.1}$$

defines the *light cone* (Figure 3.1). Events that have $s^2 > 0$ are called *timelike*; events that have $s^2 < 0$ are called *spacelike*; events on the light cone ($s^2 = 0$) are connected by signals propagating at the speed of light from the origin and are called *lightlike*. Timelike events can be causally connected to the origin with signals for which $v < c$, but spacelike events cannot be connected to the origin except by signals for which $v > c$.

The special theory of relativity is based on two postulates: (1) the velocity of light in a vacuum is constant in all inertial systems (an inertial system is a system of reference in which Newton's first law of motion holds), and (2) the laws of physics are invariant under transformations between inertial systems (*covariance*). The transformations that respect these postulates are called *Lorentz transformations*. Before discussing them we take a short notational detour.

## 3.2 Contravariant and Covariant Vectors

It is customary to introduce added notation and nomenclature that are not actually required in special relativity, but that considerably simplify general relativity. Even though we will normally require only fidelity to special relativity in our equations, we will maintain the usual practice and introduce the distinction between covariant and contravariant vectors. We will most commonly work in the previously introduced pseudo-Euclidean vector space (Minkowski space) for which the spacetime continuum is associated with the coordinates $(t,x,y,z) \equiv (x^0,x^1,x^2,x^3)$, where ~ = $c$ = 1 units have been used . It will be assumed that there exists some transformation law that yields new coordinates $(x'^0, x'^1, x'^2, x'^3)$:

$$x'^\mu = x'^\mu\left(x^0, x^1, x^2, x^3\right) \qquad \mu = 0,1,2,3, \qquad (3.2)$$

and that the transformation may be inverted. We may define tensors of rank $k$ associated with the spacetime point $x = (x^0,x^1,x^2,x^3)$ by their properties under the transformation $x \to x^0$.

A scalar is unchanged by the transformation. Two tensors of rank 1, that is *vectors*, may be distinguished. A *contravariant vector* $A = (A^0, A^1, A^2, A^3)$ transforms as

$$A'^{\mu} = \sum_{\mu} \frac{\partial x'^{\mu}}{\partial x^{\nu}} A^{\nu}. \tag{3.3}$$

Here, the Einstein summation convention is introduced: an index appearing twice on one side of the equation (once as superscript, once as subscript) implies a summation over that index. We will usually reserve Greek letters for indices ranging over 0,1,2. and 3, but Latin letters (*i,j,k,...*) for indices ranging over only 1,2, and 3. Explicitly then, for a contravariant vector

$$A'^{\mu} = \frac{\partial x'^{\mu}}{\partial x^{0}} A^{0} + \frac{\partial x'^{\mu}}{\partial x^{1}} A^{1} + \frac{\partial x'^{\mu}}{\partial x^{2}} A^{2} + \frac{\partial x'^{\mu}}{\partial x^{3}} A^{3}. \tag{3.4}$$

Likewise a covariant vector $A_\mu = (A_0, A_1, A_2, A_3)$ obeys the transformation

$$A'_{\mu} = \frac{\partial x^{\nu}}{\partial x'^{\mu}} A_{\nu}. \tag{3.5}$$

Three kinds of rank 2 tensors may be defined. A *contravariant tensor of rank 2* is composed of 16 quantities transforming as

$$F'^{\mu\nu} = \frac{\partial x'^{\mu}}{\partial x^{\gamma}} \frac{\partial x'^{\nu}}{\partial x^{\delta}} F^{\gamma\delta}, \tag{3.6}$$

while the corresponding *covariant tensor of rank 2* transforms as

$$F'_{\mu\nu} = \frac{\partial x^{\gamma}}{\partial x'^{\mu}} \frac{\partial x^{\delta}}{\partial x'^{\nu}} F_{\gamma\delta}, \tag{3.7}$$

and the *mixed second rank tensor* as

$$F'^{\mu}_{\nu} = \frac{\partial x'^{\mu}}{\partial x^{\gamma}} \frac{\partial x^{\delta}}{\partial x'^{\nu}} F^{\gamma}_{\delta}. \tag{3.8}$$

In an analogous fashion, higher order tensors may be defined.

Those definitions of covariant, contravariant, and mixed tensors are required in general relativity where the transformations may be spacetime dependent. For the flat spacetime of special relativity, the transformations between inertial systems are linear and independent of spacetime coordinates. The transformation of a contravariant coordinate vector $x^{\mu}$ is then given by

$$x'^{\mu} = \alpha_{\nu}^{\mu} x^{\nu}, \tag{3.9}$$

with $\alpha^{\mu}_{\nu}$ independent of the coordinates, so

$$\frac{\partial x'^{\mu}}{\partial x^{\nu}} = \alpha^{\mu}_{\nu}. \tag{3.10}$$

Therefore, the derivatives appearing in the previous definitions for various kinds of tensors are just constants, and we may simply define for flat spacetime the transformation laws

*Scalars*

$$A' = A, \tag{3.11}$$

*Contravariant vectors*

$$A'^{\mu} = \alpha_{\nu}^{\mu} A^{\nu}, \tag{3.12}$$

*Covariant vectors*

$$A'_{\mu} = \alpha^{\nu}_{\mu} A_{\nu}, \tag{3.13}$$

*Contravariant tensor of rank 2*

$$F_{0\mu\nu} = \alpha_{\gamma\mu}\alpha_{\delta\nu}F_{\gamma\delta}, \tag{3.14}$$

*Covariant tensor of rank 2*

$$F'_{\mu\nu} = \alpha^{\gamma}_{\mu}\alpha^{\delta}_{\nu}F_{\gamma\delta}, \tag{3.15}$$

*Mixed tensor of rank 2*

$$F'^{\mu}_{\nu} = \alpha^{\mu}_{\gamma}\alpha^{\delta}_{\nu}F^{\gamma}_{\delta}, \tag{3.16}$$

and so for higher order tensors. The coefficients $\alpha^{\nu}_{\mu}$ are elements of a Lorentz transformation and are discussed below.

The scalar product of vectors is defined by

$$A \cdot B = A^{\mu}B_{\mu}. \tag{3.17}$$

As the name implies, (3.17) defines an object that is an invariant. In analogous way, the *contractions* of higher order tensors may be defined by employing a repeated index (one upper, one lower). For example,

$$A^{\nu}_{\mu} = B^{\nu\lambda}C_{\lambda\mu}. \tag{3.18}$$

In a contraction the implied summation over any repeated index removes it as an index. The vector scalar product is just a special case of this, and tensor contraction is a generalization of taking the vector scalar product. The differential length element in a space is given by

$$(ds)^2 = g_{\mu\nu} dx^\mu dx^\nu, \tag{3.19}$$

where $g_{\mu\nu}$ is the *metric tensor* for the space. For the spacetime of special relativity the geometry may be defined in terms of the invariant interval

$$(ds)^2 = \left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2 \tag{3.20}$$

[see (3.1) and the discussion of Lorentz transformation below].

Comparing (3.19) and (3.20), we see that the metric tensor of special relativity is diagonal:

$$g_{\mu\nu} = g_{\nu\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.21}$$

($g_{\mu\nu}$ is not so simple in the curved spacetime of general relativity; there the metric tensor is not fixed in advance, but is determined by the Einstein equations). For flat spacetime the covariant and contravariant metric tensors are identical,

$$g_{\mu\nu} = g^{\mu\nu} \tag{3.22}$$

and their contraction on one index yields the four-dimensional Kronecker delta

$$g_{\mu\lambda}g^{\lambda\nu} = g_{\mu}^{\nu} = \delta_{\mu}^{\nu} = \{^{1,\, for\, \mu=\nu=0,1,2,3}_{0,\, for\, \mu\neq\nu}. \quad (3.23)$$

The general procedure for converting an index on a tensor from upper to lower, or vice-versa, is to contract it with the metric tensor. For example, to transform between covariant and contravariant vector components,

$$x_{\mu} = g_{\mu\nu}x^{\nu} \quad (3.24) \quad x^{\mu} = g^{\mu\nu}x_{\nu}. \quad (3.25)$$

Since $g_{00} = 1$ and $g_{11} = g_{22} = g_{33} = -1$, if a contravariant four-vector has components $A^{\mu} = (A^0, A^1, A^2, A^3)$, its covariant partner has components $A_0 = A^0$, $A_1 = -A^1$, $A_2 = -A^2$, and $A_3 = -A^3$. That is,

$$A^{\mu} = (A^0, \boldsymbol{A}), \quad (3.26)$$

but the corresponding covariant vector is

$$A_{\mu} = (A^0, -\boldsymbol{A}). \quad (3.27)$$

The 3-vector $A$ has components $(A^1, A^2, A^3)$, and we use a notation where boldface denotes 3-vectors, while 4-vectors are set in a normal font. In this notation the scalar product (3.17) of two 4- vectors is

$A \cdot B = A_0B_0 - \boldsymbol{A} \cdot \boldsymbol{B} = g_{\mu\nu}A_{\mu}B_{\nu} = A_0B_0 - A_1B_1 - A_2B_2 - A_3B_3$ i.e.,

$$A \cdot B = A^0B^0 + A_1B^1 + A_2B^2 + A_3B^3. \tag{3.28}$$

For a general tensor, index raising or lowering is accomplished by a contraction of the form

$$F^{\cdot\cdot\mu\cdot\cdot} = g^{\mu\nu} F^{\cdots}_{\nu} \qquad F^{\cdot\cdot}_{\cdots\mu} = g_{\mu\nu} F^{\cdot\cdot\nu}_{\cdots\cdot}. \tag{3.29}$$



The derivatives with respect to the spacetime coordinates obey

$$\frac{\partial}{\partial x'^{\mu}} = \frac{\partial x^{\nu}}{\partial x'^{\mu}} \frac{\partial}{\partial x^{\nu}}, \tag{3.30}$$

which is just the transformation law for covariant vectors (3.5). It follows that the operator implying differentiation with respect to a contravariant (covariant) coordinate vector component transforms as a component of a covariant (contravariant) vector, suggesting the notation

$$\partial^{\mu} = \frac{\partial}{\partial x_{\mu}} = (\partial^0, \partial^1, \partial^2, \partial^3) = \left(\frac{\partial}{\partial x_0}, -\boldsymbol{\nabla}\right) \tag{3.31}$$

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}} = (\partial_0, \partial_1, \partial_2, \partial_3) = \left(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla}\right). \tag{3.32}$$

The final steps are dictated by (3.26) and (3.27), and

$$\boldsymbol{\nabla} = \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = (\partial_1, \partial_2, \partial_3) = (-\partial^1, -\partial^2, -\partial^3)$$

is the 3-divergence (thus $\nabla \cdot A = \partial_k A^k$). We may then define a 4-divergence of a 4-vector $A^{\mu}$ by

$$\partial_{\mu} A^{\mu} = \partial^{\mu} A_{\mu} = \frac{\partial A^0}{\partial x^0} + \boldsymbol{\nabla} \cdot \boldsymbol{A} = \frac{\partial A^0}{\partial x^0} + \nabla_i A^i. \tag{3.33}$$

In this notation the 4-dimensional d'Alembertian operator is the contraction

$$= \partial_\mu \partial^\mu = \frac{\partial^2}{\partial x^0 \partial x_0} - \nabla^2 = \frac{\partial^2}{\partial t^2} - \nabla^2, \tag{3.34}$$

and is a scalar under Lorentz transformations. Another differential operator that we will find useful is expressed in this notation by

$$A \overleftrightarrow{\partial_\mu} B = A\left(\partial_\mu B\right) - \left(\partial_\mu A\right) B, \tag{3.35}$$

where the derivatives act only inside the parentheses.

These are the essential notational results that will be required for subsequent discussion. Fortified with this elegant mathematical machinery, we now return to our discussion of Lorentz transformations and special relativity.

## 3.3 Lorentz Transformations

**Definition 3.1.** A function $\phi : \mathrm{R}^n \to \mathrm{R}^n$ is a *Lorentz transformation* if and only if $g(\phi(x),\phi(y)) = g(x,y)$ for all $x,y$ in $\mathrm{R}^n$ where $g(x,y) = x^1y^1 - x^2y^2 - x^3y^3 - \ldots - x^ny^n$ is defined to be the Lorentz inner product of $x$ and $y$.

It follows from this definition that a Lorentz transformation is a linear orthogonal transformation, since it preserve the Lorentz inner product.

**Theorem 3.2.** *A function $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is a Lorentz transformation iff $\phi$ is linear and $\{\phi(e_1),...,\phi(e_n)\}$ is a L-orthonormal (i.e, Lorentz orthonormal) basis of $\mathbb{R}^n$, as is $\{e_1,...,e_n\}$.*

*Proof.* Suppose $\phi$ is Lorentz transformation of $\mathbb{R}^n$. We have

$$g\left(\phi(e_1),\phi(e_1)\right) = g\left(e_1,e_1\right) = 1$$

and

$$g\left(\phi(e_i),\phi(e_j)\right) = g\left(e_i,e_j\right) = -\delta_{ij},$$

for $i,j > 1$. Clearly, this implies that $\phi(e_1),...,\phi(e_n)$ are linearly independent.

Hence $\{\phi(e_1),...,\phi(e_n)\}$ is a *L-orthonormal* basis of $\mathbb{R}^n$.
Next, let $x$ be in $\mathbb{R}^n$. Then there exist coefficients $c_1,...,c_n$ in $\mathbb{R}^n$ such that

$$\phi(x) = \sum_{i=1}^{n} c_i\phi(e_i).$$

Since $\{\phi(e_1),...,\phi(e_n)\}$is a *L-orthonormal* basis, we have

$$c_1 = g\left(\phi(x),\phi(e_1)\right) = g\left(x,e_1\right) = x^1$$

and

$$-c_j = g\left(\phi(x),\phi(e_j)\right) = g\left(x,e_j\right) = -x^j,$$

for $j > 1$. Then $\phi$ is linear, since

$$\varphi\left(\sum_{i=1}^{n} x^i e_i\right) = \sum_{i=1}^{n} x^i \varphi(e_i).$$

Conversely, suppose $\phi$ is linear and $\{\phi(e_1),...,\phi(e_n)\}$ *is a L-orthonormal basis of* $\mathbb{R}^n$. *Then* $\phi$ *is a* Lorentz transformation, since

$$\begin{aligned} g(\phi(x),\phi(y)) &= g\left(\phi\left(\sum_{i=1}^{n} x_i e_i\right), \phi\left(\sum_{j=1}^{n} y_j e_j\right)\right) \\ &= g\left(\sum_{i=1}^{n} x^i \varphi(e_i), \sum_{j=1}^{n} y^j \varphi(e_j)\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x^i y^j g(\phi(e_i),\phi(e_j)) \\ &= x^1y^1 - x^2y^2 - ... - x^ny^n \\ &= g(x,y). \end{aligned}$$

□

By definition, the quadratic form $Q : M \to \mathbb{R}$, with $Q(v) = g(v,v)$ and where $M$ denotes Minkowski space, is preserved by a Lorentz transformation, and conversely, if a linear map $L : M \to \mathbb{R}$ preserves the quadratic form, then $L$ is a Lorentz

transformation.

Now, consider two inertial systems $S$ and $S^0$, moving with constant velocity $v_0$ with respect to each other, and let the origins coincide at $t = t = 0$. A light pulse is emitted from the coincident origins and it propagates as a spherical wave. In coordinate system $S$ its wave front is defined by ($c = 1$)

$$t^2 - x_1^2 - x_2^2 - x_3^2 = 0, \tag{3.36}$$

while in coordinate system $S_0$

$$t'^2 - x_1'^2 - x_2'^2 - x_3'^2 = 0, \tag{3.37}$$

where the invariance of *c* has been used.

Introducing 4-vectors and the summation convention, we may express these relations in a highly compact notation

$$x_\mu x^\mu = 0 \qquad x'_\mu x'^\mu = 0, \tag{3.38}$$

and by combining these equations the invariance of *c* is embodied in the concise algebraic expression,

$$x_\mu x^\mu = x'_\mu x'^\mu. \tag{3.39}$$

A particular physical event will be described in the two inertial systems by the respective sets of numbers $(x_0, x_1, x_2, x_3)$ and $(x'_0, x'_1, x'_2, x'_3)$. The coordinates in the two systems are connected by Lorentz transformations

$$x'_\mu = \alpha_\mu^\nu(v)\, x_\nu \qquad (\mu = 0,1,2,3), \tag{3.40}$$

where the $\alpha_\mu^\nu(v)$ are functions of the relative velocity *v* (and of the relative spatial orientation of the two systems), and are determined by imposing the constancy of *c*; that is through (3.39). We may also formulate Lorentz transformations among contravariant vectors

$$x^{0\mu} = \alpha_\nu^\mu(v) x^\nu, \tag{3.41}$$

and employ contractions with the metric tensor to raise or lower indices.

The definition (3.40) corresponds to what is more precisely termed the *homogeneous Lorentz transformations*. A more general *inhomogeneous Lorentz transformation* (*Poincare' transformation* ) is defined by

$$x'_{\mu} = \alpha^{\nu}_{\mu} x_{\nu} + b_{\mu}. \tag{3.42}$$

This corresponds to a displacement of the origin by $b_{\mu}$, as well as a rotation of the spacetime coordinate system. By Lorentz transformations we will normally mean homogenous Lorentz transformations.

The analysis of rotations in 3-dimensional space begins with the definition of rotation as an operation that leaves invariant the quantity $x^2+y^2+z^2 \equiv r^2$ (Section 1.1, above). From this, the machinery of tensor analysis may be constructed. The Lorentz transformation leaves invariant the quantity $s^2 = t^2 - x_1^2 - x_2^2 - x_3^2$ ; therefore it corresponds to a generalized rotation in 4-dimensional Minkowski space, and many well-known results from the 3-dimensional coordinate system rotation may be immediately extended to the 4-dimensional case (see Sections 1.1 and 1.2).

**Proposition 3.3.** *The Lorentz transformation may be written in matrix form*

$$x^0 = Ax, \tag{3.43}$$

*where x and $x^0$ denote 4-component vectors,* $\alpha^{\nu}_{\mu}$ constitute a 4 × 4 matrix $A$

$$\begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \alpha_{00} & \alpha_0^1 & \alpha_0^2 & \alpha_{03} \\ \alpha_{10} & \alpha_1^1 & \alpha_1^2 & \alpha_{13} \\ \alpha_{2} & \alpha_2^1 & \alpha_2^2 & \alpha_{23} \\ \alpha_{3} & \alpha_3^1 & \alpha_3^2 & \alpha_{33} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

and the coefficients satisfy

$$\alpha^{\nu}_{\lambda}\alpha^{\lambda}_{\mu} = \delta^{\nu}_{\mu}, \tag{3.44}$$

where $\delta^{\nu}_{\mu}$ is the 4-dimensional Kronecker delta and the determinant of a Lorentz transformation matrix is equal to ±1.

*Proof.* We prove the orthogonality condition (3.44) and evaluate the determinant of the matrix.

Using (3.39) and (3.40),

$$x'_{\mu}x'^{\mu} = \alpha^{\lambda}_{\mu}\alpha^{\mu}_{\epsilon}x_{\lambda}x^{\epsilon} = x_{\mu}x^{\mu},$$

which implies that $\alpha^{\lambda}_{\mu}\alpha^{\mu}_{\epsilon} = \delta^{\lambda}_{\epsilon}$. In matrix notation, the scalar product (3.28) is $a \cdot b = a^T g b$, where $a^T$ denotes the transpose of the matrix $a$ and $g$ is the matrix (3.21). Thus, (3.39) is $x^T g x = (x')^T g x'$, or on using (3.43),

$$x^T g x = x^T A^T g A x,$$

which means that $A^T gA = g$. Now, taking the determinant of both sides and using

$$\det(ab) = (\det a)(\det b) \qquad \det g = -1 \qquad \det a = \det a^T,$$

we obtain $\det A = \pm 1$. □

The continuous Lorentz transformations correspond to rotations of the coordinate system within a particular frame, and *Lorentz boots* between frames. The explicit forms of the matrices for Lorentz transformations are given in many places. Most of our discussion will require only their formal properties.

Quantities that are not changed by a Lorentz transformation [(3.11)] are Lorentz tensors of rank zero (scalars). Examples include

- the 4-dimensional volume element $d^4x$,
- the 4-vector scalar product $A \cdot B$,
- the d'Alembertian operator $= \partial_\mu \partial^\mu$.

A quantity transforming in the same way as the coordinates under the Lorentz transformations is termed a *4-vector* (Lorentz tensor of rank 1), the derivatives of which transform [see (3.30)] in flat spacetime as

$$\frac{\partial}{\partial x'_\mu} = \alpha^\mu_\nu \frac{\partial}{\partial x_\nu} = \alpha^\mu_\nu \partial^\nu \quad . \tag{3.45}$$

Hence, the 4-gradient $(\partial^0, \partial^1, \partial^2, \partial^3)$ transforms as a 4-vector. Other examples of 4-vectors include

- the 4-momentum $p^\mu = (E,\mathbf{p})$ with $E = \sqrt{\mathbf{p}^2 + m^2}$,
- the 4-current $j^\mu = (\rho,\mathbf{j})$, where $\rho$ is the charge density and $\mathbf{j}$ the ordinary spatial 3-current from the Schrödinger equation,
- the electromagnetic 4-potential $A^\mu = (A^0,\mathbf{A})$, where $A^0$ is the scalar potential and $\mathbf{A}$ is the vector potential.

Second rank tensors transform as (3.14)-(3.16).

### 3.3.1 Improper and Proper Rotations

The parity transformation corresponds to inverting the spatial coordinates but not the time, while the time-reversal transformation reverses the time but not the spatial coordinates. The matrices that accomplish this are

$$\pi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.46}$$

and

$$\tau = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.47}$$

These matrices do not correspond to generalized rotations (as shown in Chapter 1). For example, the parity operation is equivalent to a rotation by $180^0$ plus a reflection through a plane. Operations such as parity or time reversal are called improper transformations. Their transformation matrices $M$ may be characterized by $\det M = -1$, while the proper Lorentz transformation matrices $R$ are of determinant $+1$. The proper Lorentz transformations can be built by a succession of infinitesimal transformations, in contrast, the improper Lorentz transformations are discrete and cannot be obtained by compounding infinitesimal transformations.

**Proposition 3.4.** *The 4-vector scalar product $A \cdot B$ and the differential volume element $d^4x$ are invariant under Lorentz transformations.*

*Proof.* From (3.12), (3.13), (3.16) and (3.44)

$$A \cdot B = A_\mu B_\mu \qquad A_{0\mu} = \alpha_{\nu\mu} A_\nu \qquad B'_\mu = \alpha_\mu^\lambda B_\lambda$$

$$A'^{\mu}B'_{\mu} = \alpha^{\mu}_{\nu}\alpha^{\lambda}_{\mu}A^{\nu}B_{\lambda} = \delta^{\lambda}_{\nu}A^{\nu}B_{\lambda} = A^{\mu}B_{\mu} = A \cdot B.$$

So the scalar product is Lorentz invariant. The transformation to a new differential volume element is given by $d^4x^0 = Jd^4x$, where $J$ is the Jacobian determinant of the transformation matrix. But for proper Lorentz transformations $J = \det A = 1$, i.e., $d^4x^0 = d^4x$. □

**Proposition 3.5.** *(a) The d'Alembertian operator $= \partial^{\mu}\partial_{\mu}$ is invariant under Lorentz transformations. (b) If $A_{\mu}$ is a 4-vector, then $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ transforms as antisymmetric rank-2 tensor.*

*Proof.* (a) $\Box = \partial_{\mu}\partial^{\mu} \quad \partial'_{\mu} = \alpha^{\nu}_{\mu}\partial_{\nu} \quad \Box' = \partial'_{\mu}\partial'^{\mu} = \alpha^{\nu}_{\mu}\alpha^{\mu}_{\lambda}\partial_{\nu}\partial^{\lambda} = \partial_{\mu}\partial^{\mu} =$

(b) From (3.13)

$$\partial'_{\mu}A'_{\mu} = \alpha^{\lambda}_{\mu}\alpha^{\epsilon}_{\nu}\partial_{\lambda}A_{\epsilon},$$

which is the transformation law of a covariant rank-2 tensor of (3.15). Thus, $F_{\mu\nu}$ is a rank-2 tensor and it is clearly antisymmetric: $F_{\mu\nu} = -F_{\nu\mu}$. □

### 3.3.2 Group Structure of Lorentz Transformations

The homogeneous Lorentz transformations form a group called the *homogeneous Lorentz group* (often just termed the *Lorentz group*) which leaves invariant the quantity $s^2 = t^2 - x^2 - y^2 - z^2$. This group is a 6-parameter Lie group in four dimensions. The 3-dimensional orthogonal group of rotations $SO$ (3) is a subgroup of the homogeneous Lorentz group. If the group of translations is added to the homogeneous Lorentz group, we obtain the 10-parameter inhomogeneous Lorentz group or *Poincare' group*. The *proper orthochronous Poincare' group* excludes improper Lorentz transformations such as time reversal and space reflection. The laws of physics are thought to be absolutely invariant under this group. This is an elegant way of stating that nature seems to be indifferent to the choices of coordinate system origin or time origin, spatial orientation of the system, or state of uniform rectilinear motion. As is well known, fundamental conservation laws such as those for angular momentum and energy follow from this.

## 3.4 Klein-Gordon Equation

Having reviewed the theory of special relativity and Lorentz transformations,we now seek a relativistic analog to the single-particle Shr$\ddot{o}$dinger equation. From elementary quantum mechanics we know the Shr$\ddot{o}$dinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = \left[-\frac{\hbar^2}{2m_0}\nabla^2 + V(x)\right]\psi(x,t) \tag{3.48}$$

corresponds to the nonrelativistic energy relation in operator form

$$\hat{E} = \frac{\hat{p}^2}{2m_0} + V(x), \tag{3.49}$$

where

$$\hat{E} = i\hbar\frac{\partial}{\partial t}, \qquad \hat{p} = -i\hbar\nabla \tag{3.50}$$

are the operators of energy and momentum, respectively. In order to obtain a relativistic wave equation we start by considering free particles with the relativistic relation

$$p^{\mu}p_{\mu}=\frac{E^2}{c^2}-\boldsymbol{p}\cdot\boldsymbol{p}=m_0^2c^2. \tag{3.51}$$

We now replace the four-momentum $p^{\mu}$ by the four-momentum operator

$$\begin{aligned}\hat{p}^{\mu} &= i\hbar\frac{\partial}{\partial x_{\mu}}=i\hbar\left\{\frac{\partial}{\partial(ct)},-\frac{\partial}{\partial x},-\frac{\partial}{\partial y},-\frac{\partial}{\partial z}\right\}\\ &= i\hbar\left\{\frac{\partial}{\partial(ct)},-\boldsymbol{\nabla}\right\}=\{\hat{p}_0,\hat{\mathbf{p}}\}.\end{aligned} \tag{3.52}$$

Following the convention of obtaining covariant or contravariant vectors by lowering or raising indices with the help of the metric tensor, the result is in accordance with (3.50). Thus, we obtain the Klein-Gordon equation for free particles,

$$\hat{p}^{\mu}\hat{p}_{\mu}\psi=m_0^2c^2\psi. \tag{3.53}$$

Here $m_0$ is the rest mass of the particle and $c$ the velocity of light in vacuum. With the help of (3.34) we can write (3.53) in the form

$$\left(\Box+\frac{m_0^2c^2}{\hbar^2}\right)\psi=\left(\frac{\partial^2}{c^2\partial t^2}-\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}-\frac{\partial^2}{\partial z^2}+\frac{m_0^2c^2}{\hbar^2}\right)\psi=0. \tag{3.54}$$

We can immediately verify the Lorentz invariance of the Klein-Gordon, as $\hat{p}^{\mu}\hat{p}_{\mu}$ is Lorentz invariant. We also recognize (3.54) as the classical wave equation including the mass term $m_0^2c^2/\hbar^2$. Free solutions are of the form

$$\psi = \exp\left(-\frac{i}{\hbar}p_\mu x^\mu\right) = \exp\left[-\frac{i}{\hbar}\left(p_0 x^0 - \boldsymbol{p}\cdot\boldsymbol{x}\right)\right]$$
$$= \exp\left[+\frac{i}{\hbar}\left(\boldsymbol{p}\cdot\boldsymbol{x} - Et\right)\right].$$

(3.55) Now, the

insertion of (3.55) into (3.53) leads to the condition

$$\hat{p}^\mu \hat{p}_\mu \psi = m_0^2 c^2 \psi \Rightarrow p^\mu p_\mu \exp\left(-\frac{i}{\hbar}p_\mu x^\mu\right) = m_0^2 c^2 \exp\left(-\frac{i}{\hbar}p_\mu x^\mu\right)$$
$$\Rightarrow p^\mu p_\mu = m_0^2 c^2 \quad or \quad \frac{E^2}{c^2} - \boldsymbol{p}\cdot\boldsymbol{p} = m_0^2 c^2,$$

which results in

$$E = \pm c\sqrt{m_0^2 c^2 + \boldsymbol{p}^2}. \tag{3.56}$$

Thus, there exist solutions both for positive $E = +c\sqrt{m_0^2 c^2 + \boldsymbol{p}^2}$ as well as for negative$E = -c\sqrt{m_0^2 c^2 + \boldsymbol{p}^2}$ energies, respectively. We shall see later that the solutions yielding negative energies are physically connected with antiparticles. Since antiparticles can indeed be observed in nature, this gives us an indication of the value of extending the nonrelativistic theory.

Next we construct the four-current $j_\mu$ connected to (3.53), In analogy to our considerations concerning the Shr¨*o*dinger equation, we expect a conservation law for the $j_\mu$. We start from (3.54), in the form

$$\left(\hat{p}_\mu \hat{p}^\mu - m_0^2 c^2\right)\psi = 0,$$

and take the complex conjugate of this equation, i.e.,

$$\left(\hat{p}_\mu \hat{p}^\mu - m_0^2 c^2\right)\psi^\star = 0.$$

Multiplying both equations from the left, the first by $\psi^{?}$ and the second by $\psi$, and calculating the difference of the resulting two equations yields

$$\psi^{\star}\left(\hat{p}_{\mu}\hat{p}^{\mu}-m_0^2c^2\right)\psi-\psi\left(\hat{p}_{\mu}\hat{p}^{\mu}-m_0^2c^2\right)\psi^{\star}=0$$

or

$$-\psi^{\star}\left(\hbar^2\nabla_{\mu}\nabla^{\mu}+m_0^2c^2\right)\psi+\psi\left(\hbar^2\nabla_{\mu}\nabla^{\mu}+m_0^2c^2\right)\psi^{\star}=0$$

$$\Rightarrow \nabla_{\mu}\left(\psi^{?}\nabla^{\mu}\psi-\psi\nabla^{\mu}\psi^{?}\right)\equiv\nabla_{\mu}j^{\mu}=0. \tag{3.57}$$

The four-current density is therefore

$$j_{\mu}=\frac{i\hbar}{2m_0}\left(\psi^{\star}\nabla_{\mu}\psi-\psi\nabla_{\mu}\psi^{\star}\right). \tag{3.58}$$

Here we have multiplied by $i$~$/2m_0$, so that the zero component $j_0$ has the dimension of a probability density (that is $1/cm^3$). Furthermore this ensures us that we obtain the correct nonrelativistic limit (3.62)-(3.63) below. In detail, (3.57) reads

$$\frac{\partial}{\partial t}\left[\frac{i\hbar}{2m_0c^2}\left(\psi^{\star}\frac{\partial\psi}{\partial t}-\psi\frac{\partial\psi^{\star}}{\partial t}\right)\right]+div\left(\frac{-i\hbar}{2m_0}\right)\left[\psi^{\star}\left(\boldsymbol{\nabla}\psi\right)-\psi\left(\boldsymbol{\nabla}\psi^{\star}\right)\right]=0. \tag{3.59}$$

This expression possesses the form of a continuity equation

$$\frac{\partial\rho}{\partial t}+div\boldsymbol{j}=0. \tag{3.60}$$

As usual, integration over the entire configuration space yields

$$\int_V\frac{\partial\rho}{\partial t}d^3x=\frac{\partial}{\partial t}\int_V\rho d^3x=-\int_V div\boldsymbol{j}d^3x=-\int_V\boldsymbol{j}\cdot d\boldsymbol{F}=0.$$

Hence,

$$\int_V \rho d^3x = const.,$$

i.e., $\int_V \rho d^3x$ is constant in time. It would be a natural guess to interpret

$$\rho = \frac{i\hbar}{2m_0c^2}\left(\psi^\star \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^\star}{\partial t}\right) \tag{3.61}$$

as a probability density. However, there is a problem with such an interpretation: At a given time $t$ both $\psi$ and $\partial\psi/\partial t$ may have arbitrary values; therefore $\rho(x,t)$ in (3.61) may be either positive or negative. Hence, $\rho(x,t)$ is not positive definite and thus not a probability density. The deeper reason for this is that the KleinGordon equation is of second order in time, so that we must know both $\psi\ (x,t)$ and $\partial\psi\ (x,t)/\partial t$ for a given $t$. Furthermore there exist solutions for negative energy (we shall see this later). This and the difficulty of the probability interpretation were the reason that, for a long time, the Klein-Gordon equation was regarded to be physically senseless. One therefore looks for a relativistic wave equation of first order in time with positive definite probability, which was finally derived by Dirac. However, it turns out that this equation has negative solutions too. As we have previously noticed and as we shall discussed later, these solutions are connected with the existence of antiparticles (Guildry, 1991).

## 3.5 The Nonrelativistic Limit

We can study the nonrelativistic limit of the Klein-Gordon equation (3.53). In order to do this we form the relation

$$\psi\left(r,t\right)=\phi\left(r,t\right)\exp\left(-\frac{i}{\hbar}m_0c^2t\right), \tag{3.62}$$

that is, we split the time dependence of $\psi$ into two terms, one containing the rest mass. In the nonrelativistic limit the difference of total energy $E$ of the particle and the rest mass $m_0c^2$ is small. Therefore we define

$$E' = E - m_0c^2$$

$_0$ and remark that the kinetic $E' \ll m_0c^2$
energy $E$ is nonrelativistic, which means. Hence,

$$\left|i\hbar\frac{\partial\phi}{\partial t}\right|\approx E'\phi\ll m_0c^2\phi \tag{3.63}$$

holds also and with (3.62) we have

$$\frac{\partial\psi}{\partial t}=\left(\frac{\partial\phi}{\partial t}-i\frac{m_0c^2}{\hbar}\phi\right)\exp\left(-\frac{i}{\hbar}m_0c^2t\right)\approx -i\frac{m_0c^2}{\hbar}\phi\exp\left(-\frac{i}{\hbar}m_0c^2t\right)$$

$$\begin{aligned}\frac{\partial^2\psi}{\partial t^2} &= \frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t} - i\frac{m_0c^2}{\hbar}\phi\right)\exp\left(-\frac{i}{\hbar}m_0c^2t\right)\\ &\approx \left[-i\frac{m_0c^2}{\hbar}\frac{\partial\phi}{\partial t} - \frac{m_0c^2}{\hbar}\frac{\partial\phi}{\partial t} - \frac{m_0c^4}{\hbar^2}\phi\right]\exp\left(-\frac{i}{\hbar}m_0c^2t\right)\\ &= -\left[i\frac{2m_0c^2}{\hbar}\frac{\partial\phi}{\partial t} + \frac{m_0c^4}{\hbar^2}\phi\right]\exp\left(-\frac{i}{\hbar}m_0c^2t\right).\end{aligned}$$

Inserting this result into (3.53) yields

$$-\frac{1}{c^2}\left[i\frac{2m_0c^2}{\hbar}\frac{\partial\phi}{\partial t} + \frac{m_0c^4}{\hbar^2}\phi\right]\exp\left(-\frac{i}{\hbar}m_0c^2t\right)$$

$$= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{m_0^2c^2}{\hbar^2}\right)\phi\exp\left(-\frac{i}{\hbar}m_0c^2t\right)$$

or

$$i\hbar\frac{\partial\phi}{\partial t} = -\frac{\hbar^2}{2m_0}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi = -\frac{\hbar^2}{2m_0}\Delta\phi. \tag{3.64}$$

This is the Shr¨*o*dinger equation for spinless particles. As the type of particle which is described by a wave equation does not depend upon whether the particle is relativistic or nonrelativistic, we infer that the Klein-Gordon equation describes *spin zero particles*.

## 3.6 A Wave Equation for Spin-$\frac{1}{2}$ Particles: The Dirac Equation

The Dirac equation describes spin-1/2 particles; these are elementary particles at the subatomic level classified as fermions (electrons, protons, neutrons, neutrinos, etc.) whose quantum mechanical description ("wave function") or state is altered by a rotation of the system through $2\pi$ about some axis, but is returned to the original value by a rotation through $4\pi$. Quantities such as the wave function of an electron (which depend not only on the object's configuration, but also on its behavior) we call *spinorial objects* and are described mathematically by carriers of the representations of $SL(2,C)$, i.e., by spinors.

Since space and time enter relativity on the same footing, their derivatives should appear in equivalent orders for relativistic wave equation. A wave equation of first order in both time and space derivatives, such as one of the Shr¨$o$dinger form (Schro¨dinger, 1930)

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi \tag{3.65}$$

is required to avoid difficulties such as negative probabilities. Since this last equation is linear in time, it is natural to try to construct a Hamiltonian that is also linear in the spatial derivatives (equality of spatial and temporal coordinates). Hence, the desired equation (3.65) has to be of the form

$$i\hbar\frac{\partial\psi}{\partial t} = \left[\frac{\hbar c}{i}\left(\hat{\alpha}_1\frac{\partial}{\partial x^1} + \hat{\alpha}_2\frac{\partial}{\partial x^2} + \hat{\alpha}_3\frac{\partial}{\partial x^3}\right) + \hat{\beta}m_0c^2\right]\psi \equiv \hat{H}_f\psi, \tag{3.66}$$

(Greiner, 1994). Here $\hat{H}_f$ stands for Hamiltonian for free fields. The yet unknown coefficients $\hat{\alpha}_i$ cannot be simple numbers, otherwise (3.66) would not be form invariant with respect to the simple spatial rotations. We suspect that the $\hat{\alpha}_i$ are matrices and indicate this by the operator sign ∧. Then $\psi$ cannot be a simple scalar, but has to be a column vector

$$\psi = \begin{pmatrix} \psi_1(x,t) \\ \psi_2(x,t) \\ \vdots \\ \psi_N(x,t) \end{pmatrix}, \tag{3.67}$$

from which a positive definite density of the form

$$\rho(\chi) = \psi^\dagger\psi(\chi) = (\psi_1^*, \psi_2^*, \ldots, \psi_N^*)\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} = \sum_{i=1}^{N} \psi_i^* \psi_i(\chi) \tag{3.68}$$

can be constructed immediately. We still have to show that $\rho(\chi)$ is the temporal component of a four-vector (current) for which a continuity equation must exist so that the spatial integral $\int \rho d^3x$ becomes constant in time. Only then is the probability interpretation of $\rho(\chi)$ ensured. It is clear that the wave function $\psi$ in (3.67) is a column vector analogous to the spin wave function of the Pauli equation. Hence, we shall call them spinors, specifying this name later. The dimension $N$ of the spinor is not yet known, but we will be able to decide this soon.

The coefficients $\hat{\alpha}_i$ and $\hat{\beta}$ must obviously be quadratic $N \times N$. Thus the Shr¨*o*dinger-like equation (3.65) and (3.67) represents a system of $N$ coupled firstorder differential equations of the spinor components $\psi_i$, $i = 1,2,...,N$. We also indicate this point in the notation and write (3.66) in the form

$$
\begin{aligned}
i\hbar\frac{\partial\psi}{\partial t} &= \frac{\hbar c}{i}\sum_{\tau=0}^{N}\left(\hat{\alpha}_1\frac{\partial}{\partial x^1}+\hat{\alpha}_2\frac{\partial}{\partial x^2}+\hat{\alpha}_3\frac{\partial}{\partial x^3}\right)_{\sigma\tau}\psi_\tau + m_0c^2\sum_{\tau=0}^{N}\hat{\beta}_{\sigma\tau}\psi_\tau \\
&\equiv \sum_{\tau=0}^{N}\left(\hat{H}_f\right)_{\sigma\tau}\psi_\tau .
\end{aligned}
\tag{3.69}
$$

Equation (3.66) is a short form of (3.69), in which the four $N \times N$ matrices $(\hat{\alpha}_i)_{\sigma\tau}$ ($i = 1,2,3$) and $\hat{\beta}_{\sigma\tau}$ are expressed in the usual abbreviated form for matrices by $\hat{\alpha}_i$ ($i = 1,2,3$) and $\hat{\beta}$ respectively. To continue, we demand the following natural properties:

(a) the correct energy-momentum relation for a relativistic free particle

$$E^2 = p^2c^2 + m_0^2c^4, \tag{3.70}$$

(b) the continuity equation for the density (3.68), and

(c) the Lorentz covariance (i.e., Lorentz-form invariance) for (3.66) and (3.69), respectively.

To fulfill requirement (a), every single component $\psi_\sigma$ of the spinor $\psi$ has to satisfy the Klein-Gordon equation, i.e.,

$$-\hbar^2\frac{\partial^2\psi_\sigma}{\partial t^2} = \left(-\hbar^2c^2\boldsymbol{\nabla}^2 + m_0^2c^2\right)\psi_\sigma. \tag{3.71}$$

On the other hand, from (3.66) it follows by iteration that

$$-\hbar^2\frac{\partial^2\psi_\sigma}{\partial t^2} = -\hbar^2c^2\sum_{i,j=1}^{3}\frac{\hat{\alpha}_i\hat{\alpha}_j+\hat{\alpha}_j\hat{\alpha}_i}{2}\frac{\partial^2\psi}{\partial x^i\partial x^j}$$

$$+\frac{\hbar m_0c^3}{i}\sum_{i=1}^{3}\left(\hat{\alpha}_i\hat{\beta}+\hat{\beta}\hat{\alpha}_i\right)+\hat{\beta}^2m_0^2c^4\psi.$$

Comparison with (3.71) shows the following requirements for the matrices $\hat{\alpha}_i, \hat{\beta}$:

$$\hat{\alpha}_i\hat{\alpha}_j + \hat{\alpha}_j\hat{\alpha}_i = 2\delta_{ij}\mathbf{1},$$

$$\hat{\alpha}_i\hat{\beta} + \hat{\beta}\hat{\alpha}_i = \mathbf{0},$$

$$\hat{\alpha}_i^2 = \hat{\beta}^2 = \mathbf{1}. \tag{3.72}$$

These anticommutation relations define an algebra for the $\psi$ matrices. In order to establish hermiticity of the Hamiltonian $\hat{H}_f$ in (3.66) the matrices $\hat{\alpha}_i, \hat{\beta}$ also have to be Hermitian; thus,

$$\hat{\alpha}_i^\dagger = \hat{\alpha}_i, \qquad \hat{\beta}^\dagger = \hat{\beta}. \tag{3.73}$$

Therefore, the eigenvalues of the matrices are real. Since, according to (3.72), one has $\hat{\alpha}_i^2 = \mathbf{1}$ and $\hat{\beta}^2 = \mathbf{1}$, it follows that the eigenvalues can only have the values ±1.

Now, notice that $\hat{A}\psi_\alpha = \alpha\psi_\alpha$ implies that $\hat{U}\hat{A}\hat{U}^{-1}\hat{U}\psi_\alpha = \alpha\hat{U}\psi_\alpha$, where $\hat{U}$ is any unitary transformation, and therefore, $\hat{A}'\left(\hat{U}\psi_\alpha\right) = \alpha\left(\hat{U}\psi_\alpha\right)$, so the solutions of the rotated matrix $\hat{A}' = \hat{U}\hat{A}\hat{U}^{-1}$ are just the rotated vectors $\psi'_\alpha = \hat{U}\psi_\alpha$ with the same eigenvalues $\alpha$. It follows that the eigenvalues are independent of the special representation, this can best be shown in the diagonal representation of the single matrices. For example, $\hat{\alpha}_i$ in its eigenrepresentation has the form

$$\hat{\alpha}_i = \begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ 0 & 0 & A_3 & \cdots & 0 \\ \ldots & 0 & 0 & \ldots & \ldots \\ 0 & 0 & 0 & \cdots & A_N \end{pmatrix},$$

with eigenvalues $A_1, \ldots, A_N$, and (3.72) now yields

$$\hat{\alpha}_i^2 = \mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \cdots & \cdots & \cdots & \ldots \end{pmatrix} = \begin{pmatrix} A_1^2 & 0 & \cdots & \\ 0 & A_2^2 & \cdots & \\ \ldots & & & \ldots \\ \cdots & & \cdots & A_N^2 \end{pmatrix},$$

from which



$$A_k^2 = 1, \qquad i.e., \qquad A_k = \pm 1. \tag{3.74}$$

Furthermore, from the anticommutation relations (3.72) it follows that the trace of each $\hat{\alpha}_i$ and of $\hat{\beta}$ has to be zero. Namely, according to (3.72) one has

$$\hat{\alpha}_i = -\hat{\beta}\hat{\alpha}_i\hat{\beta}.$$

Because of the identity

$$tr\hat{A}\hat{B} = tr\hat{B}\hat{A},$$

one concludes that

$$tr\hat{\alpha}_i = tr\hat{\beta}^2\hat{\alpha}_i = tr\hat{\beta}\hat{\alpha}_i\hat{\beta} = -tr\hat{\alpha}_i \Longrightarrow tr\hat{\alpha}_i = 0. \tag{3.75}$$

The trace of a matrix is always equal to the sum of its eigenvalues, which can be seen if $\hat{U}$ transforms the matrix $\hat{\alpha}_i$ into the diagonal form,

$$\begin{pmatrix} A_1 & 0 & \cdots & \cdots \\ 0 & A_2 & \cdots & \\ \vdots & \vdots & \ddots & \end{pmatrix} = \hat{U}\hat{\alpha}_i\hat{U}^{-1}.$$

$$\begin{pmatrix} & & & \\ & & & \\ & & & \\ \cdots & \cdots & & A_N \end{pmatrix}$$

Then

$$tr\begin{pmatrix} A_1 & 0 & \cdots & \cdots \\ 0 & A_2 & \cdots & \cdots \\ \vdots & & \ddots & \vdots \\ \cdots & \cdots & & A_N \end{pmatrix} = \sum_{k=1}^{N} A_k = tr\hat{U}\hat{\alpha}_i\hat{U}^{-1} = tr\hat{\alpha}_i\hat{U}\hat{U}^{-1} = tr\hat{\alpha}_i,$$

which proves the above statement. Since the eigenvalues of $\hat{\alpha}_i$ and $\hat{\beta}$ are equal to $\pm 1$, each of these two matrices has to possess as many positive as negative eigenvalues, and therefore has to be of even dimension. The smallest even dimension, $N = 2$, cannot be right, because only three anticommuting matrices exist, namely the three Pauli matrices $\sigma_i$ [1]. Therefore, the smallest dimension for which the requirements (3.72) can be fulfill is $N = 4$. We now study this case in

[1] See equation (2.9), Chap.1.

more detail and indicate immediately one possible explicit representation of the Dirac matrices, i.e.,

$$\hat{\alpha}_i = \begin{pmatrix} \mathbf{0} & \sigma_i \\ \sigma_i & \mathbf{0} \end{pmatrix}, \qquad \hat{\beta} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \tag{3.76}$$

where $\sigma_i$ are Pauli's $2 \times 2$ matrices, $\mathbf{1}$ and $\mathbf{0}$ are the $2 \times 2$ unit and null matrices, respectively. With this explicit form of the Pauli matrices of (2.9), we have, in detail,

$$\hat{\alpha}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad \hat{\alpha}_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix},$$

$$\hat{\alpha}_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \hat{\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.77}$$

Indeed, we can readily check the validity of the relations (3.72), with the help of the Pauli matrix relation

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}\mathbf{1}. \tag{3.78}$$

We also notice that (3.76) describe just one possible choice of the Dirac matrices $\hat{\alpha}_i, \hat{\beta}$. Each set $\hat{\alpha}_i' = \hat{U}\hat{\alpha}_i\hat{U}^{-1}$, $\hat{\beta}' = \hat{U}\hat{\beta}\hat{U}^{-1}$, which obtains from the original $\hat{\alpha}_i, \hat{\beta}$ of (3.77) by a unitary transformation $\hat{U}$, can be used equally as well as the one introduced in (3.87) below. Physical results do not depend on the special choice of the Dirac matrices $\hat{\alpha}_i$ and $\hat{\beta}$, but the calculations can become particularly simple in a certain representation.

Next we construct the four-current density and the equation of continuity. To do that we multiply (3.66) from the left by $\psi^\dagger = (\psi_1^\star, \psi_2^\star, \psi_2^\star, \psi_4^\star)$ and obtain

$$i\hbar\psi^\dagger \frac{\partial}{\partial t}\psi = \frac{\hbar c}{i}\sum_{i=1}^{3}\psi^\dagger\hat{\alpha}_k\frac{\partial}{\partial x^k}\psi + m_0c^2\psi^\dagger\hat{\beta}\psi. \tag{3.79}$$

Furthermore, we form the Hermitian conjugate of (3.66), i.e.

$$-i\hbar\frac{\partial\psi^\dagger}{\partial t} = -\frac{\hbar}{i}\sum_{k=1}^{3}\frac{\partial\psi^\dagger}{\partial x^k}\hat{\alpha}_k^\dagger + m_0c^2\psi^\dagger\hat{\beta},$$

and multiply this equation from the right by $\psi$, taking into consideration the hermiticity of the Dirac matrices ($\hat{\alpha}_i^\dagger = \hat{\alpha}_i$, $\hat{\beta}^\dagger = \hat{\beta}$), to give

$$-i\hbar\psi^\dagger\frac{\partial}{\partial t}\psi = -\frac{\hbar}{i}\sum_{k=1}^{3}\frac{\partial\psi^\dagger}{\partial x^k}\hat{\alpha}_k^\dagger\psi + m_0c^2\psi^\dagger\hat{\beta}\psi. \tag{3.80}$$

Then, subtraction of (3.80) from (3.79) yields

$$i\hbar\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) = \frac{\hbar}{i}\sum_{k=1}^{3}\frac{\partial}{\partial x^k}\left(\psi^\dagger\hat{\alpha}_k\psi\right) \tag{3.81}$$

or

$$\frac{\partial \rho}{\partial t} + div\boldsymbol{j} = \mathbf{0}, \tag{3.82}$$

where

$$\rho = \psi^\dagger \psi = \sum_{i=1}^{4} \psi_i^\star \psi_i \tag{3.83}$$

is the positive definite density (3.68) and

$$j^k = c\psi^\dagger \hat{\alpha}^k \psi \quad or \quad j = c\psi^\dagger \hat{\alpha} \psi \tag{3.84}$$

is the *current density*. Here we have symbolically introduced the three-vector

$$\hat{\boldsymbol{\alpha}} = \left\{\hat{\alpha}^1, \hat{\alpha}^2, \hat{\alpha}^3\right\} = \{-\hat{\alpha}_1, -\hat{\alpha}_2, -\hat{\alpha}_3\} \tag{3.85}$$

(Greiner, 1994) and introduced the upper and lower indices according to our former convention. From (3.82) the conservation law follows immediately in the usual way

$$\frac{\partial}{\partial t}\int_V d^3x\psi^\dagger\psi = -\int_V div\boldsymbol{j}d^3x = -\int_V \boldsymbol{j}\cdot d\boldsymbol{F} = 0 \tag{3.86}$$

where $V$ denotes a certain volume and $F$ its surface. Since $\rho$ is positive definite and because of the conservation law (3.82) we can accept the interpretation of $\rho$ as a probability density [in contrast to the density $\rho$ obtained for the Klein-Gordon equation, see (3.61) which was not positive definite]. Accordingly, we call $j$ the *probability current density*. Here we have presumed that $j$ is a vector, i.e., that its components (3.84) transform under spatial rotations as the components of a threevector. This still has to be shown. Furthermore, $\{c\rho, j\}$ should form a four-vector. Hence, it should transform from one inertial system into another one by a Lorentz transformation.This point and, in addition, the covariance (that is, invariance) of the Dirac equation (3.66) with respect to Lorentz transformations have still to be shown, before we can regard the Dirac equation as an acceptable relativistic wave equation.

We also notice that we have achieved a special representation with (3.76). The choice of the matrices (3.76) is not unequivocal. One recognizes immediately that each unitary transformation $\hat{S}$ yields the matrices

$$\hat{\alpha}'_i = \hat{S}\hat{\alpha}_i\hat{S}^{-1}, \qquad \hat{\beta}' = \hat{S}\hat{\beta}\hat{S}^{-1} \tag{3.87}$$

which also satisfy the algebra (3.72). We check this for the first commutator (3.72), as an example:

$$\hat{S}\hat{\alpha}_i\hat{S}^{-1}\hat{S}\hat{\alpha}_j\hat{S}^{-1} + \hat{S}\hat{\alpha}_j\hat{S}^{-1}\hat{S}\hat{\alpha}_i\hat{S}^{-1} = 2\delta_{ij}\hat{S}\mathbf{1}\hat{S}^{-1}$$

$$\Longrightarrow \hat{\alpha}'_i\hat{\alpha}'_i + \hat{\alpha}'_j\hat{\alpha}'_i = 2\delta_{ij}\mathbf{1}. \tag{3.88}$$

## 3.7 Interaction with an Electromagnetic Field

The interaction of a Dirac particle with an electromagnetic field may be incorporated by the standard prescription from nonrelativistic quantum mechanics (*minimal coupling*).

$$p_\mu \longrightarrow p_\mu - \frac{q}{c}A_\mu \qquad (\textit{classical})$$

$$\longrightarrow i(\partial_\mu + iqA_\mu) \qquad (quantummechanical, c = 1) \tag{3.89}$$

where $A_\mu = (A^0, \boldsymbol{A})$ is the 4-potential, $p_\mu = (E,p)$ is the 4-momentum, and $q$ is the charge. Then the Dirac equation becomes

$$i\frac{\partial\psi}{\partial t} = \left[\boldsymbol{\alpha}\cdot(-i\boldsymbol{\nabla} - q\mathbf{A}) + qA^0 + \beta m\right]\psi. \tag{3.90}$$

The same prescription may be employed to incorporate electromagnetic interactions in the Kein-Gordon equation.

## 3.8 Covariant Notation

To investigate the covariance of the Dirac equation, it is useful to introduce a form that is more symmetric in space and time. The $\gamma$ – *matrices* are defined by

$$\gamma^0 = \beta \qquad \gamma^i = \beta\alpha^i = \gamma^0\alpha^i. \tag{3.91}$$

They satisfy the anticommutation relation

$$\{\gamma_\mu,\gamma_\nu\} = \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}, \tag{3.92}$$

with metric tensor (3.21). This can also be written as

$$\{\gamma_\mu, \gamma^\nu\} = 2\delta^\nu_\mu, \tag{3.93}$$

by (3.23). Since the $\gamma$−matrices are related to the $\alpha$− and $\beta$−matrices by (3.91) their explicit form in the representation we are using is

$$\gamma_0 = \begin{pmatrix} 1 & 0 \\ & \end{pmatrix} \qquad \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ & \end{pmatrix}, \tag{3.94}$$

$$0 \quad -1 \qquad\qquad -\sigma_i \quad 0$$

where the $\sigma_i$ are Pauli matrices (2.9). Two combinations of $\gamma$−matrices occur frequently and merit a special notation:

$$\sigma^{\mu\nu} \equiv \frac{i}{2}\left[\gamma^{\mu}, \gamma^{\nu}\right], \tag{3.95}$$

(the square brackets [] denote commutators, and the curly brackets {} anticommutators), and

$$\gamma^5 = \gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{3.96}$$

In this representation we have explicitly

$$\sigma^{ij} = \epsilon_{ijk}\begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} = \epsilon_{ijk}\Sigma_k \tag{3.97}$$

$$\sigma_{0k} = i\alpha_k = i\begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \tag{3.98}$$

$$\gamma_5 = \gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.99}$$

The scalar product (inner product) of $\gamma$−matrices and 4-vectors occurs often enough to justify a listing (omitting the *Feynman slash convention) :*

$$\gamma_\mu A^\mu = \gamma^0 A^0 - \gamma \cdot A \tag{3.100}$$

$$\gamma \cdot A \equiv \gamma^1 A^1 + \gamma^2 A^2 + \gamma^3 A^3 \tag{3.101}$$



$$\gamma_\mu p^\mu = E\gamma^0 - \gamma \cdot p \tag{3.102}$$

$$i\gamma^\mu \frac{\partial}{\partial x^\mu} = i\gamma^\mu \partial_\mu = i\gamma_0 \frac{\partial}{\partial t} + i\boldsymbol{\gamma} \cdot \boldsymbol{\nabla}. \tag{3.103}$$

With (3.102)-(3.103) the Dirac equation can be written in several compact forms:

$$i\left(\gamma^0 \frac{\partial}{\partial x^0} + \gamma^1 \frac{\partial}{\partial x^1} + \gamma^2 \frac{\partial}{\partial x^2} + \gamma^3 \frac{\partial}{\partial x^3}\right)\psi - m\psi = 0 \tag{3.104}$$

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{3.105}$$

$$(\gamma_\mu p^\mu - m)\psi = 0 \tag{3.106}$$

where $m$ is understood to be multiplied by the unit 4 × 4 matrix and $p$ is the 4-momentum operator.

## 3.9 Covariance of the Dirac Equation

The Dirac equation leads to a sensible probability current and incorporates the correct relativistic energy expression. It remains for us to demonstrate that it is form invariant under Lorentz transformations (covariant), and to deduce the meaning of the 4-component wavefunction ; first we address covariance.

Because $p^{\mu}$ and $A^{\mu}$ are 4-vectors, the addition of minimal electromagnetic coupling by the replacement (3.89) has an effect on the question of covariance and we work with the simple equation

$$\left(i\gamma^{\mu}\frac{\partial}{\partial x^{\mu}} - m\right)\psi\left(x\right) = 0 \quad . \tag{3.107}$$

The principle of relativity requires that under a Lorentz transformation to a primed coordinate system the Dirac equation retain the same form:

$$\left(i\gamma'^{\mu}\frac{\partial}{\partial x'^{\mu}} - m\right)\psi'\left(x'\right) = 0 \quad , \tag{3.108}$$

$\psi'\left(x'\right)$ $_{0}$ where $\psi$ ($x$) anddescribe the same physical state and $x$ and $x$ are related through (3.40) and (3.41). It can be shown that the $\gamma$−matrices in the two systems are equivalent up to a unitary transformation, so we drop the distinction between $\gamma$ and $\gamma^{0}$ and write for the transformed Dirac equation

$$\left(i\gamma^{\mu}\frac{\partial}{\partial x^{\mu}} - m\right)\psi'\left(x'\right) = 0 \quad . \tag{3.109}$$

We assume the transformation between $\psi$ and $\psi^{0}$ to be linear,

$$\psi'\left(x'\right) = S\psi\left(x\right) \quad , \tag{3.110}$$

where the transformation matrix $S$ is 4 × 4 and depends on the relative velocities and spatial orientation of the systems, but not on the coordinates. Neither system is privileged, so the inverse transformation must exist:

$$\psi(x) = S^{-1}\psi'\left(x'\right). \tag{3.111}$$

Using the inverse transformation, and multiplying from the left by $S$, the original Dirac equation may be written

$$\left(iS\gamma^{\mu}S^{-1}\frac{\partial}{\partial x^{\mu}} - m\right)\psi'\left(x'\right) = 0. \tag{3.112}$$

Then using Lorentz transformation (3.41) and (3.30)

$$\frac{\partial}{\partial x^{\mu}} = \alpha^{\nu}_{\mu}\frac{\partial}{\partial x'^{\nu}}, \tag{3.113}$$

(3.112) becomes

$$\left(iS\gamma^{\mu}S^{-1}\alpha^{\nu}_{\mu}\frac{\partial}{\partial x'^{\nu}} - m\right)\psi'\left(x'\right) = 0. \tag{3.114}$$

Comparing with (3.109), we obtain an identical form if

$$S\gamma^{\mu}S^{-1}\alpha^{\nu}_{\mu} = \gamma^{\nu}. \tag{3.115}$$

This may be taken as the defining equation for the matrix operator *S.* Since we originally introduced *S* as the transformation operator for Dirac wavefunction, the solution of (3.115) determines the Lorentz properties of the wave function, and the existence of a solution establishes the covariance of the Dirac equation. For the proper Lorentz transformations, the finite transformation may be built by compounding infinitesimal ones. As an example, under rotation by an angle $\theta$ about the $z$−axis,

$$\psi'\left(x'\right) = e^{\frac{i}{2}\theta\sigma^{12}}\psi(x), \tag{3.116}$$

where $\sigma^{12}$ is given by (3.97):

$$\sigma^{12} = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}. \tag{3.117}$$

For improper transformations this technique fails, but direct solution of the defining equation can be found.

Quantities obeying the law of transformation (3.110) with $S$ defined by (3.115) are called Dirac spinors, bispinors, or Lorentz spinors. By construction, the solution of the Dirac equation must have a bispinor structure if the equation is to satisfy at once the requirements of quantum mechanics and special relativity.

Because the exponent in (3.116) involves $\theta/2$, a rotation of $4\pi$ (not $2\pi$) is required to return $\psi(x)$ to its original value. This implies that physical observables in this Dirac theory must be constructed from even powers of $\psi(x)$ and $\psi^\dagger(x)$.
The combination $\psi^\dagger\gamma^0$ occurs frequently and it is convenient to define the *adjoint spinor* $\bar{\psi}(x)$ :

$$\bar{\psi}(x) \equiv \psi^\dagger\gamma^0 = (\psi_1^\star, \psi_2^\star, -\psi_3^\star, -\psi_4^\star), \tag{3.118}$$

which has the behaviour

$$\bar{\psi}'\left(x'\right) = \bar{\psi}(x)\, S^{-1} \tag{3.119}$$

under Lorentz transformations. From (3.82), (3.91), and (3.118)

$$\rho = \psi^\dagger\psi = \psi^\dagger\gamma^0\gamma^0\psi = \bar{\psi}\gamma^0\psi \tag{3.120}$$

$$j^i = \psi^\dagger \alpha^i \psi = \psi^\dagger \gamma^0 \gamma^i \psi = \bar{\psi} \gamma^i \psi, \tag{3.121}$$

where

$$\left(\gamma^0\right)^2 = 1 \tag{3.122}$$

has also been used. Then, from (3.120) and (3.121) the current 4-vector $j = (\rho, \mathbf{j})$ can be written

$$j^\mu = \bar{\psi} \gamma^\mu \psi. \tag{3.123}$$

## 3.10 Bilinear Covariants

Bilinear forms in $\psi^\dagger$ and $\psi$ transforming as Lorentz tensors are called bilinear covariants. They have the property of keeping the Dirac equation, Lagrangian (or action), and current invariant under Lorentz transformations. As an example, let us consider the Lorentz properties of the Dirac current $j^\mu$, which transforms as $j'^\mu = \bar{\psi}' \gamma^\mu \psi'$. From (3.110) and (3.119), $j'^\mu = \bar{\psi} S^{-1} \gamma^\mu S \psi$, and utilizing (3.115)

$$j'^\mu = \alpha^\mu_\nu \bar{\psi} \gamma^\nu \psi = \alpha^\mu_\nu j^\nu. \tag{3.124}$$

Therefore the Dirac current $j^\mu$ behaves as a vector under proper Lorentz transformations [see (3.12)]. To completely specify its transformation properties we must also examine the behaviour of $j^\mu$ under the improper Lorentz transformations. Under space reflection (3.46) (for which $S = \gamma^0$), $j^0 \to j^0$ and $j^k \to -j^k$, so $j^\mu$ is

a true vector rather than an axial vector (also called a pseudovector).

Sixteen linearly independent 4 × 4 matrices may be constructed from products of $\gamma$−matrices. These independent matrices may be used to construct an algebra, the *Clifford algebra*, which was known long before the Dirac equation. Any 4 × 4 matrix can be expressed in terms of this 16-component basis. Sandwiching these matrices between $\bar{\psi}$ and $\psi$ generates the bilinear covariants listed in Table 3.1.

**Proposition 3.6.** *The combination* $\bar{\psi}\psi$ *transforms as a scalar,* $\bar{\psi}\gamma^5\psi$ *a pseudoscalar, and* $\bar{\psi}\sigma^{\mu\nu}\psi$ *a rank-2 tensor under Lorentz transformations. Proof.*

From (3.110) and (3.119), under Lorentz transformations

$$\bar{\psi}'\psi' = \bar{\psi}S^{-1}S\psi = \bar{\psi}\psi.$$

Therefore, $\bar{\psi}\psi$ is a scalar. [Note that $\bar{\psi}\psi$, not $\psi^\dagger\psi$, is the scalar; $\psi^\dagger\psi$ is the timelike component of a 4-vector (3.120)-(3.123). Under proper Lorentz transformations

$$\bar{\psi}'\gamma_5\psi' = \bar{\psi}S^{-1}\gamma_5 S\psi = \bar{\psi}\gamma^5\psi,$$

where we have used $[S,\gamma^5] = 0$ for proper transformations. Under parity $P$ we have

$$\bar{\psi}P^{-1}\gamma_5 P\psi = -\bar{\psi}\gamma^5\psi$$

where $\gamma^\mu\gamma^5 + \gamma^5\gamma^\mu = 0$ and $P \sim \gamma^0$ [see (3.96)] have been used. Hence $\bar{\psi}\gamma^5\psi$ is a pseudoscalar. Finally

$$\bar{\psi}'\sigma_{\mu\nu}\psi' = \bar{\psi}S^{-1}\sigma_{\mu\nu}S\psi.$$ Using (3.95),

inserting $SS^{-1} = 1$, and employing (3.115), we find

$$\bar{\psi}'\sigma^{\mu\nu}\psi' = \alpha^{\mu}_{\lambda}\alpha^{\nu}_{\beta}\bar{\psi}\sigma^{\lambda\beta}\psi,$$

which is the transformation law (3.14) for a second-rank tensor. □

Table 3.1: Bilinear Dirac Covariants

| **Bilinear Form** | **Transforms as Lorentz** | **Components** |
|---|---|---|
| $\bar{\psi}\psi$ | Scalar | 1 |
| $\bar{\psi}\gamma_0\psi$ | Pseudoscalar | 1 |
| $\bar{\psi}\gamma_\mu\psi$ | Vector | 4 |
| $\bar{\psi}\gamma_5\gamma_\mu\psi$ | Axial vector | 4 |
| $\bar{\psi}\sigma_{\mu\nu}\psi$ | Antisymmetric tensor | 6 |

**Proposition 3.7.** *Maxwell Equations can be represented in the form of the Dirac Equation, i.e., the Maxwell equations*

*(a)* $$curl\,\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{H}}{\partial t} = 0, \quad curl\,\boldsymbol{H} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{4\pi}{c}j,$$

*can be written in the form analogous to the Dirac equation (spinor equation):*

*(b)* $$-\frac{1}{i}\sum_{j=0}^{3}\hat{\alpha}^j\frac{\partial}{\partial x^j}\psi = -\frac{4\pi}{c}\Phi.$$

*Proof.* We define the four-component column vectors

$\psi_0 \quad \varphi_0$

(1) $$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} \quad and \quad \Phi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix},$$

where $\varphi_0 = c\rho$, $\varphi_1 = j_1 = j_x$, $\varphi_2 = j_2 = j_y$, $\varphi_3 = j_3 = j_z$. Moreover, we have $x^0 = x_0 = ct$, $x^1 = -x_1 = x$, $x^2 = -x_2 = y$, $x^3 = -x_3 = z$. Now we define the components of $\psi$ as

(2) $$\psi_0 = 0,\ \psi_1 = H_1 - iE_1,\ \psi_2 = H_2 - iE_2,\ \psi_3 = H_3 - iE_3.$$

From this definition it follows that the matrices $\hat{\alpha}^j$ have partly real and partly pure imaginary matrix elements. From *(b)* we get the equation

(3) $$-\frac{1}{i}\left(\hat{\alpha}^0 \frac{1}{c}\frac{\partial}{\partial t}\psi + \hat{\alpha}^1 \frac{\partial}{\partial x}\psi + \hat{\alpha}^2 \frac{\partial}{\partial y}\psi + \hat{\alpha}^3 \frac{\partial}{\partial z}\psi\right) = -\frac{4\pi}{c}\Phi.$$

Denoting now the matrix elements of the matrices $\hat{\alpha}^j$ by $\alpha_{ik}^j$ we can write down the components of (3) explicitly. In the same manner we write the components of the Maxwell equations *(a)* and compare the coefficients of both systems of equations. In order to obtain the correct signs in the Maxwell equations using (2) we infer, as there appears a factor $-1/i$, that the $\alpha_{ik}^j$ only take values $\pm 1$ or $\pm i$.

*Important Remark*. As $\psi_0 = 0$, this procedure determines the columns 1,2,3, but not column 0. We can fix the remaining column by requiring that $\hat{\alpha}^j$ is Hermitian and that $(\hat{\alpha}^j)^2 = \mathbf{1}$. For the matrices $\hat{\alpha}^j$ we now find

(4)

$$\hat{\alpha}_0 = \mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \hat{\alpha}_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix},$$

$$\hat{\alpha}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & i \\ -1 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \qquad \hat{\alpha}_3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

and the operators $\hat{\alpha}^j$ are Hermitian. We see immediately that $tr\hat{\alpha}^j = 0$ holds for $j = 1,2,3$ and , we obtain the commutation relations

(5) $\hat{\alpha}^1\hat{\alpha}^2 + \hat{\alpha}^2\hat{\alpha}^1 = 0, \qquad \hat{\alpha}^1\hat{\alpha}^2 = i\hat{\alpha}^3,$

$\hat{\alpha}^2\hat{\alpha}^3 + \hat{\alpha}^3\hat{\alpha}^2 = 0, \qquad \hat{\alpha}^2\hat{\alpha}^3 = i\hat{\alpha}^1,$

$\hat{\alpha}^3\hat{\alpha}^1 + \hat{\alpha}^1\hat{\alpha}^3 = 0, \qquad \hat{\alpha}^3\hat{\alpha}^1 = i\hat{\alpha}^2$ and

(6) $(\hat{\alpha}^1)^2 = (\hat{\alpha}^2)^2 = (\hat{\alpha}^3)^2 = \mathbf{1}.$

Using these, we infer from *(b)* that

(7)
$$\left(-\frac{1}{i}\sum_{j=0}^{3}\hat{\alpha}^j\frac{\partial}{\partial x^j}\right)\left(-\frac{1}{i}\sum_{j=0}^{3}\hat{\alpha}^k\frac{\partial}{\partial x^k}\right)\psi$$
$$= -\left[\left(\hat{\alpha}^0\right)^2\frac{\partial^2}{\partial(ct)^2} - \sum_{i=1}^{3}\left(\hat{\alpha}^i\right)^2\frac{\partial^2}{\partial(x^i)^2}\right]\psi$$

$$= \left[\boldsymbol{\nabla}^2 - \frac{\partial^2}{c^2\partial t^2}\right]\psi = \Box\psi,$$

as the mixing are proportional to the anticommutators (5) and hence vanish. This means, from (3) and (7),

(8) $$\psi = \frac{4\pi}{ic}\sum_{j=1}^{3}\hat{\alpha}^j\frac{\partial}{\partial x^j}\Phi.$$

Hence, if there are no source terms ($\Phi = 0$) the components of $\psi$, that is the components of the electromagnetic field, obey a wave equation. If there are sources present, then for the upper component of (8) with $\psi_0 = 0$

(9) $$\sum_{j=0}^{3}\frac{\partial\phi_j}{\partial x^j} = \frac{\partial\rho}{\partial t} + div\boldsymbol{j} = 0$$

follows as a necessary condition of a solution of *(b)*. Obviously, (9) is just the continuity equation. We also can recognize the analogue to the Dirac equation in the Shr¨*o*dinger form:

(10) $$\left(-\frac{1}{i}\frac{\partial}{\partial t} - \hat{H}_0\right)\psi = -4\pi\Phi,$$

where $\hat{H}_0 = (1/i)\sum_{k=1}^{3}\hat{\alpha}^k\partial/\partial x^k$ has the same form as the Dirac Hamiltonian for vanishing mass $m_0$. □

# Chapter 4

# Classical Field Theory



## 4.1 Introduction

Now let us turn to the study of fields according to Greiner and Reinhardt (1996). Each point of (finite or infinite) region in space will be associated with some continuous field variables which we will denote by the generic name $\varphi(x,t)$. This obviously constitutes a system with an infinite number of degrees of freedom. The dynamical variables of the theory are the values of the field $\varphi(x)$ at each point of space instead of the discrete set of coordinates $q_i$ encountered in classical and quantum mechanics.

The Lagrange function here is a functional of the field, i.e., a mapping from a space of functions to the real numbers. It is customary to denote a functional dependence by square brackets:

$$L(t) = L\left[\varphi(x,t),\dot{\varphi}(x,t)\right]. \tag{4.1}$$



Being a functional, $L(t)$ simultaneously depends on the values of $\varphi$ and $\dot{\varphi}$ at all points in space. Note that the functional does not depend on the coordinate $x$

itself.

In order to apply Hamilton's principle we define the variation of a functional $F[\varphi(x)]$ as

$$\begin{aligned}\delta F[\phi] &= F[\phi+\delta\phi]-F[\phi]\\ &:= \int d^3x \frac{\delta F[\phi]}{\delta\phi(x)}\delta\phi(x)\end{aligned} \tag{4.2}$$

The functional derivative $\delta F/\delta\varphi(x)$ of the functional $F[\varphi]$ with respect to the function $\varphi$ at the space point $x$ was introduced. This tells us how the value of the functional is changed when the value of the function $\varphi$ is varied at the point $x$. It should be noted that the functional derivative obeys many of the rules of ordinary differential calculus.

Let us apply (4.2) to the Lagrangian (4.1), which depends on the functions $\varphi$ and $\dot{\varphi}$,

$$\delta L\left[\phi,\dot{\phi}\right] = \int d^3x\left(\frac{\delta L}{\delta\phi(x)}\delta\phi(x)+\frac{\delta L}{\delta\dot{\phi}(x)}\delta\dot{\phi}(x)\right). \tag{4.3}$$

Both sides of this equation have an additional time dependence which is not marked. Integration of the Lagrangian leads to the action $W\langle\varphi,\dot{\varphi}\rangle$ which is also a functional of the functions $\varphi$ and $\dot{\varphi}$. By integration over a time interval $t_1 \ldots t_2$ (often $t_1 = -\infty$ and $t_2 = +\infty$ are chosen), the variation of the action follows as

$$\begin{aligned}\delta W &= \delta \int_{t_1}^{t_2} dt L\left[\phi, \dot{\phi}\right] \\ &= \int_{t_1}^{t_2} dt d^3x \left(\frac{\delta L}{\delta \phi (x,t)} \delta \phi (x,t) + \frac{\delta L}{\delta \dot{\phi} (x,t)} \delta \dot{\phi} (x,t)\right) \\ &= \int_{t_1}^{t_2} dt d^3x \left(\frac{\delta L}{\delta \phi (x,t)} - \frac{\partial}{\partial t} \frac{\delta L}{\delta \dot{\phi} (x,t)}\right) \delta \phi (x,t) \,. \end{aligned} \tag{4.4}$$

Here an integration by parts was performed and the boundary condition $\delta\varphi(x,t_1) = \delta\varphi(x,t_2) = 0$ was used, together with the relation $\delta\dot{\varphi} = (\partial/\partial t)\delta\varphi$. Hamilton's principle of stationary action

$$\delta W\left[\phi, \dot{\phi}\right] = \delta \int_{t_1}^{t_2} dt L\left[\phi, \dot{\phi}\right] = 0 \tag{4.5}$$

thus leads to

$$\frac{\delta L}{\delta \phi} - \frac{\partial}{\partial t}\frac{\delta L}{\delta \dot{\phi}} = 0 \,, \tag{4.6}$$

which is just the Euler-Lagrange equation generalized to field theory.

To get a better understanding of the somewhat "abstract" notion of functional derivatives let us for a while return to the discretized description. We assume that space is divided into small cells of size $4V_i$. With each cell we associate the respective average value of the function $\varphi(x,t)$, i.e.,

$$\phi_i(t) := \frac{1}{\triangle V_i}\int_{\triangle V_i} d^3x \phi(x,t) \,. \tag{4.7}$$

Now $L$ depends on the discrete set of "coordinates" $\varphi_i$ and $\dot{\varphi}_i$. The variation of the Lagrangian then can be written as

$$\begin{aligned} \delta L\left(\phi_i,\dot{\phi}_i\right) &= \sum_i \left(\frac{\delta L}{\delta\phi_i}\delta\phi_i + \frac{\delta L}{\delta\dot{\phi}_i}\delta\dot{\phi}_i\right) \\ &= \sum_i \left(\frac{1}{\triangle V_i}\frac{\delta L}{\delta\phi_i}\delta\phi_i + \frac{1}{\triangle V_i}\frac{\delta L}{\delta\dot{\phi}_i}\delta\dot{\phi}_i\right)\triangle V_i \end{aligned} \tag{4.8}$$

Since variations at different spatial points are assumed to be independent, a comparison of (4.8) and (4.3) stipulates the identification

$$\begin{aligned} \frac{\delta L(t)}{\delta\phi(x,t)} &= \lim_{\triangle V_i\to 0}\frac{1}{\triangle V_i}\frac{\delta L(t)}{\delta\phi_i(t)}, \\ \frac{\delta L(t)}{\delta\dot{\phi}(x,t)} &= \lim_{\triangle V_i\to 0}\frac{1}{\triangle V_i}\frac{\delta L(t)}{\delta\dot{\phi}_i(t)}, \end{aligned} \tag{4.9}$$

where $x$ is located in the cell $4V_i$. The functional derivative thus essentially means differentiation with respect to the value the field takes at the point $x$. To be more specific, in the following we will consider *local field theories* in which the Lagrangian $L$ can be written as a volume integral over a density function L :

$$L(t) = \int d^3x\mathcal{L}\left(\phi(x,t), \nabla\phi(x,t), \dot{\phi}(x,t)\right). \tag{4.10}$$

The Lagrange density may depend on the field function $\varphi$, on its time derivative $\dot{\varphi}$, and also on the gradient $\nabla\varphi$. In principle L could depend also on higher derivatives of $\varphi$. This, however, would have undesirable consequences, since the resulting equations of motion would be of higher than second order. Furthermore we will not study *nonlocal* theories in which the Lagrange density at $x$ has an additional dependence on the value of the field at other space points $y$ 6= $x$. The restriction to the local Lagrange densities containing first-order derivatives has been found to be sufficiently general to form the basis for all present-day field theories.

When constructing a Lagrange density, one starts with considerable freedom. The choices get narrowed down by imposing the symmetry and invariance properties the theory under consideration is expected to satisfy. As the most basic requirement for any relativistic field theory the Lagrange density must be a *Lorentz scalar*. Using the Lagrange density L, the variation of $L(t)$ can be written as

$$\begin{aligned}\delta L(t) &= \int d^3x\Big[\frac{\partial\mathcal{L}}{\partial\phi(x,t)}\delta\phi(x,t)+\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\phi(x,t))}\delta\boldsymbol{\nabla}\phi(x,t)\\&\quad+\frac{\partial\mathcal{L}}{\partial\dot{\phi}(x,t)}\delta\dot{\phi}(x,t)\Big]\\&= \int d^3x\Big[\Big(\frac{\partial\mathcal{L}}{\partial\phi(x,t)}-\boldsymbol{\nabla}\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\phi(x,t))}\Big)\delta\phi(x,t)\\&\quad+\frac{\partial\mathcal{L}}{\partial\dot{\phi}(x,t)}\delta\dot{\phi}(x,t)\Big].\end{aligned} \tag{4.11}$$

Here the relation $\delta\nabla\varphi = \nabla\delta\varphi$ was used, followed by an integration by parts. To justify this procedure the fields and their derivatives on the surface have to approach zero fast enough. Comparing this with (4.3) we arrive at explicit expressions for the functional derivatives:

$$\frac{\delta L(t)}{\delta\phi(x,t)}=\frac{\partial\mathcal{L}}{\partial\phi(x,t)}-\boldsymbol{\nabla}\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\phi(x,t))},\qquad \frac{\delta L(t)}{\delta\dot{\phi}(x,t)}=\frac{\partial\mathcal{L}}{\partial\dot{\phi}(x,t)}. \tag{4.12}$$

Expressed in terms of the Lagrange density the Euler-Lagrange equation (4.6) reads

$$\frac{\partial\mathcal{L}}{\partial\phi(x,t)}-\boldsymbol{\nabla}\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\phi(x,t))}-\frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}(x,t)}=0 \tag{4.13}$$

or, using relativistic (covariant) notation[1] with $x^\mu = (x^0, x) = (t, x)$,

$$\frac{\partial \mathcal{L}}{\partial \phi(x)} - \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0 \quad . \qquad (4.14)$$

Here the four-dimensional gradient was written as usual as $\partial\varphi/\partial x^\mu \equiv (\partial_t, \nabla)\varphi$.

As soon as the Lagrange density of physical theory is known the field equations can be written down, taking the form of partial differential equations in the variables *x* and *t.* Since we have assumed that L does not depend on derivatives of higher than first order, the fields satisfy differential equations of at most second order.

If the Lagrange function depends on several independent fields $\varphi_r$, $r = 1,\ldots,N$, then (4.14) can be generalized as

$$\frac{\partial \mathcal{L}}{\partial \phi_r} - \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} = 0 \quad . \qquad (4.15)$$

Problems can arise if the components $\varphi_r$ are mutually dependent, as will be the case for particles with nonzero spin.

We finally remark that the field equation can also be deduced directly, i.e., without the formal introduction of the functional derivative, from the variational principle

$$\delta \int_{t_1}^{t_2} dt \int d^3x \mathcal{L}(\phi(x), \partial_\mu \phi(x)) = 0 \qquad (4.16)$$

[1] Where not stated otherwise we will always use relativistic units, setting $c = 1$.

by using partial integration in four dimensions.

## 4.2 The Hamilton Formalism

We apply Hamilton's formalism to a field theory in order to quantize the field under consideration. One first has to define a "momentum" that is canonically conjugate to the field variable. Define the canonically conjugate field by the functional

derivative

$$\pi(x,t) = \frac{\delta L(t)}{\delta \dot{\phi}(x,t)}. \tag{4.17}$$

According to the results of the last section this reduces to an ordinary derivative of the Lagrange density

$$\pi(x,t) = \lim_{\triangle V_i \to 0} \frac{1}{\triangle V_i} \frac{\delta L(t)}{\delta \dot{\phi}_i(t)} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x,t)}. \tag{4.18}$$

The time derivative of $\pi$ follows from the Euler-Lagrange equation (4.6)

$$\dot{\pi}(x,t) = \frac{\delta L}{\delta \phi(x,t)}. \tag{4.19}$$

Now the hamiltonian is introduced through the Legendre transformation

$$H(t) = \int d^3x \, \pi(x,t)\dot{\varphi}(x,t) - L(t). \tag{4.20}$$

This can also be written as an integral over the Hamiltonian density H($x$,$t$) :

$$H(t) = \int d^3x \mathcal{H}(x) \qquad where \qquad \mathcal{H}(x) = \pi(x)\dot{\varphi}(x) - \mathcal{L}(x). \tag{4.21}$$

Hamilton's equations of motion take a form reminiscent of the result from ordinary mechanics:

$$\dot{\phi} = \frac{\delta H}{\delta \pi}, \qquad \dot{\pi} = -\frac{\delta H}{\delta \phi}. \tag{4.22}$$

These equations are derived by taking the variation of $H$ and using the Legendre transformation (4.20)

$$\delta H = \int d^3x \left( \dot{\phi}\delta\pi + \pi\delta\dot{\phi} \right) - \delta L = \int d^3x \left( \dot{\phi}\delta\pi - \dot{\pi}\delta\phi \right), \tag{4.23}$$

since

$$\delta L = \int d^3x \left( \frac{\delta L}{\delta \phi}\delta\phi + \frac{\delta L}{\delta \dot{\phi}}\delta\dot{\phi} \right) = \int d^3x \left( \dot{\pi}\delta\phi + \pi\delta\dot{\phi} \right). \tag{4.24}$$

This directtly yields (4.22).

Since the functional $H[\varphi,\pi]$ can depend on $\varphi$, $\pi$ and their gradients $\nabla\varphi$, $\nabla\pi$ the functional derivatives in (4.22) can be expressed explicitly in terms of the Hamilton density H as

$$\frac{\delta H}{\delta \phi} = \frac{\partial \mathcal{H}}{\partial \phi} - \nabla \frac{\partial \mathcal{H}}{\partial (\nabla \phi)}, \tag{4.25}$$

$$\frac{\delta H}{\delta \pi} = \frac{\partial \mathcal{H}}{\partial \pi} - \nabla \frac{\partial \mathcal{H}}{\partial (\nabla \pi)}. \tag{4.26}$$

Finally let us study the role of the Poisson brackets in field theory. Given two functionals $F[\varphi,\pi]$ and $G[\varphi,\pi]$ we define

$$\{F, G\}_{PB} = \int d^3x \left( \frac{\delta F}{\delta \phi(x)} \frac{\delta G}{\delta \pi(x)} - \frac{\delta F}{\delta \pi(x)} \frac{\delta G}{\delta \phi(x)} \right). \tag{4.27}$$

Because of Hamilton's equations of motion (4.22), the time evolution of a functional satisfies

$$\dot{F}(t) = \int d^3x \left( \frac{\delta F}{\delta \phi(x)} \dot{\phi}(x) + \frac{\delta F}{\delta \pi(x)} \dot{\pi}(x) \right) = \{F, H\}_{PB} \tag{4.28}$$

provided that there is no explicit time dependence.

There is a seemingly trivial but very important special case which allows the immediate evaluation of a functional derivative. This is based on the observation that a function can be written as a functional depending on itself:

$$\phi(x,t) = \int d^3x' \phi\left(x', t\right) \delta^3\left(x - x'\right), \tag{4.29}$$

where *x,* like *t,* can be viewed as a parameter of the functional. The functional derivative with respect to $\phi\left(x', t\right)$ then reads

$$\frac{\delta \phi(x,t)}{\delta \phi(x',t)} = \delta^3\left(x - x'\right). \tag{4.30}$$

This plausible result follows from the defining relation for the functional derivative (4.2) through

$$\delta\phi(x,t) = \int d^3x' \frac{\delta \phi(x,t)}{\delta \phi(x',t)} \delta\phi\left(x', t\right). \tag{4.31}$$

The result also can be understood in the framework of discretized formulation involving volume elements. Since $\varphi(x,t)$ is then localized in a spatial cell taking the value $\varphi_j(t)$ the derivatives

$$\frac{\delta \phi(x,t)}{\delta \phi(x',t)} = \lim_{\triangle V_i \to 0} \frac{1}{\triangle V_i} \frac{\partial \phi_j}{\partial \phi_i} = \lim_{\triangle V_i \to 0} \frac{\delta_{ij}}{\triangle V_i} = \delta^3\left(x - x'\right). \tag{4.32}$$

Obviously the relations

$$\frac{\delta\pi(x,t)}{\delta\pi(x',t)} = \delta^3\left(x - x'\right) \tag{4.33}$$

and

$$\frac{\delta\pi(x,t)}{\delta\phi(x',t)} = \frac{\delta\phi(x,t)}{\delta\pi(x',t)} = 0 \tag{4.34}$$

also hold since $\varphi$ and $\pi$ are independent functionals.

With the use of (4.32)-(4.34) the Poisson brackets of $\varphi$ and $\pi$ with the Hamilton function can be evaluated:

$$\begin{aligned} \dot{\varphi}(x,t) &= \{\pi(x,t), H(t)\}_{PB} \\ &= \int d^3x' \frac{\delta\phi(x,t)}{\delta\phi(x',t)} \frac{\delta H(t)}{\delta\pi(x',t)} = \frac{\delta H}{\delta\pi(x,t)}, \end{aligned} \tag{4.35}$$

and similarly

$$\begin{aligned} \dot{\pi}(x,t) &= \{\pi(x,t), H(t)\}_{PB} \\ &= -\int d^3x' \frac{\delta\pi(x,t)}{\delta\pi(x',t)} \frac{\delta H(t)}{\delta\phi(x',t)} = -\frac{\delta H}{\delta\phi(x,t)}. \end{aligned} \tag{4.36}$$

These are Hamilton's equations of motion (4.22).

The mutual Poisson brackets of the fields will turn out to be of special interest. For these we find

$$\begin{aligned} \left\{\phi(x,t), \pi\left(x',t\right)\right\}_{PB} &= \int d^3x'' \frac{\delta\phi(x,t)}{\delta\phi(x'',t)} \frac{\delta\pi(x'',t)}{\delta\pi(x'',t)} \\ &= \int d^3x'' \delta^3\left(x - x''\right) \delta^3\left(x' - x''\right) \\ &= \delta^3\left(x - x'\right) \end{aligned} \tag{4.37}$$

and because of (4.34)

$$\left\{\phi\left(x,t\right),\phi\left(x^{\prime},t\right)\right\}_{PB}=\left\{\pi\left(x,t\right),\pi\left(x^{\prime},t\right)\right\}_{PB}=0 . \tag{4.38}$$

These simple relations of course only apply if both fields are taken at equal time value $t^{0} = t$.



A field theory can be quantized in the same way as a discrete mechanical system by replacing the Poisson brackets by commutation relations between operators in Hilbert space.

## 4.3 Conservation Laws in Classical Field Theories

Conservation laws, i.e., the existence of quantities which do not change in time, independent of the dynamical evolution of a system, play an all-important role in theoretical physics. The conservation of energy, momentum, and angular momentum are fundamental laws that any theory has to guarantee if it is to give a valid description of nature. In addition to these basic properties, physical systems often possess further additional conserved quantities, such as charge, isospin, or generalizations thereof. From a fundamental point of vue the conservation laws are natural consequences of the symmetry properties of a system. For each continuous transformation of the coordinates and / or the fields under which "physics does not change" the existence of a conserved quantity can be deduced. For instance, the conservation of energy, momentum, and angular momentum are based on the invariance of a theory under temporal and spatial translations and under rotations in space. Similarly the conservation of charge follows from an invariance under phase transformation.

The mathematical foundation of this connection was elucidated systematically at the beginning of last two centuries. The link between symmetry properties and conservation laws is known as *Noether's theorem* (Noether, 1918). In the following we will derive this theorem for a general classical field theory and general symmetry transformations. Subsequently the various conserved quantities mentioned earlier will be derived as special applications of the theorem.

## Noether's Theorem

Let us first consider the case that the action integral does not change if the coordinates are subject to continuous transformation. It will be sufficient to study infinitesimal transformations of the kind

$$x'_\mu = x_\mu + \delta x_\mu. \tag{4.39}$$

Let the corresponding change in the field $\varphi_r(x)$ be

$$\phi'_r\left(x'\right) = \phi_r(x) + \delta\phi_r(x), \tag{4.40}$$

with the resulting change in the Lagrange density of

$$\mathcal{L}'\left(x'\right) = \mathcal{L}(x) + \delta\mathcal{L}(x). \tag{4.41}$$

Here for brevity's sake the functional dependence of the Lagrange density $L(x) \equiv L(\varphi(x), \partial\varphi(x)/\partial x_\mu)$ was omitted. $\mathcal{L}'(x')$ is obtained by inserting the primed quantities into the original Lagrange density $\mathcal{L}'(x') = \mathcal{L}'(\phi'(x'), \partial\phi'(x')/\partial x_\mu)$.

It is important to understand that the variation defined above consists of two ingredients, namely the transformation of the coordinates from $x$ to $x^0$ and furthermore the change of the "shape" of the field function from $\varphi$ to $\varphi^0$. As an obvious example, think of a vector field that changes its direction if the coordinates system is rotated. Therefore it is useful to define a modified variation

$$\tilde{\delta}\phi_r(x) = \phi_r'(x) - \phi_r(x), \tag{4.42}$$

which keeps the value of the coordinate $x$ fixed and only takes into account the change of shape of the field. The two types of variations are related through

$$\begin{aligned} \tilde{\delta}\phi_r(x) &= \phi_r'(x) - \phi_r'\left(x'\right) + \phi_r'\left(x'\right) - \phi_r(x) \\ &= \delta\phi_r(x) - \left[\phi_r'\left(x'\right) - \phi_r'(x)\right] = \delta\phi_r(x) - \frac{\partial \phi_r'(x)}{\partial x_\mu}\delta x_\mu \\ &= \delta\phi_r(x) - \frac{\partial \phi_r}{\partial x_\mu}\delta x_\mu. \end{aligned} \tag{4.43}$$

In the second but last step the first term of the Taylor expression was inserted and finally, in lowest order, $\phi_r'(x)$, was replaced by $\varphi_r(x)$. Thus this equation and many of the following results are only valid up to first order in the variation. This is sufficient since the whole treatment in any case only applies to infinitesimal variations.

The modified variation $\tilde{\delta}$ quite conventionally has the property to commute with differentiation $\partial/\partial x_\mu$:

$$\frac{\partial}{\partial x_\mu}\tilde{\delta}\phi_r(x) = \tilde{\delta}\left(\frac{\partial \phi_r(x)}{\partial x_\mu}\right). \tag{4.44}$$

This is immediately obvious from the definition (4.42). Variations of the $\delta$ type do not share this property. When calculating the gradient one finds an additional term

$$\begin{aligned}
\frac{\partial}{\partial x_\mu}\left(\delta\phi_r(x)\right) &= \frac{\partial}{\partial x_\mu}\phi'_r\left(x'\right) - \frac{\partial}{\partial x_\mu}\phi_r(x) \\
&= \left(\frac{\partial\phi'_r\left(x'\right)}{\partial x'_\mu} - \frac{\partial\phi_r(x)}{\partial x_\mu}\right) + \frac{\partial\phi'_r\left(x'\right)}{\partial x_\mu} - \frac{\partial\phi'_r\left(x'\right)}{\partial x'_\mu} \\
&= \delta\left(\frac{\partial\phi_r(x)}{\partial x_\mu}\right) + \frac{\partial x'^\nu}{\partial x_\mu}\frac{\partial\phi'_r\left(x'\right)}{\partial x'^\nu} - \frac{\partial\phi'_r\left(x'\right)}{\partial x'_\mu} \\
&= \delta\left(\frac{\partial\phi_r(x)}{\partial x_\mu}\right) + \frac{\partial\phi'_r\left(x'\right)}{\partial x'^\nu}\frac{\partial\delta x^\nu}{\partial x_\mu} \\
&= \delta\left(\frac{\partial\phi_r(x)}{\partial x_\mu}\right) + \frac{\partial\phi_r(x)}{\partial x^\nu}\frac{\partial\delta x^\nu}{\partial x_\mu},
\end{aligned} \tag{4.45}$$

where, according to (4.39), the identity $\partial x'^\nu/\partial x_\mu = g^{\mu\nu} + \partial\delta x^\nu/\partial x_\mu$ was used and the last step is valid only to first order.

Now we study the consequences that follow if the transformations (4.39) and (4.40) leave the action integral invariant, i.e., we demand

$$\delta W = \int_{\Omega'} d^4x' \mathcal{L}'\left(x'\right) - \int_\Omega d^4x \mathcal{L}(x) \stackrel{!}{=} 0 , \tag{4.46}$$

where $\Omega'$ denotes the same volume of integration as $\Omega$, being expressed in terms of the new coordinates $x'$. We introduce the variation of the Lagrange density into (4.46)

$$\delta W = \int_{\Omega'} d^4x' \delta\mathcal{L}(x) + \int_{\Omega'} d^4x' \mathcal{L}(x) - \int_\Omega d^4x \mathcal{L}(x) . \tag{4.47}$$

Transformation of the volume of integration in (4.47) introduces a Jacobi determinant that in first order reduces to

$$\begin{aligned} d^4x' &= \left|\frac{\partial \left(x'^{\mu}\right)}{\partial \left(x^{\nu}\right)}\right| d^4x = \left| \begin{matrix} 1+\frac{\partial \delta x_0}{\partial x_0} & \frac{\partial \delta x_0}{\partial x_1} & \cdots & \cdots \\ \frac{\partial \delta x_1}{\partial x_0} & 1+\frac{\partial \delta x_1}{\partial x_1} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \cdots \\ \vdots & \vdots & \vdots & 1+\frac{\partial \delta x_3}{\partial x_3} \end{matrix} \right| d^4x \\ &= \left(1+\frac{\partial \delta x^{\mu}}{\partial x^{\mu}}\right) d^4x. \end{aligned}$$

All terms involving mixed derivatives of the variation $\delta x^{0\mu}$ are of higher order. If we use the modified variation of type (4.43), (4.47) is simplified in first order:

$$\begin{aligned} \delta W &= \int_{\Omega} d^4x \delta \mathcal{L}(x) + \int_{\Omega} d^4x \mathcal{L}(x) \frac{\partial \delta x^{\mu}}{\partial x^{\mu}}. \\ &= \int_{\Omega} d^4x \left(\tilde{\delta}\mathcal{L}(x) + \frac{\partial \mathcal{L}(x)}{\partial x^{\mu}} \delta x^{\mu}\right) + \int_{\Omega} d^4x \mathcal{L}(x) \frac{\partial \delta x^{\mu}}{\partial x^{\mu}} \\ &= \int_{\Omega} d^4x \left(\tilde{\delta}\mathcal{L}(x) + \frac{\partial}{\partial x^{\mu}} \left(\mathcal{L}(x) \delta x^{\mu}\right)\right). \end{aligned} \tag{4.48}$$

Now we express the total variation $\tilde{\delta}$ L($x$) in terms of the variations of the fields and their derivatives:

$$\begin{aligned} \tilde{\delta}\mathcal{L}(x) &= \frac{\partial \mathcal{L}(x)}{\partial \phi_r} \tilde{\delta}\phi_r(x) + \frac{\partial \mathcal{L}(x)}{\partial \left(\partial^{\mu}\phi_r\right)} \tilde{\delta}\left(\frac{\partial \phi_r(x)}{\partial x_{\mu}}\right) \\ &= \left[\frac{\partial \mathcal{L}(x)}{\partial \phi_r} \tilde{\delta}\phi_r(x) - \frac{\partial}{\partial x_{\mu}} \frac{\partial \mathcal{L}(x)}{\partial \left(\partial^{\mu}\phi_r\right)} \tilde{\delta}\phi_r(x)\right] \\ &\quad + \frac{\partial}{\partial x_{\mu}} \frac{\partial \mathcal{L}(x)}{\partial \left(\partial^{\mu}\phi_r\right)} \tilde{\delta}\phi_r(x) + \frac{\partial \mathcal{L}(x)}{\partial \left(\partial^{\mu}\phi_r\right)} \frac{\partial}{\partial x_{\mu}} \left(\tilde{\delta}\phi_r(x)\right) \\ &= \left[\frac{\partial \mathcal{L}(x)}{\partial \phi_r} - \frac{\partial}{\partial x_{\mu}} \left(\frac{\partial \mathcal{L}(x)}{\partial \left(\partial^{\mu}\phi_r\right)}\right)\right] \tilde{\delta}\phi_r(x) \\ &\quad + \frac{\partial}{\partial x_{\mu}} \left[\frac{\partial \mathcal{L}(x)}{\partial \left(\partial^{\mu}\phi_r\right)} \tilde{\delta}\phi_r(x)\right]. \end{aligned} \tag{4.49}$$

Here (4.44) was used , i.e., the fact that variation and differentiation can be interchanged. We use the summation convention not only for the Minkowski indices but also for the component index $r$. Thus for fields $\varphi_r$ with several degrees of freedom, $r = 1,...,N$, a summation over $r$ is implied whenever the index occurs twice in an expression.

Now we nearly have reached our goal. Since the range of integration $\Omega$ can be chosen arbitrarily, the integrand of (4.48) itself has to vanish if the action integral is to be invariant, as postulated in (4.46). Using (4.49) the integrand reads

$$\left[\frac{\partial \mathcal{L}(x)}{\partial \phi_r} - \frac{\partial}{\partial x_\mu}\frac{\partial \mathcal{L}(x)}{\partial(\partial^\mu \phi_r)}\right]\tilde{\delta}\phi_r(x)$$

$$+\frac{\partial}{\partial x_\mu}\left[\frac{\partial \mathcal{L}(x)}{\partial(\partial^\mu \phi_r)}\tilde{\delta}\phi_r(x) + \mathcal{L}(x)\,\delta x^\mu\right] = 0 \quad . \tag{4.50}$$

The first term is recognized as the Euler-Lagrange equation (4.14). This term therefore vanishes, provided that the field $\varphi_r$ satisfies the equation of motion. We are left with an expression with vanishing four-divergence that, using (4.43), can be written as

$$\frac{\partial}{\partial x_\mu}\left[\frac{\partial \mathcal{L}(x)}{\partial(\partial^\mu \phi_r)}\left(\delta\phi_r(x) - \frac{\partial \phi_r}{\partial x_\nu}\delta x_\nu\right) + \mathcal{L}(x)\,\delta x_\mu\right] = 0 \quad . \tag{4.51}$$

This is an equation of continuity for vector field defined by the terms in the square brackets, i.e.,

$$\frac{\partial}{\partial x_\mu} f_\mu(x) = 0 \tag{4.52}$$

with the "current" density

$$f_\mu(x) = \frac{\partial \mathcal{L}(x)}{\partial(\partial^\mu \phi_r)}\delta\phi_r(x) - \left[\frac{\partial \mathcal{L}(x)}{\partial(\partial^\mu \phi_r)}\frac{\partial \phi_r}{\partial x^\nu} - g_{\mu\nu}\mathcal{L}(x)\right]\delta x^\nu \quad . \tag{4.53}$$

As is well known, an equation of continuity is just the expression of a conservation law in terms of a differential equation. This becomes obvious if (4.52) is integrated over three-dimensional space and the theorem of Gauss is used:

$$\begin{aligned} 0 &= \int_V d^3x \frac{\partial}{\partial x_\mu} f_\mu(x) = \int_V d^3x \frac{\partial}{\partial x_0} f_0(x) + \int_V d^3x \boldsymbol{\nabla} \cdot \boldsymbol{f}(x) \\ &= \frac{d}{dx_0} \int_V d^3x f_0(x) + \oint_{\partial V} \boldsymbol{do} \cdot \boldsymbol{f}(x). \end{aligned} \tag{4.54}$$

This value of the integral over the surface $\partial V$ vanishes since the fields and their derivatives are assumed to fall off sufficiently fast at infinity. Therefore

$$G := \int_V d^3x f_0(x) \tag{4.55}$$

is a conserved quantity having a value constant in time.This is the essential result of *Noether's theorem*:

- *Each continuous symmetry transformation leads to a conservation law*. The conserved quantity $G$ can be obtained from the Lagrange density through the use of (4.53) and (4.55).

Let us now study several important applications of Noether's theorem.

## (1) Invariance Under Translation

The invariance under translations

$$x'^\mu = x^\mu + \epsilon^\mu, \tag{4.56}$$

follows from homogeneity of spacetime, where $^{\mu}$ is a translation constant increment factor. Since the shape of the fields is not supposed to change under translations,

$$\phi_r^{'}\left(x^{'}\right) = \phi_r\left(x\right), \tag{4.57}$$

the local variation vanishes, $\delta\varphi_r = 0$, and the conserved "Noether current" takes a simple form. The differential conservation law (4.52) reads (after splitting off$\epsilon^{\mu}$)

$$\frac{\partial}{\partial x_\mu}\Theta_{\mu\nu} = 0 \tag{4.58}$$

with the canonical *energy-momentum tensor*

$$\Theta_{\mu\nu} = \frac{\partial\mathcal{L}\left(x\right)}{\partial\left(\partial^\mu\phi_r\right)}\frac{\partial\phi_r}{\partial x^\nu} - g_{\mu\nu}\mathcal{L}. \tag{4.59}$$

Because $\nu = 0,...,3$ this implies four conserved quantities, which are the energy $E$ and the momentum vector $P$ of the field. In four-dimensional notation

$$P^\nu = \left(E/c, \boldsymbol{P}\right) = \frac{1}{c}\int_V d^3x\Theta^{0\nu}\left(x\right) = const. \tag{4.60}$$

Let us remark that the tensor $\Theta_{\mu\nu}$ defined in (4.59) in some cases turns out not to be symmetric. It is possible, however, to go over to a symmetrical energymomentum $T_{\mu\nu} = T_{\nu\mu}$ by adding a suitable term with vanishing four-divergence. The new tensor also satisfies the conservation law (4.58) and yields the same value of $P^\mu$. We will come back to this in a later section.

## (2) Lorentz Invariance

Four-dimensional spacetime is assumed to be not only homogeneous under translations but also *isotropic* with respect to rotations. Therefore the action is required to be invariant under Lorentz transformations. These (proper) Lorentz transformations are known to include ordinary three-dimensional rotations in space as well as velocity transformations (Lorentz boosts) which can be viewed as rotations in mixed spatio-temporal hyperplanes of Minkowski space. A general infinitesimal rotation is given by

$$x^{0\mu} = x^{\mu} + \delta\omega^{\mu\nu}x_{\nu}. \tag{4.61}$$

The matrix $\delta\omega^{\mu\nu}$ depends on the rotation angle (in four dimensions) and is antisymmetric

$$\delta\omega^{\mu\nu} = -\delta\omega^{\nu\mu}. \tag{4.62}$$

This ensures that the length of the vector $x^{\mu}$, measured with respect to the

Minkowski metric, remains invariant under the transformation. This is true since in the lowest order

$$\begin{aligned} x'^{\mu}x'_{\mu} &= \left(x^{\mu} + \delta\omega^{\mu\sigma}x_{\sigma}\right)\left(x_{\mu} + \delta\omega_{\mu}^{r}x_{r}\right) \\ &= x^{\mu}x_{\mu} + \delta\omega^{\mu\sigma}x_{\sigma}x_{\mu} + \delta\omega_{\mu}^{r}x^{\mu}x_{r} \\ &= x^{\mu}x_{\mu} + 2x_{\mu}x_{\nu}\delta\omega^{\mu\nu} = x^{\mu}x_{\mu} + x_{\mu}x_{\nu}\left(\delta\omega^{\mu\nu} + \delta\omega^{\nu\mu}\right), \end{aligned} \tag{4.63}$$

which leads to the requirement (4.62).

The transformed field function $\phi'_r(x')$ will show a linear dependence on the rotation angles and on the values $\varphi_r(x)$. We describe this dependence through the general relation

$$\phi'_r\left(x'\right) = \phi_r(x) + \frac{1}{2}\delta\omega_{\mu\nu}\left(I^{\mu\nu}\right)_{rs}\phi_s(x). \tag{4.64}$$



The quantities $I^{\mu\nu}$ are the infinitesimal generators of the Lorentz transformation.

The physical fields $\varphi_r$ transform according to an irreducible representation of the Lorentz group.. Therefore the $(I^{\mu\nu})_{rs}$ are the elements of the matrix representation of the corresponding infinitesimal generator. They describe a mixing of the various components of a multicomponent field (e.g., spinor, vector,...). The infinitesimal generators can be chosen to be antisymmetric with respect to the Lorentz indices $\mu$ and $\nu$, $I^{\mu\nu} = -I^{\nu\mu}$, since a symmetric part does not contribute to (4.64). Thus there are six independent generators. Three of them, $(\mu,\nu)$ = (1,2), (1,3), (2,3), correspond to spatial rotations whereas the remaining three, $(\mu,\nu)$ = (0,1), (0,2), (0,3), describe Lorentz boosts, i.e., velocity transformations along the various coordinate axes.

For completeness we remark that the infinitesimal generators satisfy the following commutator relations:

$$[I_{\mu\nu}, I_{\sigma\tau}] = +g_{\nu\sigma}I_{\mu\tau} + g_{\mu\tau}I_{\nu\sigma} - g_{\nu\tau}I_{\mu\sigma} - g_{\mu\sigma}I_{\nu\tau}. \tag{4.65}$$

This Lie algebra defines the structure of the Lorentz group and is valid in any representation. The matrix indices $r,s,...$ have been omitted for brevity in (4.65).

The transformation relations for the coordinates (4.61) and for the fields (4.64) can now be inserted into Noether's theorem. According to (4.53) this leads to the conserved field

$$f_{\mu}(x) = \frac{\partial \mathcal{L}}{\partial(\partial^{\mu}\phi_r)}\frac{1}{2}\delta\omega_{\nu\lambda}\left(I^{\nu\lambda}\right)_{rs}\phi_s(x) - \Theta_{\mu\nu}\delta\omega^{\nu\lambda}x_{\lambda}, \tag{4.66}$$

where $\Theta_{\mu\nu}$ is the energy-momentum tensor introduced in (4.59). If we use the antisymmetry of $\delta\omega^{\nu\lambda}$, the last term can be written as

$$\Theta_{\mu\nu}\delta\omega^{\nu\lambda}x_{\lambda} = \frac{1}{2}\delta\omega^{\nu\lambda}\left(\Theta_{\mu\nu}x_{\lambda} - \Theta_{\mu\lambda}x_{\nu}\right), \tag{4.67}$$

so that (4.66) becomes

$$f_{\mu}(x) = \frac{1}{2}\delta\omega^{\nu\lambda}M_{\mu\nu\lambda}(x) \tag{4.68}$$

with the abbreviation

$$M_{\mu\nu\lambda}(x) = \Theta_{\mu\lambda}x_{\nu} - \Theta_{\mu\nu}x_{\lambda} + \frac{\partial \mathcal{L}}{\partial(\partial^{\mu}\phi_r)}\left(I^{\nu\lambda}\right)_{rs}\phi_s(x). \tag{4.69}$$

The corresponding integral conservation law tells us that the antisymmetric tensor

$$M_{\nu\lambda} = \int d^3x\left[\Theta_{0\lambda}x_{\nu} - \Theta_{0\nu}x_{\lambda} + \frac{\partial \mathcal{L}}{\partial(\partial^{0}\phi_r)}\left(I^{\nu\lambda}\right)_{rs}\phi_s(x)\right] \tag{4.70}$$

is a constant of motion. Here $M_{\nu\lambda}$ plays the role of the tensor of *angular momentum*. This becomes obvious for the spatial rotations, i.e., when taking the values 1,2,3 for $\nu$ and $\lambda$. Then (4.70) is split into the following two parts:

$$M_{nl} = L_{nl} + S_{nl} \tag{4.71}$$

where

$$\begin{aligned} L_{nl} &= \int d^3x \left(x_n \Theta_{0l} - x_l \Theta_{0n}\right) \\ &= \int d^3x \frac{\partial \mathcal{L}}{\partial \left(\partial^0 \phi_r\right)} \left(x_n \frac{\partial}{\partial x^l} - x_l \frac{\partial}{\partial x^n}\right) \phi_r (x) \end{aligned}, \quad (4.72)$$

and

$$S_{nl} = \int d^3x \frac{\partial \mathcal{L}}{\partial \left(\partial^0 \phi_r\right)} \left(I^{\nu\lambda}\right)_{rs} \phi_s (x). \quad (4.73)$$

The first part, $L_{nl}$, contains one component of the cross product between the position vector and the momentum (or rather the momentum density). Therefore $L_{nl}$ has the meaning of the *orbital angular momentum* orthogonal to the plane spanned by $x_n$ and $x_l$. The second contribution, $S_{nl}$, contains the generators $\left(I^{nl}\right)_{rs}$ and thus depends on the intrinsic transformation properties of the field $\varphi_r$. It describes the internal or *spin angular momentum* and will look very differently for, scalar, spinor, or vector fields, for example. In addition to describing the angular momentum the tensor $M_{\nu\lambda}$ gives three further conserved properties with mixed spatial-temporal indices. These quantities are related to the relativistic generalization of the *center of mass.*

The space component of the angular momentum tensor, $M_{nl}$, has three independent components and can be mapped to the three-dimensional angular momentum vector $J$. This is achieved by contraction with the antisymmetric unit tensor

$$M_{nl} = \epsilon_{nlk} J^k. \quad (4.74)$$

The cartesian components of $J$ are then given by

$$J_1 = M_{23}, \qquad J_2 = M_{31}, \qquad J_3 = M_{12} \tag{4.75}$$

or, written in closed form,

$$J^m = \frac{1}{2}\epsilon_{mnl}M_{nl}. \tag{4.76}$$

This follows from the contraction rule for the three-dimensional Levi-Civita tensor

$$\epsilon_{nlk}\epsilon_{nlm} = 2\delta_{km}. \tag{4.77}$$

(Note that here the summation convention was applied to three-vectors for which no distinction between covariant and contravariant components was needed).

Thus the symmetries investigated have all been related coordinate transformations (four-dimensional translation and rotation). This has led to conservation laws which are of fundamental importance and have to be valid for any physical system.

## (3) Internal Symmetries

A third application consists of internal symmetries of the system. This yields the conserved current

$$j_\mu(x) = (-i)\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi(x) - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\star)}\phi^\star(x)\right], \tag{4.78}$$

where a convenient normalization factor was introduced. The Noether current $j_\mu$ turns out to be identical with the usual *electromagnetic current density*.

### 4.3.1 The Symmetrized Energy-Momentum Tensor

Noether's theorem leads to equations of continuity, i.e., to conservation laws in differential form. The densities and currents obtained in this way are not fixed uniquely since it is possible to add certain four-dimensional divergence terms without influencing the equation of continuity. In the following, we illustrate this for the *canonical energy-momentum tensor* $\Theta_{\mu\nu}$. We define a modified tensor through

$$T_{\mu\nu} = \Theta_{\mu\nu} + \partial_\sigma \chi_{\sigma\mu\nu}, \tag{4.79}$$

where the tensor $\chi_{\sigma\mu\nu}$ is arbitrary except for the requirement of being antisymmetric with respect to the first two indices:

$$\chi_{\sigma\mu\nu,} = -\chi_{\mu\sigma\nu}. \tag{4.80}$$

This condition guarantees that the conservation law remains unchanged:

$$\begin{aligned} \partial_\mu T_{\mu\nu} &= \partial_\mu \Theta_{\mu\nu} + \partial_\mu \partial_\sigma \chi_{\sigma\mu\nu} \\ &= \partial^\mu \Theta_{\mu\nu} + \frac{1}{2}\partial^\mu \partial^\sigma \left(\chi_{\sigma\mu\nu} + \chi_{\mu\sigma\nu}\right) = \partial^\mu \Theta_{\mu\nu} = 0. \end{aligned} \tag{4.81}$$

Also the total energy and momentum are not affected by the transformation (4.79):

$$\tilde{P}_\nu = \int d^3x T_{0\nu} = \int d^3x \left(\Theta_{0\nu} + \partial^0 \chi_{00\nu} + \partial^k \chi_{0k\nu}\right) = \int d^3x \Theta_{0\nu} = P_\nu. \tag{4.82}$$

Here $\chi_{00\nu}$ vanishes because of (4.80) and we have assumed that $\chi_{0k\nu}$ falls off fast enough at large distances to assure that the surface integral in Gauss' theorem can be neglected.

The additional freedom of choice expressed by (4.79) allows the construction of a modified energy-momentum tensor $T_{\mu\nu}$ which is symmetric under permutation of the indices:

$$T_{\mu\nu} = T_{\nu\mu}. \tag{4.83}$$

As a side effect, the energy-momentum tensor constructed in this way can be used to find a simpler formulation of the conservation law for angular momentum. We define the *modified angular-momentum tensor*

$$\widetilde{M}_{\mu\nu\lambda} = T_{\mu\lambda}x_\nu - T_{\mu\nu}x_\lambda, \tag{4.84}$$

which does not contain the additional term found in (4.69). Then the following differential conservation law for $\widetilde{M}_{\mu\nu\lambda}$ is derived immediately:

$$\begin{aligned} \partial^\mu \widetilde{M}_{\mu\nu\lambda} &= (\partial^\mu T_{\mu\lambda})\, x_\nu + g^\mu_\nu T_{\mu\lambda} - (\partial^\mu T_{\mu\nu})\, x_\lambda - g^\mu_\lambda T_{\mu\nu} \\ &= T_{\nu\lambda} - T_{\lambda\nu} = 0. \end{aligned} \tag{4.85}$$

The tensor $\widetilde{M}_{\mu\nu\lambda}$ again has to agree with the canonical angular-momentum tensor $M_{\mu\nu\lambda}$ up to a four-divergence:

$$\widetilde{M}_{\mu\nu\lambda} = M_{\mu\nu\lambda} + \partial^\sigma \eta_{\sigma\mu\nu\lambda}, \tag{4.86}$$

where $\eta$ has to be antisymmetric in the first two indices:

$$\eta_{\sigma\mu\nu\lambda} = -\eta_{\mu\sigma\nu\lambda}. \tag{4.87}$$

The same argument as in (4.82) shows that the conserved quantity remains unchanged:

$$\widetilde{M}_{\nu\lambda} = \int d^3x \widetilde{M}_{0\nu\lambda} = \int d^3x M_{0\nu\lambda} = M_{\nu\lambda}. \tag{4.88}$$

Now we will construct the transformation function $\chi_{\mu\nu\lambda}$. Inserting the explicit expression for $\widetilde{M}_{\mu\nu\lambda}$ and $M_{\mu\nu\lambda}$ into (4.86) gives

$$\begin{aligned}(\Theta_{\mu\nu} + \partial_\sigma \chi_{\sigma\mu\lambda})x_\nu - (\Theta_{\mu\nu} + \partial_\sigma \chi_{\sigma\mu\nu})x_\lambda \\ = \Theta_{\mu\lambda}x_\nu - \Theta_{\mu\nu}x_\lambda + \frac{\partial \mathcal{L}}{\partial\left(\partial^\mu \phi_r\right)}\left(I_{\nu\lambda}\right)_{rs}\phi_s + \partial^\sigma \eta_{\sigma\mu\nu\lambda}.\end{aligned} \tag{4.89}$$

Now we make a special choice for the function $\eta_{\sigma\mu\nu\lambda}$ in such a way that the expression (4.89) gets simplified:

$$\eta_{\sigma\mu\nu\lambda} = x_\nu \chi_{\sigma\mu\lambda} - x_\lambda \chi_{\sigma\mu\nu}, \tag{4.90}$$

which, because of (4.80), satisfies the condition of antisymmetry (4.87). Equation (4.89) becomes

$$\begin{aligned}(\partial_\sigma \chi_{\sigma\mu\lambda})x_\nu - (\partial_\sigma \chi_{\sigma\mu\nu})x_\lambda - \partial_\sigma\left(x_\nu \chi_{\sigma\mu\lambda} - x_\lambda \chi_{\sigma\mu\nu}\right) \\ = \frac{\partial \mathcal{L}}{\partial\left(\partial^\mu \phi_r\right)}\left(I_{\nu\lambda}\right)_{rs}\phi_s\end{aligned} \tag{4.91}$$

or

$$\chi_{\nu\mu\lambda} - \chi_{\lambda\mu\nu} = -\frac{\partial \mathcal{L}}{\partial\left(\partial^\mu \phi_r\right)}\left(I_{\nu\lambda}\right)_{rs}\phi_s. \tag{4.92}$$

The quantities $I_{\nu\lambda}$ are antisymmetric in $\nu, \lambda$ : $I_{\nu\lambda} = -I_{\lambda\nu}$. The relation (4.92) serves to determine only that part of $\chi_{\nu\mu\lambda}$ which is antisymmetric with respect to $\nu \leftrightarrow \lambda$. Thus the general solution of (4.92) reads

$$\chi_{\nu\mu\lambda} = -\frac{1}{2}\frac{\partial \mathcal{L}}{\partial\left(\partial^{\mu}\phi_r\right)}\left(I_{\nu\lambda}\right)_{rs}\phi_s + a_{\nu\mu\lambda}, \tag{4.93}$$

which contains an arbitrary part symmetric in $\nu, \lambda$ :

$$a_{\nu\mu\lambda} = a_{\lambda\mu\nu}. \tag{4.94}$$

This freedom can now be put to use to satisfy the original requirement (4.80). We choose

$$\chi_{\nu\mu\lambda} = \frac{1}{2}\left[-\frac{\partial \mathcal{L}}{\partial\left(\partial^{\mu}\phi_r\right)}\left(I_{\nu\lambda}\right)_{rs} + \frac{\partial \mathcal{L}}{\partial\left(\partial^{\nu}\phi_r\right)}\left(I_{\mu\lambda}\right)_{rs} + \frac{\partial \mathcal{L}}{\partial\left(\partial^{\lambda}\phi_r\right)}\left(I_{\mu\nu}\right)_{rs}\right]\phi_s \tag{4.95}$$

where the added term obviously satisfies the condition (4.94). The required antisymmetry with respect to the interchange $\nu \leftrightarrow \mu$ is readily checked

$$\chi_{\mu\nu\lambda} = \frac{1}{2}\left[-\frac{\partial \mathcal{L}}{\partial\left(\partial^{\nu}\phi_r\right)}\left(I_{\mu\lambda}\right)_{rs} + \frac{\partial \mathcal{L}}{\partial\left(\partial^{\mu}\phi_r\right)}\left(I_{\nu\lambda}\right)_{rs} + \frac{\partial \mathcal{L}}{\partial\left(\partial^{\lambda}\phi_r\right)}\left(I_{\nu\mu}\right)_{rs}\right]\phi_s$$

i.e.,

$$\chi_{\mu\nu\lambda} = -\chi_{\nu\mu\lambda} \tag{4.96}$$

by using $(I_{\mu\nu})_{rs} = -(I_{\nu\mu})_{rs}$ and interchanging the first terms in (4.95).

Thus we have found that although the canonical energy-momentum tensor in general (the exception is the scalar field) is not symmetric, it can be symmetrized by adding the divergence of (4.95). This has the added benefit of leading to a modified angular momentum tensor having the simple form (4.84). The symmtrized form of the energy-momentum tensor, $T_{\mu\nu}$, is often viewed as more "fundamental" than the canonical tensor $\Theta_{\mu\nu}$.

# Chapter 5

# Free Motion of a Dirac Particle



## 5.1 Introduction

In order to study in detail the free motion of spin-$\frac{1}{2}$ particles also known as Dirac particles, we first introduce the subject of spin-$\frac{1}{2}$ fields, which are described by the Dirac equation. We will start by considering the Dirac wave function as a classical field and derive the corresponding wave equation from a Lagrange function (Greiner and Reinhardt, 1996).

With the covariant relativistic notation, the Dirac equation for massive spin$\frac{1}{2}$ particles reads ($\hbar = c = 1$)

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0 \tag{5.1}$$

where the four Dirac matrices $\gamma^{\mu}$, $\mu = 0,...,3$, satisfy the algebra

$$\gamma_{\mu}\gamma_{\nu} + \gamma_{\nu}\gamma_{\mu} = 2g_{\mu\nu}. \tag{5.2}$$



The wave function $\psi$ has four components and satisfies the transformation law of a relativistic spinor. The dimension of the Dirac field is $dim[\psi] = length^{-3/2}$, i.e., it has the natural dimension $d = +3/2$.

What will be the Lagrangian density that leads to (5.1) as an equation of motion? The Lagrange density will be a bilinear function composed of the fields $\psi$, $\dot{\psi}$, $\nabla\psi$ and the hermitian conjugate fields $\psi^\dagger$, $\dot{\psi}^\dagger$, $\nabla\psi^\dagger$. It has to transform as a Lorentz scalar density and can contain only derivatives of first order since the Dirac equation itself is of first order. This strongly narrows down the choices of possible candidates for the Lagrange density. As a first attempt we consider the relation

$$L = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi = i\psi^\dagger\dot{\psi} + i\psi^\dagger\alpha \cdot \nabla\psi - m\psi^\dagger\beta\psi. \tag{5.3}$$

Here in the second step the matrices $\beta = \gamma^0$, $\beta^2 = \mathbf{1}$, $\alpha = \gamma^0\gamma$ and the adjoint spinor $\bar{\psi} = \psi^\dagger\gamma^0$ were used. We will treat the spinors $\psi$ and $\psi^\dagger$ as independent fields, each having four components. Differentiation of L with respect to these fields and to their time and space derivatives gives

$$\frac{\partial\mathcal{L}}{\partial\dot{\psi}} = i\psi^\dagger, \qquad \frac{\partial\mathcal{L}}{\partial\dot{\psi}^\dagger} = 0$$

$$\frac{\partial\mathcal{L}}{\partial(\nabla\psi)} = i\psi^\dagger\alpha, \qquad \frac{\partial\mathcal{L}}{\partial(\nabla\psi^\dagger)} = 0,$$

$$\frac{\partial\mathcal{L}}{\partial\psi} = -m\psi^\dagger\beta, \quad \frac{\partial\mathcal{L}}{\partial\psi^\dagger} = i\dot{\psi} + i\alpha\cdot\nabla\psi - m\beta\psi. \tag{5.4}$$

Of course these equations are to be interpreted as matrix equations. For example the last but one equation reads explicitly

$$\frac{\partial\mathcal{L}}{\partial\psi_\rho} = -m\sum_\tau \psi^\dagger_\tau\beta_{\tau\rho}. \tag{5.5}$$

Fortunately in most situations the spinor indices can be dropped without causing confusion. In general we will use this convention since it makes for a much compact notation. Variation of the action with respect to $\psi^\dagger$ leads to the Euler-Lagrange equation

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\psi}^\dagger} = \frac{\partial \mathcal{L}}{\partial \psi^\dagger} - \nabla \frac{\partial \mathcal{L}}{\partial (\nabla \psi^\dagger)}, \tag{5.6}$$

which by use of (5.4) simplifies to

$$i\dot{\psi} + i\alpha \cdot \nabla\psi - m\beta\psi = 0. \tag{5.7}$$

Multiplying by $\gamma^0 = \beta$, we obtain the standard form of the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{5.8}$$

as desired. Variation with respect to $\psi$ using (5.4) leads to the differential equation

$$i\dot{\psi}^\dagger = -m\psi^\dagger - i\nabla\psi^\dagger\alpha. \tag{5.9}$$

This can be identified as the hermitian conjugate of (5.7)

$$\bar{\psi}\left(i\gamma^\mu \overleftarrow{\partial}_\mu + m\right) = 0, \tag{5.10}$$

where the arrow indicates that the partial derivative acts on the function to the left.

As a side remark we note that the Lagrange density (5.3) and thus also the action $W$ of the Dirac field happens to vanish, L = 0, if we insert solutions of the Dirac equation for the wave function $\psi$. This property has no particular consequence, however, since what matters is not the numerical value of the action but rather its responses to variations.

To proceed further we need the canonically conjugate fields $\pi_\psi$ and $\pi_{\psi^\dagger}$. As in the case of the Shr¨*o*dinger field these turn out to dependent quantities. We find

$$\pi_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger, \qquad \pi_{\psi^\dagger} = \frac{\partial \mathcal{L}}{\partial \dot{\psi}^\dagger} = 0 . \tag{5.11}$$

Thus there are two independent degrees of freedom of the Dirac field, which are given by $\psi$ and $\psi^\dagger$.

As usual the Hamilton density is obtained through the Legendre transformation

$$\begin{aligned} \mathrm{H} &= \pi_\psi \dot{\psi} + \pi_{\psi^\dagger}\dot{\psi}^\dagger - \mathrm{L} = i\psi^\dagger\dot{\psi} - i\psi^\dagger\dot{\psi} - i\psi^\dagger\alpha\cdot\nabla\psi + m\psi^\dagger\beta\psi \\ &= \psi^\dagger(-i\alpha\cdot\nabla + \beta m)\psi. \end{aligned} \tag{5.12}$$

The Hamiltonian is thus given by the expectation value of Dirac's differential operator $H_D = \alpha \cdot p + \beta m$ :

$$H = \int d^3x \psi^\dagger(x,t)(-i\alpha\cdot\nabla + \beta m)\psi(x,t). \tag{5.13}$$

Finally let us derive the conserved quantities that follow from Noether's theorem. The canonical energy-momentum tensor $\Theta_{\mu\nu}$ reads

$$\begin{aligned} \Theta_{\mu\nu} &= \frac{\partial \mathcal{L}}{\partial(\partial^\mu\psi)}\partial_\nu\psi + \frac{\partial \mathcal{L}}{\partial(\partial^\mu\psi^\dagger)}\partial_\nu\psi^\dagger - g_{\mu\nu}\mathcal{L} \\ &= \bar{\psi} i\gamma_\mu\partial_\nu\psi - g_{\mu\nu}\bar{\psi}(i\gamma^\sigma\partial_\sigma - m)\psi, \end{aligned} \tag{5.14}$$

which leads to the conserved four-momentum vector

$$P_\nu = \int d^3x \Theta_{0\nu} = \int d^3x \left[\bar{\psi} i\gamma_0 \partial_\nu \psi - g_{0\nu} \bar{\psi} \left(i\gamma^\sigma \partial_\sigma - m\right) \psi\right]. \tag{5.15}$$

The time component of this vector is the energy

$$\begin{aligned} P_0 &= \int d^3x \bar{\psi} \left(i\gamma_0\partial_0 - i\gamma_0\partial_0 - i\gamma \cdot \nabla + m\right)\psi \\ &= \int d^3x \psi^\dagger \left(-i\alpha \cdot \nabla + \beta m\right)\psi. \end{aligned} \tag{5.16}$$

As expected this is just the Hamiltonian $H$ of (5.13). The spatial momentum vector reads

$$P = i\int d^3x \psi^\dagger \nabla \psi. \tag{5.17}$$

The transformation law of Dirac spinors under infinitesimal Lorentz transformations is given by

$$\psi^{'}\left(x^{'}\right) = \psi\left(x\right) - \frac{i}{4}\delta\omega_{\mu\nu}\sigma^{\mu\nu}\psi\left(x\right), \tag{5.18}$$

where $\sigma^{\mu\nu} = \frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right]$. Comparing this with (4.64) we read off the infinitesimal generators

$$\left(I^{\mu\nu}\right)_{\alpha\beta} = -\frac{i}{2}\left(\sigma^{\mu\nu}\right)_{\alpha\beta}, \tag{5.19}$$

where $\mu,\nu = 0,...,3$ are Lorentz indices and $\alpha,\beta = 1,...,4$ are Dirac indices. In particular $\sigma^{ij}$ and $\sigma^{0i}$ are obtained as in (3.97) and (3.98), respectively. The generalized angular momentum tensor (4.70) of the Dirac field has the form

$$M_{\nu\lambda} = L_{\nu\lambda} + S_{\nu\lambda}, \tag{5.20}$$

$$L_{\nu\lambda} = \int d^3x \left(\Theta_{0\lambda}x_\nu - \Theta_{0\nu}x_\lambda\right), \tag{5.21}$$

$$S_{\nu\lambda,} = \int d^3x \frac{\partial \mathcal{L}}{\partial\left(\partial^0\psi\right)} I_{\nu\lambda}\psi = \frac{1}{2}\int d^3x \psi^\dagger \sigma_{\nu\lambda}\psi. \tag{5.22}$$

The three-dimensional vectors of orbital and spin angular momentum are



$$\mathbf{L} = -i\int d^3x \psi^\dagger \mathbf{x} \times \nabla\psi, \tag{5.23}$$

$$\mathbf{S} = \frac{1}{2}\int d^3x \psi^\dagger \mathbf{\Sigma}\psi. \tag{5.24}$$

In the standard representation of the Dirac matrices, $\mathbf{\Sigma}$ is the "double" Pauli matrix

$$\mathbf{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma} \end{pmatrix}. \tag{5.25}$$

The Lagrange density (5.3) is invariant under global phase transformations $\psi \to \psi e^{i\chi}$ and $\psi^\dagger \to \psi^\dagger e^{-i\chi}$. This implies the existence of conserved current-density vector $j_\mu(x,t)$ :

$$j_\mu = -ie\left(\frac{\partial\mathcal{L}}{\partial\left(\partial^\mu\psi\right)}\psi - \frac{\partial\mathcal{L}}{\partial\left(\partial^\mu\psi^\dagger\right)}\psi^\dagger\right) = e\bar{\psi}\gamma_\mu\psi. \tag{5.26}$$

We have included the electrical elementary charge $e$ as a factor in the definition of $j_\mu$ since equation (5.26) is just the electrical current density of the Dirac field. The resulting conserved quantity is the total charge

$$Q = \int d^3x j_0(x,t) = e\int d^3x \psi^\dagger\psi. \tag{5.27}$$

The constancy of $Q$ is seen to coincide with the conservation of the norm of the Dirac field.

Our relation for the Lagrangian of the Dirac fields looks very natural and so far has lead to reasonable results. However the expression (5.3) has one annoying feature. It can be readily shown that L is not a real number since,

$$\begin{aligned}\mathcal{L}^\star \;&\equiv\; \mathcal{L}^\dagger = \left[\bar{\psi}\left(i\gamma^\mu\overrightarrow{\partial_\mu} - m\right)\psi\right]^\dagger = \psi^\dagger\left(-i\left(\gamma^\mu\right)^\dagger\overleftarrow{\partial_\mu} - m\right)\left(\gamma^0\right)^\dagger\psi \\ &= \psi^\dagger\left(-i\gamma^0\gamma^\mu\gamma^0\overleftarrow{\partial_\mu} - m\right)\gamma^0\psi = \bar{\psi}\left(-i\gamma^\mu\overleftarrow{\partial_\mu} - m\right)\psi,\end{aligned} \tag{5.28}$$

where the properties of the $\gamma$ matrices $(\gamma^0)^\dagger = \gamma^0$ and $\gamma^0\,(\gamma^\mu)^\dagger\,\gamma^0 = \gamma^\mu$ have been used . The expression (5.28) in general will not agree with (5.3), L? 6= L. This deficiency, however, can be cured by defining the real Lagrange density as

$$\begin{aligned}\mathcal{L}' \,:\; &= \frac{1}{2}\left(\mathcal{L} + \mathcal{L}^\dagger\right) = \frac{1}{2}\bar{\psi}\left(i\gamma^\mu\overrightarrow{\partial_\mu} - m\right)\psi + \frac{1}{2}\bar{\psi}\left(-i\gamma^\mu\overleftarrow{\partial_\mu} - m\right)\psi \\ &= \frac{1}{2}\bar{\psi}i\gamma^\mu\overleftrightarrow{\partial_\mu}\psi - m\bar{\psi}\psi.\end{aligned} \tag{5.29}$$

This symmetrized Lagrange density essentially leads to the same equations of motion and to the same conserved quantities as the simpler relation (5.3).

**Proposition 5.1.** *The symmetrized Lagrange density (5.29) leads to the same equations of motion and the same conservation laws as the unsymmetrized expression (5.3).*

*Proof.* This can be seen even without performing the explicit calculation since L0 and L are identical up to a four-divergence:

$$\mathcal{L}' = \mathcal{L} + \delta\mathcal{L} = \mathcal{L} + \frac{1}{2}(-i)\,\partial^{\mu}\left(\bar{\psi}\gamma_{\mu}\psi\right) \tag{5.30}$$

This additional term $\delta$L only leads to a surface term in the action integral and thus does not show up in the Euler-Lagrange equation.

To compute the conserved quantities the canonically conjugate fields are needed:

$$\pi_{\psi} = \frac{\partial\mathcal{L}'}{\partial\dot{\psi}} = \frac{i}{2}\psi^{\dagger} \equiv \pi \quad and \quad \pi_{\psi^{\dagger}} = \frac{\partial\mathcal{L}'}{\partial\dot{\psi}^{\dagger}} = -\frac{i}{2}\psi \equiv \pi^{\dagger}. \tag{5.31}$$

The energy-momentum tensor then becomes

$$\begin{aligned}\Theta'_{0\nu} &= \pi\partial_{\nu}\psi + \pi^{\dagger}\partial_{\nu}\psi^{\dagger} - g_{0\nu}\mathcal{L}' \\ &= \frac{i}{2}\psi^{\dagger}\overleftrightarrow{\partial_{\nu}}\psi - g_{0\nu}\left(\frac{i}{2}\bar{\psi}\gamma^{\mu}\overleftrightarrow{\partial_{\mu}}\psi - m\bar{\psi}\psi\right).\end{aligned} \tag{5.32}$$

The difference between $\Theta'_{0\nu}$ and the original result $\Theta_{0\nu}$ given in (5.14) is

$$\delta\Theta_{0\nu} = \Theta'_{0\nu} - \Theta_{0\nu} = -\frac{i}{2}\partial_{\nu}\left(\psi^{\dagger}\psi\right) + \frac{i}{2}g_{0\nu}\partial_{\mu}\left(\bar{\psi}\gamma^{\mu}\psi\right). \tag{5.33}$$

This implies that there is no difference in the field energy

$$\begin{aligned}\delta P_0 &= \int d^3x\,\delta\Theta_{00} = \int d^3x\left[-\frac{i}{2}\partial_0\left(\psi^{\dagger}\psi\right) + \frac{i}{2}\partial_0\left(\bar{\psi}\gamma^0\psi\right) + \frac{i}{2}\,\nabla\cdot\left(\bar{\psi}\gamma\psi\right)\right] \\ &= \frac{i}{2}\int d^3x\,\nabla\cdot\left(\psi^{\dagger}\alpha\psi\right) = 0.\end{aligned} \tag{5.34}$$

In the same way the space components ($n$ = 1,2,3) of the momentum differ only by a vanishing surface term

$$\delta P_n = -\frac{i}{2}\int d^3x\,\partial_n\left(\psi^{\dagger}\psi\right) = 0. \tag{5.35}$$

The generalized angular momentum tensor becomes

$$M'_{\nu\lambda} = L'_{\nu\lambda} + S'_{\nu\lambda,} \tag{5.36}$$

where

$$\begin{aligned} L'_{\nu\lambda} &= \int d^3x \left(\Theta'_{0\lambda}x_\nu - \Theta'_{0\nu}x_\lambda\right) \\ &= L_{\nu\lambda} + \int d^3x \left(\delta\Theta_{0\lambda}x_\nu - \delta\Theta_{0\nu}x_\lambda\right) = L_{\nu\lambda} + \delta L_{\nu\lambda}. \end{aligned} \tag{5.37}$$

The difference in the spatial orbital angular momentum tensor ($\nu = n, \lambda = l$) again vanishes:

$$\begin{aligned} \delta L_{nl} &= \int d^3x \left[-\frac{i}{2}\partial_l \left(\psi^\dagger\psi\right) x_n + \frac{i}{2}\partial_n \left(\psi^\dagger\psi\right) x_l\right] \\ &= \int d^3x \left[-\frac{i}{2}\partial_l \left(\psi^\dagger\psi x_n\right) - \frac{i}{2}\psi^\dagger\psi g_{ln} + \frac{i}{2}\partial_n \left(\psi^\dagger\psi x_l\right) + \frac{i}{2}\psi^\dagger\psi g_{nl}\right] \\ &= 0. \end{aligned} \tag{5.38}$$

The mixed spacetime components, however, do not vanish:

$$\begin{aligned} \delta L_{0l} &= \int d^3x \left(\delta\Theta_{0l}x_0 - \delta\Theta_{00}x_l\right) = -\frac{i}{2}\int d^3x \left[\partial_l \left(\psi^\dagger\psi\right) x_0 + \partial_k \left(\bar{\psi}\gamma^k\psi\right) x_l\right] \\ &= -\frac{i}{2}\int d^3x \left[\partial_k \left(\bar{\psi}\gamma^k\psi x_l\right) - \bar{\psi}\gamma^k\psi g_{kl}\right] = \frac{i}{2}\int d^3x \bar{\psi}\gamma_l\psi \neq 0. \end{aligned} \tag{5.39}$$

To compute the spin contribution $S'_{\nu\lambda}$ in (5.36), the hermitian conjugate of (5.18) is needed:

$$\psi'^\dagger\left(x'\right) = \psi^\dagger(x) + \frac{1}{2}\Delta\omega^{\mu\nu}\psi^\dagger(x) I^\dagger_{\mu\nu}, \tag{5.40}$$

which according to (5.19) leads to

$$\begin{aligned} S'_{\nu\lambda} &= \int d^3x \left( \pi I_{\nu\lambda}\psi + \psi^\dagger I^\dagger_{\nu\lambda}\pi^\dagger \right) \\ &= \frac{1}{4}\int d^3x \left( \psi^\dagger\sigma_{\nu\lambda}\psi + \psi^\dagger\sigma^\dagger_{\nu\lambda}\psi \right) \end{aligned} \tag{5.41}$$

The space components satisfy

$$\begin{aligned} \sigma^\dagger_{nl} &= \left[\frac{i}{2}\left(\gamma_n\gamma_l - \gamma_l\gamma_n\right)\right]^\dagger = -\frac{i}{2}\left(\gamma^\dagger_l\gamma^\dagger_n - \gamma^\dagger_n\gamma^\dagger_l\right) \\ &= \sigma_{nl}, \end{aligned} \tag{5.42}$$

which implies

$$S'_{nl} = \frac{1}{2}\int d^3x\psi^\dagger\sigma_{nl}\psi = S_{nl}. \tag{5.43}$$

The mixed components, on the other hand, vanish identically since,

$$\delta S_{0l} = -\frac{1}{2}\int d^3x\psi^\dagger\sigma_{0l}\psi = -\frac{1}{2}\int d^3x\psi^\dagger i\gamma_0\gamma_l\psi \tag{5.44}$$

$$\sigma^\dagger_{0l} = \left[\frac{i}{2}\left(\gamma_0\gamma_l - \gamma_l\gamma_0\right)\right]^\dagger = -\sigma_{0l} \tag{5.45}$$

and thus

$$S'_{0l} = S_{0l} + \delta S_{0l} = 0. \tag{5.46}$$

Therefore the differences in the orbital and spin angular momentum are found to cancel each other:

$$\delta M_{0l} = \delta L_{0l} + \delta S_{0l} = \frac{i}{2}\int d^3x\bar{\psi}\gamma_l\psi - \frac{1}{2}\int d^3x\psi^\dagger i\gamma_0\gamma_l\psi = 0. \tag{5.47}$$

□

Let us now examine the solution of the free Dirac equation (3.66) (that is, the Dirac equation without potentials) and again write it in the form

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}_f\psi = \left(c\hat{\alpha}\cdot\hat{\boldsymbol{p}} + m_0c^2\hat{\beta}\right)\psi. \tag{5.48}$$

Its stationary states are found with the relation

$$\psi(x,t) = \psi(x)\exp[-(i/\hbar)\varepsilon t], \tag{5.49}$$

which transforms (3.66) into

$$\varepsilon\psi(x) = \hat{H}_f\psi(x). \tag{5.50}$$

Again the quantity $\varepsilon$ describes the time evolution of the stationary state $\psi(x)$. For many applications it is useful to split up the four-component spinor into twocomponent spinors $\phi$ and $\chi$, i.e.,

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{5.51}$$

with

$$\phi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \text{ and } \chi = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}. \tag{5.52}$$

Using the explicit form (3.76) for the $\hat{\alpha}$ and $\hat{\beta}$ matrices (5.50) can be written as

$\phi$ 0

$\varepsilon\begin{pmatrix} \phi \\ \chi \end{pmatrix} = c$

$\chi$ $\hat\sigma$ $\hat\sigma$ $\phi$ $\mathbf{1}$

$\hat p$ $+ m_0c^2$

or $\chi$ $0$

$0$

$$\varepsilon\phi = c\hat\sigma\cdot\hat p\chi + m_0c^2\phi,$$

$$\varepsilon\chi = c\hat\sigma\cdot\hat p\phi - m_0c^2\chi$$

States with definite momentum $p$ are

$$\begin{pmatrix}\phi\\ \chi\end{pmatrix} = \begin{pmatrix}\phi_0\\ \chi_0\end{pmatrix}\exp[(i/\hbar)p\cdot x]. \tag{5.54}$$

The equations (5.53) are transformed into the same equations for $\phi_0$ and $\chi_0$, but replacing the operators $\hat p$ by the eigenvalues $p$. Ordering with respect to $\phi_0$ and $\chi_0$ results in the system of equations

$$\begin{aligned}\left(\varepsilon - m_0c^2\right)\mathbf{1}\varphi_0 - c\hat\sigma\cdot\hat{\boldsymbol p}\chi_0 &= 0,\\ c\hat\sigma\cdot\hat{\boldsymbol p}\varphi_0 + \left(\varepsilon + m_0c^2\right)\mathbf{1}\chi_0 &= 0.\end{aligned} \tag{5.55}$$

This linear homogeneous system of equations for $\phi_0$ and $\chi_0$ has nontrivial solutions only in the case of a vanishing determinant of the coefficients, that is

$$\begin{vmatrix}\left(\varepsilon - m_0c^2\right)\mathbf{1} & -c\hat\sigma\cdot\boldsymbol p\\ -c\hat\sigma\cdot\boldsymbol p & \left(\varepsilon + m_0c^2\right)\mathbf{1}\end{vmatrix} = 0. \tag{5.56}$$

Now, we prove (Greiner and Müller, 1994) the following relation of a computation rule for Pauli matrices given by

$$(\hat{\sigma} \cdot A)(\hat{\sigma} \cdot B) = A \cdot B\mathbf{1} + i\hat{\sigma} \cdot (A \times B), \tag{5.57}$$

Let $A$ and $B$ be arbitrary vectors. The commutation relations of the $\hat{\sigma}_i$ are $\hat{\sigma}_i\hat{\sigma}_j = i\varepsilon_{ijk}\hat{\sigma}_k + \delta_{ij}$, where

$$\varepsilon_{ijk} = \begin{cases} 1 & \textit{even permutation of } 1,2,3 \\ -1 & \textit{odd permutation of } 1,2,3 \\ 0 & \textit{otherwise.} \end{cases}$$

Addition (or subtraction) then gives

$$\begin{aligned} \hat{\sigma}_i\hat{\sigma}_j - \hat{\sigma}_j\hat{\sigma}_i &= 2i\varepsilon_{ijk}\hat{\sigma}_k, \\ \hat{\sigma}_i\hat{\sigma}_j + \hat{\sigma}_j\hat{\sigma}_i &= 2\delta_{ij}. \end{aligned}$$

We write out the scalar product

$$(\hat{\sigma} \cdot \boldsymbol{A})(\hat{\sigma} \cdot \boldsymbol{B}) = \left(\sum_{i=1}^{3} \hat{\sigma}_i A_i\right)\left(\sum_{j=1}^{3} \hat{\sigma}_j B_j\right).$$

For the individual components we can write

$$\hat{\sigma}_i A_i \hat{\sigma}_j B_j = A_i B_j \left(i\varepsilon_{ijk}\hat{\sigma}_k + \delta_{ij}\right),$$

and

$$\sum_{ij} A_i B_j \delta_{ij} = \sum_i A_i B_i$$

is just the scalar product $A \cdot B$. In the first term, the sum can be expanded over $k$ without making any changes:

$$\sum_{i,j} \varepsilon_{ijk} A_i B_j \hat{\sigma}_k = \sum_{i,j,k} \varepsilon_{ijk} A_i B_i \hat{\sigma}_k,$$

because, for example, for $i,j = 1,2$, $k$ has to be equal to 3 and the additional term for the supplementary summation over $k$ with $k = 1,2$ vanish identically. Now $\varepsilon_{ijk} A_i B_j$ are just the components of the vector product $A \times B$. Therefore we have

$$\sum_{i,j} \varepsilon_{ijk} A_i B_j \hat{\sigma}_k = \sum_{i,j,k} \varepsilon_{ijk} A_i B_i \hat{\sigma}_k = \sum_k (A \times B)_k \hat{\sigma}_k = \hat{\sigma} \cdot (A \times B).$$

Altogether we get

$$(\hat{\sigma} \cdot A)(\hat{\sigma} \cdot B) = A \cdot B\mathbf{1} + i\hat{\sigma} \cdot (A \times B),$$

as required.

Using equation (5.57), (5.56) transforms into

$$\left(\varepsilon^2 - m_0^2 c^4\right) \mathbf{1} - c^2 \left(\hat{\sigma} \cdot \boldsymbol{p}\right) \left(\hat{\sigma} \cdot \boldsymbol{p}\right) = 0,$$

$$\varepsilon^2 = m_0^2 c^4 + c^2 \boldsymbol{p}^2,$$

from which follows

$$\varepsilon = \pm E_p, \qquad E_p = +c\sqrt{\boldsymbol{p}^2 + m_0^2 c^2}\,. \tag{5.58}$$

The two signs of the time-evolution factor $\varepsilon$ correspond to two types of solutions of the Dirac equations. We call them positive and negative solutions, respectively. From (5.55), for fixed $c$,

$$\chi_0 = \frac{c\left(\hat{\sigma}\cdot\boldsymbol{p}\right)}{m_0c^2+\varepsilon}\varphi_0. \tag{5.59}$$

Let us denote the two-spinor $\phi_0$ in the form

$$\phi_0 = U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}, \tag{5.60}$$

with the normalization

$$U^{\dagger}U = U_1^{\star}U_1 + U_2^{\star}U_2 = 1,$$

where $U_1$, $U_2$ are complex. Using (5.54) and (5.49) we obtain the complete set of positive and negative free solutions of the Dirac equation as

$$\Psi_{p\lambda}\left(x,t\right) = N\begin{pmatrix} U \\ \frac{c(\hat{\sigma}\cdot\boldsymbol{p})}{m_0c^2+\lambda E_p}U \end{pmatrix}\frac{\exp\left[i\left(\boldsymbol{p}\cdot\boldsymbol{x}-\lambda E_p t\right)/\hbar\right]}{\sqrt{2\pi\hbar}^3}. \tag{5.61}$$

Here $\lambda = \pm 1$ characterizes the positive and negative solutions with the time evolution factor $\varepsilon = \lambda E_p$. The normalization factor $N$ is determined from the condition

$$\int \Psi^{\dagger}_{p\lambda}\left(x,t\right)\Psi_{p'\lambda'}\left(x,t\right)d^3x = \delta_{\lambda\lambda'}\delta\left(\boldsymbol{p}-\boldsymbol{p}'\right). \tag{5.62}$$

Hence,

$$N^2\left[U^\dagger U + U^\dagger \frac{c^2\left(\hat{\sigma}\cdot\boldsymbol{p}\right)\left(\hat{\sigma}\cdot\boldsymbol{p}\right)}{\left(m_0c^2+\lambda E_p\right)^2}U\right]=1$$

or, using (5.57)

$$N^2\left[1+\frac{c^2\hat{\sigma}^2\cdot\boldsymbol{p^2}}{\left(m_0c^2+\lambda E_p\right)^2}\right]=1$$

,

this implies

$$\begin{aligned}
N &= \left[\frac{\left(m_0c^2+\lambda E_p\right)^2}{\left(m_0c^2+\lambda E_p\right)^2+c^2\boldsymbol{p}^2}\right]^{1/2}\\
&= \left[\frac{\left(m_0c^2+\lambda E_p\right)^2}{\left(m_0^2c^4+c^2\boldsymbol{p}^2\right)^2+2m_0c^2\lambda E_p+E_p^2}\right]^{1/2}\\
&= \left[\frac{\left(m_0c^2+\lambda E_p\right)^2}{2\left(m_0c^2+\lambda E_p\right)\lambda E_p}\right]^{1/2}\\
&= \left(\frac{m_0c^2+\lambda E_p}{2\lambda E_p}\right)^{1/2}.
\end{aligned} \tag{5.63}$$

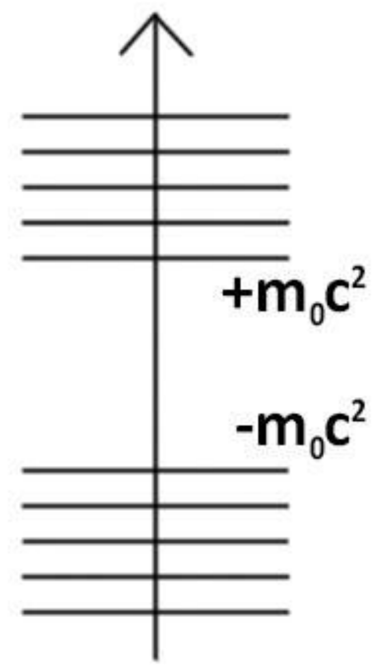


Figure 5.1: Spectrum of the energies of free Dirac equation.

The spectrum of $\varepsilon_{p\lambda} = \lambda E_p$, corresponding to the spinors $\psi_{p\lambda}(x,t)$, is shown in Figure 5.1. There appears, as in the case of the Klein-Gordon equation, a domain of positive and negative frequencies ("energy eigenvalues"). The interpretation of the states with $\lambda = +1$ alone, is our interest in this work. We now recognize that all states (5.61) are eigenfunctions of momentum

$$\hat{p}\Psi_{p\lambda} = p\Psi_{p\lambda}(x,t). \tag{5.64}$$

For every momentum $p$ there are two different kinds of solutions, those with $\lambda = +1$ ($\varepsilon = +E_p$) and those with $\lambda = -1$ ($\varepsilon = -E_p$). We will now show that another quantum number, the helicity, can be used to classify the free one particle states (5.61). First we note that the operator

$$\hat{\Sigma}\cdot\hat{p} = \begin{pmatrix} \hat{\sigma}\cdot\hat{p} & 0 \\ 0 & \hat{\sigma}\cdot\hat{p} \end{pmatrix} \tag{5.65}$$

commutes with the Dirac-Hamiltonian operator $\hat{H}_f$ [cf. (3.66)]. Here

$$\hat{S} = \frac{\hbar}{2}\hat{\Sigma} = \frac{\hbar}{2}\begin{pmatrix} \hat{\sigma} & 0 \\ 0 & \hat{\sigma} \end{pmatrix} \tag{5.66}$$

is the four-dimensional generalization of the spin vector operator. We calculate

$$\left[\hat{H}_f, \hat{\Sigma}\cdot\hat{p}\right] = \left[c\hat{\alpha}\cdot\hat{p} + \hat{\beta}m_0c^2, \hat{\Sigma}\cdot\hat{p}\right] = c\left[\hat{\alpha}\cdot\hat{p}, \hat{\Sigma}\cdot\hat{p}\right],$$

as $\hat{\beta}$ is diagonal matrix and hence $\left[\hat{\beta}, \hat{\Sigma}\right] = 0$. Furthermore, we obtain

$$
\begin{aligned}
(\hat{\alpha}\cdot\hat{p})\left(\hat{\Sigma}\cdot\hat{p}\right)-\left(\hat{\Sigma}\cdot\hat{p}\right)(\hat{\alpha}\cdot\hat{p}) &= \begin{pmatrix} 0 & \hat{\sigma}\cdot\hat{p} \\ \hat{\sigma}\cdot\hat{p} & 0 \end{pmatrix}\begin{pmatrix} \hat{\sigma}\cdot\hat{p} & 0 \\ 0 & \hat{\sigma}\cdot\hat{p} \end{pmatrix} \\
&\quad - \begin{pmatrix} \hat{\sigma}\cdot\hat{p} & 0 \\ 0 & \hat{\sigma}\cdot\hat{p} \end{pmatrix}\begin{pmatrix} 0 & \hat{\sigma}\cdot\hat{p} \\ \hat{\sigma}\cdot\hat{p} & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & (\hat{\sigma}\cdot\hat{p})^2 \\ (\hat{\sigma}\cdot\hat{p})^2 & 0 \end{pmatrix} \\
&\quad - \begin{pmatrix} 0 & (\hat{\sigma}\cdot\hat{p})^2 \\ (\hat{\sigma}\cdot\hat{p})^2 & 0 \end{pmatrix} = 0.
\end{aligned}
$$



Hence,

$$\left[\hat{H}_f, \hat{\Sigma}\cdot\hat{p}\right] = 0 \tag{5.67}$$

and naturally

$$\left[\hat{p}, \hat{\Sigma}\cdot\hat{p}\right] = 0. \tag{5.68}$$

This means that $\hat{\Sigma}\cdot\hat{p}$, $\hat{H}_f$ and $\hat{p}$ can be diagonalized together. The same holds for the helicity operator

$$\hat{\Lambda}_S = \frac{\hbar}{2}\hat{\Sigma}\cdot\frac{\hat{p}}{|\boldsymbol{p}|} = \hat{S}\cdot\frac{\hat{p}}{|\boldsymbol{p}|} \tag{5.69}$$

as we can immediately see by repeating the calculations that led us to (5.67)(5.68). Physically, helicity is the projection of the fermion spin onto the momentum direction, which is a constant of motion of a free Dirac particle, and is invariant under spatial rotations (see Figure 5.2).

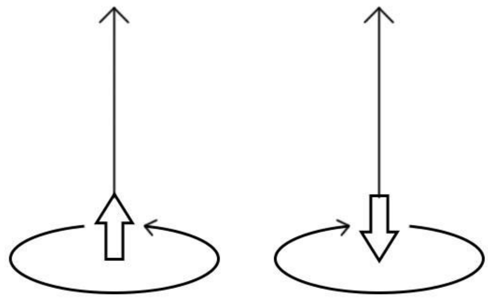

Figure 5.2: Electrons with positive and negative helicity: spin-up and spin-down

If the electron wave propagates into the direction of the $z$ -axis, we have

$$p = (0,0,p)$$

and because of (5.69),

$$\hat{\Lambda}_S = \hat{S}_z = \frac{\hbar}{2}\hat{\Sigma}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{5.70}$$

with the eigenvalues $\pm\hbar/2$. Clearly, the eigenvectors of $\hat{\Lambda}_S$ are

$$\begin{pmatrix} u_1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} u_{-1} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ u_1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ u_{-1} \end{pmatrix} \tag{5.71}$$

with

$$u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad and \quad u_{-1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

0 1

Now we can classify completely the free Dirac waves propagating in the $z$–direction (Greiner, 1994) with positive energy ($\lambda$ = +1) ; we denote them by $\Psi_{pz,\lambda,S_z}(x,t)$ and write explicitly

$$\Psi_{p,\lambda,+1/2} = N \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \frac{c\hat{\sigma}_z p}{m_0c^2+\lambda E_p} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} \exp\left[i\left(pz - \lambda E_p t\right)/\hbar\right] , \tag{5.72}$$

$$\Psi_{p,\lambda,-1/2} = N \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{c\hat{\sigma}_z p}{m_0c^2+\lambda E_p} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} \exp\left[i\left(pz - \lambda E_p t\right)/\hbar\right] . \tag{5.73}$$

From (5.62) one immediately recognizes the validity of the orthogonality relations

$$\int \Psi^{\dagger}_{p_z,\lambda,S_z} \Psi_{p_z',\lambda',S_z'} d^3x = \delta_{\lambda\lambda'} \delta_{S_z S_z'} \delta\left(p_z - p_z'\right) . \tag{5.74}$$

## 5.2 Single-Particle Interpretation of the Plane (Free) Dirac Waves

In this section, by Greiner (1994) and Guildry (1991), we shall examine the singleparticle interpretation of the Dirac equation and its solutions in great detail.

First we discuss the interpretation of the energy. From (5.58) we know that the eigenvalues of the free Dirac Hamiltonian $\hat{H}_f$ in (3.66) are $\varepsilon = \pm E_p$. We have to find out whether the time evolution factors can be interpreted as energies. In order to answer we use the canonical (Lagrange) formalism. It will be shown that the Lagrange density, leading to the free Dirac equation (3.66) is given by

$$\mathcal{L} = i\hbar\Psi^\dagger \frac{\partial}{\partial t}\Psi + i\hbar c\Psi^\dagger \nabla \cdot \hat{\alpha}\Psi - m_0 c^2 \Psi^\dagger \hat{\beta}\Psi, \tag{5.75}$$

and using this result, we will also calculate the energy related to the free solutions (5.72)-(5.73) of the Dirac equation. The result is that

$$E = \varepsilon = \pm E_p \tag{5.76}$$

is identical with the energy of the states. What is the interpretation of these positive and negative energy solutions since there is no precedent in the framework of a one-particle theory? With the proposal of the hole theory, Dirac showed the following way out of this difficulty: let us assume that real electrons are described only by positive energy states (5.76), these are the states with $E = +E_p$. All states of negative energy are occupied by electrons, one electron in each state of negative energy and given spin projection. This is illustrated in Figure 5.3.

In this way a real electron of positive energy is also prevented from falling into energetically lower and lower states by radiation emission. A radiation catastrophe of this kind is averted by the effective Pauli exclusion principle -which requires that two identical fermions (for example, two electrons) cannot occupy the same

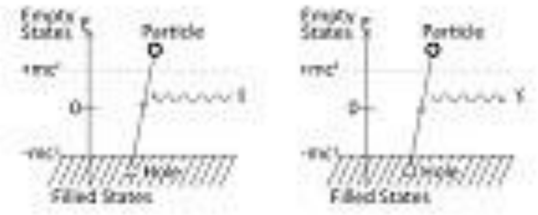


Figure 5.3: Gap between the filled negative-energy states and the positiveenergy states.

state- which simply does not allow these transitions. On the other hand the question arises as to the meaning of a hole in the occupied "sea of negative states". This set of filled negative-energy states is called the *Fermi sea*. If the sea were quiescent it would not be difficult to believe that such a uniform background would have no observable consequences. However, this sea cannot be quiescent because of quantum-fluctuations of the vacuum. From the free Dirac equation the gap between the filled negative-energy states and the positive-energy states is twice the fermion rest mass of $2m_0c^2$, as illustrated in Figure 5.3. Therefore, energy supplied to the vacuum either by virtual fluctuations, or by real physical processes such as incident radiation, may cause a fermion to be promoted from the negative-energy sea to an unoccupied positive-energy state, leaving a hole in the sea (Figure 5.3). Such a particle-hole excitation of the vacuum produces a positive-energy fermion and a hole in the sea that will behave, because of the negative sign on the energy, as fermion of opposite charge - an antifermion. This leads to a meaningful concept of the *positron*, as the antiparticle of the electron. Thus, the hole-theory interpretation shows explicitly that the Dirac equation is generally not a single particle equation; because of the particle-hole excitation of the vacuum, it is inherently a many-body equation.

However, it is legitimate to use the Klein-Gordon or Dirac equations as singleparticle equations only for weak, slowly varying fields for which a broad gap ' $2m_0c^2$ remains between positive-energy and negative-energy states.

We therefore pursue the single-particle aspect further and devote ourselves to the investigation of modified one-particle operators and the Ehrenfest's theorems (Ehrenfest, 1927) according to which: *in nonrelativistic quantum mechanics there is always a correspondence between a relation of operators and that of classical objects (measurable values)*. As an example, Newton's classical equation of motion corresponds to the operator equation $\frac{d\hat{p}}{dt} = \frac{1}{i\hbar}\left[\hat{p}, \hat{H}\right] = -\nabla U,$ with $\hat{H} = \hat{p}^2/2m + U(x)$. Another example is given by the operator relation

$$\frac{d\boldsymbol{x}}{dt} = \frac{1}{i\hbar}\left[\boldsymbol{x}, \hat{H}\right] = \frac{\hat{p}}{m_0}, \tag{5.77}$$

which corresponds to the classical relation between the velocity and linear momentum. This way, perhaps, a consistent single-particle description can be obtained. For this it is helpful to introduce the *sign operator* $\hat{\Lambda}$,

$$\hat{\Lambda} = \frac{\hat{H}_f}{\sqrt{\hat{H}_f^2}} = \frac{c\hat{\alpha}\cdot\hat{p} + \hat{\beta}m_0c^2}{c\sqrt{\hat{p}^2 + m_0^2c^2}}. \tag{5.78}$$

Of course it commutes with the free Dirac Hamiltonian $\hat{H}_f$. Moreover, $\hat{\Lambda}$ is Hermitian and unitary, i.e.,

$$\hat{\Lambda} = \hat{\Lambda}^{\dagger} = \hat{\Lambda}^{-1}. \tag{5.79}$$

In momentum representation $\hat{\Lambda}$ has an especially simple form, that is

$$\hat{\Lambda} = \frac{c\left(\hat{\alpha}\cdot\hat{p}\right) + \hat{\beta} m_0 c^2}{E_p}. \tag{5.80}$$

The name "*sign operator*" comes from the fact that

$$\hat{\Lambda}\Psi_{p,\lambda,S_z} = \frac{\varepsilon}{E_p}\Psi_{p,\lambda,S_z} = \frac{\lambda E_p}{E_p}\Psi_{p,\lambda,S_z} = \lambda\Psi_{p,\lambda,S_z}. \tag{5.81}$$

$\hat{\Lambda}$ has as an eigenvalue the sign $\lambda$ $(= \pm 1)$ of the time evolution factor, with $\lambda = +1$ meaning positive-energy states and, $\lambda = -1$ negative-energy states. An arbitrary state with fixed $\lambda$ can be written in the form

$$\Psi_{\lambda} = \sum_{S_z}\int A_{S_z}\left(\boldsymbol{p}\right)\Psi_{p,\lambda,S_z} d^3\boldsymbol{p}. \tag{5.82}$$

Then

$$\hat{\Lambda}\Psi_{\lambda} = \sum_{S_z}\int A_{S_z}\left(\boldsymbol{p}\right)\frac{\hat{H}_f}{E_p}\Psi_{p,\lambda,S_z} d^3\boldsymbol{p} = \lambda\Psi_{\lambda}. \tag{5.83}$$

We can use $\hat{\Lambda}$ in order to introduce the projection operators $\hat{\Lambda}_{\pm}$ by

$$\hat{\Lambda}_{\pm} = \frac{1}{2}\left(1 \pm \hat{\Lambda}\right) \tag{5.84}$$

with the usual properties

$$\hat{\Lambda}_{+}\Psi_{\lambda=+1} = \Psi_{\lambda=+1},$$

$$\hat{\Lambda}_{+}\Psi_{\lambda=-1} = 0,$$

$$\hat{\Lambda}_- \Psi_{\lambda=+1} = 0,$$

$$\hat{\Lambda}_- \Psi_{\lambda=-1} = \Psi_{\lambda=-1}. \qquad (5.85)$$

The operators $\hat{\Lambda}_\pm$ split off the positive (or negative) parts of the states to which they are applied. We call operators "even" or "odd" if they transform positive (negative) functions into positive (negative) or negative (positive) functions, respectively. The product of two even or two odd operators is always an even operator, and the product of an even and an odd operator is always an odd operator. Since all positive functions ($\lambda$ = +1) are orthogonal with respect to all negative ($\lambda$ = −1) functions, the expectation value of an odd operator with states of fixed $\lambda$ is always zero; hence,

$$\left\langle \Psi_\lambda \,\middle|\, \hat{A}_{odd} \,\middle|\, \Psi_\lambda \right\rangle = 0.$$

A consistent one-particle theory can only use states with specified sign (either $\lambda$ = +1 or $\lambda$ = −1), because the energy can only be defined meaningfully in that way. However, from that it follows that in a consistent one-particle theory all physical quantities must necessarily be defined by even operators. In the following we shall see that for Dirac's theory, Ehrenfest's theorems follow under this condition too, i.e., the quantum-mechanical operator equations and the corresponding classical equations become identical. This means that the mean values comply with the classical equations, a fact which is quite significant. We formalize these considerations by splitting up every operator $\hat{A}$ into an even $[\hat{A}]$ and an odd $\{\hat{A}\}$ part

$$\hat{A} = \mathrm{h}\hat{A}\mathrm{i} + \mathrm{n}\hat{A}\mathrm{o}. \text{ If we} \tag{5.86}$$

simply write in short form $\Psi_{\pm}$ for $\Psi_{\lambda=\pm 1}$ we obtain

$$\hat{A}\Psi_{+} = \mathrm{h}\hat{A}\mathrm{i}\Psi_{+} + \mathrm{n}\hat{A}\mathrm{o}\Psi_{+}, \tag{5.87}$$



$$\hat{A}\Psi_{-} = \mathrm{h}\hat{A}\mathrm{i}\Psi_{-} + \mathrm{n}\hat{A}\mathrm{o}\Psi_{-}, \tag{5.88}$$

$$\hat{\Lambda}\hat{A}\hat{\Lambda}\hat{\Psi}_{+} = \hat{\Lambda}\hat{A}\hat{\Psi}_{+} = \mathrm{h}\hat{A}\mathrm{i}\Psi_{+} - \mathrm{n}\hat{A}\mathrm{o}\Psi_{+}, \tag{5.89}$$

$$\hat{\Lambda}\hat{A}\hat{\Lambda}\hat{\Psi}_{-} = -\hat{\Lambda}\hat{A}\hat{\Psi}_{-} = \mathrm{h}\hat{A}\mathrm{i}\Psi_{-} - \mathrm{n}\hat{A}\mathrm{o}\Psi_{-}. \tag{5.90}$$

Hence , it follows by, say, addition and subtraction of the first and the third or, also the second and the fourth in this set of last four equations that

$$\left[\hat{A}\right] = \frac{1}{2}\left(\hat{A} + \hat{\Lambda}\hat{A}\hat{\Lambda}\right), \qquad \left\{\hat{A}\right\} = \frac{1}{2}\left(\hat{A} - \hat{\Lambda}\hat{A}\hat{\Lambda}\right). \tag{5.91}$$

We immediately recognize that the free Dirac Hamiltonian $\hat{H}_f$ is an even operator since

$$\hat{\Lambda}\hat{H}_f\hat{\Lambda} = \hat{H}_f \quad and therefore \quad \left[\hat{H}_f\right] = \frac{1}{2}\left(\hat{H}_f + \hat{\Lambda}\hat{H}_f\hat{\Lambda}\right).$$

### 5.2.1 Lagrange Density and Energy-Momentum Tensor of the Free Dirac Equation

We now determine the Lagrangian density of the free Dirac field , we calculate the energy-momentum tensor and interpret the individual results.

First we claim that the Dirac Lagrange density has the form

$$\mathcal{L} = \bar{\psi}\left(ci\hbar\gamma^{\mu}\partial_{\mu} - m_0c^2\right)\psi. \tag{5.92}$$

$\bar{\psi} := \psi^\dagger\gamma^0$ is called the spinor adjoint to $\psi$ and the abbreviation $\gamma^\mu$ stands for $\gamma^0 = \hat{\beta}$, $\hat{\gamma}^i = \hat{\beta}\hat{\alpha}^i$ [cf. Chap.3, (3.91)]. Also $\partial_\mu := \partial/\partial x^\mu$ is a short hand notation. Straight way we introduce these $\gamma^\mu$ matrices instead of the $\hat{\alpha}^i$ and $\hat{\beta}$, but will revert to the $\hat{\alpha}^i$, $\hat{\beta}$ representation in all important sections. The $\gamma^\mu$ are appropriate for the covariant formulation of the Dirac equation. The Lagrangian density (5.92) can therefore be rewritten as

$$\mathcal{L} = \psi^\dagger i\hbar\partial_t\psi + c\psi^\dagger i\hbar\gamma^0\gamma^i\partial_i\psi - m_0c^2\psi^\dagger\gamma^0\psi, \tag{5.93}$$

since $\gamma_0 = \gamma^0$. Because of $\partial_i = \partial/\partial x^i = (\nabla)_i$, this yields

$$\begin{aligned}\mathcal{L} &= \psi^\dagger\left(i\hbar\partial_t + i\hbar c\hat{\alpha}\cdot\boldsymbol{\nabla} - m_0c^2\gamma^0\right)\psi \\ &= \psi^\dagger\left(i\hbar\partial_t - c\hat{\alpha}\cdot\boldsymbol{p} - \hat{\beta}m_0c^2\right)\psi,\end{aligned} \tag{5.94}$$

which uses the $\hat{\alpha}^i$, $\hat{\beta}$ matrices instead of the $\gamma^\mu$. We realize that this is the correct Lagrangian density if we determine the equations of motion by variation. Variation with respect to $\bar{\psi}$ yields

$$\begin{aligned}\frac{\delta\int\mathcal{L}d^4x}{\delta\bar{\psi}} &= 0 \Longrightarrow \frac{\partial\mathcal{L}}{\partial\bar{\psi}} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\bar{\psi}\right)} = 0 \\ &\Longrightarrow \left(ci\hbar\gamma^\mu\partial_\mu - m_0c^2\right)\psi = 0.\end{aligned} \tag{5.95}$$

This is Dirac's equation

$$i\hbar\partial_t\psi = \left(c\hat{\alpha}\cdot\hat{p} + \hat{\beta}m_0c^2\right)\psi \tag{5.96}$$

$$\equiv \hat{H}_f\psi \tag{5.97}$$

with the free Hamiltonian

$$\hat{H}_f = \left( c\hat{\alpha} \cdot \hat{p} + \hat{\beta} m_0 c^2 \right). \tag{5.98}$$

We recognize that for the solutions of the equations of motion (5.98)

$$\delta L(\psi) \equiv 0. \tag{5.99}$$



Variation with respect to $\psi$ yields

$$\begin{aligned} \frac{\delta \int \mathcal{L} d^4x}{\delta \psi} &= 0 \Longrightarrow \frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} = 0 \\ &\Longrightarrow \bar{\psi} \left( ci\hbar \gamma^\mu \overleftarrow{\partial_\mu} + m_0 c^2 \right) = 0, \end{aligned} \tag{5.100}$$

where $\overleftarrow{\partial_\mu}$ acts to the left on $\bar{\psi}$. One can readily calculate the canonical energymomentum tensor from the Lagrangian density L :

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} \partial_\nu \psi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \bar{\psi})} \partial_\nu \bar{\psi} - \delta^\mu_\nu \mathcal{L}, \tag{5.101}$$

which follows explicitly with (5.92) as

$$T^\mu_\nu = \bar{\psi} i\hbar c \gamma^\mu \partial_\mu \psi - \delta^\mu_\nu \bar{\psi} i\hbar c \gamma^\sigma \partial_\sigma \psi + \delta^\mu_\nu m_0 c^2 \bar{\psi} \psi. \tag{5.102}$$

From that one obtains the energy density $T^0_0$,

$$\begin{aligned} T^0_0 &= -\psi^\dagger i\hbar \hat{\alpha} \cdot \boldsymbol{\nabla} c\psi + m_0 c^2 \bar{\psi} \psi \\ &= \psi^\dagger \left( \hat{\alpha} \cdot \hat{p} c + \hat{\beta} m_0 c^2 \right) \psi = \psi^\dagger \hat{H}_f \psi. \end{aligned} \tag{5.103}$$

Consequently

$$\int T_0^0 d^3x = \left\langle \psi \mid \hat{H}_f \mid \psi \right\rangle \tag{5.104}$$

that is, the energy is just the expectation value of $\hat{H}_f$ in the state $\psi$. By analogy the momentum density $T_i^0$ is given by

$$T_i^0 = \bar{\psi} i\hbar c\gamma^0\partial_i\psi = \psi^\dagger (\hat{p})_i c\psi, \tag{5.105}$$

in other words, $p_i = (1/c)\int T_i^0 d^3x = \langle\psi \mid (\hat{p})_i \mid \psi\rangle$ is the expectation value of the momentum operator in the state $\psi$. The components

$$T_j^i = \bar{\psi} i\hbar c\left(\gamma^j\partial_j - \delta_j^i\gamma^\mu\partial_\mu\right)\psi + \delta_j^i m_0 c^2\bar{\psi}\psi, \tag{5.106}$$

which can simply be written for each solution of the equation of motion as

$$T_j^i = \bar{\psi} i\hbar c\gamma^i\partial_j\psi = -\bar{\psi}\gamma^i\hat{p}_j c\psi = -\psi^\dagger\hat{\alpha}^i\hat{p}_j c\psi, \tag{5.107}$$

are called the components of the *stress-strain tensor*. The trace of $T_\nu^\mu$ is given by

$$\begin{aligned} T \equiv T_\mu^\mu &= \bar{\psi} i\hbar c\left(\gamma^\mu\partial_\mu - 4\gamma^\sigma\partial_\sigma\right)\psi + 4m_0c^2\bar{\psi}\psi \\ &= -3\bar{\psi}\left(i\hbar c\gamma^\mu\partial_\mu - m_0c^2\right)\psi + m_0c^2\bar{\psi}\psi \end{aligned} \tag{5.108}$$

which, for every solution of the equation of motion, just becomes

$$T = m_0c^2\bar{\psi}\psi. \tag{5.109}$$

One should notice that this is not proportional to the charge density

$$\rho = e\bar{\psi}\gamma^0\psi = e\psi^\dagger\psi. \tag{5.110}$$

### 5.2.2 Energies of the Solutions to the Free Dirac Equation in the Canonical Formalism

We calculate in the framework of the canonical formalism the energies of the solutions (5.72)-(5.73) of the Dirac equation.

To start, the energy can be determined as the integral

$$E = \int_V T_0^0 d^3x. \tag{5.111}$$

If we insert (5.103) as well as the solutions (5.72) into this, we obtain

$$\begin{aligned} E &= N^2 \int_V \left(1, 0, \frac{pc}{m_0c^2 + \lambda E_p}, 0\right) \\ &\quad \times \left(\hat{\alpha} \cdot \hat{p}c + \hat{\beta} m_0 c^2\right) \begin{pmatrix} 1 \\ 0 \\ \frac{pc}{m_0c^2+\lambda E_p} \\ 0 \end{pmatrix} d^3x. \end{aligned} \tag{5.112}$$

Since for (5.72)-(5.73) $p = (0,0,p)$ holds, we get by calculation

$$
\begin{aligned}
E &= N^2 V \left(1, 0, \frac{pc}{m_0c^2 + \lambda E_p}, 0\right) \\
&\quad \times \left[ pc \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} + m_0c^2 \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \right] \begin{pmatrix} \\ 0 \\ \frac{pc}{m_0c^2+\lambda E_p} \\ 0 \end{pmatrix} \quad 1 \\
&= N^2 V \left(1, 0, \frac{pc}{m_0c^2 + \lambda E_p}, 0\right) \\
&\quad \times \left[ pc \begin{pmatrix} \frac{pc}{m_0c^2+\lambda E_p} \\ 0 \\ 1 \\ 0 \end{pmatrix} + m_0c^2 \begin{pmatrix} \\ 0 \\ \frac{-pc}{m_0c^2+\lambda E_p} \\ 0 \end{pmatrix} \right] \quad 1 \\
&= N^2 V \left[ \frac{2p^2c^2}{m_0c^2 + \lambda E_p} + m_0c^2 \frac{m_0c^2p^2c^2}{(m_0c^2 + \lambda E_p)^2} \right] \\
&= N^2 V (m_0c^2 + \lambda E_p)^{-2} (2p^2c^2m_0c^2 + 2p^2c^2\lambda E_p \\
&\quad + m_0^3c^6 + 2m_0^2c^4\lambda E_p + m_0^3c^6 + m_0c^2p^2c^2 - m_0c^2p^2c^2) \\
&= N^2 V (m_0c^2 + \lambda E_p)^{-2} [2(\lambda E_p)^3 + 2m_0^3c^6 + 2m_0c^2p^2c^2] \\
&= N^2 V (m_0c^2 + \lambda E_p)^{-1} 2\lambda^2 E_p^2. \qquad (5.113)
\end{aligned}
$$

Here we have used wave functions which are normalized with respect to a finite spherical volume *V*. If we furthermore insert the normalization factor *N* from (5.63)

$$
N^2 = V^{-1} \left( \frac{2\lambda E_P}{m_0c^2 + \lambda E_p} \right)^{-1}, \qquad (5.114)
$$

we finally obtain

$$E = \lambda E_p,$$

that means

$$E = +\sqrt{p^2c^2 + m_0^2c^4}$$

for the upper energy continuum ($\lambda = +1$) and

$$E = -\sqrt{p^2c^2 + m_0^2c^4}$$

for the lower energy continuum ($\lambda = -1$). Hence, the free Dirac equation leads to states with positive and negative energy.

### 5.2.3 Calculation of the Dirac Hamiltonian and the Velocity Operator

As for the Dirac Hamiltonian $\hat{H}_f$, for the momentum operator $\hat{p}$ also

$$\hat{\Lambda}\hat{p}\hat{\Lambda} = \hat{p}, \quad \mathit{so\,that} \quad [\hat{p}] = \hat{p}.$$

Let us also determine the even part of the $\hat{\alpha}$ operator. We have

$$\hat{\Lambda}\hat{\alpha}_i\hat{\Lambda} = \frac{c\hat{\alpha}\cdot\hat{p} + \hat{\beta}m_0c^2}{c\sqrt{\hat{p}^2 + m_0^2c^2}}\hat{\alpha}_i\frac{c\hat{\alpha}\cdot\hat{p} + \hat{\beta}m_0c^2}{c\sqrt{\hat{p}^2 + m_0^2c^2}} = -\hat{\alpha}_i + 2c\hat{p}_i\frac{\hat{\Lambda}}{c\sqrt{\hat{p}^2 + m_0^2c^2}},$$

and therefore

$$[\hat{\alpha}_i] = \frac{1}{2}\left(\hat{\alpha}_i + \hat{\Lambda}\hat{\alpha}_i\hat{\Lambda}\right) = c\hat{p}_i\frac{\hat{\Lambda}}{c\sqrt{\hat{p}^2 + m_0^2c^2}}. \tag{5.115}$$

In similar way one calculates

$$\left[\hat{\beta}\right] = \frac{m_0c^2}{E_p}\hat{\Lambda} = m_0c^2\frac{\hat{\Lambda}}{c\sqrt{\hat{p}^2 + m_0^2c^2}}. \tag{5.116}$$

Now we determine the *velocity operator* in the Dirac theory. Since we are concerned with an equation of Shrˆodinger type (3.66), the theorems for the time derivatives of operators which are formulated in (5.77) are also valid here and we obtain

$$\begin{aligned}\frac{d\hat{x}}{dt} &= \frac{1}{i\hbar}\left[\hat{x}, \hat{H}_f\right] = \frac{1}{i\hbar}\left[\hat{x}, c\hat{\alpha}\cdot\hat{p} + \hat{\beta}m_0c^2\right] \\ &= \frac{c}{i\hbar}\left[\hat{x}, \hat{\alpha}\cdot\hat{p}\right] = c\hat{\alpha} \equiv \hat{v}.\end{aligned} \tag{5.117}$$

Notice that the eigenvalues of $\hat{\alpha}$ have the values ±1, and so we obtain here the *"seeming" paradoxical result* that the absolute value of the velocity of a relativistic spin-$\frac{1}{2}$ particle always equals the velocity of light.[1] Of course it may be argued that since the $\hat{\alpha}_i$ do not commute with each other one may observe the components of the velocity $dx_i/dt$ would not be simultaneously measurable. However, quite interesting, in the examination (below) of the nonrelativistic limit of the Dirac equation, this seeming paradoxical result is once more confirmed, obtains support, and thus requires scrutiny.

[1] This was first recognized by Gregory Breit in 1928.

### 5.2.4 Nonrelativistic Limit of the Dirac Equation

It is important to check whether the Dirac equation yields physically reasonable results in the nonrelativistic limiting case. First we study the case of an electron at rest; in this case we obtain the Dirac equation by setting $\hat{p}\psi = 0$ in (3.66),

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{\beta} m_0 c^2 \psi. \tag{5.118}$$

In the particular representation (3.76) with

$$\hat{\beta} = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}$$

we are able to write down the four solutions

$$\psi^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \exp\left[-i\left(m_0c^2/\hbar\right)t\right], \quad \psi^{(2)} = \begin{pmatrix} \\ \\ \\ \end{pmatrix} \exp\left[-i\left(m_0c^2/\hbar\right)t\right],$$

$$\psi^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \exp\left[+i\left(m_0c^2/\hbar\right)t\right], \quad \psi^{(4)} = \begin{pmatrix} \\ \\ \\ \end{pmatrix} \exp\left[+i\left(m_0c^2/\hbar\right)t\right] \tag{5.119}$$

0

1

0

0

0

0

0

1

The first two wave functions correspond to positive, the last two to negative energy values. The interpretation of the solutions with negative energy will not be considered here. We restrict ourselves to solutions with positive energy. In order to show that the Dirac equation reproduces the two-component Pauli equation in the nonrelativistic limit, we introduce the electromagnetic four-potential

$$A^{\mu} = (A_0(x), \boldsymbol{A}(x)) \tag{5.120}$$

into the Dirac equation (3.66). It is known that the minimal coupling

$$\hat{p}^{\mu} \rightarrow \hat{p}^{\mu} - \frac{e}{c}A^{\mu} \equiv \hat{\Pi}^{\mu}$$

ensures gauge invariance of the theory, where $\hat{\Pi}^{\mu}$ is the kinetic momentum and $\hat{p}^{\mu}$ the canonical momentum. So we are inevitably guided to the Dirac equation with electromagnetic potentials

$$c\left(i\hbar\frac{\partial}{\partial ct} - \frac{e}{c}A_0\right)\psi = \left[c\hat{\alpha}\cdot\left(\hat{p} - \frac{e}{c}\boldsymbol{A}\right) + \hat{\beta}m_0c^2\right]\psi$$

or

$$i\hbar\frac{\partial\psi}{\partial t} = \left[c\hat{\alpha}\cdot\left(\hat{p}-\frac{e}{c}\boldsymbol{A}\right)+eA_0+\hat{\beta}m_0c^2\right]\psi. \tag{5.121}$$

This contains the interaction with the electromagnetic field

$$\hat{H}' = -\frac{e}{c}c\hat{\alpha}\cdot\boldsymbol{A}+eA_0 = -\frac{e}{c}\hat{v}\cdot\boldsymbol{A}+eA_0, \tag{5.122}$$

where

$$\hat{v} = c\hat{\alpha} \tag{5.123}$$

is the relativistic velocity operator. The expression (5.122) corresponds to the classical interaction of a moving charged point-like particle with the electromagnetic field. Once again notice that the velocity operator in (5.123) is the formal operator $\hat{v}$ from (5.117), which is equal to the speed of light.

The nonrelativistic limiting case of (5.121) can be most efficiently studied in the representation

$$\psi = \begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix}, \tag{5.124}$$

where the four-component spinor $\psi$ is decomposed into two two-component spinors $\tilde{\varphi}$ and $\tilde{\chi}$. Then the Dirac equation (5.121) becomes

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix} = \begin{pmatrix}c\hat{\sigma}\cdot\hat{\Pi} & \hat{\chi}\\ c\hat{\sigma}\cdot\hat{\Pi} & \hat{\varphi}\end{pmatrix}+eA_0\begin{pmatrix}\tilde{\varphi}\\ \tilde{\chi}\end{pmatrix}+m_0c^2\begin{pmatrix}\tilde{\varphi}\\ -\tilde{\chi}\end{pmatrix}, \tag{5.125}$$

the $\hat{\alpha}$ as well as the $\hat{\beta}$ matrices having been inserted according to (3.76). If the rest energy $m_0c^2$, as the largest occurring energy, is additionally separated by

$$\begin{pmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{pmatrix} = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \exp\left[-i\left(m_0c^2/\hbar\right)t\right] \quad , \tag{5.126}$$

then (5.125) takes the form

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \begin{pmatrix} c\hat{\sigma}\cdot\hat{\Pi} & \chi \\ c\hat{\sigma}\cdot\hat{\Pi} & \varphi \end{pmatrix} + eA_0\begin{pmatrix} \varphi \\ \chi \end{pmatrix} - 2m_0c^2\begin{pmatrix} 0 \\ \chi \end{pmatrix}. \tag{5.127}$$

Let us consider first the lower (second) of the above equations. For the conditions $|i\hbar\partial\chi/\partial t| \ll |m_0c^2\chi|$ and $|eA_0\chi| \ll |m_0c^2\chi|$ (i.e., if the kinetic energy as well as the potential are small compared to the rest energy) we obtain from the lower component of (5.127)

$$\chi = \frac{\hat{\sigma}\cdot\hat{\Pi}}{2m_0c}\varphi. \tag{5.128}$$

This mean that $\chi$ represents the small components of the wavefunction $\psi$, a result we already know from (5.59), while $\phi$ represents the large component $\chi$ $\chi \sim (v/2c)\phi$. Insertion of (5.128) into the first equation (5.127) results in a nonrelativistic wave function for $\psi$

$$i\hbar\frac{\partial\varphi}{\partial t} = \frac{\left(\hat{\sigma}\cdot\hat{\Pi}\right)\left(\hat{\sigma}\cdot\hat{\Pi}\right)}{2m_0}\varphi + eA_0\varphi. \tag{5.129}$$

With the help of (5.57) we continue the calculation,

$$\begin{aligned}
\left(\hat{\sigma}\cdot\hat{\Pi}\right)\left(\hat{\sigma}\cdot\hat{\Pi}\right) &= \hat{\Pi}^2 + i\hat{\sigma}\cdot\left(\hat{\Pi}\cdot\hat{\Pi}\right) \\
&= \left(\hat{p} - \frac{e}{c}\boldsymbol{A}\right)^2 + i\hat{\sigma}\cdot\left[\left(-i\hbar\boldsymbol{\nabla} - \frac{e}{c}\boldsymbol{A}\right)\times\left(-i\hbar\boldsymbol{\nabla} - \frac{e}{c}\boldsymbol{A}\right)\right] \\
&= \left(\hat{p} - \frac{e}{c}\boldsymbol{A}\right)^2 - \frac{e}{c}\hbar\hat{\sigma}\cdot\left(\boldsymbol{\nabla}\times\boldsymbol{A}\right)
\end{aligned}$$

$$= \left(\hat{p} - \frac{e}{c}\boldsymbol{A}\right)^2 - \frac{e\hbar}{c}\hat{\sigma} \cdot B$$

and finally obtain (5.129) in the form

$$i\hbar\frac{\partial\varphi}{\partial t} = \left[\left(\hat{p} - \frac{e}{c}\boldsymbol{A}\right)^2 /2m_0 - \frac{e\hbar}{2m_0 c}\hat{\sigma} \cdot \boldsymbol{B} + eA_0\right]\varphi. \tag{5.130}$$

Equation (5.130) is known as the *Pauli equation*. The two components of $\phi$, therefore, describe the spin degree of freedom, which we have already dealt with in the section concerning the free Dirac wave [cf. (5.65)-(5.73)]. It is well known that this Pauli equation yields the correct gyromagnetic factor of $g$ = 2 for a free electron. This can be demonstrated once again by turning on a weak, homogeneous magnetic field

$$B = curl A, \qquad A = \frac{1}{2}\boldsymbol{B} \times \boldsymbol{x},$$

where the quadratic terms of $A$ in (5.130) have been neglected. With

$$\begin{aligned}\left(\hat{p} - \frac{e}{c}\boldsymbol{A}\right)^2 &= \left(\hat{p} - \frac{e}{c}\boldsymbol{B} \times \boldsymbol{x}\right)^2 \approx \hat{p}^2 - \frac{e}{c}\left(\boldsymbol{B} \times \boldsymbol{x}\right) \cdot \hat{p} \\ &= \hat{p}^2 - \frac{e}{c}\boldsymbol{B} \cdot \boldsymbol{L},\end{aligned}$$

where $\hat{L} = x \times \hat{p}$ is the operator of orbital angular momentum, and

$$\hat{S} = \frac{1}{2}\hbar\hat{\sigma}$$

is the spin operator, it follows for the Pauli equation (5.130) that

$$i\hbar\frac{\partial\varphi}{\partial t} = \left[\frac{\hat{p}^2}{2m_0} - \frac{e}{2m_0 c}\left(\hat{L} + 2\hat{S}\right) \cdot \boldsymbol{B} + eA_0\right]\varphi. \tag{5.131}$$

This form shows explicitly the $g$ factor 2. However, in the nonrelativistic limit, the Dirac equation transforms into the Pauli equation, i.e., to the proper nonrelativistic wave equation for spin-1/2 particles. Since spin exists both at low as well as at high velocities, this implies that the Dirac equation describes particles with spin−1/2.

At this point, we reach the limit of our review of results sufficient for the purpose of this work. In the following, for our investigation, we propose, with no essential lost of generality, to restrict ourselves to realistic one-particle systems with positive energy, and we use the free Dirac waves propagating in the $z$−direction.

## 5.3 Superluminal Motion of Free Spin-1/2 Particles

We utilize the symmetrized Dirac Lagrange density. Our interest here concerns two results of Proposition 5.1; they are (5.39) of variations in mixed space-time components of orbital angular momentum, and (5.44) of spin angular momentum. Consider waves propagating in the $z$−direction, so that the index $l$ = 3; we rewrite, using functional notation (Gelfand et al., 2000), these equations as

$$\delta L_{03}\left[\psi\left(x\right),\psi^{\dagger}\left(x\right)\right]=S_{03}\left[\psi\left(x\right),\psi^{\dagger}\left(x\right)\right]=\frac{1}{2}\int d^{3}x\psi^{\dagger}\left(x\right)\sigma_{03}\psi\left(x\right) \qquad (5.132)$$

$$\delta S_{03}\left[\psi\left(x\right),\psi^{\dagger}\left(x\right)\right]=-S_{03}\left[\psi\left(x\right),\psi^{\dagger}\left(x\right)\right]=-\frac{1}{2}\int d^{3}x\psi^{\dagger}\left(x\right)\sigma_{03}\psi\left(x\right). \qquad (5.133)$$

Inserting the solution of the Dirac equation (5.72) (with $\lambda$ = +1) into (5.132), computing and integrating, yield

$$
\begin{aligned}
\delta L_{03}(x) &= \frac{1}{2}\int d^3xN\begin{pmatrix} 1 & 0 & \frac{c\hat{\sigma}_z p}{m_0c^2+E_p} & 0 \end{pmatrix}\exp\left[-i\left(pz-E_pt\right)/\hbar\right] \\
&\quad\times\sigma_{03}N\begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \frac{c\hat{\sigma}_z p}{m_0c^2+E_p}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix}\exp\left[i\left(pz-E_pt\right)/\hbar\right] \\
&= \frac{1}{2}\sigma_{03}x_\mu = \frac{1}{2}\sigma_{03}\begin{pmatrix} t \\ z \\ y \\ x \end{pmatrix}.
\end{aligned}
\tag{5.134}
$$

Thus, $\delta L_{03}(x)$ is a linear dependence on the 4-vector $x_\mu = \begin{pmatrix} t \\ z \\ y \\ x \end{pmatrix}$ alone, along the $z$ – *axis.* Similarly, (5.133) simplifies to

$$
\delta S_{03}(x) = -\frac{1}{2}\sigma_{03}x_\mu = \frac{1}{2}\sigma_{03}\left(-x_\mu\right) = \frac{1}{2}\sigma_{03}\begin{pmatrix} -t \\ -z \\ -y \\ -x \end{pmatrix}. \tag{5.135}
$$

So, variation in spin angular momentum in (5.135) leads to negative temporal component ($t \rightarrow -t$), depicting backward in time motion (Figure 3.1): Clearly, physically interpreted, this is causality violation. Shore (2006) and Modanese (2013) discuss the velocity requirements for the possibility of such an occurrence. There exist suitable reference systems in relative motion where the temporal order



$_0$ of events appears to be reversed, i.e., the delay $\Delta t$ becomes negative; this is mathematically due to the singularity of the Lorentz transformations for $\beta \rightarrow 1^-$. But then, a legitimate question arises as to how fast must such a relative motion be, for the moving observer to see the temporal order of the two events reversed.

We show, by a simple exercise on Lorentz transformations that, if the two events appear to be connected by a superluminal signal with velocity $V$ (i.e., $\Delta x/\Delta t = V > c$ in the rest reference system), then the causality violation is observed in reference systems moving with relative velocity $v$ larger than $c^2/V$ (Modanese, 2013).

Namely, suppose that the superluminal signal is emitted at the origin ($x = 0$, $t = 0$) of a spacetime reference system. (Coordinates $y$ and $z$ are omitted, only boosts along $x$ are considered.) The signal is received at a point $R(x,t)$ outside the light-cone. We have $t < x/c$, and more exactly $t = x/V$ , where $V$ is the signal velocity ($V > c$). All this is in a reference system where emitter and receiver are at rest.

Now make a Lorentz boost with parameters $\beta, \gamma$ along $x$; the new coordinates in the "moving" system are $t^0=\gamma(t-\beta x/c)$, $x^0=\gamma(x-\beta ct)$. We look for values of $\beta$

0

such that $t < 0$, i.e., the event $R$ is seen in the new reference system as antecedent to the emission. (Note that the emission remains at (0,0) in the new coordinates).

0

Putting $t = 0$ we find $t = \beta x/c$ and remembering that $t = x/V$, we obtain $\beta = c/V$. This is the limit value for passive causality violation; for $\beta > c/V$ we will have $t_0 < 0$.

Moreover, extensions of Special Relativity exist, compatible with the relativity principle, which provide a complete framework for the kinematics of superluminal particles. The usual definitions of energy and momentum are extended by introducing an imaginary mass $m = iM$ :

$$E = \frac{mc^2}{\sqrt{1 - V^2/c^2}} = \frac{Mc^2}{\sqrt{V^2/c^2 - 1}}, \tag{5.136}$$

$$\mathbf{P} = \frac{mV}{\sqrt{1 - V^2/c^2}} = \frac{MV}{\sqrt{V^2/c^2 - 1}}. \tag{5.137}$$

Combining these last two equations one obtains

$$\frac{E}{\mathbf{P}} = \frac{c^2}{V}. \tag{5.138}$$

By this last equation, for an observer at rest in the reference system moving with relative velocity $v$, the kinetic energy is $E = (1/2)mv^2$ and momentum $\mathbf{P} = mv$, so that in (5.138), $E/P = v/2 = c^2/V \Longrightarrow v = 2c^2/V > c^2/V$. This completes our proof.

As a result, it is legitimate to regard the phenomenon described by (5.135) as a superluminal propagation. In contrast, variation in orbital angular momentum does not violate causality and so accords with special relativistic principles of subluminal motion.

### 5.3.1 Evaluation of the Dirac spinor

From the transformation law of Dirac spinors under infinitesimal Lorentz transformations given through (5.18)-(5.19) above, we have in particular

$$\sigma^{0k} = \frac{i}{2}\left(\gamma^0\gamma^k - \gamma^k\gamma^0\right) = i\gamma^0\gamma^k = i\alpha^k = i\begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \quad (5.139)$$

$$\sigma^{ij} = \epsilon^{ijk}\Sigma_k = \epsilon^{ijk}\begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}, \quad (5.140)$$

where we employ an obvious shorthand notation (each entry of the matrices is itself a 2 × 2 matrix), $\epsilon^{ijk}$ is the antisymmetric tensor defined in Section 4.1, and $\sigma_k$ are Pauli spin matrices (with $k$ = 1,2,3 ). In (5.139) when $k$ = 3, we have

$\sigma_{03} = i\alpha_3 = i\begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix}$ where $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ represents spin around the $z-axis$, so that

$$\sigma_{03} = i\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \quad (5.141)$$

?

0 −1 0

Next, if a contravariant 4−vector has components $A^{\mu} = (A^0,A^1,A^2,A^3)$, its covariant partner has components $A_0 = A^0$, $A_1 = -A^1$, $A_2 = -A^2$, and $A_3 = -A^3$. That is, if $A^{\mu} = (A^0,\mathbf{A})$ then the corresponding covariant vector is $A_{\mu} = (A^0,-\mathbf{A})$.

The 3-vector $\mathbf{A}$ (in boldface notation ) has components $(A^1,A^2,A^3)$. It follows, symbolically, concerning the 3-vector $\alpha^k$ in (5.139), that $\alpha^k = \{\alpha^1,\alpha^2,\alpha^3\} = \{-\alpha_1,-\alpha_2,-\alpha_3\}$, (Greiner, 1994). Now, from the contravariant equation $\sigma^{03} = i\alpha^3$ in (5.141), we have the corresponding covariant equation $\sigma_{03} = i\alpha_3 = -i\alpha^3$, that is

$$\sigma_{03} = i\left(-\alpha^3\right) = i\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{5.142}$$

Furthermore, the *polar form* of the complex number $(x,y)$ or $x+iy$ in the complex plane is $w = x + iy = r(\cos\theta + i\sin\theta)$ with $x = r\cos\theta$ and $y = r\sin\theta$, where

$r = \sqrt{x^2 + y^2} = |x + iy|$ is called the *modulus* of $w$ and $\theta$ the *amplitude* or *argument* of $w$. Specifically, the complex number $i = 1 \cdot [\cos(\pi/2) + i\sin(\pi/2)]$ in polar form or, $i = e^{i\pi/2}$, using Euler's formula, is unimodular and has amplitude or argument $\theta = \pi/2$. In this form, when it is applied (in multiplication) to any plane vector $(x,y)$, this represents a unimodular rotation transformation of angle $\theta = \pi/2$ about the origin O of the $xy$−coordinate system. Obviously, its application to a 3-vector $(x,y,z)$ in 3-space corresponds to a unimodular 3-dimensional rotation transformation of angle $\theta = \pi/2$ about the $z$−axis; so that the matrix of the transformation is unimodular and given by

$$\begin{pmatrix} \cos(\pi/2) & \sin(\pi/2) & 0 \\ -\sin(\pi/2) & \cos(\pi/2) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is an element of the group $SO(3)$ of $3 \times 3$ unimodular matrices of determinant 1

1. Notice that $SO(3)$ is a subgroup of $SO(1,3)$ that denotes the group of all unimodular $4 \times 4$ rotation matrices of determinant unity, also called the *Lorentz group*. Hence, in an extension to 4-dimensional space-time (1-time, 3-space), multiplication by the complex number $i$ corresponds to multiplying by the $4 \times 4$ matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\pi/2) & \sin(\pi/2) & 0 \\ 0 & -\sin(\pi/2) & \cos(\pi/2) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

1 0

Substituting this into (5.142), we obtain the Dirac spinor in the $z$−direction:

$$\sigma_{03} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{5.143}$$

We can now compute the matrix of the variation $\delta S_{03}$. We obtain

$$M_{\delta S_{03}} = -\frac{1}{2}\sigma_{03} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix}. \tag{5.144}$$

Similarly the matrix of the variation $\delta L_{03}$ is

$$M_{\delta L_{03}} = \frac{1}{2}\sigma_{03} = \begin{pmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix}. \quad (5.145)$$

### 5.3.2 Determination of the expectation of trajectory of the free spin-1/2 field

The propagation of the fermion field is 'expected' to be a helical motion and helicity is invariant under spatial rotations, (Greiner, 1994, Guildry, 1991, Ohanian, 1986) (see Figure 5.2). Based on Ehrenfest's theorems, the *expectation of position* or, precisely, the *expectation of trajectory* of the spin-1/2 field corresponds to the classical 3-space circular (right-handed) helix parametrized by the equations:

$$x_1 = a\cos t, \quad x_2 = a\sin t, \quad x_3 = bt, \quad a,b \neq 0, \quad -\infty < t < \infty, \quad (5.146)$$

which lead to the parametric derivative equations:

$$dx_1/dt = -a\sin t, \quad dx_2/dt = a\cos t, \quad dx_3/dt = b. \quad (5.147)$$

Observe here that the constant $b$ represents boost (or velocity) in the $z$−direction.

### 5.3.3 Calculation of the Expectation Value of the Overall Relative Linear Velocity Component

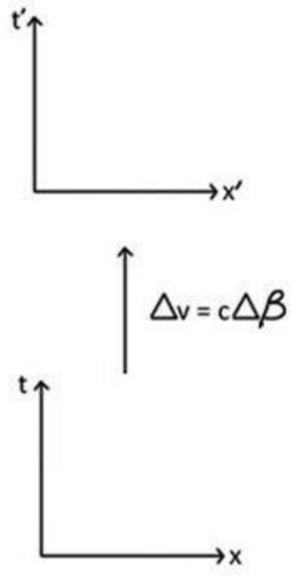


Figure 5.4: Two reference systems: Lorentz boost in the $z-$direction.

Let $V_S$ and $v_L$ be the expectation values of the linear velocity components under the transformations $M_{\delta S_{03}}$ amd $M_{\delta L_{03}}$, respectively. We must understood that every transformation from one reference system to another (in the light cone) is a velocity transformation or Lorentz boost (see Figure 5.4), from which the constant transformation matrices $M_{\delta S_{03}}$ and $M_{\delta S_{03}}$ originate. The matrices in (5.144) must be regarded as solutions to some generalized velocity transformations which remained to be determined. Since the variations $\delta S_{03}\,(x)$ and $\delta L_{03}\,(x)$ are lin-

$_0$ ear, they carry a vector $M_{\delta S_{03}}\mathbf{x}'_{\mu}$ $M_{\delta L_{03}}\mathbf{x}'_{\mu}$ $\mathbf{x}_{\mu}$
into its scalar multiplesandunder

their corresponding matrices $M_{\delta S_{03}}$ and $M_{\delta L_{03}}$ both of which have the same eigenvalue of $\pm 1/2$. Furthermore, to get the clear picture of the transformation of $t$

matrix $M_{\delta S_{03}}$, observe that $M_{\delta S_{03}}$ maps the 4-vector $x_\mu = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$ unto its image

$$x'_\mu = \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \frac{1}{2}y \\ \frac{1}{2}t \\ \frac{1}{2}z \\ -\frac{1}{2}x \end{pmatrix},$$

i.e., $t \mapsto t' = \frac{1}{2}y$, $x \mapsto x' = \frac{1}{2}t$, $y \mapsto y' = \frac{1}{2}z$, and $z \mapsto z' = -\frac{1}{2}x$. So the position of each coordinate of the image vector is found on a different coordinate axis, as if the coordinate axes have been altered or somehow 'interchanged' under the effect of $M_{\delta S_{03}}$. This scenario describes a quadratic (or a rotation) transformation of the coordinate axes in four dimensions. A prediction can be made of this situation from (5.144) where the matrix $M_{\delta S_{03}}$ has been obtained as a rotation matrix since, it is the product of the Dirac Spinor $\sigma_{03}$ which is a spin (i.e., rotation) matrix in 4-dimensions and the factor 1/2 which emanates from the *giro factor* of intrinsic spin. Bearing in mind our previous remarks, we deduce that the nonzero entries of this matrix, which are all equal in the absolute value of 1/2, are expressible in a unique sine or cosine function of argument $V_S/c$ (representing 'boost'), where $c$ is the speed of light. This is also valid for the transformation $M_{\delta L_{03}}$. By inspection and based on our previous discussions, due to the superluminal propagation under spin, we will assign the cosine function to $M_{\delta S_{03}}$ and the sine function to the subluminal variation $\delta L_{03}$.

(1) Explicitly, for variation under spin angular momentum, we have

$$M_{\delta S_{03}} = \begin{pmatrix} 0 & 0 & \cos(V_S/c) & 0 \\ \cos(V_S/c) & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos(V_S/c) \\ 0 & -\cos(V_S/c) & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix}, \tag{5.148}$$

that is,

$$\cos(V_S/c) = \frac{1}{2},$$

$$\Longrightarrow V_S = \left(\frac{\pi}{3} + 2k\pi\right) \times c \tag{5.149}$$

where $k = 0,1,2,...$ Hence, the expectation value of the relative linear velocity component produced by spin angular momentum is superluminal, and its minimal value is

$$V_{S_{min}} = (\pi/3)c = (1.047....) \times c, \tag{5.150}$$

which also exceeds the speed of light.

(2) Similarly, for variation under orbital angular momentum, we have

$$M_{\delta L03} = \begin{pmatrix} 0 & 0 & -\sin(v_L/c) & 0 \\ -\sin(v_L/c) & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sin(v_L/c) \\ 0 & \sin(v_L/c) & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix}, \tag{5.151}$$

that is,

$$\sin(v_L/c) = \frac{1}{2},$$

$$\Longrightarrow v_L = \left(\frac{\pi}{6}\right) \times c \simeq (0.524) \times c. \tag{5.152}$$

Notice here that we must ignore the additive $\pm 2k\pi$ term in order to keep $v_L$ subluminal.

(3) Furthermore, let $V$ denote the expectation value of the overall relative linear velocity component of the field. It should also be understood that the transformation of the generalized angular momentum in mixed space-time components (let us denote it by F ) is the composition of the two transformations spin $M_{\delta S_{03}}$ and orbital $M_{\delta L_{03}}$. Based on our previous discussions, we have

$$\begin{aligned} F(V) &= M_{\delta L_{03}} \circ M_{\delta S_{03}} = M_{\delta L_{03}} \cdot M_{\delta S_{03}} \\ &= \begin{pmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & -\frac{1}{4} \\ 0 & 0 & -\frac{1}{4} & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

0 0

$$0 \quad = \begin{pmatrix} & & & -\cos(V/c) \\ 0 & 0 & -\cos(V/c) & 0 \\ 0 & \cos(V/c & & \\ \cos(V/c & & & \end{pmatrix}, \quad (5.153)$$

) 0 0

) 0 0 0

that is,

$$\cos\left(V/c\right) = \frac{1}{4}$$

$$\Longrightarrow V \simeq \left(\frac{21\pi}{50} + 2k\pi\right) \times c, \quad (k = 0, 1, 2, \ldots) \, . \tag{5.154}$$

Thus, the expectation value of the overall relative lnear velocity component of a spin-1/2 particle is superluminal and, its minimal value is

$$V_{min} = (1.318\ldots) \times c, \tag{5.155}$$

this also exceeds light speed, as expected, completing our investigation.

Observe that $V$ takes discrete values (one at a time with respect to $k = 0,1,2\ldots$). Each value of $V$ may be thought of as corresponding to the level at which the electron is spinning. But then, a legitimate query arises here: Must $k = 0,1,2,\ldots$ be left to run indefinitely (that is, from zero to infinity)? In other words: Is there any constraint on the velocity of the superluminal electron? For an answer to this question, see, for example, Stecker (2014) who studies and discusses the limits on superluminal electron and neutrino velocities.

## 5.4 Superluminal Motion of Free Spinning Fermions in the Fiber Bundle Formalism

The invariance under spatial rotations of the spin-1/2 fields relates to the Invariance Reduction Theorem (IRT) which renders the existence and uniqueness problems of invariant fields into a problem of Topological Dynamics involving techniques from Ergodic Theory and Homotopy Theory. We will now, briefly, give a fiber bundle formalism to our theory of superluminal free electron. To accomplish this, we employ a method developed in the 1980s by R. Zimmer and others for Dynamical-Systems theory. This allows us to generalize the notions of particle-spin motion and field motion (Heinemann et al., 2015). We employ a discrete-time formalism (but a continuous-time treatment would do as well). Four major theorems are presented, the Decomposition Theorem, which allows one to compare different invariant fields, the Invariant Reduction Theorem, which gives new insights into the existence and uniqueness problems of invariant fields (and in particular invariant spin fields), the Cross Section Theorem which supplements the Invariant Reduction Theorem, and the Normal Form Theorem which ties invariant fields with the notion of normal form. It turns out that the well established notions of invariant polarization field and invariant spin field are generalized to invariant $(E,l)$-fields. Here the notation $(E,l)$ will mean that $E$ is a topological space and that the function $l : SO(3)\times E \to E$ is a continuous $SO(3)$–action, i.e., $l(I;x) = x$ and $l(r_1r_2;x) = l(r_1;l(r_2;x))$, where $I$ is the identity 3-dimensional (3-D) rotation element and $r_i$ $(i = 1,2)$ are 3-D rotation matrices. With the flexibility in the choice of $(E,l)$ we have a unified way to study the dynamics of spin-1/2 particles. Accordingly the special cases ($(E, l) = (\mathbb{R}^3, l_{1/2})$ is discussed in some detail.

**Definition 5.2.** A fiber bundle is a structure $(E,B,\pi,F)$, where $E$, $B$, and $F$ are topological spaces and $\pi : E \to B$ is a continuous surjection satisfying a local triviality condition outlined below. The space $B$ is called the *base space* of the bundle, $E$ the *total space*, and $F$ the *fiber*. The map $\pi$ is called the *projection map* (or *bundle projection*). We shall assume in what follows that the base space $B$ is connected.

We require that for every $x$ in $E$, there is an open neighborhood $U \subset B$ of $\pi$ $(x)$ (which will be called a *trivializing neighborhood*) such that there is a homeomorphism $\phi$: $\pi^{-1}(U) \to U \times F$ (where $U \times F$ is the product space) in such a way that $\pi$ agrees with the projection onto the first factor. That is, the following diagram

(Figure 5.5) should commute: where $proj_1 : U \times F \to U$ is the natural projection and $\phi$: $\pi^{-1}(U) \to U \times F$ is a homeomorphism. The set $\{(U_i,\phi_i)\}$ of all $(U_i,\phi_i)$ is

called a *local trivialization* of the bundle.

$$\begin{tikzcd}
\pi^{-1}(U) \arrow[r, "\varphi"] \arrow[d, "\pi"'] & U \times F \arrow[dl, "\mathrm{proj}_1"] \\
U &
\end{tikzcd}$$

Figure 5.5: Commutation diagram

Thus for any $p$ in $B$, the preimage $\pi^{-1}(\{p\})$ is homeomorphic to $F$ (since $proj_1(\{p\})$ clearly is) and is called the fiber over $p$. Every fiber bundle $\pi : E \to B$ is an open map, since projections of products are open maps. Therefore $B$ carries the quotient topology determined by the map $\pi$.

A fiber bundle $(E,B,\pi,F)$ is often denoted $F \longrightarrow E \xrightarrow{\pi} B$ which, in analogy with a short exact sequence, indicates which space is the *fiber*, *total space* and *base space*, as well as the map from total to base space.

A smooth fiber bundle is a fiber bundle in the category of smooth manifolds. That is, $E$, $B$, and $F$ are required to be smooth manifolds and all the functions above are required to be smooth maps.

Fiber bundles are the natural language for the description of a gauge theory. Locally a *principal bundle* **P** is the product of a *structure group* (i,e., gauge group) with the base space (space-time), but globally twists can appear given by the translation functions. For example, a cylinder is the trivial product of a line with a circle. A M¨*o*bius strip would be a nontrivial bundle with one $180^o$ twist as the line went around the circle. Locally a M¨*o*bius strip and a circle are identical but globally quite different.

### 5.4.1 Homotopy in the Rotation and Lorentz Groups

The description of intrinsic spin, whether for bosons or for fermions, is in terms of fiber bundles with an $SO(3)$ structure group. It is clearly of some interest then to understand the "loop structure", i.e., the fundamental group, of $\mathbb{R}^3 \times SO(3)$. Notice that $SO(3)$ does indeed have a natural topology. The entries of a $3 \times 3$ matrix can be strung out into a column matrix which can be viewed as a point in $\mathbb{R}^9$ (Naber, 1992). Thus, $SO(3)$ can be viewed as a subset of $\mathbb{R}^9$ and therefore inherits a topology as a subspace of $\mathbb{R}^9$. A considerably more informative "picture" of $SO(3)$ can be obtained as follows: Every rotation of $\mathbb{R}^3$ can be uniquely specified by an axis of rotation, an angle and a sense of rotation about the axis. All of this information can be codified in a single object, namely, a vector in $\mathbb{R}^3$ of magnitude at most $\pi$. Then the axis of rotation is the line along $\vec{n}$, the angle of rotation is $|\vec{n}|$ and the sense is determined by the "right hand rule". Notice that a rotation along $\vec{n}$ through an angle $\theta$ with $\pi \leq \theta \leq 2\pi$ is equivalent to a rotation along $-\vec{n}$ through $2\pi-\theta$ so the restriction on $|\vec{n}|$ is necessary (although not quite sufficient) to ensure that the correspondence between rotations and vectors be one-to-one. The set of vectors $\vec{n}$ in $\mathbb{R}^3$ with $|\vec{n}| \leq \pi$ is just the closed ball of radius $\pi$ about the origin. However, a rotation about $\vec{n}$ through $\pi$ is the same as a rotation about $-\vec{n}$ through $\pi$ so *antipodal points* on the boundary of this ball represent the same rotation and therefore must be identified in order that this correspondence with rotations be bijective. Carrying out this identification yields real projective 3−space (topologically, the radius of the ball is irrelevant, of course).

Having found the structure group of our fiber bundle, let us consider the appropriate base space. Let us consider the simple case of static particles with spin. A very natural choice for the base space is thus $\mathrm{R}^3$ minus the point at the origin where the particle is assumed to be. Thus we take our base space to be $\mathrm{R}^3 - \{0\}$. Note that Minkowski space minus the world line of a particle is contractible to $\mathrm{R}^3 - \{0\}$ which, in turn, can be retracted to $S^2$ sphere without changing the fiber bundle. Thus we are led to a principal fiber bundle, **P**, with structure group $SO\,(3)$ and base space the two-sphere, $S^2$, in connection with describing static particles with spin. The present work has nothing to do with static spinning fermions, of course.

In our present case of moving spin-1/2 particles, we consider the torus $T^d$ as the locus of the position $z$ of our particle with spin-1/2. It is known that $T^d = S^d \times S^d$, where $S^d$ is the $d$−sphere and, generally $d$ = 1,2,3. The "unreduced" principal bundle underlying our formalism is a product principal bundle $(T^d \times SO(3), p_d, T^d, R_d)$ with bundle space $T^d \times SO(3)$, base space $T^d$, bundle projection $p_d(z,r) := z$ (with $z$ in $T^d$ and $r \in SO(3)$), and structure group $SO(3)$. So $R_d : SO(3) \times T^d \times SO(3) \rightarrow T^d \times SO(3)$ is an $SO(3)$−action defined by $R_d(r; z^0, r^0) := (z^0,\ r^0 r^t)$, where the upper index $t$ means "transpose". The reductions are just the principal subbundles of the unreduced bundle. So they are uniquely determined by their bundle space $X$ which of course is a subgroup of $T^d \times SO(3)$.

### 5.4.2 Particle-spin Motion

For given $(E,l)$ each particle carries, in addition to its position $z$ on the torus $T^d$ an $E$−valued quantity $x$ we call spin. The one-turn particle-spin map is the function $P[j,A] : T^d \times E \to T^d \times E$, defined by

$$P[j,A](z,x) = (j(z), l(A(z),x)), \tag{5.156}$$

where $j \in Homeo(T^d)$ is the one-turn particle map (e.g., linear translation on the torus) and $A \in C(T^d, SO(3))$ is the one-turn spin transfer matrix. Here $Homeo(T^d)$ denotes the set of homeomorphisms on $T^d$, $C(X,Y)$ denotes the set of continuous functions from $X$ to $Y$ (for the spinor formalism our formalism is obtained by simply replacing $SO(3)$ by $SU(2)$. In the present formalism, (5.156) is the most general description of particle-spin dynamics and the choice of $(E,l)$ depends on the situation, e.g., $(E,l) = (\mathbb{R}^3, l_{1/2})$ for spin-1/2 particles. We work in the framework of topological dynamical systems and therefore $A$, $j$, $l$ are continuous functions. This condition could be strengthened to $A$, $j$, $l$ being smooth functions.

### 5.4.3 Field Motion and Invariant Fields

We are primarily interested in the field dynamics induced by the particle-spin dynamics. Let $f : T^d \to E$ be an $E$−valued field on $T^d$ and set $x = f(z)$ in (5.156). Then after one turn $z$ becomes $j(z)$ and the field value at $j(z)$ becomes $l(A(z);f(z))$.

Observe that not after one turn as in Heinemann et al. (2015) but rather after two turns (to generalize the case to spin-1/2 particles, as well ) the field $f$ becomes the field $f^0 : T^d \to E$ where $f^0(z) := l(A(j^{-1}(z); f(j^{-1}(z))))$. Thus we have the field map

$$f \mapsto f' = l\left(A \circ j^{-1}; f \circ j^{-1}\right), \tag{5.157}$$

where $\circ$ denotes the composition of functions. We call $f \in C\left(T^d, E\right)$ an "invariant $(E,l)$-field of $(j,A)$" if it is mapped by (5.157) into itself, i.e., if

$$f \circ j = l(A,f). \tag{5.158}$$

We call (5.158) the $(E,l)$-*stationarity equation of* $(j,A)$. Our main focus is on the existence of solutions of (5.158) as this is what describes the spin equilibrium of a bunch. In the important case where $(E,l) = \left(\mathbb{R}^3, l_{1/2}\right)$, an invariant $(E,l)$−field $f$ such that k$f$ k= 1 is called an *invariant spin field* (ISF). This completes our introduction to the formalism.

### 5.4.4 The Set $\Sigma_x[f]$ and Its Invariance

Let $E_x := \{l(r;x) : r \in SO(3)\}$; so, clearly, $E_x = \{S \in \mathbb{R}^3 : kSk = kxk\}$ is a sphere centered at (0,0,0). Then the $E_x$ partition $E$ and each set $T^d \times E_x$ is invariant under the particle-spin motion of (5.156) and so we have "decomposed" $T^d \times E$. Let

$$\Sigma_x[f] := \left\{z \in T^d : f(z) \in E_x\right\}. \tag{5.159}$$

The nonempty sets among the $\Sigma_x[f]$ form a partition of $T^d$ and tell us how the values of $f$ are distributed, i.e., $z \in \Sigma_x[f]$ iff $f(z) \in E_x$. It follows from the

$_0$ definition of $\Sigma_x[f]$ and (5.157) that $\Sigma_x[f$ ].$=j(\Sigma_x[f])$. Thus if $f$ is invariant then every $\Sigma_x[f]$ is invariant under $j$ and $T^d$ is partitioned into $f$-dependent invariant sets for the particle dynamics, an interesting fact in its own right.

We can now state three facts related to the existence of invariant fields. Firstly, if there exists an $x$ such that $\Sigma_x[f]$ is not invariant then $f$ is not an invariant field. Secondly, if $\Sigma_x[f]$ is nonempty, let $f_x \in C(\Sigma_x[f],E_x)$ where $f_x(z)= f(z)$. Then $f$ is invariant iff $f_x(j(z)) = l(A(z); f_x(z))$ for every nonempty $\Sigma_x[f]$.

Finally, suppose that $j$ is topologically transitive [e.g., off orbital resonance (see definition below)]. This means that a $z_0 \in T^d$ exists such that $B:=\{j^n(z_0): n=0,\pm1, \pm2,...\}$ is dense in $T^d$, i,e., that the closure $\bar{B}$ of $B$ equals $T^d$. Let $f$ be invariant and pick $x$ such that $z_0 \in \Sigma_x[f]$ then $B \subset \Sigma_x[f]$. Assume $E$ is Hausdorff (e.g., the topology is from a metric). Then it follows that $\Sigma_x[f]$ is closed and , since $\bar{B} = T^d$, we have $\Sigma_x[f] = T^d$. Thus topological transitivity and the Hausdorff property imply an invariant $f$ takes values only in one $E_x$. The so-called ISF-conjecture claims, for $(E,l)=(\mathbb{R}^3,l_{1/2})$, that if $j$ is topologically transitive, then ISF exists.

To show the formalism at work in our present case of spin-1/2 particles, we now present, four theorems (Heinemann et al., 2015): the Decomposition Theorem (DT), the Invariant Reduction Theorem (IRT), the Cross Section Theorem (CST), and the Normal Form Theorem (NFT) .

### 5.4.5 The Decomposition Theorem (DT)

Let $E$ be Hausdorff. It is natural to ask about the relation between the dynamics on two distinct invariant sets $T^d \times E_x$, $T^d \times E_y$. Consider the particle-spin trajectories defined by $(z(n+1), x(n+1)) = P[j,A](z(n),x(n))$ where $(z(0),x(0)) = (z_0,x_0)$ is given with $x_0 \in E_x$. Suppose there exists $\beta \in C(E_x,E_y)$ such that for every particle-spin trajectory $(z(n),x(n)) \in T^d \times E_x$, the function $(z(n),\beta(x(n))) \in T^d \times E_y$ is particle-spin trajectory. A necessary and sufficient condition for $\beta$ to have this property is that $\beta(l(r;\xi)) = l(r;\beta(\xi))$ for all $r \in SO(3)$, $\xi \in E_x$ and this is true iff $r_0 \in SO(3)$ exists such that $\{r_0 r r_0^t : r \in H_x\} \subset H_y$. Here, the subgroup $H_\eta$ of $SO(3)$ is defined by $H_\eta := \{r \in SO(3) : l(r;\eta) = \eta\}$ for every $\eta \in E$. The proof of this is constructive showing that $\beta$ can be defined by $\beta(l(r;x)) := l(r r_0^t; y)$. Furthermore, it can be shown that if $f$ is an invariant $(E,l)$−field of $(j,A)$ which takes values only in $E_x$, then $g \in C(T^d,E)$, defined by $g(z) := \beta(f(z))$, is an invariant field taking values only in $E_y$. In summary, the DT classifies invariant fields in terms of the functions $\beta$, i.e., in terms of the subgroups $H_x$ of $SO(3)$.

### 5.4.6 The Invariant Reduction Theorem (IRT) and the Cross Section Theorem (CST)

Let $f \in C(T^d,E)$, $x \in E$, $\check{\Sigma}_x[f] := \{(z,r) \in (T^d \times SO(3)) : l(r,x) = f(z)\}$ and $\check{P}[j,A] \in Homeo(T^d \times SO(3))$ with $\check{P}[j,A](z,r) := (j(z), A(z)r)$. Then the IRT states that $f$ satisfies (5.158) iff $\check{\Sigma}_x[f]$ is invariant under $\check{P}[j,A]$ for every $x\in E$. As its name suggests, the IRT identifies the dynamical invariance of the reductions with the invariance of the labeling fields $f$. An important aspect of $\check{\Sigma}_x[f]$ is the CST which gives valuable information to be used when applying the IRT to (5.158). Let $f\in C(T^d,E)$ take values only in $E_x$. Thus for every $z\in T^d$, $f(z)\in E_x$ and so there exists a function $\Gamma : T^d \to SO(3)$ such that $f(z)=l(\Gamma(z),x)$ but in general there may not be a continuous $\Gamma$. The CST asserts that a continuous $\Gamma$ exists iff the continuous function $p_x[f]:\check{\Sigma}_x[f] \to T^d$, defined by $p_x[f](z,r):=z$, has a cross section $\sigma$, i.e., iff a continuous $\sigma : T^d \to \check{\Sigma}_x[f]$ exists such that $p_x[f](\sigma(z))=z$. Most importantly, if $p_x[f]$ has a cross section then the CST gives a natural homeomorphism on $T^d \times SO(3)$ which maps $\check{\Sigma}_x[f]$ unto $T^d \times H_x$.

### 5.4.7 The Normal Form Theorem (NFT)

To address the case where a cross section exists we use the NFT, which is clearly related to the CST. Let $\Gamma\in C(T^d,SO(3))$, $x\in E$ and $f(z):=l(\Gamma(z),x)$. Then $f$ is an invariant field iff $l(\Gamma^t(z)A(z)\Gamma(z), x)=x$ for all $z$. This motivates the definition of $H_x$ above and we state the NFT as: $\Gamma^t(j(z))A(z)\Gamma(z)\in H_x$ for all $z\in T^d$ iff $f(j(z))=l(A(z);f(z))$ for all $z\in T^d$. The moral of the NFT is that $f$ and $\Gamma$ are effectively the same, i.e., that one can view invariant fields from two different perspectives. The special case of rotation around the $z$−axis, where $(E,l)$ and $x$ are such that

$$\cos(2\pi n v) \quad -\sin(2\pi n v) \quad 0 \quad \cos(2\pi n v) \quad 0 \qquad (5.160)$$

$H_x = G_v := \sin(2\pi n v)$ 0 $\pm 1, \pm 2, \ldots, 0 : n =$ 0, 1

is the case where spin tunes (defined as the number of spin precession), $v$, exist. The subcase where $v = 0$ describes a spin-orbit resonance (spin-orbit resonances are attained when $v$ takes integral values in (5.160). The terminology NFT is justified as follows: if $\Gamma \in C(T^d, SO(3))$ and if $A^t(z) := \Gamma^t(z)A(z)\Gamma(z)$ belongs to a subgroup $H$ of $SO(3)$ then one calls $(j, A^t)$ an $H$−normal form of $(j, A)$. The notion of normal form gives new view on spin tunes and spin-orbit resonances.

## 5.5 Bundle Interpretation of our Special Case of Free Spin-1/2 Particles in Spacetime

For spin-1/2 particles in 4-dimensional spacetime, the most important $(E, l)$ is given by

$$E = \mathrm{R}^3 \quad and \quad l_{1/2}(r; S) := rS. \tag{5.161}$$

(Here $r \in SO(1,3)$ is a Lorentz rotation matrix and $S \in \mathrm{R}^{1,3}$ is a 4-vector). Clearly $E = \mathrm{R}^{1,3}$ is Hausdorff (since $\mathrm{R}^3$ is) and the $E_x$ are concentric spheres centered at (0,0,0,0) (Figure 5.6) and the field map (5.157) gives $f^0(z) = A(j^{-1}(z))f(j^{-1}(z))$.

The invariant $(\mathbb{R}^{1,3}, l_{1/2})$ −fields $f$ are just the invariant polarization fields describing the spin equilibrium of a bunch and, calculating the norm of our field (using

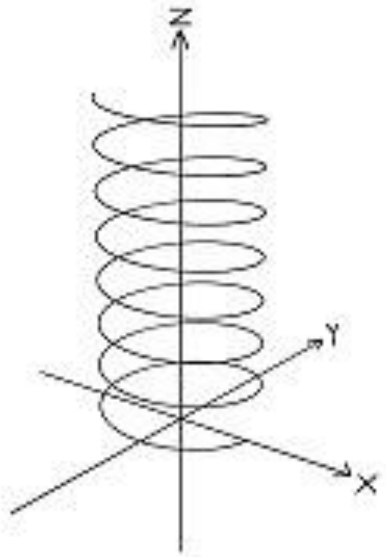


Figure 5.6: Invariant helical motion of free spin-half fields.

(5.72)), we have

$$\begin{aligned} \parallel f \parallel^2 &= \Psi^{\dagger}_{p,\lambda,+1/2} \cdot \Psi_{p,\lambda,+1/2} \\ &= N \begin{pmatrix} 1 & 0 & \frac{c\hat{\sigma}_z p}{m_0c^2+E_p} & 0 \end{pmatrix} \exp\left[-i\left(pz - E_p t\right)/\hbar\right] \\ &\quad \times N \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \frac{c\hat{\sigma}_z p}{m_0c^2+E_p} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} \exp\left[i\left(pz - E_p t\right)/\hbar\right] \end{aligned}$$

= 1 =k *f* k; (5.162)

which shows they are the invariant spin fields. Coming to the NFT with *x* = (0,0,0,1) and using

$$
\begin{aligned}
H_x : &= \left\{ r \in SO(1,3) : l_{1/2}(r;x) = x \right\} \\
&= \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(2\pi n v) & -\sin(2\pi n v) & 0 \\ 0 & \sin(2\pi n v) & \cos(2\pi n v) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} : v \in \mathbb{R} \right\} \\
&=: SO(1,2), \qquad (5.163)
\end{aligned}
$$

we find that a $\Gamma \in C(T^d, SO(1,3))$ satisfies $\Gamma^t(j(z))A(z)\Gamma(z)) \in H_x$ for all $z \in T^d$ iff the fourth column of $\Gamma$ is an invariant spin field. Then $\Gamma$ is called an invariant frame field. Thus the notion of normal form gives a new view on the notion of invariant frame field and generalizes it from the group $SO(1,2)$ to an arbitrary subgroup of $SO(1,3)$.

Next, notice that equations (5.134) and (5.135) of orbital and spin angular momenta give identification of antipodal points (mentioned earlier) as $x_\mu =$

$t$

$x$

$y$

$z$

$\frac{1}{2}\sigma_{03} =$

and $-x_\mu = \begin{pmatrix} -t \\ -x \\ -y \\ -z \end{pmatrix}$ under the same constant transformation of matrix $\begin{pmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix}$. This leads to the matrix of the generalized angular momentum of the system:

$$\begin{pmatrix} 0 & 0 & 0 & -\frac{1}{4} \\ 0 & 0 & -\frac{1}{4} & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -\cos(V/c) \\ 0 & 0 & -\cos(V/c) & 0 \\ 0 & \cos(V/c) & 0 & 0 \\ \cos(V/c) & 0 & 0 & 0 \end{pmatrix},$$

which culminates into the derivation of the expectation value of the overall relative linear superluminal velocity component of the free spin-1/2 field:

$$V \simeq \left(\frac{21\pi}{50} + 2k\pi\right) \times c, \quad (k = 0, 1, 2, ...),$$

and brings to an end a fiber bundle interpretation of our result.

# Chapter 6

# Conclusions and Recommendations

## 6.1 Conclusions

It has been shown, by restricting the homomorphism (i.e., the spinor map) between the special linear group of order 2 [$SL(2,\mathrm{C})$] and the Lorentz group to three dimensions, that spin transformations of the 3-dimensional real space result precisely in $2k\pi$ rotations. This confirms the evidence that spinning mechanism is a continuous rotating process in a time interval. However, the question resides in how to determine mathematically precisely this result. This is what this thesis has firstly made an attempt to answer. Subsequently, our calculations show, in the strict limit of realistic one-particle systems, using positive energy only, and waves propagating in the $z-$direction, that the expectation value of the relative linear velocity component of a free spin-1/2 field exceeds the speed of light. Under these restrictions, it can be noted that what so far appears to be a disaster





(namely, the violation of causality by spin angular momentum) can prove to be a highly desirable aspect of spin-1/2 field theory. This work shows explicitly that the phenomenon of spin exists, not only at low and relativistic velocities, but at superluminal propagation, as well. Finally, the demonstration of superluminal free fermions fits in a fiber bundle formalism.

## 6.2 Recommendations

The present explanation of free spin-half particles superluminal motion is completely consistent with the standard interpretation of quantum mechanics and field theories. I hereby proclaim that my calculations provide a new physical insight into the spin of fermions and so, must receive a wide recognition. My objective is to revive the interest of scientists in the existence of particles capable of moving beyond the speed of light. The time has come for a finite velocity conception of dynamical systems to give place to the perspective of infinite possibilities offered by superluminal electron and neutrino, especially in the context of interstellar and intergalactic mechanically manned propulsion. I therefore leave the practical demonstration of this explanation in the hands of future experimentation.

*References*

# References

Belinfante, F. (1940). On the covariant derivative of tensor-undors. *Physica*, 7(4):305–324.

Breit, G. (1932). Dirac's equation and the spin-spin interactions of two electrons. *Physical Review*, 39(4):616.

Carmeli, M. (1977). *Group theory and general relativity: representations of the Lorentz group and their applications to the gravitational field*. World Scientific.

Dirac, P. A. M. (1981). *The principles of quantum mechanics*. Number 27. Oxford university press.

Ehrenfest, P. (1927). Bemerkung u¨ber die angen¨aherte gu¨ltigkeit der klassischen mechanik innerhalb der quantenmechanik. *Zeitschrift fu¨r Physik A Hadrons and Nuclei*, 45(7):455–457.

Gelfand, I. M., Gelfand, I. M., Fomin, S. V., and Silverman, R. A. (2000). *Calculus of variations*. Courier Corporation.

Greiner, W. (1994). Relativistic quantum mechanics. wave equations. 2. corr.

Greiner, W. and Mu¨ller, B. (1994). *Quantum Mechanics: With 128 Worked Examples and Problems*, volume 2. Springer Science & Business Media.

Greiner, W. and Reinhardt, J. (1996). *Field quantization*. Springer Science & Business Media.

Guildry, M. (1991). Gauge Field Theories: An Introduction with Applications.

Heinemann, K., Barber, D., Ellison, J., and Vogt, M. (2015). An informal summary of a new formalism for classifying spin-orbit systems using tools distilled from the theory of bundles. *arXiv physics.acc-ph arXiv:1502.00538v1*.

Modanese, G. (2013). Velocity requirements for causality violation. *arXiv preprint arXiv:1304.5522*.

*References*

Naber, G. L. (1992). The Geometry of Minkowski Spacetime. An Introduction to the Mathematics of the Special Theory of Relativity, Sect. 1.4.

Noether, E. (1918). Invarianten beliebiger differentialausdru¨cke. *Nachrichten von der Gesellschaft der Wissenschaften zu G¨ottingen, mathematisch-physikalische Klasse*, 1918:37–44.

Ohanian, H. C. (1986). What is spin. *Am. J. Phys*, 54(6):500–505.

Pauli, W. (1993). *Wissenschaftlicher Briefwechsel Mit Bohr, Einstein, Heisenberg UA, 1930-1939: Scientific Correspondence*. Springer-Verlag.

Schro¨dinger, E. (1930). *Uber die kr¨aftefreie Bewegung in der relativistischen Quan-¨ tenmechanik*. Akademie der wissenschaften in kommission bei W. de Gruyter u. Company.

Shore, G. M. (2006). Causality and superluminal light. *arXiv preprint grqc/0302116*.

Stecker, F. W. (2014). Limiting superluminal electron and neutrino velocities using the 2010 crab nebula flare and the icecube pev neutrino events. *Astroparticle Physics*, 56:16–18.

Van der Waerden, B. (1960). Theoretical physics in the twentieth century.