Quantum Field Theory

Priv.-Doz. Dr. Axel Pelster

Department of Physics

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Lecture in Winter Term 2020/2021

Kaiserslautern, February 12, 2021

Preface

The present manuscript represents the material of the lecture **Quantum Field Theory**, which I held as part of the *Master of Science in Advanced Quantum Physics* program at the Department of Physics of the Technische Universität Kaiserslautern during the winter term 2020/2021. Due to the Corona pandemic the 28 lectures were both live streamed and recorded. The corresponding videos are available at the platform Panopto at the national network of the state Rhineland Palatinate. After each lecture both the corresponding hand-written notes as well as this manuscript were made available to the students.

The presented material is partially inspired by former lectures of Prof. Dr. Ulrich Weiß at the Universität Stuttgart from I which benefitted during my own physics study. And I had the pleasure to learn from the lectures of Prof. Dr. Hagen Kleinert at the Freie Universität Berlin during my postdoc time, when I supervised the corresponding exercises. The manuscript is mainly based on the two valuable books:

- W. Greiner and J. Reinhardt, *Field Quantization*, Springer, 2008
- H. Kleinert, *Particles and Quantum Fields*, World Scientific, 2016

which I recommend for self-study. Additional useful information, partially including historical insights, can be found in the following selected books:

- J. Gleick, *Genius - The Life and Science of Richard Feynman*, Vintage Books, 1991
- W. Greiner, *Relativistic Quantum Mechanics - Wave Equations*, Springer, 2000
- W. Greiner and J. Reinhardt, *Quantum Electrodynamics*, Springer, 2008
- W. Greiner and J. Reinhardt, *Field Quantization*, Springer, 2008
- H. Kleinert, *Particles and Quantum Fields*, World Scientific, 2016
- M.E. Peskin and D.V. Schröder, *An Introduction to Quantum Field Theory*, Cambridge University Press, 1995
- S. Schweber, *QED and the men who made it - Feynman, Schwinger, and Tomonaga*, Princeton University Press, 1994

- M. Veltman, *Diagrammatica - The Path to Feynman Diagrams*, Cambridge University Press, 1994

- T. Lancaster and S.J. Blundell, *Quantum Field Theory for the Gifted Amateur*, Oxford University Press, 2014

With this I thank the student workers Alexander Guthmann, Joshua Krauss, and Andra Pirktina for typing and, partially, translating the original hand-written German lecture notes. They originated from previous lectures held at Freie Universität Berlin (2000), Universtität Potsdam (2010, 2011), and Technische Universität Kaiserslautern (2014, 2017/2018). And, finally, I am also grateful to Dr. Milan Randonjić for contributing to the success of the lecture in the winter term 2020/2021 by supervising the exercises. The corresponding problem sets are available at the homepage

`http://www-user.rhrk.uni-kl.de/~apelster/Vorlesungen/WS2021/qft.html`

Contents

Chapter 1

Introduction

This lecture provides a hands-on insight into quantum electrodynamics, which represents an important building block of the standard model of elementary particles physics. To this end we proceed in three steps. At first, we introduce the concept of second quantization, which allows to deal with an arbitrary number of quantum particles, by the example of non-relativistic many-body theory. Then we discuss the relativistic wave equations as representations of the Poincaré symmetry of space-time. And, ultimately, we work out how to perturbatively calculate scattering cross sections of fundamental quantum electrodynamic processes by using the technique of Feynman diagrams.

1.1 Standard Model

The standard model of elementary particle physics describes quite successfully the basic structure of matter and three of the overall four fundamental interactions. All its predictions agree precisely with all experimental measurements performed so far within the respective error bars. The basic concept of the standard model is local gauge invariance. This means that the physics does not change provided that the particle wave functions acquire local phase factors, which change continuously from space-time point to space-time point. This represents a quite hard restriction. Free massive particles do not fulfill this condition, as their wave functions only allow for a global change of the phase factors. But if we postulate to have in addition for massive particles also an invariance with respect to a local change of their phase factors, we can deduce how these massive particles interact. Within such a local gauge theory it turns out that the interaction between the massive particles is mediated by an exchange of gauge bosons, which represent the quantized excitations of the corresponding gauge fields. In this way the three interactions of the standard model can be classified as is summarized in Tab. 1.1.

The first and by far most successful theory of fundamental interactions is quantum electrodynamics. Its $U(1)$ gauge theory was later on extended to the description of the other two interactions of the standard model, which lead to the electroweak theory, unifying both the

interaction	gauge symmetry	gauge bosons
electromagnetic	$U(1)$	photon
weak	$SU(2)$	intermediate vector bosons
strong	$SU(3)$	gluon

Table 1.1: Overview of the three types of interactions of the standard model of elementary particle physics together with their gauge symmetries and gauge bosons.

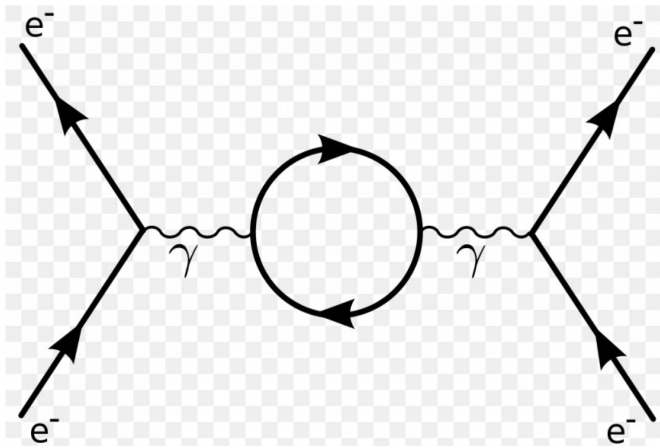


Figure 1.1: Due to the presence of the vacuum the scattering of two electrons also involves the creation and annihilation of virtual electrons and positrons.

eletrodynamic and weak interactions, as well as quantum chromodynamics, the quantum theory of the strong interaction. Furthermore, quantum electrodynamics is the theory in all natural sciences, whose predictions agree most precisely with experimental results. According to a comparison of Richard Feynman, its precision of 10 orders of magnitude corresponds to a resolution, where the thickness of a single hair is resolvable by looking from the West to the East Coast of America.

1.2 Non-Relativistic Quantum Many-Body Theory

Within a quantum electrodynamic scattering process not only real particles are involved. The physical vacuum is not empty but, instead, consists of a sea of virtual particles, which are also involved in a scattering process, see Fig. 1.1. Therefore, it is necessary to work out a quantum mechanical formalism which is capable of describing an arbitrary number of particles. The formalism of *first quantization* is not appropriate for that as there the number of particles remains conserved. With the first quantization it is possible to calculate, for instance, for the hydrogen atom the stationary energy states and the respective transition probabilities between them. But the fundamental processes of the absorption of a photon and the corresponding excitation of an electron as well as the later relaxation of the electron to the ground state

bosons	fermions
integer spin	half-integer spin
particles mediating forces	matter particles
Bose-Einstein statistics	Fermi-Dirac statistics
symmetric many-body wave function	anti-symmetric many-body wave function

Table 1.2: According to the spin-statistic theorem there exist with bosons and fermions two kinds of indistinguishable particles.

spin	0	1/2	1	3/2	2
mass > 0	Higgs	leptons, quarks	intermediate vector bosons	Δ resonances	
mass $= 0$			photon, gluon		graviton

Table 1.3: Classification of elementary particles according to their spin and mass.

and the corresponding emission of a photon are not describable within the first quantization formalism as they violate the particle number conservation.

The description of identical particles, which have exactly the same physical properties as, for instance, mass, spin and charge, turns out to be problematic in the realm of quantum mechanics. In classical mechanics identical particles are distinguishable so that the trajectory of each particle can always be identified. All experiments suggest, however, that this principle of distiguishability can no longer be maintained in quantum mechanics. Due to the Heisenberg uncertainty relation the probability densities of identical particles overlap so that the identification of a single particle is not possible. Despite of this fundamental principle of indistinguishability of identical particles in quantum mechanics one is nevertheless forced, due to calculational purposes, to enumerate the particles. But this artificial particle enumeration has to be performed in such a way that physically observable results turn out to be invariant with respect to any change of this particle labeling. From this definition of indistinguishability then follows that a many-particle wave function must obey special symmetry requirements. To this end Wolfgang Pauli derived 1940 the spin-statistic theorem of relativistic quantum field theory. By unifying the basic principles of special relativity with those of quantum mechanics he showed that there are in three dimensions exactly two kinds of indistinguishable identical particles, namely bosons and fermions. Their respective properties are summarized in Tab. 1.2.

It turns out that concrete calculations with (anti-)symmetric many-body wave functions are quite cumbersome. Therefore, one has worked out a quite elegant formalism for quantum many-body systems, which is capable of dealing with an arbitrary number of particles and is called *second quantization*. In Part I of the lecture we work out the so-called canonical field quantization which deals with creation and annihilation operators for particles. Note that the Bose-Einstein and Fermi-Dirac statistics is automatically taken into account by defining appropriate commutation relations for the creation and annihilation operators. Because of illustrative purposes and in view of applications in the realm of solid-state physics we restrict ourselves in Part I to elaborate this second quantization formalism in the realm of non-relativistic quantum many-body theory. Thus, this amounts to quantize the first quantized Schrödinger theory, which is possible to perform separately for both bosons and fermions.

1.3 Relativistic Fields and Their Quantization

In Part II we discuss at first the Poincaré group as the fundamental space-time symmetry in the absence of gravity. By the concrete examples of rotations, boosts, and translations we introduce the concepts of Lie groups and Lie algebras as well as their respective representations. In particular, the Casimir operators of the Poincaré group are of importance, i.e. those operators which commute with all generators of rotations, boosts, and translations. Namely, it turns out that all states of relativistic quantum field theory can be classified with respect to the eigenvalues of the Casimir operators of the Poincaré group, which are the spin and the mass of the elementary particles, respectively, see Tab. 1.3.

Thus, one can understand relativistic quantum field theory as the representation theory of the Poincaré group. From this group-theoretical point of view we discuss in detail the examples of both the Maxwell and the Dirac field. To this end we determine the respective free solutions with their different helicity and polarization states. But instead of directly solving the respective Maxwell and Dirac equation, we take group theory to our advantage. For the massless (massive) spin 1 (1/2) particles we solve the underlying Maxwell (Dirac) equation in a particular reference frame (the inertial frame) and rotate (boost) then the solution to an arbitrary reference (inertial) frame. Afterwards, we second quantize the Maxwell as well as the Dirac theory and construct their respective free propagators. Furthermore, we discuss the fundamental relations between symmetries and conservation laws in terms of the seminal Noether theorem. As a concrete example we deal with all conservation laws of quantum electrodynamics.

1.4 Quantum Electrodynamics

In Part III we finally turn to quantum electrodynamics. At first we derive the light-matter interaction by postulating the aforementioned local gauge invariance. Based on the formalism of second quantization we then perform a systematic perturbation theory around the free theory and expand with respect to the light-matter interaction strength. In particular, we demonstrate that, although using the non-covariant Coulomb gauge for the Maxwell field, we finally yield covariant perturbative corrections, which can be graphically represented in terms of Feynman diagrams. In order to construct all Feynman diagrams order by order we introduce a graphical recursion relation, which is based on cutting the lines of Feynman diagrams of lower orders and

Mott scattering	$e^-Z - e^-Z$
Møller scattering	$e^- - e^-$
Bhabha scattering	$e^- - e^+$
Compton scattering	$e^-\gamma - e^-\gamma$

Table 1.4: Examples of scattering processes in quantum electrodynamics.

gluing them together with new interaction vertices. Based on the Feynman diagrams we are then able to calculate the cross sections for individual scattering processes, see the examples mentioned in Tab. 1.4.

In lowest order the respective cross sections are generically finite, but in higher orders notorious infinities appear. These infinities prevent to make any concrete quantitative prediction for an experimental measurement of a cross section. In quantum electrodynamics it turns out that these infinities can be systematically removed order by order with a so-called renormalization scheme. In a first step one regularizes the infinite integrals, i.e. one introduces an additional calculational degree of freedom in such a way that these integrals become finite. For instance, one introduces an ultraviolet cut-off Λ in momentum space or one follows the notion of Gerard 't Hooft and calculates the momentum integrals in $D = 4 - \varepsilon$ dimensions. By construction the infinities of the integrals then emerge in the limit $\Lambda \to \infty$ or $\varepsilon \to 0$. In a second step one shows then that the infinities can be absorbed by the few parameters of the theory as the mass, the coupling constant, and the fields. Depending on the available time we plan to perform this renormalization scheme in quantum electrodynamics explicitly in the lowest perturbative order. The general proof, that quantum electrodynamics is renormalizable to all orders of perturbation theory, is due to Freeman Dyson.

Part I:

Non-Relativistic Quantum Many-Body Theory

Chapter 2

Identical Particles

Here we deal with identical particles, thus they have exactly the same properties like mass, spin or charge. From all experiments performed so far in the realm of quantum mechanics one can deduce that such identical particles are indistinguishable. Nevertheless we start with describing a quantum many-particle system in Section 2.1 as if their identical particles would be distinguishable. Based on that we investigate then in Section 2.2 the consequences for postulating the indistinguishability of identical particles. Namely, it turns out in three spatial dimensions that identical particles are either bosons or fermions, which are characterized by a symmetric and anti-symmetric many-body wave function, respectively. We illustrate the corresponding complications in concrete calculations by the illustrative example of non-interacting identical particles in Section 2.3.

2.1 Distinguishable Particles

A many-body system of identical nonrelativistic particles of mass M is classically described by the Lagrange function

$$L(\mathbf{x}_1, \ldots, \mathbf{x}_n; \dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_n) = \sum_{\nu=1}^{n} \frac{M}{2} \dot{\mathbf{x}}_\nu^2 - V(\mathbf{x}_1, \ldots, \mathbf{x}_n) \,. \tag{2.1}$$

The n-particle potential $V(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is usually additive in both the 1-particle potentials $V_1(\mathbf{x}_\nu)$ and the 2-particle potentials $V_2(\mathbf{x}_\nu - \mathbf{x}_\mu)$:

$$V(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{\nu=1}^{n} V_1(\mathbf{x}_\nu) + \frac{1}{2} \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} V_2(\mathbf{x}_\nu - \mathbf{x}_\mu) \,. \tag{2.2}$$

Note that the latter must obey the symmetry

$$V_2(\mathbf{x}_\nu - \mathbf{x}_\mu) = V_2(\mathbf{x}_\mu - \mathbf{x}_\nu) \tag{2.3}$$

due to the Newton axiom "action = - reactio". The Euler-Lagrange equations

$$\frac{\partial L}{\partial \mathbf{x}_\nu} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{x}}_\nu} = 0 \tag{2.4}$$

corresponding to the Lagrange function (2.1), (2.2) lead to the Newton equations of motion:

$$M\ddot{\mathbf{x}}_\nu = -\frac{\partial V_1(\mathbf{x}_\nu)}{\partial \mathbf{x}_\nu} - \sum_{\mu=1}^{n} \frac{\partial V_2(\mathbf{x}_\nu - \mathbf{x}_\mu)}{\partial \mathbf{x}_\nu} \,. \tag{2.5}$$

The transition to the Hamilton formalism is implemented by introducing the canonically conjugated momenta

$$\mathbf{p}_\nu = \frac{\partial L}{\partial \dot{\mathbf{x}}_\nu} = M\dot{\mathbf{x}}_\nu \tag{2.6}$$

and by performing the Legendre transformation

$$H(\mathbf{p}_1, \dots, \mathbf{p}_n; \mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{\nu=1}^{n} \mathbf{p}_\nu \dot{\mathbf{x}}_\nu - L(\mathbf{x}_1, \dots, \mathbf{x}_n; \dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_n) \tag{2.7}$$

which yields the Hamilton function

$$H(\mathbf{p}_1, \dots, \mathbf{p}_n; \mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{\nu=1}^{n} \frac{\mathbf{p}_\nu^2}{2M} + \sum_{\nu=1}^{n} V_1(\mathbf{x}_\nu) + \frac{1}{2} \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} V_2(\mathbf{x}_\nu - \mathbf{x}_\mu) \,. \tag{2.8}$$

The corresponding Hamilton equations

$$\begin{aligned} \dot{\mathbf{x}}_\nu &= \frac{\partial H}{\partial \mathbf{p}_\nu} = \frac{\mathbf{p}_\nu}{M} \,, & (2.9) \\ \dot{\mathbf{p}}_\nu &= -\frac{\partial H}{\partial \mathbf{x}_\nu} = -\frac{\partial V_1(\mathbf{x}_\nu)}{\partial \mathbf{x}_\nu} - \sum_{\mu=1}^{n} \frac{\partial V_2(\mathbf{x}_\nu - \mathbf{x}_\mu)}{\partial \mathbf{x}_\nu} & (2.10) \end{aligned}$$

turn out to be equivalent to the Newton equations of motion (2.5).

The transition from classical mechanics to quantum mechanics is achieved by assigning operators to observables:

$$\mathbf{x}_\nu \to \hat{\mathbf{x}}_\nu \,, \quad \mathbf{p}_\nu \to \hat{\mathbf{p}}_\nu \,, \quad H(\mathbf{p}_1, \dots, \mathbf{p}_n; \mathbf{x}_1, \dots, \mathbf{x}_n) \to H(\hat{\mathbf{p}}_1, \dots, \hat{\mathbf{p}}_n; \hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_n) \,. \tag{2.11}$$

In order to obey the Heisenberg uncertainty relation, we postulate here the following canonical commutation relations

$$\left[\hat{x}_{j\nu}, \hat{x}_{k\mu}\right]_- = \left[\hat{p}_{j\nu}, \hat{p}_{k\mu}\right]_- = 0 \,, \qquad \left[\hat{p}_{j\nu}, \hat{x}_{k\mu}\right]_- = \frac{\hbar}{i} \, \delta_{jk} \, \delta_{\nu\mu} \,, \tag{2.12}$$

where the commutator between two quantum mechanical operators $\hat{A}$ and $\hat{B}$ is defined by

$$\left[\hat{A}, \hat{B}\right]_- = \hat{A}\hat{B} - \hat{B}\hat{A} \,. \tag{2.13}$$

The time evolution of a quantum mechanical state vector $|\psi(t)\rangle$ is described by the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \,. \tag{2.14}$$

In order to convert this representation independent formulation of quantum mechanics to the spatial representation, one chooses as a basis the eigenstates $|\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle$ of the coordinate operators $\hat{\mathbf{x}}_\nu$. They fulfill the eigenvalue problem

$$\hat{\mathbf{x}}_\nu |\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle = \mathbf{x}_\nu |\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle \tag{2.15}$$

as well as the orthonormality relation

$$\langle \mathbf{x}_1, \ldots, \mathbf{x}_n | \mathbf{x}_1', \ldots, \mathbf{x}_n' \rangle = \delta(\mathbf{x}_1 - \mathbf{x}_1') \cdot \ldots \cdot \delta(\mathbf{x}_n - \mathbf{x}_n') \tag{2.16}$$

and the completeness relation

$$\int d^3x_1 \cdot \ldots \cdot \int d^3x_n \, |\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle\langle \mathbf{x}_1, \ldots, \mathbf{x}_n| = 1 \,. \tag{2.17}$$

The spatial representation of the momentum operators $\hat{\mathbf{p}}_\nu$ is given by the Jordan rule:

$$\langle \mathbf{x}_1, \ldots, \mathbf{x}_n | \hat{\mathbf{p}}_\nu = \frac{\hbar}{i} \frac{\partial}{\partial \mathbf{x}_\nu} \langle \mathbf{x}_1, \ldots, \mathbf{x}_n| \,. \tag{2.18}$$

Evolving the quantum mechanical state vector $|\psi(t)\rangle$ with respect to this basis yields due to the completeness relation (2.17)

$$|\psi(t)\rangle = \int d^3x_1 \cdot \ldots \cdot \int d^3x_n \, \psi(\mathbf{x}_1, \ldots, \mathbf{x}_n; t) \, |\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle \,, \tag{2.19}$$

where the expansion coefficients represent the n-particle wave function

$$\psi(\mathbf{x}_1, \ldots, \mathbf{x}_n; t) = \langle \mathbf{x}_1, \ldots, \mathbf{x}_n | \psi(t)\rangle \,. \tag{2.20}$$

Multiplying (2.14) from the left with the bra-vector $\langle \mathbf{x}_1, \ldots, \mathbf{x}_n|$ leads for the n-particle wave function (2.20) to the n-particle Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{x}_1, \ldots, \mathbf{x}_n; t) = \hat{H} \psi(\mathbf{x}_1, \ldots, \mathbf{x}_n; t) \,. \tag{2.21}$$

Here the spatial representation of the Hamilton operator $\hat{H}$ follows due to (2.11), (2.15), and (2.18) from the Hamilton function H as follows:

$$\hat{H} = H \left(\frac{\hbar}{i} \frac{\partial}{\partial \mathbf{x}_1}, \ldots, \frac{\hbar}{i} \frac{\partial}{\partial \mathbf{x}_n}; \mathbf{x}_1, \ldots, \mathbf{x}_n \right) . \tag{2.22}$$

In case of the standard Hamilton function (2.8) we get

$$\hat{H} = \sum_{\nu=1}^{n} \left\{ -\frac{\hbar^2}{2M} \Delta_\nu + V_1(\mathbf{x}_\nu) \right\} + \frac{1}{2} \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} V_2(\mathbf{x}_\nu - \mathbf{x}_\mu) \,. \tag{2.23}$$

As we have assumed here that both the 1- and the 2-particle potential V_1 and V_2 do not explicitly depend on time, one can perform for the n-particle wave function the separation ansatz

$$\psi(\mathbf{x}_1, \ldots, \mathbf{x}_n; t) = \psi_E(\mathbf{x}_1, \ldots, \mathbf{x}_n) \, e^{-iEt/\hbar} \,. \tag{2.24}$$

This reduces the time-dependent Schrödinger equation (2.21) to the time-independent Schrödinger equation:

$$\hat{H} \psi_E(\mathbf{x}_1, \ldots, \mathbf{x}_n) = E \psi_E(\mathbf{x}_1, \ldots, \mathbf{x}_n) \,, \tag{2.25}$$

where E denotes the energy eigenvalue.

2.2 Bosons and Fermions

The quantum mechanical laws summarized in the last section are only valid for identical particles, which are assumed to be distinguishable. But experimentally it has turned out that identical particles always happen to behave in the same way so that no objective measurement allows to distinguish one from the other. Thus, in the realm of quantum many-body theory the fundamental principle of the indistinguishability of identical particles has to be taken into account.

Physically relevant are only expectation values of observables. The principle of the indistinguishability of identical particles means in this context concretely that the expectation value of any operator $\hat{A}$ must not change when the enumeration of two particles is swapped within the n-particle wave function:

$$\begin{aligned} & \int d^3x_1 \cdot \ldots \cdot \int d^3x_n \, \psi^*(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_n) \hat{A} \, \psi(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_n) \\ = & \int d^3x_1 \cdot \ldots \cdot \int d^3x_n \, \psi^*(\mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) \hat{A} \, \psi(\mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) \, . \end{aligned} \tag{2.26}$$

From this definition of indistinguishability of identical particles we now derive various characteristic properties for both the operators $\hat{A}$ and the n-particle wave functions $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Note that restricting the equality of expectation values (2.26) to just two particles is not a principle limitation as any permutation $\hat{P}$ can always be represented as a certain product of transpositions $\hat{P}_{jk}$

$$\hat{P} = \prod \hat{P}_{jk} \, . \tag{2.27}$$

Here the action of $\hat{P}_{jk}$ is defined by exchanging the particle coordinates j and k in the n-particle wave function:

$$\hat{P}_{jk} \psi(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_n) = \psi(\mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) \, . \tag{2.28}$$

From (2.28) it is self-evident that the transposition $\hat{P}_{jk}$ is involutoric, i.e. applying it twice yields back the original n-particle wave function:

$$\hat{P}_{jk} \, \hat{P}_{jk} = 1 \qquad \Longrightarrow \qquad \hat{P}_{jk} = \hat{P}_{jk}^{-1} \, . \tag{2.29}$$

With the help of the transposition operator $\hat{P}_{jk}$ the defining equation (2.26) of the indistinguishability of identical particles can be converted from the spatial representation into the representation-free formulation:

$$\langle \psi | \hat{A} | \psi \rangle = \langle \hat{P}_{jk} \psi | \hat{A} | \hat{P}_{jk} \psi \rangle = \langle \psi | \hat{P}_{jk}^\dagger \hat{A} \hat{P}_{jk} | \psi \rangle \, . \tag{2.30}$$

From the straight-forward decomposition

$$\begin{aligned} \langle \phi | \hat{A} | \psi \rangle \;=\; & \frac{1}{4} \Big[\langle \phi + \psi | \hat{A} | \phi + \psi \rangle - \langle \phi - \psi | \hat{A} | \phi - \psi \rangle \\ & - i \langle \phi + i\psi | \hat{A} | \phi + i\psi \rangle + i \langle \phi - i\psi | \hat{A} | \phi - i\psi \rangle \Big] \, , \end{aligned} \tag{2.31}$$

which takes advantage from the sesquilinearity property

$$\langle\phi|\hat{A}|i\psi\rangle = i\langle\phi|\hat{A}|\psi\rangle\,, \qquad \langle i\psi|\hat{A}|\phi\rangle = -i\langle\psi|\hat{A}|\phi\rangle\,, \tag{2.32}$$

follows then together with (2.30) a useful identity for any matrix element:

$$\langle\phi|\hat{A}|\psi\rangle = \langle\phi|\hat{P}_{jk}^{\dagger}\hat{A}\hat{P}_{jk}|\psi\rangle\,. \tag{2.33}$$

Due to the arbitrariness of the states $|\phi\rangle$ and $|\psi\rangle$ we thus conclude the operator identity

$$\hat{A} = \hat{P}_{jk}^{\dagger}\hat{A}\hat{P}_{jk}\,. \tag{2.34}$$

Evaluating (2.34) for the special case $\hat{A} = \hat{P}_{jk}$ we read off due to (2.29) that the transposition operator $\hat{P}_{jk}$ turns out to be both hermitian

$$\hat{P}_{jk} = \hat{P}_{jk}^{\dagger} \tag{2.35}$$

and unitary

$$\hat{P}_{jk}^{-1} = \hat{P}_{jk}^{\dagger}\,. \tag{2.36}$$

Furthermore, we conclude from (2.34) and (2.36) that any operator $\hat{A}$ commutes with a transposition $\hat{P}_{jk}$:

$$\left[\hat{P}_{jk}, \hat{A}\right]_{-} = \hat{P}_{jk}\hat{A} - \hat{A}\hat{P}_{jk} = 0\,. \tag{2.37}$$

As the latter identity holds in particular for the Hamilton operator $\hat{A} = \hat{H}$ we know that there exist states, which are at the same time eigenstates of both the Hamilton operator $\hat{H}$ and all transposition operators $\hat{P}_{jk}$:

$$\hat{H}|\psi\rangle = E|\psi\rangle\,, \qquad \hat{P}_{jk}|\psi\rangle = p_{jk}|\psi\rangle\,. \tag{2.38}$$

Due to the hermiticity (2.35) of the transposition operators $\hat{P}_{jk}$ their respective eigenvalues p_{jk} must be real. And from the involutoric property (2.29) follows furthermore

$$p_{jk}^2 = 1 \tag{2.39}$$

Thus, the eigenvalues of the transposition operators $\hat{P}_{jk}$ are either $p_{jk} = 1$ or $p_{jk} = -1$. Moreover, it is reasonable that an n-particle wave function $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, which is an eigenfunction of all transposition operators $\hat{P}_{jk}$, must always have one and the same eigenvalue. In order to show this we consider the following identity:

$$\begin{aligned}
&\hat{P}_{1j}\hat{P}_{2k}\hat{P}_{12}\hat{P}_{2k}\hat{P}_{1j}\psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_n)\\
&= \hat{P}_{1j}\hat{P}_{2k}\hat{P}_{12}\psi(\mathbf{x}_j, \mathbf{x}_k, \ldots, \mathbf{x}_1, \ldots, \mathbf{x}_2, \ldots, \mathbf{x}_n)\hat{P}_{1j}\hat{P}_{2k}\psi(\mathbf{x}_k, \mathbf{x}_j, \ldots, \mathbf{x}_1, \ldots, \mathbf{x}_2, \ldots, \mathbf{x}_n)\\
&= \psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) = \hat{P}_{jk}\psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_n)\,.
\end{aligned} \tag{2.40}$$

From this we conclude the operator identity:

$$\hat{P}_{jk} = \hat{P}_{1j}\hat{P}_{2k}\hat{P}_{12}\hat{P}_{2k}\hat{P}_{1j} \,, \tag{2.41}$$

so we obtain for the corresponding eigenvalues due to (2.39)

$$p_{jk} = (p_{1j})^2 \, (p_{2k})^2 \, p_{12} \qquad \Longrightarrow \qquad p_{jk} = p_{12} \,. \tag{2.42}$$

Therefore, identical particle possess either a symmetric ($\epsilon = +1$) or an anti-symmetric ($\epsilon = -1$) wave function with the property

$$\hat{P}_{jk}|\psi^\epsilon\rangle = \epsilon|\psi^\epsilon\rangle \,. \tag{2.43}$$

Using (2.35) and (2.43) we get that symmetric and anti-symmetric wave function are always orthogonal with respect to each other:

$$\begin{aligned} \langle\psi^-|\psi^+\rangle &= \langle\psi^-|\hat{P}_{jk}\psi^+\rangle = \langle\psi^-|\hat{P}^\dagger_{jk}\psi^+\rangle = \langle\hat{P}_{jk}\psi^-|\psi^+\rangle = -\langle\psi^-|\psi^+\rangle \\ &\Longrightarrow \qquad \langle\psi^-|\psi^+\rangle = 0 \,. \end{aligned} \tag{2.44}$$

Furthermore, it turns out that identical particles maintain their symmetry character for all times. To this end we state that the time evolution operator $\hat{U}(t_2, t_1)$ transforms an initial state of definite symmetry $|\psi^{\epsilon_1}(t_1)\rangle$ into a final state of definite symmetry $|\psi^{\epsilon_2}(t_2)\rangle$ via

$$|\psi^{\epsilon_2}(t_2)\rangle = \hat{U}(t_2, t_1)|\psi^{\epsilon_1}(t_1)\rangle \,. \tag{2.45}$$

Thus, taking (2.37) and (2.43) into account we conclude

$$\begin{aligned} \epsilon_2|\psi^{\epsilon_2}(t_2)\rangle = \hat{P}_{jk}|\psi^{\epsilon_2}(t_2)\rangle &= \hat{P}_{jk}\hat{U}(t_2, t_1)|\psi^{\epsilon_1}(t_1)\rangle = \hat{U}(t_2, t_1)\hat{P}_{jk}|\psi^{\epsilon_1}(t_1)\rangle = \epsilon_1|\psi^{\epsilon_1}(t_2)\rangle \\ &\Longrightarrow \qquad \epsilon_1 = \epsilon_2 \,. \end{aligned} \tag{2.46}$$

As a result we state that the Hilbert space of identical particles consists of either only symmetric or only anti-symmetric wave functions. In relativistic quantum field theory it is shown which Hilbert space is appropriate for which sort of particles. According to the spin-statistic theorem of Pauli identical particles with integer (half-integer) spin are bosons (fermions) and have symmetric (anti-symmetric) wave functions, see Tab. 1.2.

2.3 Non-Interacting Identical Particles

In general it is quite cumbersome to calculate n-particle wave functions by taking into account the symmetry property. We illustrate this by the example of non-interacting identical particles. According to (2.23), (2.25) and a vanishing 2-particle potential $V_2(\mathbf{x}_\nu - \mathbf{x}_\mu) = 0$ the following time-independent Schrödinger equation has to be solved:

$$\sum_{\nu=1}^{n} \left\{ -\frac{\hbar^2}{2M}\,\Delta_\nu + V_1(\mathbf{x}_\nu) \right\} \psi_E(\mathbf{x}_1, \ldots, \mathbf{x}_n) = E\,\psi_E(\mathbf{x}_1, \ldots, \mathbf{x}_n) \,. \tag{2.47}$$

In the following we assume that the 1-particle wave functions $\psi_{E_\alpha}(\mathbf{x})$ with the vector of quantum numbers α are known as solutions of the time-independent 1-particle Schrödinger equation

$$\left\{-\frac{\hbar^2}{2M}\,\Delta + V_1(\mathbf{x})\right\}\psi_{E_\alpha}(\mathbf{x}) = E_\alpha\,\psi_{E_\alpha}(\mathbf{x})\,. \tag{2.48}$$

Thus they represent an orthonormal basis obeying both the orthonormality relation

$$\int d^3x\,\psi^*_{E_\alpha}(\mathbf{x})\psi_{E_{\alpha'}}(\mathbf{x}) = \delta_{\alpha,\alpha'} \tag{2.49}$$

and the completeness relation

$$\sum_\alpha \psi^*_{E_\alpha}(\mathbf{x})\psi_{E_\alpha}(\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}')\,. \tag{2.50}$$

In case that the particles would be distinguishable, then a solution of the time-independent n-particle Schrödinger equation (2.47) factorizes into 1-particle wave functions:

$$\psi_E(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \psi_{E_{\alpha_1},\ldots,E_{\alpha_n}}(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \prod_{\nu=1}^n \psi_{E_{\alpha_\nu}}(\mathbf{x}_\nu) \tag{2.51}$$

and the total energy is the sum of the respective 1-particle energies

$$E = \sum_{\nu=1}^n E_{\alpha_\nu}\,. \tag{2.52}$$

Furthermore, the orthonormality and completeness relations of the 1-particle wave functions (2.49) and (2.50) imply corresponding relations for the n-particle wave functions

$$\int d^3x_1\cdots\int d^3x_n\,\psi^*_{E_{\alpha_1},\ldots,E_{\alpha_n}}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\psi_{E_{\alpha'_1},\ldots,E_{\alpha'_n}}(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \prod_{\nu=1}^n \delta_{\alpha_\nu,\alpha'_\nu}\,, \tag{2.53}$$

$$\sum_{\alpha_1}\cdots\sum_{\alpha_n}\psi^*_{E_{\alpha_1},\ldots,E_{\alpha_n}}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\psi_{E_{\alpha_1},\ldots,E_{\alpha_n}}(\mathbf{x}'_1,\ldots,\mathbf{x}'_n) = \prod_{\nu=1}^n \delta(\mathbf{x}_\nu-\mathbf{x}'_\nu)\,. \tag{2.54}$$

But, as identical particles are indistinguishable, the n-particle wave functions must either be symmetric or anti-symmetric. To this end we introduce the (anti-)symmetrization operator

$$\hat{S}^\epsilon = \sum_{\hat{P}} \epsilon^p\hat{P}\,, \tag{2.55}$$

which consists of a sum over all permutation operators $\hat{P}$ and p denotes the number of transpositions of a certain permutation corresponding to the decomposition (2.27). Multiplying a permutation $\hat{P}$ in the sum (2.55) with a single transposition $\hat{P}_{jk}$, one obtains another permutation $\hat{P}' = \hat{P}_{jk}\hat{P}$ with $p' = p\pm 1$. This has due to $\epsilon=\pm1$ the following consequence:

$$\hat{P}_{jk}\hat{S}^\epsilon = \sum_{\hat{P}}\epsilon^p\hat{P}_{jk}\hat{P} = \sum_{\hat{P}'}\epsilon^{p'\mp1}\hat{P}' = \epsilon\sum_{\hat{P}'}\epsilon^{p'}\hat{P}' = \epsilon\hat{S}^\epsilon\,. \tag{2.56}$$

With the prescription

$$\psi^{\epsilon}_{\{E_\alpha\}}(\mathbf{x}_1,\dots,\mathbf{x}_n) = N^{\epsilon}_{\{E_\alpha\}} \hat{S}^{\epsilon} \prod_{\nu=1}^{n} \psi_{E_{\alpha_\nu}}(\mathbf{x}_\nu) \tag{2.57}$$

we construct for each wave function (2.51) of n distinguishable particles a corresponding symmetric ($\epsilon = 1$) or anti-symmetric $\epsilon = -1$ n-particle wave function, which obeys (2.43) by taking (2.56) into account. Due to the indistinguishability property the (anti-)symmetrized n-particle wave function (2.57) turns out to be independent of the concrete order of the 1-particle energies $E_{\alpha_1},\dots,E_{\alpha_1}$. In order to emphasize that within our notation, we have introduced in (2.57) the index $\{E_\alpha\}$.

At first, we remark that the (anti-)symmetrized n-particle wave function (2.57) obeys the time-independent Schrödinger equation (2.47) with the energy eigenvalue (2.52). This follows from (2.37) as well as the circumstance (2.27) that each permutation operator $\hat{P}$ in the sum (2.55) can be represented as a product of transposition operators:

$$\hat{H}|\psi_E\rangle = E|\psi_E\rangle \qquad \Longrightarrow \qquad \hat{H}\hat{S}^{\epsilon}|\psi_E\rangle = \hat{S}^{\epsilon}\hat{H}|\psi_E\rangle = E\hat{S}^{\epsilon}|\psi_E\rangle\,. \tag{2.58}$$

Furthermore, we read off from (2.55) and (2.57) an important observation for the anti-symmetric n-particle wave function, which is characterized by $\epsilon = -1$:

$$\psi^{-}_{\{E_\alpha\}}(\mathbf{x}_1,\dots,\mathbf{x}_n) = N^{-}_{\{E_\alpha\}} \sum_{\hat{P}} (-1)^p\, \psi_{E_{\alpha_1}}(\mathbf{x}_{P(1)}) \cdots \psi_{E_{\alpha_n}}(\mathbf{x}_{P(n)})\,. \tag{2.59}$$

Thus, the anti-symmetric n-particle wave function can be represented in form of a Slater determinant:

$$\psi^{-}_{\{E_\alpha\}}(\mathbf{x}_1,\dots,\mathbf{x}_n) = N^{-}_{\{E_\alpha\}} \begin{vmatrix} \psi_{E_{\alpha_1}}(\mathbf{x}_1) & \psi_{E_{\alpha_1}}(\mathbf{x}_2) & \cdots & \psi_{E_{\alpha_1}}(\mathbf{x}_n) \\ \vdots & \vdots & & \vdots \\ \psi_{E_{\alpha_n}}(\mathbf{x}_1) & \psi_{E_{\alpha_n}}(\mathbf{x}_2) & \cdots & \psi_{E_{\alpha_n}}(\mathbf{x}_n) \end{vmatrix}\,. \tag{2.60}$$

In the case of an equality of two rows, i.e. $\alpha_j = \alpha_k$, or two columns, i.e. $\mathbf{x}_j = \mathbf{x}_k$, the anti-symmetric n-particle wave function (2.60) vanishes and with this the probability to have such a wave function. This just represents the fundamental Pauli exclusion principle that two fermions can not be neither in the same state nor at the same space point. A corresponding restriction does not exist for bosons. This means that there can be more than one boson in one state or at one space point. In order not to overload the following combinatorial considerations, we consider from now on only those bosonic wavefunctions, where a state or a space point is occupied at most by one boson.

It remains to determine the normalization constant $N^{\epsilon}_{\{E_\alpha\}}$ in (2.57). To this end we apply (2.27), that each permutation operator $\hat{P}$ can be represented by transpositions $\hat{P}_{jk}$, and conclude that iterating (2.56) yields

$$\hat{P}\hat{S}^{\epsilon} = \epsilon^p \hat{S}^{\epsilon}\,. \tag{2.61}$$

Taking into account (2.55) the scalar product between two (anti-)symmetric n-particle wave functions (2.57) reads at first

$$\langle\psi^{\epsilon}_{\{E_\alpha\}}|\psi^{\epsilon}_{\{E_{\alpha'}\}}\rangle = N^{\epsilon}_{\{E_\alpha\}}\sum_{\hat{P}}\epsilon^{p}\langle\psi_{E_{\alpha_1}}\cdots\psi_{E_{\alpha_n}}|\hat{P}^{\dagger}\psi^{\epsilon}_{\{E_{\alpha'}\}}\rangle\,. \tag{2.62}$$

Due to (2.27) and (2.35) as well as (2.57) and (2.61) this reduces to

$$\langle\psi^{\epsilon}_{\{E_\alpha\}}|\psi^{\epsilon}_{\{E_{\alpha'}\}}\rangle = N^{\epsilon}_{\{E_\alpha\}}\sum_{\hat{P}}\epsilon^{2p}\langle\psi_{E_{\alpha_1}}\cdots\psi_{E_{\alpha_n}}|\psi^{\epsilon}_{\{E_{\alpha'}\}}\rangle\,. \tag{2.63}$$

As we have $\epsilon=\pm1$, the summand turns out to be independent of the respective permutations $\hat{P}$, so the sum reduces to the factor $n!$, which is the number of all possible permutations. Taking into account again (2.57) and (2.61) we get

$$\langle\psi^{\epsilon}_{\{E_\alpha\}}|\psi^{\epsilon}_{\{E_{\alpha'}\}}\rangle = N^{\epsilon}_{\{E_\alpha\}}N^{\epsilon}_{\{E_{\alpha'}\}}n!\sum_{\hat{P}'}\epsilon^{p'}\langle\psi_{E_{\alpha_1}}\cdots\psi_{E_{\alpha_n}}|\hat{P}'\psi_{E_{\alpha'_1}}\cdots\psi_{E_{\alpha'_n}}\rangle\,. \tag{2.64}$$

And with the orthonormality (2.53) of the 1-particle wavefunctions we finally obtain for the scalar product the expression

$$\langle\psi^{\epsilon}_{\{E_\alpha\}}|\psi^{\epsilon}_{\{E_{\alpha'}\}}\rangle = N^{\epsilon}_{\{E_\alpha\}}N^{\epsilon}_{\{E_{\alpha'}\}}n!\sum_{\hat{P}'}\epsilon^{p'}\delta_{\alpha_1,\alpha'_{P'(1)}}\cdots\delta_{\alpha_1,\alpha'_{P'(n)}}\,. \tag{2.65}$$

We now demand the orthonormality relation

$$\langle\psi^{\epsilon}_{\{E_\alpha\}}|\psi^{\epsilon}_{\{E_{\alpha'}\}}\rangle = \delta^{\epsilon}_{\alpha_1,\ldots,\alpha_n;\alpha'_1,\ldots\alpha'_n} \tag{2.66}$$

with the (anti-)symmetrized Kronecker symbol

$$\delta^{\epsilon}_{\alpha_1,\ldots,\alpha_n;\alpha'_1,\ldots\alpha'_n} = \sum_{\hat{P}}\epsilon^{p}\,\delta_{\alpha_1,\alpha'_{P(1)}}\cdots\delta_{\alpha_1,\alpha'_{P(n)}}\,. \tag{2.67}$$

As we restrict ourselves both for bosons and fermions to the case that all single-particle states differ from each other, i.e. $\alpha_\mu\neq\alpha_\nu$ for $\mu\neq\nu$, in (2.66) and (2.67) only the identity permutation $\hat{P}=1$ survives, which fixes the normalization constant $N^{\epsilon}_{\{E_\alpha\}}$ according to

$$N^{\epsilon}_{\{E_\alpha\}} = \frac{1}{\sqrt{n!}}\,. \tag{2.68}$$

Finally, we show that one can span the whole Hilbert space of (anti-)symmetrized n-particle wave functions with (2.57). To this end we start from the completeness relation (2.54) of the n-particle wave function and apply twice the (anti-)symmetrization operator (2.55), once upon the space coordinates $\mathbf{x}_1,\ldots,\mathbf{x}_n$ and once upon the space coordinates $\mathbf{x}'_1,\ldots,\mathbf{x}'_n$:

$$\begin{aligned}\sum_{\hat{P}}\sum_{\hat{P}'}\epsilon^{p+p'}\sum_{\alpha_1}\cdots\sum_{\alpha_n}\psi^{*}_{E_{\alpha_1}}(\mathbf{x}_{P(1)})\cdots\psi^{*}_{E_{\alpha_n}}(\mathbf{x}_{P(n)})\psi_{E_{\alpha_1}}(\mathbf{x}'_{P'(1)})\cdots\psi_{E_{\alpha_n}}(\mathbf{x}'_{P'(n)})\\ =\sum_{\hat{P}}\sum_{\hat{P}'}\epsilon^{p+p'}\delta(\mathbf{x}_{P(1)}-\mathbf{x}'_{P'(1)})\cdots\delta(\mathbf{x}_{P(n)}-\mathbf{x}'_{P'(n)})\,.\end{aligned} \tag{2.69}$$

At the left-hand side the space coordinates $\mathbf{x}_{P(1)},\ldots,\mathbf{x}_{P(n)}$ and $\mathbf{x}'_{P'(1)},\ldots,\mathbf{x}'_{P'(n)}$ are rearranged in their respective standard order $\mathbf{x}_1,\ldots,\mathbf{x}_n$ and $\mathbf{x}'_1,\ldots,\mathbf{x}'_n$. As a consequence the quantum numbers $\alpha_1,\ldots\alpha_n$ are rearranged to $\alpha_{P(1)},\ldots\alpha_{P(n)}$ and $\alpha_{P'(1)},\ldots\alpha_{P'(n)}$, respectively. A corresponding reordering on the right-hand side from $\mathbf{x}_{P(1)},\ldots,\mathbf{x}_{P(n)}$ to $\mathbf{x}_1,\ldots,\mathbf{x}_n$ rearranges then $\mathbf{x}'_{P'(1)},\ldots,\mathbf{x}'_{P'(n)}$ to $\mathbf{x}'_{P'(P(1))},\ldots,\mathbf{x}'_{P'(P(n))}$, yielding

$$\sum_{\alpha_1}\cdots\sum_{\alpha_n}\Big\{\sum_{\hat{P}}\epsilon^p\psi^*_{E_{\alpha_{P(1)}}}(\mathbf{x}_1)\cdots\psi^*_{E_{\alpha_{P(n)}}}(\mathbf{x}_n)\Big\}\Big\{\sum_{\hat{P}'}\epsilon^{p'}\psi^*_{E_{\alpha_{P'(1)}}}(\mathbf{x}'_1)\cdots\psi^*_{E_{\alpha_{P'(n)}}}(\mathbf{x}'_n)\Big\}$$
$$=\sum_{\hat{P}}\sum_{\hat{P}'}\epsilon^{p+p'}\delta(\mathbf{x}_1-\mathbf{x}'_{P'(P(1))})\cdots\delta(\mathbf{x}_n-\mathbf{x}'_{P'(P(n))})\,. \tag{2.70}$$

At the left-hand side we now use (2.55), (2.57), and (2.68), whereas at the right-hand side the sum over all permutations $\hat{P}'$ is substituted by an equivalent sum over all permutations $\hat{Q}=\hat{P}'\hat{P}$ with $q=p'+p$, so that afterwards the sum over $\hat{P}$ can straight-forwardly be performed. With this we finally obtain the completeness relation

$$\sum_{\alpha_1}\cdots\sum_{\alpha_n}\psi^{\epsilon *}_{\{E_\alpha\}}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\psi^{\epsilon}_{\{E_\alpha\}}(\mathbf{x}'_1,\ldots,\mathbf{x}'_n)=\delta^\epsilon(\mathbf{x}_1,\ldots,\mathbf{x}_n;\mathbf{x}'_1,\ldots,\mathbf{x}'_n)\,, \tag{2.71}$$

where we have introduced analogous to (2.67) the (anti-)symmetrized delta function

$$\delta^\epsilon(\mathbf{x}_1,\ldots,\mathbf{x}_n;\mathbf{x}'_1,\ldots,\mathbf{x}'_n)=\sum_{\hat{Q}}\epsilon^q\delta(\mathbf{x}_1-\mathbf{x}'_{Q(1)})\cdots\delta(\mathbf{x}_n-\mathbf{x}'_{Q(n)})\,. \tag{2.72}$$

The considerations of the present section have the purpose to generate a basis of the Hilbert space of indistinguishable identical particles via a(n) (anti-)symmetrization of the known basis of the Hilbert space of distinguishable identical particles. So far the starting point has been the eigenvalue problem (2.47) of the underlying Hamilton operator. But another basis results from considering the eigenvalue problem (2.15) of the coordinate operators as the starting point. Then the eigenfunctions $|\mathbf{x}_1,\ldots,\mathbf{x}_n\rangle$ with the continuous eigenvalues $\mathbf{x}_1,\ldots,\mathbf{x}_n$ span the Hilbert space of distinguishable identical particles. The subsequent (anti-)symmetrization is performed analogous to (2.55), (2.57), and (2.68), yielding another basis in the Hilbert space of indistinguishable identical particles:

$$|\mathbf{x}_1,\ldots,\mathbf{x}_n\rangle^\epsilon=\frac{1}{\sqrt{n!}}\sum_{\hat{P}}\epsilon^p|\mathbf{x}_{P(1)},\ldots,\mathbf{x}_{P(n)}\rangle\,. \tag{2.73}$$

Both the orthonormality relation and the completeness relation corresponding to (2.66) and (2.71) read then

$$^\epsilon\langle\mathbf{x}_1,\ldots,\mathbf{x}_n|\mathbf{x}'_1,\ldots,\mathbf{x}'_n\rangle^\epsilon=\delta^\epsilon(\mathbf{x}_1,\ldots,\mathbf{x}_n;\mathbf{x}'_1,\ldots,\mathbf{x}'_n)\,, \tag{2.74}$$
$$\int d^3x_1\cdots\int d^3x_n\,|\mathbf{x}_1,\ldots,\mathbf{x}_n\rangle^\epsilon\;{}^\epsilon\langle\mathbf{x}_1,\ldots,\mathbf{x}_n|=1\,. \tag{2.75}$$

For the purpose of illustration we consider the spatial representation for two particles. The basis for two distinguishable identical particles reads in coordinate representation according to (2.16) and (2.20)

$$\psi_{\mathbf{x}_1,\mathbf{x}_2}(\mathbf{z}_1,\mathbf{z}_2)=\langle\mathbf{z}_1,\mathbf{z}_2|\mathbf{x}_1,\mathbf{x}_2\rangle=\delta(\mathbf{z}_1-\mathbf{x}_1)\delta(\mathbf{z}_2-\mathbf{x}_2)\,. \tag{2.76}$$

Correspondingly, the coordinate representation for two indistinguishable particles follows from (2.20):

$$\psi^{\epsilon}_{\mathbf{x}_1,\mathbf{x}_2}(\mathbf{z}_1,\mathbf{z}_2) = \langle \mathbf{z}_1,\mathbf{z}_2|\mathbf{x}_1,\mathbf{x}_2\rangle^{\epsilon}\,, \tag{2.77}$$

which reduces due to (2.16) and (2.73) to

$$\psi^{\epsilon}_{\mathbf{x}_1,\mathbf{x}_2}(\mathbf{z}_1,\mathbf{z}_2) = \frac{1}{\sqrt{2}}\Big\{\delta(\mathbf{z}_1-\mathbf{x}_1)\delta(\mathbf{z}_2-\mathbf{x}_2)+\epsilon\delta(\mathbf{z}_1-\mathbf{x}_2)\delta(\mathbf{z}_2-\mathbf{x}_1)\Big\}\,. \tag{2.78}$$

Note that (2.78) also follows from an (anti-)symmetrization (2.73) from (2.76) as defined by (2.55), (2.57), and (2.68). With this one obtains for the orthonormality relation (2.74) by taking into account (2.72)

$$\int d^3z_1 \int d^3z_2\, \psi^{\epsilon *}_{\mathbf{x}_1,\mathbf{x}_2}(\mathbf{z}_1,\mathbf{z}_2)\psi^{\epsilon}_{\mathbf{x}_1',\mathbf{x}_2'}(\mathbf{z}_1,\mathbf{z}_2) = \delta^{\epsilon}(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1',\mathbf{x}_2') \tag{2.79}$$

and correspondingly the completeness relation (2.75) reads together with (2.74) and (2.76)

$$\int d^3x_1 \int d^3x_2\, \psi^{\epsilon *}_{\mathbf{x}_1,\mathbf{x}_2}(\mathbf{z}_1,\mathbf{z}_2)\psi^{\epsilon}_{\mathbf{x}_1,\mathbf{x}_2}(\mathbf{z}_1',\mathbf{z}_2') = \delta^{\epsilon}(\mathbf{z}_1,\mathbf{z}_2;\mathbf{z}_1',\mathbf{z}_2')\,. \tag{2.80}$$

Chapter 3

Second Quantization

The formulation of quantum many-body systems introduced so far dealt first with distinguishable particles and necessitated then to perform afterwards a(n) (anti-)symmetrization of wave functions in order to describe indistinguishable particles in form of bosons (fermions). Usually this procedure turns out to be quite cumbersome due the huge number of particles involved in a quantum many-body system. Therefore, one has worked out second quantization as an alternative formulation for describing quantum many-body systems, which has the advantage that it automatically takes into account the (anti-)symmetrization of wave functions. It is based on the ladder formalism, which allows an algebraic treatment of the first quantized harmonic oscillator and is therefore initially reviewed. Afterwards, we heuristically formulate second quantization, which represents the technical basis for non-relativistic quantum many-body theory. Due to the introduction of creation and annihilation operators for identical particles we are able to describe interacting bosonic and fermionic systems involving an arbitrary number of particles. This is relevant for concrete applications in the realm of solid-state physics like the description of Bose-Einstein condensation and superfluidity as well as the Bardeen-Cooper-Schrieffer theory of superconductivity, which is not the content of this lecture. But, a similar second quantization formalism is later on used to quantize relativistic fields like the Maxwell and the Dirac field and, thus, represents the very basis for quantum electrodynamics.

3.1 Harmonic Oscillator

The harmonic oscillator represents a standard quantum mechanical model with which it is possible to describe quite successfully, for instance, collective oscillations in molecules or in solids. The Hamilton operator of a one-dimensional harmonic oscillator with mass M and frequency ω reads

$$\hat{H} = \frac{\hat{p}^2}{2M} + \frac{M}{2}\,\omega^2\hat{x}^2\,, \tag{3.1}$$

where one demands non-trivial commutation relations between the coordinate operator $\hat{q}$ and the momentum operator $\hat{p}$ analogous to (2.12):

$$\left[\hat{x},\hat{x}\right]_- = \left[\hat{p},\hat{p}\right]_- = 0\,, \qquad \left[\hat{p},\hat{x}\right]_- = \frac{\hbar}{i}\,. \tag{3.2}$$

The problem is now to solve the eigenvalue problem of the Hamilton operator

$$\hat{H}|\alpha\rangle = E_\alpha|\alpha\rangle\,, \tag{3.3}$$

i.e. to determine how the energy eigenvalues E_α and the energy eigenfunctions $|\alpha\rangle$ depend on the quantum number α. Usually this representation-free eigenvalue problem (3.3) is transformed into the coordinate representation, so it amounts to solve the corresponding Schrödinger equation by taking into account the appropriate Dirichlet boundary condition. In the following, however, we proceed differently by solving the representation-free eigenvalue problem (3.3) directly by taking into account the cummator relations (3.2).

At first, the two hermitian operators $\hat{x}$ and $\hat{p}$ are transformed into two new operators $\hat{a}^\dagger$ and $\hat{a}$, which are adjoint with respect to each other:

$$\hat{a}^\dagger = \sqrt{\frac{M\omega}{2\hbar}}\left(\hat{x} - \frac{i}{M\omega}\,\hat{p}\right)\,, \qquad \hat{a} = \sqrt{\frac{M\omega}{2\hbar}}\left(\hat{x} + \frac{i}{M\omega}\,\hat{p}\right)\,. \tag{3.4}$$

The inverse transformation reads correspondingly

$$\hat{x} = \sqrt{\frac{\hbar}{2M\omega}}\,\left(\hat{a}^\dagger + \hat{a}\right)\,, \qquad \hat{p} = \sqrt{\frac{\hbar M\omega}{2}}\,i\left(\hat{a}^\dagger - \hat{a}\right)\,. \tag{3.5}$$

Here the physical dimension of the coordinate operator $\hat{x}$ is provided by the oscillator length $\sqrt{\hbar/(2M\omega)}$, whereas the corresponding one $\sqrt{\hbar M\omega/2}$ of the momentum operator $\hat{p}$ is related to the oscillator length via the Heisenberg uncertainty relation. Inserting (3.5) into (3.1), the Hamilton operator of the harmonic oscillator can be expressed in terms of the new operators $\hat{a}^\dagger$ and $\hat{a}$, yielding

$$\hat{H} = \frac{\hbar\omega}{2}\left(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger\right)\,. \tag{3.6}$$

Furthermore, the transformation (3.4) allows to deduce from (3.2) the commutation relations between the new operators $\hat{a}^\dagger$ and $\hat{a}$:

$$\left[\hat{a},\hat{a}\right]_- = \left[\hat{a}^\dagger,\hat{a}^\dagger\right]_- = 0\,, \qquad \left[\hat{a},\hat{a}^\dagger\right]_- = 1\,. \tag{3.7}$$

Using (3.7) the Hamilton operator of the harmonic oscillator (3.6) reduces to

$$\hat{H} = \hbar\omega\left(\hat{n} + \frac{1}{2}\right)\,, \tag{3.8}$$

where the zero-point energy $\hbar\omega/2$ and the operator

$$\hat{n} = \hat{a}^\dagger\hat{a} \tag{3.9}$$

appear. In order to calculate commutators the following identity turns out to be quite useful

$$\left[\hat{A}\hat{B},\hat{C}\right]_{-}=\hat{A}\left[\hat{B},\hat{C}\right]_{-}+\left[\hat{A},\hat{C}\right]_{-}\hat{B}\,, \tag{3.10}$$

which follows immediately from the definition of the commutator (2.13). Indeed, applying (3.10) we obtain the commutation relations for the operator (3.9):

$$\left[\hat{n},\hat{a}^{\dagger}\right]_{-} = \hat{a}^{\dagger}\,, \tag{3.11}$$

$$\left[\hat{n},\hat{a}\right]_{-} = -\hat{a}\,. \tag{3.12}$$

Let us now consider the eigenvalue problem of the operator (3.9):

$$\hat{n}|\lambda\rangle=\lambda|\lambda\rangle \tag{3.13}$$

As the operator (3.9) is hermitian, its eigenvalues λ must be real. Furthermore, the commutation relations (3.11) and (3.12) allow to investigate which consequences occur once the operators $\hat{a}^{\dagger}$ and $\hat{a}$ are applied to the eigenfunctions $|\lambda\rangle$. On the one hand we read off from (3.11) and (3.13)

$$\hat{n}\hat{a}^{\dagger}|\lambda\rangle=\left(\hat{a}^{\dagger}\hat{n}+\hat{a}^{\dagger}\right)|\lambda\rangle=(\lambda+1)\hat{a}^{\dagger}|\lambda\rangle \qquad \Longrightarrow \qquad \hat{a}^{\dagger}|\lambda\rangle\sim|\lambda+1\rangle\,, \tag{3.14}$$

on the other hand we conclude from (3.12) and (3.13)

$$\hat{n}\hat{a}|\lambda\rangle=\left(\hat{a}\hat{n}-\hat{a}\right)=(\lambda-1)\hat{a}|\lambda\rangle \qquad \Longrightarrow \qquad \hat{a}|\lambda\rangle\sim|\lambda-1\rangle\,. \tag{3.15}$$

Thus, the operators $\hat{a}^{\dagger}$ and $\hat{a}$ can be considered as ladder operators, which allow to climb up or down the ladder of eigenfunctions $|\lambda\rangle$. Applying the raising (lowering) ladder operator $\hat{a}^{\dagger}$ ($\hat{a}$) to $|\lambda\rangle$ yields an eigenfunction corresponding to an eigenvalue which is increased (decreased) by one, see Fig. 3.1

Furthermore, one can show that the eigenvalues λ of the operator $\hat{N}$ are always positive by taking into account (3.9) and (3.13) and by assuming without loss of generality that the eigenfunctions $|\lambda\rangle$ are normalized:

$$0\leq\langle\hat{a}\lambda|\hat{a}\lambda\rangle=\langle\lambda|\hat{a}^{\dagger}\hat{a}|\lambda\rangle=\langle\lambda|\hat{n}|\lambda\rangle=\lambda\langle\lambda|\lambda\rangle=\lambda\,. \tag{3.16}$$

From (3.15) and (3.16) we conclude that the eigenvalues λ are given by positive integer number including zero:

$$\lambda=n=0,1,2,\ldots\,. \tag{3.17}$$

If there were a positive, non-integer eigenvalue λ, one could apply iteratively the lowering ladder operator $\hat{a}$ and reduce in this way the eigenvalue due to (3.15) until it would become negative. But this would then contradict the inequality (3.16). Thus, due to this contradiction proof, there must be a ground state $|0\rangle$ with the property

$$\hat{a}|0\rangle=0 \qquad \Longleftrightarrow \qquad \langle 0|\hat{a}^{\dagger}=0\,. \tag{3.18}$$

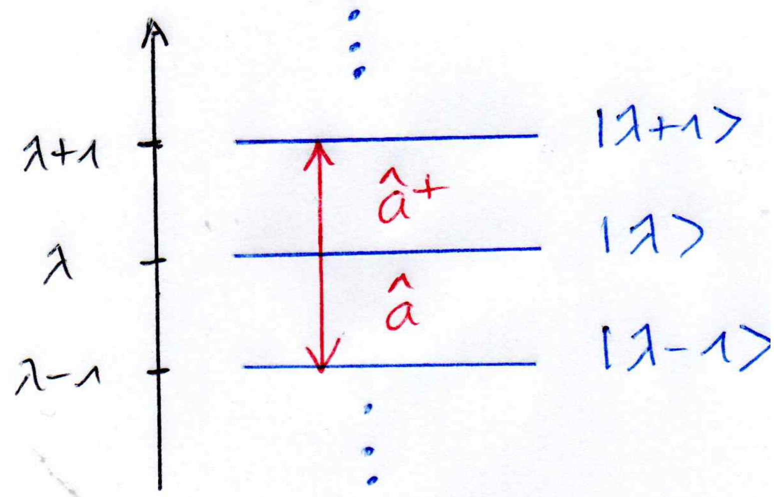


Figure 3.1: Raising (lowering) operator $\hat{a}^\dagger$ ($\hat{a}$) increases (decreases) the quantum number λ of the harmonic oscillator by one.

Normalized eigenfunctions $|n\rangle$ can then be constructed as follows. At first, we deduce from (3.7), (3.9), (3.13), and (3.17):

$$\langle \hat{a}^\dagger n|\hat{a}^\dagger n\rangle = \langle n|\hat{a}\hat{a}^\dagger|n\rangle = \langle n|\big(\hat{a}^\dagger\hat{a}+1\big)|n\rangle = \langle n|\big(\hat{n}+1\big)|n\rangle = n+1\,. \tag{3.19}$$

From (3.14), (3.17), and (3.19) follows a rule how applying the raising ladder operator $\hat{a}^\dagger$ upon the normalized eigenfunction $|n\rangle$ yields the next normalized eigenfunction $|n+1\rangle$:

$$\hat{a}^\dagger|n\rangle = C_n|n+1\rangle \quad\Longrightarrow\quad C_n^2\langle n+1|n+1\rangle = (n+1) \quad\Longrightarrow\quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{3.20}$$

And then iterating (3.20) yields a prescription how the eigenfunctions $|n\rangle$ can be constructed from the ground state $|0\rangle$ defined by (3.18):

$$|n\rangle = \frac{1}{\sqrt{n}}\,\hat{a}^\dagger|n-1\rangle = \frac{1}{\sqrt{n(n-1)}}\,\big(\hat{a}^2\big)^2|n-2\rangle = \ldots \quad\Longrightarrow\quad |n\rangle = \frac{1}{\sqrt{n!}}\,\big(\hat{a}^\dagger\big)^n|0\rangle\,. \tag{3.21}$$

For the sake of completeness we also determine the action of the lowering ladder operator $\hat{a}$ upon the eigenfunction $|n\rangle$. At first we obtain from (2.68), (3.13), and (3.17)

$$\langle \hat{a}n|\hat{a}n\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle n|\hat{n}|n\rangle = n\,. \tag{3.22}$$

Thus, we conclude from (3.15) and (3.22)

$$\hat{a}|n\rangle = D_n|n-1\rangle \quad\Longrightarrow\quad D_n^2\langle n-1|n-1\rangle = n \quad\Longrightarrow\quad \hat{a}|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{3.23}$$

Furthermore, we read off from (3.8), (3.9), (3.13), and (3.17) the energy eigenvalues of the harmonic oscillator

$$E_n = \hbar\omega\left(n+\frac{1}{2}\right)\,. \tag{3.24}$$

3.2 Creation and Annihilation Operators for Bosons

This ladder formalism for the algebraic treatment of the first quantized harmonic oscillator is now used in the realm of second quantization for describing indistinguishable identical bosons. We outline heuristically this basic idea by working out the analogy step by step:

- Whereas n describes the quantum number of the 1-particle system, we denote from now on with $n_{\mathbf{x}}$ the number of bosons at space point $\mathbf{x}$.

- The ladder operators $\hat{a}^\dagger$ and $\hat{a}$, which are defined by the commutator relations (3.7), allow to increase and decrease the quantum number n of the harmonic oscillator. Correspondingly we introduce operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ via the commutator relations

$$\left[\hat{a}_{\mathbf{x}}, \hat{a}_{\mathbf{x}'}\right]_- = \left[\hat{a}^\dagger_{\mathbf{x}}, \hat{a}^\dagger_{\mathbf{x}'}\right]_- = 0\,, \qquad \left[\hat{a}_{\mathbf{x}}, \hat{a}^\dagger_{\mathbf{x}'}\right]_- = \delta(\mathbf{x}-\mathbf{x}')\,. \tag{3.25}$$

  With these commutator relations at hand, we can now proceed and deduce similar conclusions for the second quantized description of many bosons as we have just obtained for the first quantized harmonic oscillator. In particular, this allows to determine a concrete physical interpretation for the operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$.

- The operator $\hat{n} = \hat{a}^\dagger\hat{a}$ has turned out to have the eigenvalues n, which follows ultimately from the commutator relations (3.11) and (3.12). Analogously we define the particle number operator

$$\hat{N} = \int d^3x'\, \hat{a}^\dagger_{\mathbf{x}'}\hat{a}_{\mathbf{x}'} \tag{3.26}$$

  which obeys due to (3.10), (3.25), and (3.26) the commutator relations

$$\left[\hat{N}, \hat{a}^\dagger_{\mathbf{x}}\right]_- = \hat{a}^\dagger_{\mathbf{x}}\,, \tag{3.27}$$

$$\left[\hat{N}, \hat{a}_{\mathbf{x}}\right]_- = -\hat{a}_{\mathbf{x}}\,. \tag{3.28}$$

  Note that we have deliberately introduced in the commutator relations (3.25) a delta function in order to obtain for the particle number operator (3.26) commutator relations (3.27), (3.28) in analogy to (3.11) and (3.12). This has the consequence that the operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ can be interpreted as a creation and annihilator operator as they create and annihilate a boson at space point $\mathbf{x}$, respectively.

- The first quantized harmonic oscillator has a ground state $|0\rangle$, which is introduced according to (3.18). In a similar way we define in second quantization a vacuum state $|0\rangle$ via

$$\hat{a}_{\mathbf{x}}|0\rangle = 0 \qquad \Longleftrightarrow \qquad \langle 0|\hat{a}^\dagger_{\mathbf{x}} = 0\,. \tag{3.29}$$

- Similar to (3.21) an iterative application of creation operators to the vacuum state yields the basis states of the underlying Hilbert space for describing bosons

$$|\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle^{+1} = \hat{a}^\dagger_{\mathbf{x}_1} \cdots \hat{a}^\dagger_{\mathbf{x}_n} |0\rangle \,, \tag{3.30}$$

where we assume that the space coordinates differ pairwise, i.e. $\mathbf{x}_i \neq \mathbf{x}_j$ for all $i \neq j$. For the sake of illustration we exemplary verify the identity of (2.73) and (3.30) for $n = 1$ and $n = 2$ bosons in the coordinate representation. From (2.72), (3.25), (3.29), and (3.30) we obtain at first

$$\begin{aligned} {}^{+1}\langle \mathbf{x}_1 | \mathbf{x}_1' \rangle^{+1} &= \langle \hat{a}^\dagger_{\mathbf{x}_1} 0 | \hat{a}^\dagger_{\mathbf{x}_1'} 0 \rangle = \langle 0 | \hat{a}_{\mathbf{x}_1} \hat{a}^\dagger_{\mathbf{x}_1'} | 0 \rangle \\ &= \langle 0 | \hat{a}^\dagger_{\mathbf{x}_1'} \hat{a}_{\mathbf{x}_1} + \delta(\mathbf{x}_1 - \mathbf{x}_1') | 0 \rangle = \delta(\mathbf{x}_1 - \mathbf{x}_1') = \delta^{+1}(\mathbf{x}_1; \mathbf{x}_1') \,. \end{aligned} \tag{3.31}$$

Correspondingly, we get then

$$\begin{aligned} &{}^{+1}\langle \mathbf{x}_1, \mathbf{x}_2 | \mathbf{x}_1', \mathbf{x}_2' \rangle^{+1} = \langle \hat{a}^\dagger_{\mathbf{x}_1} \hat{a}^\dagger_{\mathbf{x}_2} 0 | \hat{a}^\dagger_{\mathbf{x}_1'} \hat{a}^\dagger_{\mathbf{x}_2'} 0 \rangle = \langle 0 | \hat{a}_{\mathbf{x}_2} \hat{a}_{\mathbf{x}_1} \hat{a}^\dagger_{\mathbf{x}_1'} \hat{a}^\dagger_{\mathbf{x}_2'} | 0 \rangle \\ &= \langle 0 | \hat{a}_{\mathbf{x}_2} \left[\hat{a}^\dagger_{\mathbf{x}_1'} \hat{a}_{\mathbf{x}_1} + \delta(\mathbf{x}_1 - \mathbf{x}_1') \right] \hat{a}^\dagger_{\mathbf{x}_2'} | 0 \rangle = \delta(\mathbf{x}_1 - \mathbf{x}_1') \langle 0 | \hat{a}^\dagger_{\mathbf{x}_2'} \hat{a}_{\mathbf{x}_2} + \delta(\mathbf{x}_2 - \mathbf{x}_2') | 0 \rangle \\ &+ \langle 0 | \left[\hat{a}^\dagger_{\mathbf{x}_1'} \hat{a}_{\mathbf{x}_2} + \delta(\mathbf{x}_1' - \mathbf{x}_2) \right] \left[\hat{a}^\dagger_{\mathbf{x}_2'} \hat{a}_{\mathbf{x}_1} + \delta(\mathbf{x}_1 - \mathbf{x}_2') \right] | 0 \rangle = \delta(\mathbf{x}_1 - \mathbf{x}_1') \delta(\mathbf{x}_2 - \mathbf{x}_2') \\ &+ \delta(\mathbf{x}_1 - \mathbf{x}_2') \delta(\mathbf{x}_2 - \mathbf{x}_1') = \delta^{+1}(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_1', \mathbf{x}_2') \,. \end{aligned} \tag{3.32}$$

3.3 Schrödinger Equation for Interacting Bosons

Introducing local creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ has not only the advantage of constructing many-particle states, which automatically have the correct symmetry. In addition one obtains a universal form of the time-dependent Schrödinger equation, which turns out to be independent of the particle number n. In its representation-independent form it reads

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \,. \tag{3.33}$$

Here $|\psi(t)\rangle$ denotes some many-particle state in the second-quantized Hilbert space, which is spanned by the basis states (3.30). The second-quantized Hamilton operator $\hat{H}$ consists of two terms:

$$\hat{H} = \hat{H}_1 + \hat{H}_2 \,. \tag{3.34}$$

The local Hamilton operator $\hat{H}_1$ is determined the 1-particle Hamilton operator of non-interacting bosons

$$-\frac{\hbar^2}{2M} \Delta + V_1(\mathbf{x}) \,. \tag{3.35}$$

Due to the sandwich principle the first-quantized Hamilton operator (3.35) is multiplied with the local creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ to the left and to the right, respectively, so a subsequent integration over the coordinate $\mathbf{x}$ yields the corresponding second-quantized 1-particle Hamilton operator:

$$\hat{H}_1 = \int d^3x \, \hat{a}^\dagger_{\mathbf{x}} \left\{ -\frac{\hbar^2}{2M} \, \Delta + V_1(\mathbf{x}) \right\} \hat{a}_{\mathbf{x}} \, . \tag{3.36}$$

Correspondingly the bi-local Hamilton operator $\hat{H}_2$ is constructed with the help of the 2-particle interaction $V_2(\mathbf{x} - \mathbf{x}')$:

$$\hat{H}_2 = \frac{1}{2} \int d^3x \int d^3x' \, \hat{a}^\dagger_{\mathbf{x}} \hat{a}^\dagger_{\mathbf{x}'} \, V_2(\mathbf{x} - \mathbf{x}') \hat{a}_{\mathbf{x}'} \, \hat{a}_{\mathbf{x}} \, . \tag{3.37}$$

Note that in both terms (3.36) and (3.37) the creation and annihilation operators appear at the left and at the right, respectively. This particular ordering of second-quantized operators is called normal ordering. It has the consequence that the vacuum energy of the Hamilton operator defined by (3.34), (3.36), and (3.37) vanishes due to the definition of the vacuum state in (3.29):

$$\hat{H}|0\rangle = 0 \qquad \Longleftrightarrow \qquad \langle 0|\hat{H} = 0 \, . \tag{3.38}$$

In the following we demonstrate that the operator character of $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ is essential for the fact that the Schrödinger equation (3.33) describes a many-body problem. To this end we multiply (3.33) from the left with the adjoint of the basis state (3.30)

$$^{+1}\langle \mathbf{x}_1, \ldots, \mathbf{x}_n| = \langle 0|\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1} \, . \tag{3.39}$$

With this we get at first

$$i\hbar \, \frac{\partial}{\partial t} \, \langle 0|\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}|\psi(t)\rangle = \langle 0|\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}\hat{H}|\psi(t)\rangle \, . \tag{3.40}$$

Due to (3.38) we can express the right-hand side of (3.40) in terms of a commutator

$$i\hbar \, \frac{\partial}{\partial t} \, \langle 0|\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}|\psi(t)\rangle = \langle 0| \left[\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}, \hat{H}\right]_- |\psi(t)\rangle \, . \tag{3.41}$$

Taking into account both contributions (3.36) and (3.37) of the Hamilton operator this leads to the expression

$$\begin{aligned} &\int d^3y \int d^3z \, \delta(\mathbf{y} - \mathbf{z}) \left\{ -\frac{\hbar^2}{2M} \, \Delta_{\mathbf{z}} + V_1(\mathbf{z}) \right\} \langle 0| \left[\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}, \hat{a}^\dagger_{\mathbf{y}} \hat{a}_{\mathbf{z}}\right]_- |\psi(t)\rangle + \frac{1}{2} \int d^3y_1 \int d^3y_2 \\ &\cdot \int d^3z_1 \int d^3z_2 \, \delta(\mathbf{y}_1 - \mathbf{z}_1)\delta(\mathbf{y}_2 - \mathbf{z}_2) \, V_2(\mathbf{z}_1 - \mathbf{z}_2)\langle 0| \left[\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}, \hat{a}^\dagger_{\mathbf{y}_1} \hat{a}^\dagger_{\mathbf{y}_2} \hat{a}_{\mathbf{z}_2} \hat{a}_{\mathbf{z}_1}\right]_- |\psi(t)\rangle \, . \end{aligned} \tag{3.42}$$

In order to evaluate the first commutator in (3.42) we use an identity similar to (3.10)

$$\left[\hat{A}, \hat{B}\hat{C}\right]_- = \left[\hat{A}, \hat{B}\right]_- \hat{C} + \hat{B}\left[\hat{A}, \hat{C}\right]_- \, , \tag{3.43}$$

which yields

$$\left[\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}, \hat{a}^\dagger_{\mathbf{y}} \hat{a}_{\mathbf{z}}\right]_- = \left[\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}, \hat{a}^\dagger_{\mathbf{y}}\right]_- \hat{a}_{\mathbf{z}} + \hat{a}^\dagger_{\mathbf{y}} \left[\hat{a}_{\mathbf{x}_n} \cdots \hat{a}_{\mathbf{x}_1}, \hat{a}_{\mathbf{z}}\right]_- \, . \tag{3.44}$$

Note that here the second term vanishes as the annihilation operators commute with respect to each other due to (3.25). Applying now recursively the identity (3.10), we get

$$\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}}\right]_- = \sum_{\nu=1}^{n}\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_{\nu+1}}\left[\hat{a}_{\mathbf{x}_\nu},\hat{a}^\dagger_{\mathbf{y}}\right]_-\hat{a}_{\mathbf{x}_{\nu-1}}\cdots\hat{a}_{\mathbf{x}_1}\,, \tag{3.45}$$

where the remaining commutators yield a delta function $\delta(\mathbf{x}_\nu-\mathbf{y})$ due to (3.25):

$$\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}}\right]_- = \sum_{\nu=1}^{n}\delta(\mathbf{y}-\mathbf{x}_\nu)\,\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_{\nu+1}}\hat{a}_{\mathbf{x}_{\nu-1}}\cdots\hat{a}_{\mathbf{x}_1}\,. \tag{3.46}$$

Thus, the first expectation value in (3.42) yields

$$\langle 0|\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}}\hat{a}_{\mathbf{z}}\right]_-|\psi(t)\rangle = \sum_{\nu=1}^{n}\delta(\mathbf{x}_\nu-\mathbf{y})\langle 0|\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_{\nu+1}}\hat{a}_{\mathbf{z}}\hat{a}_{\mathbf{x}_{\nu-1}}\cdots\hat{a}_{\mathbf{x}_1}|\psi(t)\rangle\,. \tag{3.47}$$

In a similar manner we proceed also for the second commutator in (3.42) by applying the identity (3.43) twice, yielding

$$\begin{aligned}&\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\left(\hat{a}^\dagger_{\mathbf{y}_1}\hat{a}^\dagger_{\mathbf{y}_2}\right)\left(\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\right)\right]_- = \left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}_1}\hat{a}^\dagger_{\mathbf{y}_2}\right]_-\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\\ &= \left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}_1}\right]_-\hat{a}^\dagger_{\mathbf{y}_2}\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}+\hat{a}^\dagger_{\mathbf{y}_1}\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}_2}\right]_-\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\end{aligned} \tag{3.48}$$

Thus, taking into account (3.46) reduces (3.48) to

$$\begin{aligned}&\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}_1}\hat{a}^\dagger_{\mathbf{y}_2}\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\right]_- = \sum_{\nu=1}^{n}\delta(\mathbf{x}_\nu-\mathbf{y}_1)\,\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_{\nu+1}}\hat{a}_{\mathbf{x}_{\nu-1}}\cdots\hat{a}_{\mathbf{x}_1}\hat{a}^\dagger_{\mathbf{y}_2}\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\\ &+\sum_{\nu=1}^{n}\delta(\mathbf{x}_\nu-\mathbf{y}_2)\,\hat{a}^\dagger_{\mathbf{y}_1}\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_{\nu+1}}\hat{a}_{\mathbf{x}_{\nu-1}}\cdots\hat{a}_{\mathbf{x}_1}\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\,.\end{aligned} \tag{3.49}$$

Now we determine the second expectation value in (3.42) from (3.49). Due to (3.29) we observe that the second term in (3.49) then vanishes and the first term can be rewritten as a commutator:

$$\begin{aligned}&\langle 0|\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}_1}\hat{a}^\dagger_{\mathbf{y}_2}\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\right]_-|\psi(t)\rangle\\ &=\sum_{\nu=1}^{n}\delta(\mathbf{x}_\nu-\mathbf{y}_1)\;\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_{\nu+1}}\hat{a}_{\mathbf{x}_{\nu-1}}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}_2}\right]_-\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}|\psi(t)\rangle\,.\end{aligned} \tag{3.50}$$

Using again (3.46) we can then evaluate (3.50):

$$\begin{aligned}&\langle 0|\left[\hat{a}_{\mathbf{x}_n}\cdots\hat{a}_{\mathbf{x}_1},\hat{a}^\dagger_{\mathbf{y}_1}\hat{a}^\dagger_{\mathbf{y}_2}\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{z}_1}\right]_-|\psi(t)\rangle\\ &=\sum_{\nu=1}^{n}\sum_{\mu=1}^{n}\delta(\mathbf{x}_\nu-\mathbf{y}_1)\delta(\mathbf{x}_\mu-\mathbf{y}_2)\langle 0|\hat{a}_{\mathbf{x}_n}\hat{a}_{\mathbf{x}_{\nu+1}}\hat{a}_{\mathbf{z}_1}\hat{a}_{\mathbf{x}_{\nu-1}}\cdots\hat{a}_{\mathbf{x}_{\mu+1}}\hat{a}_{\mathbf{z}_2}\hat{a}_{\mathbf{x}_{\mu-1}}\cdots\hat{a}_{\mathbf{x}_1}|\psi(t)\rangle\,.\end{aligned} \tag{3.51}$$

Finally, inserting the intermediate results (3.47) and (3.51) into the projected Schrödinger equation (3.40) and the expectation value of the Hamilton operator (3.42) as well as performing the integrations over the delta functions yields the n-particle Schrödinger equation (2.21) with (2.23). Here we take into account that the n-particle wave function $\psi(\mathbf{x}_1,\ldots\mathbf{x}_n;t)$ follows from projecting the state $|\psi(t)\rangle$ upon the basis state (3.30) similar to (2.20):

$$\psi^{+1}(\mathbf{x}_1,\ldots\mathbf{x}_n;t) = {}^{+1}\langle\mathbf{x}_1,\ldots,\mathbf{x}_n|\psi(t)\rangle\,. \tag{3.52}$$

3.4 Field Operators in Heisenberg Picture

So far the non-relativistic many-body theory was formulated in the Schrödinger picture as the local particle creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ were time-independent, whereas the many-body state $|\psi(t)\rangle$ from the second quantized Hilbert space was time-dependent. Now we perform the transformation to the Heisenberg picture, where the many-body state is time-independent and the whole time dependence is carried by so-called field operators.

At first we repeat the general procedure in first quantization. To this end we start with the Schrödinger picture and restrict ourselves for the sake of simplicity to the case of a time-independent Hamilton operator $\hat{H}_{\rm S}$. The corresponding equations of motion for both the time-dependent state $|\psi_{\rm S}(t)\rangle$ and a time-independent operator $\hat{O}_{\rm S}$ read

$$i\hbar \frac{\partial}{\partial t} |\psi_{\rm S}(t)\rangle = \hat{H}_{\rm S} |\psi_{\rm S}(t)\rangle \,, \tag{3.53}$$

$$i\hbar \frac{\partial}{\partial t} \hat{O}_{\rm S} = 0 \,. \tag{3.54}$$

The formal solution of the Schrödinger equation (3.53) is given by

$$|\psi_{\rm S}(t)\rangle = e^{-i\hat{H}_{\rm S}t/\hbar} |\psi_{\rm S}(0)\rangle \,. \tag{3.55}$$

Here we identify the initial state $|\psi_{\rm S}(0)\rangle$ in the Schrödinger picture with the state $|\psi_{\rm H}\rangle$ in the Heisenberg picture:

$$|\psi_{\rm S}(0) = |\psi_{\rm H}\rangle \,. \tag{3.56}$$

Thus, the transformations from the Schrödinger to the Heisenberg picture and vice versa are defined according to the relations

$$|\psi_{\rm S}(t)\rangle = e^{-i\hat{H}_{\rm S}t/\hbar} |\psi_{\rm H}\rangle \qquad \Longleftrightarrow \qquad |\psi_{\rm H}\rangle = e^{i\hat{H}_{\rm S}t/\hbar} |\psi_{\rm S}(t)\rangle \,. \tag{3.57}$$

From (3.53) and (3.57) we then read off that the state in the Heisenberg picture $|\psi_{\rm H}\rangle$ is time-independent:

$$i\hbar \frac{\partial}{\partial t} |\psi_{\rm H}\rangle = -\hat{H}_{\rm S}\, e^{i\hat{H}_{\rm S}t/\hbar} |\psi_{\rm S}(t)\rangle + e^{i\hat{H}_{\rm S}t/\hbar}\, i\hbar \frac{\partial}{\partial t} |\psi_{\rm S}(t)\rangle = 0 \,. \tag{3.58}$$

In order to determine the operator $\hat{O}_{\rm H}(t)$ in the Heisenberg picture, we demand that the expectation values do not change once we perform a transformation from the Schrödinger to the Heisenberg picture:

$$\langle \psi_{\rm S}(t)|\hat{O}_{\rm S}|\psi_{\rm S}(t)\rangle = \langle \psi_{\rm H}|\hat{O}_{\rm H}(t)|\psi_{\rm H}\rangle \,. \tag{3.59}$$

Inserting (3.57) into (3.59) we determine, indeed, formally the time dependence of the operator $\hat{O}_{\rm H}(t)$ in the Heisenberg picture:

$$\begin{aligned} \langle e^{-i\hat{H}_{\rm S}t/\hbar}\, \psi_{\rm H}|\hat{O}_{\rm S}|e^{-i\hat{H}_{\rm S}t/\hbar}\, \psi_{\rm H}\rangle &= \langle \psi_{\rm H}|e^{i\hat{H}_{\rm S}t/\hbar}\, \hat{O}_{\rm S} e^{-i\hat{H}_{\rm S}t/\hbar}|\psi_{\rm H}\rangle = \langle \psi_{\rm H}|\hat{O}_{\rm H}(t)|\psi_{\rm H}\rangle \,. \\ &\Longrightarrow \quad \hat{O}_{\rm H}(t) = e^{i\hat{H}_{\rm S}t/\hbar}\, \hat{O}_{\rm S}\, e^{-i\hat{H}_{\rm S}t/\hbar} \,. \end{aligned} \tag{3.60}$$

Thus, multiplying an operator in the Schrödinger picture $\hat{O}_{\rm S}$ from the left with $e^{i\hat{H}_{\rm S}t/\hbar}$ and from the right with $e^{-i\hat{H}_{\rm S}t/\hbar}$ yields the corresponding operator in the Heisenberg picture $\hat{O}_{\rm H}(t)$. For instance, for the Hamilton operator $\hat{O}_{\rm S} = \hat{H}_{\rm S}$ we obtain from (3.60) the result that it does not change its form when we perform the transformation from the Schrödinger to the Heisenberg picture:

$$\hat{H}_{\rm H}(t) = e^{i\hat{H}_{\rm S}t/\hbar}\,\hat{H}_{\rm S}\,e^{-i\hat{H}_{\rm S}t/\hbar} = \hat{H}_{\rm S}\,. \tag{3.61}$$

Furthermore, for the operator in the Heisenberg picture $\hat{O}_{\rm H}(t)$ we determine from (3.54), (3.60), and (3.61) the Heisenberg equation of motion:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{O}_{\rm H}(t) &= e^{i\hat{H}_{\rm S}t/\hbar}\left\{-\hat{H}_{\rm S}\hat{O}_{\rm S}+\hat{O}_{\rm S}\hat{H}_{\rm S}\right\}e^{-i\hat{H}_{\rm S}t/\hbar} + e^{i\hat{H}_{\rm S}t/\hbar}\,i\hbar\frac{\partial}{\partial t}\hat{O}_{\rm S}\,e^{-i\hat{H}_{\rm S}t/\hbar} \\ &\Longrightarrow\; i\hbar\frac{\partial}{\partial t}\hat{O}_{\rm H}(t) = \left[\hat{O}_{\rm H}(t),\hat{H}_{\rm S}\right]_- = \left[\hat{O}_{\rm H}(t),\hat{H}_{\rm H}(t)\right]_-\,. \end{aligned} \tag{3.62}$$

Now we transfer this procedure to the second quantization. To this end we assign analogous to (3.60) to the local particle creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ in the Schrödinger picture corresponding time-dependent fields operators in the Heisenberg picture:

$$\hat{\psi}^\dagger(\mathbf{x},t) = \hat{a}^\dagger_{\mathbf{x}\rm H}(t) = e^{i\hat{H}t/\hbar}\,\hat{a}^\dagger_{\mathbf{x}}\,e^{-i\hat{H}t/\hbar}\,, \qquad \hat{\psi}(\mathbf{x},t) = \hat{a}_{\mathbf{x}\rm H}(t) = e^{i\hat{H}t/\hbar}\,\hat{a}_{\mathbf{x}}\,e^{-i\hat{H}t/\hbar}\,. \tag{3.63}$$

At first we determine from (3.25) and (3.63) the equal-time commutator relations of these field operators:

$$\left[\hat{\psi}(\mathbf{x},t),\hat{\psi}(\mathbf{x}',t)\right]_- = \left[\hat{\psi}^\dagger(\mathbf{x},t),\hat{\psi}^\dagger(\mathbf{x}',t)\right]_- = 0\,, \qquad \left[\hat{\psi}(\mathbf{x},t),\hat{\psi}^\dagger(\mathbf{x}',t)\right]_- = \delta(\mathbf{x}-\mathbf{x}')\,. \tag{3.64}$$

Thus, the field operators $\hat{\psi}^\dagger(\mathbf{x},t)$, $\hat{\psi}(\mathbf{x},t)$ in the Heisenberg picture fulfill at each time instant t the same commutator relations (3.25) as the local creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$, $\hat{a}_{\mathbf{x}}$ in the Schrödinger picture. This means that $\hat{\psi}^\dagger(\mathbf{x},t)$ and $\hat{\psi}(\mathbf{x},t)$ have the physical interpretation to create and annihilate a boson at space point $\mathbf{x}$ at time t.

Now we transform the Hamilton operator (3.34), (3.36), and (3.37) from the Schrödinger to the Heisenberg picture. Analogous to (3.60) we multiply the Hamilton operator

$$\hat{H} = \int d^3x\,\hat{a}^\dagger_{\mathbf{x}}\left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}\hat{a}_{\mathbf{x}} + \frac{1}{2}\int d^3x\int d^3x'\,\hat{a}^\dagger_{\mathbf{x}}\hat{a}^\dagger_{\mathbf{x}'}\,V_2(\mathbf{x}-\mathbf{x}')\hat{a}_{\mathbf{x}'}\,\hat{a}_{\mathbf{x}} \tag{3.65}$$

from the left with $e^{i\hat{H}t/\hbar}$ and from the right with $e^{-i\hat{H}t/\hbar}$:

$$\begin{aligned} \hat{H}_{\rm H}(t) &= \int d^3x\,e^{i\hat{H}t/\hbar}\,\hat{a}^\dagger_{\mathbf{x}}\,e^{-i\hat{H}t/\hbar}\left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}e^{i\hat{H}t/\hbar}\,\hat{a}_{\mathbf{x}}\,e^{-i\hat{H}t/\hbar} \\ &+\frac{1}{2}\int d^3x\int d^3x'\,e^{i\hat{H}t/\hbar}\,\hat{a}^\dagger_{\mathbf{x}}e^{-i\hat{H}t/\hbar}\,e^{i\hat{H}t/\hbar}\,\hat{a}^\dagger_{\mathbf{x}'}e^{-i\hat{H}t/\hbar}\,V_2(\mathbf{x}-\mathbf{x}')e^{i\hat{H}t/\hbar}\,\hat{a}_{\mathbf{x}'}\,e^{-i\hat{H}t/\hbar}\,e^{i\hat{H}t/\hbar}\,\hat{a}_{\mathbf{x}}e^{-i\hat{H}t/\hbar}\,. \end{aligned} \tag{3.66}$$

Using the field operators (3.63) the Hamilton operator reads in the Heisenberg picture:

$$\begin{aligned} \hat{H}_{\rm H}(t) &= \int d^3x\,\hat{\psi}^\dagger(\mathbf{x},t)\left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}\hat{\psi}(\mathbf{x},t) \\ &+\frac{1}{2}\int d^3x\int d^3x'\,\hat{\psi}^\dagger(\mathbf{x},t)\hat{\psi}^\dagger(\mathbf{x}',t)\,V_2(\mathbf{x}-\mathbf{x}')\hat{\psi}(\mathbf{x}',t)\,\hat{\psi}(\mathbf{x},t)\,. \end{aligned} \tag{3.67}$$

With this Hamilton operator in the Heisenberg picture we can determine from (3.62) the Heisenberg equation of motion of the field operator $\hat{\psi}(\mathbf{x},t)$:

$$i\hbar \frac{\partial \hat{\psi}(\mathbf{x},t)}{\partial t} = \left[\hat{\psi}(\mathbf{x},t), \hat{H}_{\mathrm{H}}(t)\right]_- . \tag{3.68}$$

At first we get

$$i\hbar \frac{\partial \hat{\psi}(\mathbf{x},t)}{\partial t} = \int d^3x' \int d^3x''\, \delta(\mathbf{x}-\mathbf{x}') \left\{-\frac{\hbar^2}{2M}\Delta'' + V_1(\mathbf{x}'')\right\} \left[\hat{\psi}(\mathbf{x},t), \hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}(\mathbf{x}'',t)\right]_-$$
$$+\frac{1}{2}\int d^3x' \int d^3x''\, V_2(\mathbf{x}'-\mathbf{x}'') \left[\hat{\psi}(\mathbf{x},t), \hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}^\dagger(\mathbf{x}'',t)\hat{\psi}(\mathbf{x}'',t)\hat{\psi}(\mathbf{x}',t)\right]_- . \tag{3.69}$$

Here the respective commutators can be evaluated with the help of the identity (3.43) and the commutator relations (3.64), yielding

$$\left[\hat{\psi}(\mathbf{x},t), \hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}(\mathbf{x}'',t)\right]_- = \delta(\mathbf{x}-\mathbf{x}')\,\hat{\psi}(\mathbf{x}'',t) \tag{3.70}$$

and, correspondingly,

$$\left[\hat{\psi}(\mathbf{x},t), \hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}^\dagger(\mathbf{x}'',t)\hat{\psi}(\mathbf{x}'',t)\hat{\psi}(\mathbf{x}',t)\right]_-$$
$$= \left\{\delta(\mathbf{x}-\mathbf{x}')\,\hat{\psi}^\dagger(\mathbf{x}'',t) + \delta(\mathbf{x}-\mathbf{x}'')\,\hat{\psi}^\dagger(\mathbf{x}',t)\right\}\hat{\psi}(\mathbf{x}'',t)\hat{\psi}(\mathbf{x}',t) . \tag{3.71}$$

Inserting (3.70) and (3.71) in (3.69) we finally obtain

$$i\hbar \frac{\partial \hat{\psi}(\mathbf{x},t)}{\partial t} = \left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}\hat{\psi}(\mathbf{x},t) + \int d^3x'\, V_2(\mathbf{x}-\mathbf{x}')\,\hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}(\mathbf{x}',t)\hat{\psi}(\mathbf{x},t) . \tag{3.72}$$

In the same way also the Heisenberg equation of motion of the adjoint field operator

$$i\hbar \frac{\partial \hat{\psi}^\dagger(\mathbf{x},t)}{\partial t} = \left[\hat{\psi}^\dagger(\mathbf{x},t), \hat{H}_{\mathrm{H}}(t)\right]_- \tag{3.73}$$

is evaluated:

$$-i\hbar \frac{\partial \hat{\psi}^\dagger(\mathbf{x},t)}{\partial t} = \left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}\hat{\psi}^\dagger(\mathbf{x},t) + \hat{\psi}^\dagger(\mathbf{x},t)\int d^3x'\, V_2(\mathbf{x}-\mathbf{x}')\,\hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}(\mathbf{x}',t) . \tag{3.74}$$

This is, indeed, the adjoint of the Heisenberg equation of motion (3.72). The operator-valued integro-differential equations (3.72) and (3.74) are nonlinear. Due to their complexity it is not possible to obtain exact analytic solutions. Therefore, one has to reside to develop physically reasonable approximate solutions.

3.5 Creation and Annihilation Operators for Fermions

So far we have shown that the symmetric many-body states for bosons can be practically realized with the help of local creation and annihilation operators $\hat{a}_{\mathbf{x}}^\dagger$ and $\hat{a}_{\mathbf{x}}$ in the Schrödinger picture.

Here the symmetry of the many-body states of bosons was ultimately a direct consequence of the commutation relations (3.25). Therefore, the question arises whether there exists a similar formalism also in view of the anti-symmetric many-body states for fermions. To this end we aim for creating an anti-symmetric many-body state for fermions similar to (3.30) via

$$|\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle^{-1} = \hat{a}^\dagger_{\mathbf{x}_1} \cdots \hat{a}^\dagger_{\mathbf{x}_n} |0\rangle \,. \tag{3.75}$$

But then we have to demand instead of the commutation relations (3.25) corresponding anti-commutation relations

$$\left[\hat{a}_{\mathbf{x}}, \hat{a}_{\mathbf{x}'}\right]_+ = \left[\hat{a}^\dagger_{\mathbf{x}}, \hat{a}^\dagger_{\mathbf{x}'}\right]_+ = 0 \,, \qquad \left[\hat{a}_{\mathbf{x}}, \hat{a}^\dagger_{\mathbf{x}'}\right]_+ = \delta(\mathbf{x} - \mathbf{x}') \,. \tag{3.76}$$

where the anti-commutator between two quantum mechanical operators $\hat{A}$ and $\hat{B}$ is defined by

$$\left[\hat{A}, \hat{B}\right]_+ = \hat{A}\hat{B} + \hat{B}\hat{A} \,. \tag{3.77}$$

As in the bosonic case in (3.29) we define in addition the vacuum state $|0\rangle$ by the condition that it does not contain any particles:

$$\hat{a}_{\mathbf{x}}|0\rangle = 0 \qquad \Longleftrightarrow \qquad \langle 0|\hat{a}^\dagger_{\mathbf{x}} = 0 \,. \tag{3.78}$$

Indeed, (3.77) and (3.78) turn out to guarantee for the anti-symmetric many-body states (3.75) the orthonormality relations (2.74) for $n = 1$ and $n = 2$ fermions, which are characterized by $\epsilon = -1$. From (2.72), (3.75), (3.76), and (3.78) we obtain at first

$$\begin{aligned} {}^{-1}\langle \mathbf{x}_1|\mathbf{x}'_1\rangle^{-1} &= \langle \hat{a}^\dagger_{\mathbf{x}_1} 0|\hat{a}^\dagger_{\mathbf{x}'_1} 0\rangle = \langle 0|\hat{a}_{\mathbf{x}_1} \hat{a}^\dagger_{\mathbf{x}'_1}|0\rangle \\ &= \langle 0| - \hat{a}^\dagger_{\mathbf{x}'_1} \hat{a}_{\mathbf{x}_1} + \delta(\mathbf{x}_1 - \mathbf{x}'_1)|0\rangle = \delta(\mathbf{x}_1 - \mathbf{x}'_1) = \delta^{-1}(\mathbf{x}_1; \mathbf{x}'_1) \,. \end{aligned} \tag{3.79}$$

Correspondingly, we get then

$$\begin{aligned} &{}^{-1}\langle \mathbf{x}_1, \mathbf{x}_2|\mathbf{x}'_1, \mathbf{x}'_2\rangle^{-1} = \langle \hat{a}^\dagger_{\mathbf{x}_1} \hat{a}^\dagger_{\mathbf{x}_2} 0|\hat{a}^\dagger_{\mathbf{x}'_1} \hat{a}^\dagger_{\mathbf{x}'_2} 0\rangle = \langle 0|\hat{a}_{\mathbf{x}_2} \hat{a}_{\mathbf{x}_1} \hat{a}^\dagger_{\mathbf{x}'_1} \hat{a}^\dagger_{\mathbf{x}'_2}|0\rangle \\ &= \langle 0|\hat{a}_{\mathbf{x}_2}\big(- \hat{a}^\dagger_{\mathbf{x}'_1} \hat{a}_{\mathbf{x}_1} + \delta(\mathbf{x}_1 - \mathbf{x}'_1)\big)\hat{a}^\dagger_{\mathbf{x}'_2}|0\rangle = \delta(\mathbf{x}_1 - \mathbf{x}'_1)\langle 0| - \hat{a}^\dagger_{\mathbf{x}'_2} \hat{a}_{\mathbf{x}_2} + \delta(\mathbf{x}_2 - \mathbf{x}'_2)|0\rangle \\ &-\langle 0|\big(- \hat{a}^\dagger_{\mathbf{x}'_1} \hat{a}_{\mathbf{x}_2} + \delta(\mathbf{x}'_1 - \mathbf{x}_2)\big)\big(- \hat{a}^\dagger_{\mathbf{x}'_2} \hat{a}_{\mathbf{x}_1} + \delta(\mathbf{x}_1 - \mathbf{x}'_2)\big)|0\rangle = \delta(\mathbf{x}_1 - \mathbf{x}'_1)\delta(\mathbf{x}_2 - \mathbf{x}'_2) \\ &-\delta(\mathbf{x}_1 - \mathbf{x}'_2)\delta(\mathbf{x}_2 - \mathbf{x}'_1) = \delta^{-1}(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) \,. \end{aligned} \tag{3.80}$$

As two local creation operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}^\dagger_{\mathbf{x}'}$ anti-commute due to (3.76), we conclude that then the square of the fermionic creation operator $\hat{a}^\dagger_{\mathbf{x}}$ vanishes:

$$\left(\hat{a}^\dagger_{\mathbf{x}}\right)^2 = 0 \,. \tag{3.81}$$

For the anti-symmetric many-body state (3.75) this has the consequence that it vanishes provided that two space coordinates $\mathbf{x}_i$ and $\mathbf{x}_j$ for $i \neq j$ coincide:

$$|\mathbf{x}_1, \ldots, \mathbf{x}_n\rangle^{-1} = 0 \,, \qquad \text{if } \mathbf{x}_i = \mathbf{x}_j \text{ for } i \neq j \,. \tag{3.82}$$

Thus, the anti-commutation relations (3.76) contain automatically the Pauli exclusion principle that two fermions can not be at the same space point.

The properties (3.75) and (3.76) are also sufficient in order to formulate with the help of the second quantized Hamilton operator

$$\hat{H} = \int d^3x\, \hat{a}^\dagger_{\mathbf{x}} \left\{ -\frac{\hbar^2}{2M}\,\Delta + V_1(\mathbf{x}) \right\} \hat{a}_{\mathbf{x}} + \frac{1}{2} \int d^3x \int d^3x'\, \hat{a}^\dagger_{\mathbf{x}} \hat{a}^\dagger_{\mathbf{x}'}\, V_2(\mathbf{x}-\mathbf{x}') \hat{a}_{\mathbf{x}'}\, \hat{a}_{\mathbf{x}} \tag{3.83}$$

the second quantized Schrödinger equation for a fermionic many-body state $|\psi(t)\rangle$:

$$i\hbar\, \frac{\partial}{\partial t}\, |\psi(t)\rangle = \hat{H} |\psi(t)\rangle\,. \tag{3.84}$$

Projecting (3.84) to the anti-symmetric basis states (3.75) yields, like in the bosonic case, the corresponding n-body Schrödinger equation (2.21) with (2.23) for the n-particle wave function

$$\psi^{-1}(\mathbf{x}_1, \ldots \mathbf{x}_n; t) = {}^{-1}\langle \mathbf{x}_1, \ldots, \mathbf{x}_n | \psi(t)\rangle\,. \tag{3.85}$$

We leave the detailed proof to the reader, which follows a consideration similar to Section 3.3.

Furthermore, transforming the fermionic creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ from the Schrödinger to the Heisenberg picture yields fermionic field operators

$$\hat{\psi}^\dagger(\mathbf{x}, t) = e^{i\hat{H}t/\hbar}\, \hat{a}^\dagger_{\mathbf{x}}\, e^{-i\hat{H}t/\hbar}\,, \qquad \hat{\psi}(\mathbf{x}, t) = e^{i\hat{H}t/\hbar}\, \hat{a}_{\mathbf{x}}\, e^{-i\hat{H}t/\hbar}\,, \tag{3.86}$$

which fulfill due to (3.76) equal-time anti-commutation relations:

$$\left[\hat{\psi}(\mathbf{x}, t), \hat{\psi}(\mathbf{x}', t)\right]_+ = \left[\hat{\psi}^\dagger(\mathbf{x}, t), \hat{\psi}^\dagger(\mathbf{x}', t)\right]_+ = 0\,, \quad \left[\hat{\psi}(\mathbf{x}, t), \hat{\psi}^\dagger(\mathbf{x}', t)\right]_+ = \delta(\mathbf{x}-\mathbf{x}')\,. \tag{3.87}$$

Furthermore, we remark that the Hamilton operator in the Heisenberg picture

$$\begin{aligned} \hat{H}_{\mathrm{H}}(t) = & \int d^3x\, e^{i\hat{H}t/\hbar}\, \hat{a}^\dagger_{\mathbf{x}}\, e^{-i\hat{H}t/\hbar} \left\{ -\frac{\hbar^2}{2M}\,\Delta + V_1(\mathbf{x}) \right\} e^{i\hat{H}t/\hbar}\, \hat{a}_{\mathbf{x}}\, e^{-i\hat{H}t/\hbar} \\ & + \frac{1}{2} \int d^3x \int d^3x'\, e^{i\hat{H}t/\hbar}\, \hat{a}^\dagger_{\mathbf{x}} e^{-i\hat{H}t/\hbar}\, e^{i\hat{H}t/\hbar}\, \hat{a}^\dagger_{\mathbf{x}'} e^{-i\hat{H}t/\hbar}\; V_2(\mathbf{x}-\mathbf{x}') e^{i\hat{H}t/\hbar}\, \hat{a}_{\mathbf{x}'}\, e^{-i\hat{H}t/\hbar}\, e^{i\hat{H}t/\hbar}\, \hat{a}_{\mathbf{x}} e^{-i\hat{H}t/\hbar}\,. \end{aligned} \tag{3.88}$$

turns out to have the same form as in the bosonic case, see (3.67):

$$\begin{aligned} \hat{H}_{\mathrm{H}}(t) \;=\; & \int d^3x\, \hat{\psi}^\dagger(\mathbf{x}, t) \left\{ -\frac{\hbar^2}{2M}\,\Delta + V_1(\mathbf{x}) \right\} \hat{\psi}(\mathbf{x}, t) \\ & + \frac{1}{2} \int d^3x \int d^3x'\, \hat{\psi}^\dagger(\mathbf{x}, t) \hat{\psi}^\dagger(\mathbf{x}', t)\, V_2(\mathbf{x}-\mathbf{x}') \hat{\psi}(\mathbf{x}', t)\, \hat{\psi}(\mathbf{x}, t)\,. \end{aligned} \tag{3.89}$$

With this the Heisenberg equations of motion of the field operators $\hat{\psi}(\mathbf{x}, t)$ and $\hat{\psi}^\dagger(\mathbf{x}, t)$

$$i\hbar\, \frac{\partial \hat{\psi}(\mathbf{x}, t)}{\partial t} = \left[\hat{\psi}(\mathbf{x}, t), \hat{H}_{\mathrm{H}}(t)\right]_-\,, \tag{3.90}$$

$$i\hbar\, \frac{\partial \hat{\psi}^\dagger(\mathbf{x}, t)}{\partial t} = \left[\hat{\psi}^\dagger(\mathbf{x}, t), \hat{H}_{\mathrm{H}}(t)\right]_- \tag{3.91}$$

are evaluated and yield

$$i\hbar \frac{\partial\hat{\psi}(\mathbf{x},t)}{\partial t} = \left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}\hat{\psi}(\mathbf{x},t) + \int d^3x'\, V_2(\mathbf{x}-\mathbf{x}')\,\hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}(\mathbf{x}',t)\hat{\psi}(\mathbf{x},t) \tag{3.92}$$

as well as its adjoint

$$-i\hbar \frac{\partial\hat{\psi}^\dagger(\mathbf{x},t)}{\partial t} = \left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}\hat{\psi}^\dagger(\mathbf{x},t) + \hat{\psi}^\dagger(\mathbf{x},t)\int d^3x'\, V_2(\mathbf{x}-\mathbf{x}')\,\hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}(\mathbf{x}',t) \tag{3.93}$$

corresponding to the bosonic case, see (3.72) and (3.74). Note that obtaining (3.92) and (3.93) necessitates the operator identity (3.43) and the complementary one

$$\left[\hat{A},\hat{B}\hat{C}\right]_- = \left[\hat{A},\hat{B}\right]_+\hat{C} - \hat{B}\left[\hat{A},\hat{C}\right]_+ , \tag{3.94}$$

which directly follows from the definitions of both the commutator (2.13) and the anti-commutator (3.77).

3.6 Occupation Number Representation

Let us finally consider the case that the 2-particle interaction vanishes, i.e. $V_2(\mathbf{x}-\mathbf{x}')=0$, from the point of view of second quantization. We show in this section that then identical particles are described within the so-called occupation number representation. To this end we start with the second quantized Hamilton operator in the Schrödinger picture for non-interacting identical particles

$$\hat{H} = \int d^3x\, \hat{a}^\dagger_{\mathbf{x}} \left\{-\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x})\right\}\hat{a}_{\mathbf{x}} . \tag{3.95}$$

As we deal at the same time with bosons and fermions, the creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$, $\hat{a}_{\mathbf{x}}$ fulfill either canonical commutation or canonical anti-commutation relations:

$$\left[\hat{a}_{\mathbf{x}},\hat{a}_{\mathbf{x}'}\right]_\mp = \left[\hat{a}^\dagger_{\mathbf{x}},\hat{a}^\dagger_{\mathbf{x}'}\right]_\mp = 0 , \qquad \left[\hat{a}_{\mathbf{x}},\hat{a}^\dagger_{\mathbf{x}'}\right]_\mp = \delta(\mathbf{x}-\mathbf{x}') . \tag{3.96}$$

In the following we assume again that the 1-particle wavefunctions $\psi_{E_\alpha}(\mathbf{x})$ with the quantum numbers α are known as solutions of the time-independent 1-particle Schrödinger equation (2.48), obeying both the orthonormality relation (2.49) and the completeness relation (2.50). Due to the latter the creation and annihilation operators $\hat{a}^\dagger_{\mathbf{x}}$, $\hat{a}_{\mathbf{x}}$ can be expanded in the 1-particle basis:

$$\hat{a}_{\mathbf{x}} = \sum_\alpha \psi_{E_\alpha}(\mathbf{x})\,\hat{a}_\alpha \qquad \Longleftrightarrow \qquad \hat{a}^\dagger_{\mathbf{x}} = \sum_\alpha \psi^*_{E_\alpha}(\mathbf{x})\,\hat{a}^\dagger_\alpha . \tag{3.97}$$

Both expansions are inverted with the help of the orthonormality relation (2.49), yielding

$$\hat{a}_\alpha = \int d^3x\, \psi^*_{E_\alpha}(\mathbf{x})\,\hat{a}_{\mathbf{x}} \qquad \Longleftrightarrow \qquad \hat{a}^\dagger_\alpha = \int d^3x\, \psi_{E_\alpha}(\mathbf{x})\,\hat{a}^\dagger_{\mathbf{x}} . \tag{3.98}$$

With this we deduce the commutation and anti-commutation relations for the operator-valued expansion coefficients $\hat{a}^\dagger_\alpha$, $\hat{a}_\alpha$ by taking into account (3.96):

$$\left[\hat{a}_\alpha, \hat{a}_{\alpha'}\right]_\mp = \left[\hat{a}^\dagger_\alpha, \hat{a}^\dagger_{\alpha'}\right]_\mp = 0\,, \qquad \left[\hat{a}_\alpha, \hat{a}^\dagger_{\alpha'}\right]_\mp = \delta_{\alpha,\alpha'}\,. \tag{3.99}$$

Inserting the expansions of the creation and annihilation operators (3.97) in the second quantized Hamilton operator (3.95), we can express it via the operator-valued expansion coefficients $\hat{a}^\dagger_\alpha$, $\hat{a}_\alpha$ due to (2.48) and (2.49) and end up with

$$\hat{H} = \sum_\alpha E_\alpha\, \hat{n}_\alpha\,, \tag{3.100}$$

where we have introduced the particle number operator

$$\hat{n}_\alpha = \hat{a}^\dagger_\alpha \hat{a}_\alpha\,. \tag{3.101}$$

Note that the useful operator identity

$$\left[\hat{A}\hat{B}, \hat{C}\right]_- = \hat{A}\left[\hat{B}, \hat{C}\right]_\mp \pm \left[\hat{A}, \hat{C}\right]_\mp \hat{B}\,, \tag{3.102}$$

which follows from the definitions of both the commutator (2.13) and the anti-commutator (3.77), complements the bosonic version (3.10) with a corresponding fermionic one. With (3.43) and (3.102) we can then show that the particle operators $\hat{n}_\alpha$ and $\hat{n}_{\alpha'}$ for two quantum numbers α and α' commute:

$$\begin{aligned}\left[\hat{n}_\alpha, \hat{n}_{\alpha'}\right]_- &= \left[\hat{n}_\alpha, \hat{a}^\dagger_{\alpha'}\hat{a}_{\alpha'}\right]_- = \left[\hat{a}^\dagger_\alpha\hat{a}_\alpha, \hat{a}^\dagger_{\alpha'}\right]_- \hat{a}_{\alpha'} + \hat{a}^\dagger_{\alpha'}\left[\hat{a}^\dagger_\alpha\hat{a}_\alpha, \hat{a}_{\alpha'}\right]_- \\ &= \left(\hat{a}^\dagger_\alpha\left[\hat{a}_\alpha, \hat{a}^\dagger_{\alpha'}\right]_\mp \pm \left[\hat{a}^\dagger_\alpha, \hat{a}^\dagger_{\alpha'}\right]_\mp \hat{a}_\alpha\right)\hat{a}_{\alpha'} + \hat{a}^\dagger_{\alpha'}\left(\hat{a}^\dagger_\alpha\left[\hat{a}_\alpha, \hat{a}_{\alpha'}\right]_\mp \pm \left[\hat{a}^\dagger_\alpha, \hat{a}_{\alpha'}\right]_\mp \hat{a}_\alpha\right) = 0\,.\end{aligned} \tag{3.103}$$

Thus, we conclude that the particle number operator (3.101) commutes with the Hamilton operator (3.100):

$$\left[\hat{n}_\alpha, \hat{H}\right]_- = \sum_{\alpha'} E_{\alpha'}\left[\hat{n}_\alpha, \hat{n}_{\alpha'}\right]_- = 0\,. \tag{3.104}$$

Due to (3.103) and (3.104) we know that there must exist a set of states, which are eigenstates for both all particle number operators (3.101)and the Hamilton operator (3.100):

$$\hat{n}_\alpha|\ldots, n_\alpha, \ldots\rangle = n_\alpha\,|\ldots, n_\alpha, \ldots\rangle\,, \tag{3.105}$$

$$\hat{H}|\ldots, n_\alpha, \ldots\rangle = \sum_\alpha E_\alpha n_\alpha\,|\ldots, n_\alpha, \ldots\rangle\,. \tag{3.106}$$

In the case of bosons we already know from Section 3.2 that the commutation relations for the operators $\hat{a}^\dagger_\alpha$, $\hat{a}_\alpha$ imply that the eigenvalues of the particle operator $\hat{n}_\alpha$ can have any integer value including zero:

$$\text{bosons:} \qquad n_\alpha = 0, 1, 2, \ldots\,. \tag{3.107}$$

But for fermions it turns out that the anti-commutation relations for the operators $\hat{a}_\alpha^\dagger$, $\hat{a}_\alpha$ lead to an essential restriction for the eigenvalues of the particle operators. Namely we read off from (3.81) and (3.101):

$$(\hat{n}_\alpha)^2 = \hat{a}_\alpha^\dagger \hat{a}_\alpha \hat{a}_\alpha^\dagger \hat{a}_\alpha = \hat{a}_\alpha^\dagger \hat{a}_\alpha - \left(\hat{a}_\alpha^\dagger\right)^2 (\hat{a}_\alpha)^2 = \hat{n}_\alpha \,. \tag{3.108}$$

Applying (3.108) to the eigenstates $|\ldots, n_\alpha, \ldots\rangle$ we conclude due to the eigenvalue problem (3.105):

$$n_\alpha^2 = n_\alpha \,, \tag{3.109}$$

which yields straightforwardly

$$\text{fermions:} \qquad n_\alpha = 0, 1 \,. \tag{3.110}$$

Thus, each state characterized by the quantum number α can be occupied with at most one fermion in accordance with the Pauli exclusion principle.

Chapter 4

Canonical Field Quantization for Bosons

The equal-time commutation relations (3.64) of the field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$ have so far been introduced heuristically in order to describe a non-relativistic quantum many-body problem. In the following we show that these equal-time commutation relations (3.64) can be systematically derived from first principles within the canonical field quantization formalism for bosons. To this end we have to generalize the recipe how to quantize a system with a finite number of degrees of freedom to a continuum of degrees of freedom. But prior to that it is essential to work out the field theory of non-relativistic quantum mechanics.

4.1 Action of Schrödinger Field

We start with considering the complex Schrödinger field $\psi(\mathbf{x},t)$ and its adjoint $\psi^*(\mathbf{x},t)$ as two independent fields with their respective equations of motion:

$$i\hbar \frac{\partial \psi(\mathbf{x},t)}{\partial t} = \left\{ -\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x}) \right\} \psi(\mathbf{x},t)\,, \tag{4.1}$$

$$-i\hbar \frac{\partial \psi^*(\mathbf{x},t)}{\partial t} = \left\{ -\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x}) \right\} \psi^*(\mathbf{x},t)\,. \tag{4.2}$$

Now we derive a variational principle with an underlying action so that these equations of motion emerge from applying the corresponding Hamilton principle. To this end we multiply (4.1) and (4.2) with the variations $\delta\psi^*(\mathbf{x},t)$ and $\delta\psi(\mathbf{x},t)$, respectively, add both equations together, yielding the spatio-temporal integral

$$\int dt \int d^3x \left\{ i\hbar \left[\delta\psi^*(\mathbf{x},t)\,\frac{\partial \psi(\mathbf{x},t)}{\partial t} - \delta\psi(\mathbf{x},t)\,\frac{\partial \psi^*(\mathbf{x},t)}{\partial t} \right] \right. \tag{4.3}$$
$$\left. + \frac{\hbar^2}{2M}\left[\delta\psi^*(\mathbf{x},t)\Delta\,\delta\psi(\mathbf{x},t) + \delta\psi(\mathbf{x},t)\Delta\,\delta\psi^*(\mathbf{x},t) \right] - V_1(\mathbf{x})\,\delta\Big[\psi^*(\mathbf{x},t)\psi(\mathbf{x},t) \Big] \right\} = 0\,.$$

Note we have used in the last term the product rule for field variations:

$$\delta\Big[\psi^*(\mathbf{x},t)\psi(\mathbf{x},t)\Big] = \delta\psi^*(\mathbf{x},t)\,\psi(\mathbf{x},t) + \delta\psi(\mathbf{x},t)\,\psi^*(\mathbf{x},t)\,. \tag{4.4}$$

Both terms in (4.3) with the temporal and the spatial derivatives are now partially integrated appropriately. Here we implicitly assume that the variations of the fields $\delta\psi(\mathbf{x},t)$ and $\delta\psi^*(\mathbf{x},t)$ vanish at the respective integration boundaries. Furthermore, we apply the calculational rule that a variation and a partial derivative are independent from each other, so they can be interchanged. With this one partial integration in time leads to

$$\int dt\left[\delta\psi^*(\mathbf{x},t)\frac{\partial\psi(\mathbf{x},t)}{\partial t} - \delta\psi(\mathbf{x},t)\frac{\partial\psi^*(\mathbf{x},t)}{\partial t}\right] = \delta\int dt\,\psi^*(\mathbf{x},t)\frac{\partial\psi(\mathbf{x},t)}{\partial t} \tag{4.5}$$

and two partial integrations in space give, correspondingly:

$$\int d^3x\Big[\delta\psi^*(\mathbf{x},t)\boldsymbol{\nabla}^2\,\delta\psi(\mathbf{x},t) + \delta\psi(\mathbf{x},t)\boldsymbol{\nabla}^2\,\delta\psi^*(\mathbf{x},t)\Big] = -\delta\int d^3x\,\boldsymbol{\nabla}\psi^*(\mathbf{x},t)\,\cdot\,\boldsymbol{\nabla}\psi(\mathbf{x},t)\,. \tag{4.6}$$

Inserting (4.5) and (4.6) into (4.3) we obtain a variational principle of the form

$$\delta\mathcal{A}\left[\psi^*(\bullet,\bullet);\psi(\bullet,\bullet)\right] = 0\,. \tag{4.7}$$

Note that we use here a bullet $\bullet$ in order to emphasize that the action is a functional of both Schrödinger fields $\psi^*(\mathbf{x},t)$ and $\psi(\mathbf{x},t)$. The action $\mathcal{A}$ is defined as a temporal integral over a Lagrange function L according to

$$\mathcal{A} = \int dt\,L\left[\psi^*(\bullet,t),\frac{\partial\psi^*(\bullet,t)}{\partial t};\psi(\bullet,t),\frac{\partial\psi(\bullet,t)}{\partial t}\right] \tag{4.8}$$

and the Lagrange function L represents a spatial integral over the Lagrange density

$$L = \int d^3x\,\mathcal{L}\left(\psi^*(\mathbf{x},t),\boldsymbol{\nabla}\psi^*(\mathbf{x},t),\frac{\partial\psi^*(\mathbf{x},t)}{\partial t};\psi(\mathbf{x},t),\boldsymbol{\nabla}\psi(\mathbf{x},t),\frac{\partial\psi(\mathbf{x},t)}{\partial t}\right)\,. \tag{4.9}$$

In case of the Schrödinger field the Lagrange density reads

$$\mathcal{L} = i\hbar\,\psi^*(\mathbf{x},t)\,\frac{\partial\psi(\mathbf{x},t)}{\partial t} - \frac{\hbar^2}{2M}\,\boldsymbol{\nabla}\psi^*(\mathbf{x},t)\,\cdot\,\boldsymbol{\nabla}\psi(\mathbf{x},t) - V_1(\mathbf{x})\psi^*(\mathbf{x},t)\psi(\mathbf{x},t)\,. \tag{4.10}$$

Conversely, it is also possible to rederive the original equations of motion (4.1) and (4.2) from a variational principle, which is based on the action (4.8)–(4.10). But this necessitates to introduce before the technique of functional derivatives, which we now introduce concisely without mathematical rigour.

4.2 Functional Derivative: Definition

At first we consider a function f of a finite number of degrees of freedom:

$$f = f(q_1,\ldots,q_N)\,. \tag{4.11}$$

The partial derivative of f with respect to the variable q_j, i.e.

$$\frac{\partial f(q_1, \ldots, q_N)}{\partial q_j}, \tag{4.12}$$

then denotes the change of the function with respect to the variable q_j, where all other variables $q_1, \ldots, q_{j-1}, q_{j+1}, \ldots, q_N$ remain constant. The total change of the function f

$$df(q_1, \ldots, q_N) = \sum_{j=1}^{N} \frac{\partial f(q_1, \ldots, q_N)}{\partial q_j} \, dq_j \tag{4.13}$$

is then additive in all possible changes of the function, where always only one variable changes and all the other variables remain constant. Specializing (4.13) to an infinitesimal change in one variable, i.e. $dq_j = \epsilon \delta_{ij}$, yields

$$f(q_1, \ldots, q_i + \epsilon, \ldots, q_N) - f(q_1, \ldots, q_i, \ldots, q_n) = df(q_1, \ldots, q_N) = \epsilon \, \frac{\partial f(q_1, \ldots, q_N)}{\partial q_i} . \tag{4.14}$$

Thus, the partial derivative follows from the limit of a difference quotient:

$$\frac{\partial f(q_1, \ldots, q_N)}{\partial q_i} = \lim_{\epsilon \to 0} \frac{f(q_1, \ldots, q_i + \epsilon, \ldots, q_N) - f(q_1, \ldots, q_i, \ldots, q_n)}{\epsilon} . \tag{4.15}$$

Now we generalize this concept of differentiation from a finite number to a continuum of variables. Therefore, we regard now a functional

$$F = F\big[\phi(\bullet)\big] , \tag{4.16}$$

i.e. a mapping of a field $\phi(x)$ to a real or a complex number. The functional derivative

$$\frac{\delta F\big[\phi(\bullet)\big]}{\delta \phi(x)} \tag{4.17}$$

should then describe how the functional F changes provided that the function $\phi(x)$ is only changed at a single point x. Thus, the functional derivative (4.17) becomes in this way an ordinary function, which depends on the variable x. In analogy to (4.13) the total change of the functional F is defined via

$$\delta F\big[\phi(\bullet)\big] = \int dx \, \frac{\delta F\big[\phi(\bullet)\big]}{\delta \phi(x)} \, \delta \phi(x) , \tag{4.18}$$

so it is additive with respect to all local changes of the function $\phi(x)$ at all space points x. Similar to the case of a partial derivative also the functional derivative can be determined from the limit of a difference quotient. To this end we introduce a local perturbation of the field $\phi(x)$ at space point y with strength ϵ:

$$\delta \phi(x) = \epsilon \delta(x - y) . \tag{4.19}$$

and determine from (4.18)

$$F[\phi(\bullet) + \epsilon \delta(\bullet - y)] - F[\phi(\bullet)] = \delta F[\phi(\bullet)] = \int dx \, \frac{\delta F[\phi(\bullet)]}{\delta \phi(x)} \, \delta \phi(x) = \epsilon \, \frac{\delta F[\phi(\bullet)]}{\delta \phi(y)} . \tag{4.20}$$

In the limit $\epsilon \to 0$ we obtain

$$\frac{\delta F[\phi(\bullet)]}{\delta \phi(y)} = \lim_{\epsilon\to 0} \frac{F[\phi(\bullet) + \epsilon\delta(\bullet - y)] - F[\phi(\bullet)]}{\epsilon} . \tag{4.21}$$

From this definition of the functional derivative as a limit of a difference quotient follow several useful calculation rules. At first, we obtain from (4.21) the trivial functional derivative

$$\frac{\delta\phi(x)}{\delta\phi(y)} = \lim_{\epsilon\to 0} \frac{\phi(x) + \epsilon\delta(x-y) - \phi(x)}{\epsilon} = \delta(x-y) . \tag{4.22}$$

Then we determine from (4.21) the product rule

$$\begin{aligned}
\frac{\delta\{F[\phi(\bullet)]G[\phi(\bullet)]\}}{\delta\phi(y)} &= \lim_{\epsilon\to 0} \frac{F[\phi(\bullet)+\epsilon\delta(\bullet-y)]G[\phi(\bullet)+\epsilon\delta(\bullet-y)] - F[\phi(\bullet)]G[\phi(\bullet)]}{\epsilon} \\
&= \lim_{\epsilon\to 0} \left\{ \frac{F[\phi(\bullet)+\epsilon\delta(\bullet-y)] - F[\phi(\bullet)]}{\epsilon} G[\phi(\bullet)] + F[\phi(\bullet)] \frac{G[\phi(\bullet)+\epsilon\delta(\bullet-y)] - G[\phi(\bullet)]}{\epsilon} \right\} \\
&= \frac{\delta F[\phi(\bullet)]}{\delta\phi(y)} G[\phi(\bullet)] + F[\phi(\bullet)] \frac{\delta G[\phi(\bullet)]}{\delta\phi(y)} .
\end{aligned} \tag{4.23}$$

And, finally, combining (4.21) and (4.22) yields the chain rule:

$$\frac{\delta f(\phi(x))}{\delta\phi(y)} = \lim_{\epsilon\to 0} \frac{f\left(\phi(x)+\epsilon\delta(x-y)\right) - f(\phi(x))}{\epsilon} = \frac{\partial f(\phi(x))}{\partial\phi(x)} \delta(x-y) = \frac{\partial f(\phi(x))}{\partial\phi(x)} \frac{\delta\phi(x)}{\delta\phi(y)} . \tag{4.24}$$

4.3 Functional Derivative: Application

Now we work out several non-trivial applications of the functional derivative in the realm of second quantization, where it turns out to be a useful tool in order to determine commutators between second quantized operators. We start with the observation that the commutator (3.46) can also be determined from a functional derivative via

$$\left[\hat a_{\mathbf{x}_n} \cdots \hat a_{\mathbf{x}_1}, \hat a^\dagger_{\mathbf{x}}\right]_- = \sum_{\nu=1}^{n} \delta(\mathbf{x}-\mathbf{x}_\nu)\, \hat a_{\mathbf{x}_n} \cdots \hat a_{\mathbf{x}_{\nu+1}} \hat a_{\mathbf{x}_{\nu-1}} \cdots \hat a_{\mathbf{x}_1} = \frac{\delta}{\delta \hat a_{\mathbf{x}}} \hat a_{\mathbf{x}_n} \cdots \hat a_{\mathbf{x}_1} . \tag{4.25}$$

Let us consider then an arbitrary functional $F[\hat a_\bullet]$ of the annihilation operator $\hat a_{\mathbf{x}}$:

$$F[\hat a_\bullet] = \sum_{n=1}^{\infty} \int d^3x_1 \cdots \int d^3x_n\, F_n(\mathbf{x}_1, \ldots \mathbf{x}_n)\, \hat a_{\mathbf{x}_n} \cdots \hat a_{\mathbf{x}_1} . \tag{4.26}$$

Then the functional derivative of this functional (4.26) with respect to the annihilation operator $\hat a_{\mathbf{x}}$ can be efficiently determined via a functional derivative due to (4.25):

$$\begin{aligned}
\left[F[\hat a_\bullet], \hat a^\dagger_{\mathbf{x}}\right]_- &= \sum_{n=1}^{\infty} \int d^3x_1 \cdots \int d^3x_n\, F_n(\mathbf{x}_1, \ldots \mathbf{x}_n)\, \left[\hat a_{\mathbf{x}_n} \cdots \hat a_{\mathbf{x}_1}, \hat a^\dagger_{\mathbf{x}}\right]_- \\
&= \sum_{n=1}^{\infty} \int d^3x_1 \cdots \int d^3x_n\, F_n(\mathbf{x}_1, \ldots \mathbf{x}_n)\, \frac{\delta}{\delta\hat a_{\mathbf{x}}} \hat a_{\mathbf{x}_n} \cdots \hat a_{\mathbf{x}_1} = \frac{\delta}{\delta \hat a_{\mathbf{x}}} F[\hat a_\bullet] .
\end{aligned} \tag{4.27}$$

In a similar manner one also proves

$$\left[\hat{a}_{\mathbf{x}}, F[\hat{a}^{\dagger}_{\bullet}]\right]_{-} = \frac{\delta}{\delta \hat{a}^{\dagger}_{\mathbf{x}}} F[\hat{a}^{\dagger}_{\bullet}] \,. \tag{4.28}$$

In particular, (4.27) and (4.28) allow to reproduce the non-trivial commutation relation in (3.25) with functional derivatives:

$$\left[\hat{a}_{\mathbf{x}}, \hat{a}^{\dagger}_{\mathbf{x}'}\right]_{-} = \frac{\delta \hat{a}_{\mathbf{x}}}{\delta \hat{a}_{\mathbf{x}'}} = \frac{\delta \hat{a}^{\dagger}_{\mathbf{x}'}}{\delta \hat{a}^{\dagger}_{\mathbf{x}}} = \delta(\mathbf{x} - \mathbf{x}') \,. \tag{4.29}$$

And it is even possible to show that both calculational rules (4.27) and (4.28) can also be applied to functionals, which contain creation and annihilation operators in normal order:

$$\left[F[\hat{a}^{\dagger}_{\bullet}, \hat{a}_{\bullet}], \hat{a}^{\dagger}_{\mathbf{x}}\right]_{-} = F[\hat{a}^{\dagger}_{\bullet}, \hat{a}_{\bullet}] \, \overleftarrow{\frac{\delta}{\delta \hat{a}_{\mathbf{x}}}} \,, \tag{4.30}$$

$$\left[\hat{a}_{\mathbf{x}}, F[\hat{a}^{\dagger}_{\bullet}, \hat{a}_{\bullet}]\right]_{-} = \overrightarrow{\frac{\delta}{\delta \hat{a}^{\dagger}_{\mathbf{x}}}} \, F[\hat{a}^{\dagger}_{\bullet}, \hat{a}_{\bullet}] \,. \tag{4.31}$$

Here the arrows over the functional derivatives indicate from which side the normal ordered functional of creation and annihilation operators has to be differentiated. With this it is also possible to reproduce the trivial commutation relation in (3.25) with functional derivatives:

$$\left[\hat{a}^{\dagger}_{\mathbf{x}}, \hat{a}^{\dagger}_{\mathbf{x}'}\right]_{-} = \hat{a}^{\dagger}_{\mathbf{x}} \overleftarrow{\frac{\delta}{\delta \hat{a}_{\mathbf{x}'}}} = 0 \,, \qquad \left[\hat{a}_{\mathbf{x}}, \hat{a}_{\mathbf{x}'}\right]_{-} = \overrightarrow{\frac{\delta}{\delta \hat{a}^{\dagger}_{\mathbf{x}}}} \hat{a}_{\mathbf{x}'} = 0 \,. \tag{4.32}$$

Furthermore, the calculational rules (4.30) and (4.31) in the Schödinger picture can be extended correspondingly to the Heisenberg picture:

$$\left[F\left[\hat{\psi}^{\dagger}(\bullet,\bullet)\hat{\psi}(\bullet,\bullet)\right], \hat{\psi}^{\dagger}(\mathbf{x},t)\right]_{-} = F\left[\hat{\psi}^{\dagger}(\bullet,\bullet)\hat{\psi}(\bullet,\bullet)\right] \overleftarrow{\frac{\delta}{\delta \hat{\psi}(\mathbf{x},t)}} \,, \tag{4.33}$$

$$\left[\hat{\psi}(\mathbf{x},t), F\left[\hat{\psi}^{\dagger}(\bullet,\bullet)\hat{\psi}(\bullet,\bullet)\right]\right]_{-} = \overrightarrow{\frac{\delta}{\delta \hat{\psi}^{\dagger}(\mathbf{x},t)}} F\left[\hat{\psi}^{\dagger}(\bullet,\bullet)\hat{\psi}(\bullet,\bullet)\right] . \tag{4.34}$$

With this the Heisenberg equations of motion (3.62) of the fields operators $\hat{\psi}^{\dagger}(\mathbf{x},t)$, $\hat{\psi}(\mathbf{x},t)$ can be formulated with the help of functional derivatives:

$$i\hbar \frac{\partial \hat{\psi}(\mathbf{x},t)}{\partial t} = \left[\hat{\psi}(\mathbf{x},t), \hat{H}_{\mathrm{H}}(t)\right] = \overrightarrow{\frac{\delta}{\delta \hat{\psi}^{\dagger}(\mathbf{x},t)}} \hat{H}_{\mathrm{H}}(t) \,, \tag{4.35}$$

$$i\hbar \frac{\partial \hat{\psi}^{\dagger}(\mathbf{x},t)}{\partial t} = \left[\hat{\psi}^{\dagger}(\mathbf{x},t), \hat{H}_{\mathrm{H}}(t)\right] = -\hat{H}_{\mathrm{H}}(t) \overleftarrow{\frac{\delta}{\delta \hat{\psi}(\mathbf{x},t)}} \,. \tag{4.36}$$

Thus, we conclude that all commutators between second-quantized operators in Sections 3.2–3.4, which have been evaluated via the operator identities (3.10) and (3.43), can also be calculated with appropriate functional derivatives.

4.4 Euler-Lagrange Equations

After this technical excursion to the definition and application of functional derivatives we now return to the question how to determine the underlying equations of motion from the variational principle (4.7). Applying (4.18) to (4.7) we get

$$\delta\mathcal{A} = \int dt \int d^3x \left\{ \frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x},t)}\,\delta\psi^*(\mathbf{x},t) + \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)}\,\delta\psi(\mathbf{x},t) \right\} = 0\,. \tag{4.37}$$

As the variations of the fields $\delta\psi^*(\mathbf{x},t)$ and $\delta\psi(\mathbf{x},t)$ are considered to be independent, we obtain from (4.37) the following two conditions:

$$\frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x},t)} = 0\,, \qquad \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} = 0\,. \tag{4.38}$$

Thus, the Hamilton principle in Lagrangian field theory states that the fields $\psi^*(\mathbf{x},t)$ and $\psi(\mathbf{x},t)$ are determined from extremizing the action. It remains to explicitly determine the functional derivatives of the action $\mathcal{A}$ with respect to the fields $\psi^*(\mathbf{x},t)$ and $\psi(\mathbf{x},t)$. Due to (4.8) we have to consider the spatial coordinates $\mathbf{x}$ to be fixed and only take only variations with respect to the functional dependencies in time t into account. With the chain rule of functional differentiation (4.24) we get

$$\frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x},t)} = \int dt' \left\{ \frac{\delta L}{\delta\psi^*(\mathbf{x},t')}\,\frac{\delta\psi^*(\mathbf{x},t')}{\delta\psi^*(\mathbf{x},t)} + \frac{\delta L}{\delta\dfrac{\partial\psi^*(\mathbf{x},t')}{\partial t'}}\,\frac{\delta\dfrac{\partial\psi^*(\mathbf{x},t')}{\partial t'}}{\delta\psi^*(\mathbf{x},t)} \right\}. \tag{4.39}$$

Interchanging variation and partial derivative allows for a partial integration, where the boundary terms can be ignored, yielding

$$\frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x},t)} = \int dt' \left\{ \frac{\delta L}{\delta\psi^*(\mathbf{x},t')} - \frac{\partial}{\partial t'}\,\frac{\delta L}{\delta\dfrac{\partial\psi^*(\mathbf{x},t')}{\partial t'}} \right\} \frac{\delta\psi^*(\mathbf{x},t')}{\delta\psi^*(\mathbf{x},t)}\,. \tag{4.40}$$

From the trivial function derivative (4.22) follows

$$\frac{\delta\psi^*(\mathbf{x},t')}{\delta\psi^*(\mathbf{x},t)} = \delta(t-t')\,, \tag{4.41}$$

so we read off from (4.40)

$$\frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x},t)} = \frac{\delta L}{\delta\psi^*(\mathbf{x},t)} - \frac{\partial}{\partial t}\,\frac{\delta L}{\delta\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}}\,. \tag{4.42}$$

Correspondingly we obtain

$$\frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} = \frac{\delta L}{\delta\psi(\mathbf{x},t)} - \frac{\partial}{\partial t}\,\frac{\delta L}{\delta\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}}\,. \tag{4.43}$$

Thus, we conclude that (4.38) together with (4.42) and (4.43) represent the underlying Euler-Langrange equations. It remains to determine the respective functional derivatives of the Lagrange function (4.9). To this end we consider, conversely the time t to be fixed and only take only variations with respect to the functional dependencies in the spatial coordinates $\mathbf{x}$ into account. Applying similar techniques of the functional differentiation as before, we obtain

$$\frac{\delta L}{\delta\psi^*(\mathbf{x},t)} = \frac{\partial\mathcal{L}}{\partial\psi^*(\mathbf{x},t)} - \boldsymbol{\nabla}\,\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi^*(\mathbf{x},t)}\,, \qquad \frac{\delta L}{\delta\,\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}} = \frac{\partial\mathcal{L}}{\partial\,\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}}\,, \tag{4.44}$$

$$\frac{\delta L}{\delta\psi(\mathbf{x},t)} = \frac{\partial\mathcal{L}}{\partial\psi(\mathbf{x},t)} - \boldsymbol{\nabla}\,\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi(\mathbf{x},t)}\,, \qquad \frac{\delta L}{\delta\,\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}} = \frac{\partial\mathcal{L}}{\partial\,\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}}\,. \tag{4.45}$$

Thus, combining (4.38) with (4.42)–(4.45) yields ultimately the Euler-Lagrange equations of classical field theory:

$$\frac{\partial\mathcal{L}}{\partial\psi^*(\mathbf{x},t)} - \boldsymbol{\nabla}\,\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi^*(\mathbf{x},t)} - \frac{\partial}{\partial t}\,\frac{\partial\mathcal{L}}{\partial\,\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}} = 0\,, \tag{4.46}$$

$$\frac{\partial\mathcal{L}}{\partial\psi(\mathbf{x},t)} - \boldsymbol{\nabla}\,\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi(\mathbf{x},t)} - \frac{\partial}{\partial t}\,\frac{\partial\mathcal{L}}{\partial\,\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}} = 0\,. \tag{4.47}$$

Although we have derived these field equations in two variational steps by taking into account (4.8) and (4.9), they can also be directly determined by considering the action $\mathcal{A}$ as a spatio-temporal integral over the Lagrange density $\mathcal{L}$:

$$\mathcal{A} = \int dt \int d^3x\, \mathcal{L}\left(\psi^*(\mathbf{x},t), \boldsymbol{\nabla}\psi^*(\mathbf{x},t), \frac{\partial\psi^*(\mathbf{x},t)}{\partial t}; \psi(\mathbf{x},t), \boldsymbol{\nabla}\psi(\mathbf{x},t), \frac{\partial\psi(\mathbf{x},t)}{\partial t}\right)\,. \tag{4.48}$$

Now it remains in (4.46) and (4.47) to evaluate the respective partial derivatives of the Lagrange density $\mathcal{L}$ of the Schrödinger field theory defined in (4.10):

$$\frac{\partial\mathcal{L}}{\partial\psi^*(\mathbf{x},t)} = -V_1(\mathbf{x})\psi(\mathbf{x},t) + i\hbar\,\frac{\partial\psi(\mathbf{x},t)}{\partial t}\,, \quad \frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi^*(\mathbf{x},t)} = -\frac{\hbar^2}{2M}\,\boldsymbol{\nabla}\psi(\mathbf{x},t)\,, \quad \frac{\partial\mathcal{L}}{\partial\,\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}} = 0\,, \tag{4.49}$$

$$\frac{\partial\mathcal{L}}{\partial\psi(\mathbf{x},t)} = -V_1(\mathbf{x})\psi^*(\mathbf{x},t)\,, \quad \frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi(\mathbf{x},t)} = -\frac{\hbar^2}{2M}\,\boldsymbol{\nabla}\psi^*(\mathbf{x},t)\,, \quad \frac{\partial\mathcal{L}}{\partial\,\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}} = i\hbar\psi^*(\mathbf{x},t)\,. \tag{4.50}$$

Inserting these intermediate results (4.49) and (4.50) into (4.46) and (4.47) yields, indeed, the equations of motion of the Schrödinger theory (4.1) and (4.2).

4.5 Hamilton Field Theory

Now we go over from the Lagrange to the Hamilton formulation of classical field theory. To this end we have to determine at first the momenta fields $\pi^*(\mathbf{x},t)$, $\pi(\mathbf{x},t)$, which are canonically conjugated to the Schödinger fields $\psi^*(\mathbf{x},t)$, $\psi(\mathbf{x},t)$. In close analogy to a classical system with a finite number of degrees of freedom we obtain from (4.44), (4.49), and (4.50):

$$\pi^*(\mathbf{x},t) = \frac{\delta L}{\delta \dfrac{\partial \psi^*(\mathbf{x},t)}{\partial t}} = \frac{\partial \mathcal{L}}{\partial \dfrac{\partial \psi^*(\mathbf{x},t)}{\partial t}} = 0\,, \tag{4.51}$$

$$\pi(\mathbf{x},t) = \frac{\delta L}{\delta \dfrac{\partial \psi(\mathbf{x},t)}{\partial t}} = \frac{\partial \mathcal{L}}{\partial \dfrac{\partial \psi(\mathbf{x},t)}{\partial t}} = i\hbar\psi^*(\mathbf{x},t)\,, \tag{4.52}$$

Thus, we conclude that $\psi^*(\mathbf{x},t)$ represents the canonically conjugated momentum field of $\psi(\mathbf{x},t)$. The Hamilton function follows via a Legendre transformation from the Lagrange function:

$$H = \int d^3x \left\{ \pi^*(\mathbf{x},t)\,\frac{\partial \psi^*(\mathbf{x},t)}{\partial t} + \pi(\mathbf{x},t)\,\frac{\partial \psi(\mathbf{x},t)}{\partial t} \right\} - L\,. \tag{4.53}$$

Inserting therein (4.9), (4.10) and (4.51), (4.52) the Hamilton function turns out to be of the form

$$H = \int d^3x\, \mathcal{H}\left(\pi(\mathbf{x},t), \boldsymbol{\nabla}\pi(\mathbf{x},t); \psi(\mathbf{x},t), \boldsymbol{\nabla}\psi(\mathbf{x},t)\right)\,, \tag{4.54}$$

where the Hamilton density $\mathcal{H}$ is given by

$$\mathcal{H} = \frac{\hbar}{2Mi}\,\boldsymbol{\nabla}\pi(\mathbf{x},t)\cdot\boldsymbol{\nabla}\psi(\mathbf{x},t) + \frac{V_1(\mathbf{x})}{i\hbar}\pi(\mathbf{x},t)\psi(\mathbf{x},t)\,. \tag{4.55}$$

Thus, taking into account the relation (4.52) between $\pi(\mathbf{x},t)$ and $\psi^*(\mathbf{x},t)$ yields

$$H = \int d^3x \left\{ \frac{\hbar^2}{2M}\,\boldsymbol{\nabla}\psi^*(\mathbf{x},t)\cdot\boldsymbol{\nabla}\psi(\mathbf{x},t) + V_1(\mathbf{x})\psi^*(\mathbf{x},t)\psi(\mathbf{x},t) \right\}\,, \tag{4.56}$$

where a partial integration leads to the standard form

$$H = \int d^3x\, \psi^*(\mathbf{x},t) \left\{ -\frac{\hbar^2}{2M}\,\Delta + V_1(\mathbf{x}) \right\} \psi(\mathbf{x},t)\,, \tag{4.57}$$

Also the Hamilton equations of motion can be obtained in close analogy to the classical mechanics of a finite number of degrees of freedom. To this end one has to consider the action $\mathcal{A}$ as a functional of the fields $\pi(\mathbf{x},t)$ and $\psi(\mathbf{x},t)$. Then the Hamilton principle

$$\delta\mathcal{A}\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right] = \int dt \int d^3x \left\{ \frac{\delta\mathcal{A}}{\delta\pi(\mathbf{x},t)}\,\delta\pi(\mathbf{x},t) + \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)}\,\delta\psi(\mathbf{x},t) \right\} = 0\,. \tag{4.58}$$

leads because of the arbitrariness of the variations $\delta\pi(\mathbf{x},t)$ and $\delta\psi(\mathbf{x},t)$ to

$$\frac{\delta\mathcal{A}}{\delta\pi(\mathbf{x},t)} = 0\,, \qquad \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} = 0\,. \tag{4.59}$$

Due to (4.8) and (4.53) the action $\mathcal{A}$ depends on the Hamilton function H as follows:

$$\mathcal{A} = \int dt \int d^3x\, \pi(\mathbf{x},t)\,\frac{\partial \psi(\mathbf{x},t)}{\partial t} - \int dt\, H\left[\pi(\bullet,t);\psi(\bullet,t)\right]\,. \tag{4.60}$$

With this we can now evaluate the functional derivatives in (4.59), yielding the Hamilton equations of motion of classical field theory:

$$\frac{\delta \mathcal{A}}{\delta \pi(\mathbf{x},t)} = \frac{\partial \psi(\mathbf{x},t)}{\partial t} - \frac{\delta H}{\delta \pi(\mathbf{x},t)} = 0\,, \tag{4.61}$$

$$\frac{\delta \mathcal{A}}{\delta \psi(\mathbf{x},t)} = -\frac{\partial \pi(\mathbf{x},t)}{\partial t} - \frac{\delta H}{\delta \psi(\mathbf{x},t)} = 0\,. \tag{4.62}$$

As the Hamilton function H is of the form (4.56), the respective functional derivatives in (4.61) and (4.62) yield

$$\frac{\delta H}{\delta \pi(\mathbf{x},t)} = \frac{\partial \mathcal{H}}{\partial \pi(\mathbf{x},t)} - \boldsymbol{\nabla} \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \pi(\mathbf{x},t)}\,, \tag{4.63}$$

$$\frac{\delta H}{\delta \psi(\mathbf{x},t)} = \frac{\partial \mathcal{H}}{\partial \psi(\mathbf{x},t)} - \boldsymbol{\nabla} \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \psi(\mathbf{x},t)}\,, \tag{4.64}$$

Thus, inserting (4.63) and (4.64) into (4.61), (4.62) the Hamilton equations of classical field theory have the form

$$\frac{\partial \psi(\mathbf{x},t)}{\partial t} = \frac{\partial \mathcal{H}}{\partial \pi(\mathbf{x},t)} - \boldsymbol{\nabla} \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \pi(\mathbf{x},t)}\,, \tag{4.65}$$

$$\frac{\partial \pi(\mathbf{x},t)}{\partial t} = -\frac{\partial \mathcal{H}}{\partial \psi(\mathbf{x},t)} + \boldsymbol{\nabla} \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \psi(\mathbf{x},t)}\,. \tag{4.66}$$

Due to the Hamilton density of the Schrödinger theory (4.55) the respective partial derivatives read

$$\frac{\partial \mathcal{H}}{\partial \pi(\mathbf{x},t)} = \frac{V_1(\mathbf{x})}{i\hbar}\psi(\mathbf{x},t)\,, \qquad \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \pi(\mathbf{x},t)} = \frac{\hbar}{2Mi}\,\boldsymbol{\nabla}\psi(\mathbf{x},t)\,, \tag{4.67}$$

$$\frac{\partial \mathcal{H}}{\partial \psi(\mathbf{x},t)} = \frac{V_1(\mathbf{x})}{i\hbar}\pi(\mathbf{x},t)\,, \qquad \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \psi(\mathbf{x},t)} = \frac{\hbar}{2Mi}\,\boldsymbol{\nabla}\pi(\mathbf{x},t)\,. \tag{4.68}$$

Thus, we recover from (4.65)–(4.68) due to (4.52) the equations of motion of the Schrödinger theory (4.1) and (4.2).

4.6 Poisson Brackets

And, finally, we analyze the role of Poisson brackets in classical field theory. To this end we define for two functionals $F\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right]$ and $G\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right]$ their Poisson bracket via

$$\Big\{F,G\Big\}_- = \int d^3x \left(\frac{\delta F}{\delta \psi(\mathbf{x},t)}\,\frac{\delta G}{\delta \pi(\mathbf{x},t)} - \frac{\delta F}{\delta \pi(\mathbf{x},t)}\,\frac{\delta G}{\delta \psi(\mathbf{x},t)} \right)\,. \tag{4.69}$$

This allows to reexpress the Hamilton equations (4.61), (4.62) with the help of Poisson brackets:

$$\Big\{\psi(\mathbf{x},t),H\Big\}_- = \int d^3x' \left(\frac{\delta \psi(\mathbf{x},t)}{\delta \psi(\mathbf{x}',t)}\,\frac{\delta H}{\delta \pi(\mathbf{x}',t)} - \frac{\delta \psi(\mathbf{x},t)}{\delta \pi(\mathbf{x}',t)}\,\frac{\delta H}{\delta \psi(\mathbf{x}',t)} \right) = \frac{\delta H}{\delta \pi(\mathbf{x},t)} = \frac{\partial \psi(\mathbf{x},t)}{\partial t} \tag{4.70}$$

$$\Big\{\pi(\mathbf{x},t),H\Big\}_- = \int d^3x' \left(\frac{\delta\pi(\mathbf{x},t)}{\delta\psi(\mathbf{x}',t)}\frac{\delta H}{\delta\pi(\mathbf{x}',t)} - \frac{\delta\pi(\mathbf{x},t)}{\delta\pi(\mathbf{x}',t)}\frac{\delta H}{\delta\psi(\mathbf{x}',t)}\right) = -\frac{\delta H}{\delta\psi(\mathbf{x},t)} = \frac{\partial\pi(\mathbf{x},t)}{\partial t} \quad (4.71)$$

Also the temporal change of a functional $F\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right]$ can be formulated with the help of Poisson brackets. At first we obtain with the chain rule of functional differentiation:

$$\frac{\partial F}{\partial t} = \int d^3x \left(\frac{\partial\pi(\mathbf{x},t)}{\partial t}\frac{\delta F}{\delta\pi(\mathbf{x},t)} + \frac{\partial\psi(\mathbf{x},t)}{\partial t}\frac{\delta F}{\delta\psi(\mathbf{x},t)}\right), \quad (4.72)$$

which reduces due to the Hamilton equations (4.61), (4.62) and the Poisson brackets (4.69) to

$$\frac{\partial F}{\partial t} = \int d^3x \left(-\frac{\delta H}{\delta\psi(\mathbf{x},t)}\frac{\delta F}{\delta\pi(\mathbf{x},t)} + \frac{\delta H}{\delta\pi(\mathbf{x},t)}\frac{\delta F}{\delta\psi(\mathbf{x},t)}\right) = \Big\{F,H\Big\}_- . \quad (4.73)$$

Thus, the formulation of the Hamilton equations in form of Poisson brackets according to (4.70) and (4.71) follows immediately from (4.73). Furthermore, we obtain for the fundamental Poisson brackets of the Schrödinger field $\psi(\mathbf{x},t)$ and its canonical momentum field $\pi(\mathbf{x},t)$ at equal times:

$$\Big\{\psi(\mathbf{x},t),\psi(\mathbf{x}',t)\Big\}_- = \int d^3x'' \left(\frac{\delta\psi(\mathbf{x},t)}{\delta\psi(\mathbf{x}'',t)}\frac{\delta\psi(\mathbf{x}',t)}{\delta\pi(\mathbf{x}'',t)} - \frac{\delta\psi(\mathbf{x},t)}{\delta\pi(\mathbf{x}'',t)}\frac{\delta\psi(\mathbf{x}',t)}{\delta\psi(\mathbf{x}'',t)}\right) = 0 . \quad (4.74)$$

$$\Big\{\pi(\mathbf{x},t),\pi(\mathbf{x}',t)\Big\}_- = \int d^3x'' \left(\frac{\delta\pi(\mathbf{x},t)}{\delta\psi(\mathbf{x}'',t)}\frac{\delta\pi(\mathbf{x}',t)}{\delta\pi(\mathbf{x}'',t)} - \frac{\delta\pi(\mathbf{x},t)}{\delta\pi(\mathbf{x}'',t)}\frac{\delta\pi(\mathbf{x}',t)}{\delta\psi(\mathbf{x}'',t)}\right) = 0 . \quad (4.75)$$

$$\Big\{\psi(\mathbf{x},t),\pi(\mathbf{x}',t)\Big\}_- = \int d^3x'' \left(\frac{\delta\psi(\mathbf{x},t)}{\delta\psi(\mathbf{x}'',t)}\frac{\delta\pi(\mathbf{x}',t)}{\delta\pi(\mathbf{x}'',t)} - \frac{\delta\psi(\mathbf{x},t)}{\delta\pi(\mathbf{x}'',t)}\frac{\delta\pi(\mathbf{x}',t)}{\delta\psi(\mathbf{x}'',t)}\right) = \delta(\mathbf{x}-\mathbf{x}') . \quad (4.76)$$

4.7 Canonical Field Quantization

On the basis of having worked out the classical field theory to such an extent, we can now perform the canonical field quantization in the Heisenberg picture. To this end we associate to the complex Schrödinger field $\psi(\mathbf{x},t)$ and its canonically conjugated momentum field $\pi(\mathbf{x},t)$ corresponding second quantized field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\pi}(\mathbf{x},t)$. Furthermore, in close analogy to the quantum mechanics for a finite number of degrees of freedom, we postulate that the Poisson bracket between two functionals F and G goes over into a commutator between their corresponding second quantized operators $\hat{F}$ and $\hat{G}$ as follows:

$$\Big\{F,G\Big\}_- \qquad \Longrightarrow \qquad \frac{1}{i\hbar}\Big[\hat{F},\hat{G}\Big]_- . \quad (4.77)$$

In this way, the fundamental Poisson brackets (4.74)–(4.76) go over into equal-time commutation relations

$$\Big[\hat{\psi}(\mathbf{x},t),\hat{\psi}(\mathbf{x}',t)\Big]_- = \Big[\hat{\pi}(\mathbf{x},t),\hat{\pi}(\mathbf{x}',t)\Big]_- = 0 , \qquad \Big[\hat{\psi}(\mathbf{x},t),\hat{\pi}(\mathbf{x}',t)\Big]_- = i\hbar\,\delta(\mathbf{x}-\mathbf{x}') . \quad (4.78)$$

As (4.52) implies that the momentum field operator $\hat{\pi}(\mathbf{x},t)$ is given by the adjoint field operator $\hat{\psi}^\dagger(\mathbf{x},t)$ via

$$\hat{\pi}(\mathbf{x},t) = i\hbar\,\hat{\psi}^\dagger(\mathbf{x},t) , \quad (4.79)$$

we recognize that the previous equal-time commutation relations (3.64) between the field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$ follow from (4.78). Furthermore, the postulate (4.77) converts the Hamilton equations (4.70), (4.71) into

$$i\hbar \frac{\partial \hat{\psi}(\mathbf{x},t)}{\partial t} = \left[\hat{\psi}(\mathbf{x},t), \hat{H}\right]_- , \tag{4.80}$$

$$i\hbar \frac{\partial \hat{\pi}(\mathbf{x},t)}{\partial t} = \left[\hat{\pi}(\mathbf{x},t), \hat{H}\right]_- . \tag{4.81}$$

Due to (4.79) they turn out to agree with the Heisenberg equations of motion of the fields operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$ in (3.68) and (3.73). And the Hamilton function (4.57) is converted within the canonical field quantization to the Hamilton operator (3.67) in the Heisenberg picture without the 2-particle interaction.

Chapter 5

Canonical Field Quantization for Fermions

In the previous chapter we worked out with the help of the functional derivative the classical field theory for bosons. Its canonical field quantization then allowed to derive the equal-time commutation relations for the bosonic field operators. Here we show that a corresponding derivation is also possible in view of the equal-time anti-commutation relations for the fermionic field operators. But in order to obtain a proper classical field theory for fermions, one needs anti-commuting Grassmann fields. Therefore, we start this chapter with introducing the concept of anti-commuting Grassmann numbers and fields, which was developed by the mathematician Hermann Grassmann in the middle of the 19th century.

5.1 Grassmann Fields

The classical analogue of the Pauli exclusion principle is not realizable within the realm of the usual numbers like real or complex numbers but needs the new mathematical concept of Grassmann numbers.

5.1.1 Grassmann Numbers

The entity of anti-commuting Grassmann numbers is called the Grassmann algebra. Each element of a Grassmann algebra of dimension n can be represented by a set of n generators or Grassmann variables ψ_i, where the index i runs from 1 to n. The Grassmann algebra is defined by postulating the anti-commutation relations

$$\Big[\psi_i, \psi_j\Big]_+ = \psi_i\,\psi_j + \psi_j\,\psi_i = 0 \tag{5.1}$$

for all $i, j = 1, \ldots, n$. As a special case of (5.1) we read off that the square and all higher powers of a generator have to vanish:

$$\psi_i^2 = 0 \,. \tag{5.2}$$

This has the consequence that each element of the Grassmann algebra can be expanded in a finite sum over products of generators as follows:

$$\begin{aligned} f(\psi_1, \ldots, \psi_n) &= f^{(0)} + \sum_{i=1}^{n} f_i^{(1)} \, \psi_i + \sum_{i=1}^{n} \sum_{j=1}^{i-1} f_{ij}^{(2)} \, \psi_i \, \psi_j + \ldots \\ &\quad + \sum_{i=1}^{n} f_i^{(n-1)} \, \psi_1 \, \cdots \, \psi_{i-1} \, \psi_{i+1} \, \cdots \, \psi_n + f^{(n)} \, \psi_1 \, \cdots \, \psi_n \,, \end{aligned} \tag{5.3}$$

where all coefficients $f^{(0)}, f_i^{(1)}, f_{ij}^{(2)}, \ldots, f_i^{(n-1)}, f^{(n)}$ are complex numbers. Due to the anti-commutation relations (5.1) it is sufficient in the sum (5.3) that the indices of the generators appear in ascending order, i.e. $i > j$ in the third term. This reduces correspondingly the number of independent products of p generators to the binomial coefficients

$$n_p = \binom{n}{p} \,. \tag{5.4}$$

For instance, one obtains for $p = 0, 1, 2, \ldots, n-1, n$:

$$\begin{aligned} n_0 = 1 = \binom{n}{0} \,, \quad n_1 = \sum_{i=1}^{n} 1 = n = \binom{n}{1} \,, \quad n_2 = \sum_{i=1}^{n} \sum_{j=1}^{i-1} 1 = \frac{1}{2} \, n(n-1) = \binom{n}{2} \\ , \ldots , \quad n_{n-1} = \sum_{i=1}^{n} 1 = n = \binom{n}{n-1} \,, \quad n_n = 1 = \binom{n}{n} \,. \end{aligned} \tag{5.5}$$

The dimension of the Grassmann algebra, i.e. the maximal number of linear independent terms in the expansion (5.3) amounts to

$$\sum_{p=0}^{n} n_p = \sum_{p=0}^{n} 1^p \, 1^{n-p} \binom{n}{p} = 2^n \tag{5.6}$$

as (5.4) has to be taken into account.

5.1.2 Grassmann Functions

A Grassmann function maps a Grassmann number (5.3) to another Grassmann number (5.3). Consider as an example the Grassmann algebra of degree 2 with the generators ψ_1 and ψ_2, which has the dimension $2^2 = 4$. A Grassmann number f is then represented as

$$f(\psi_1, \psi_2) = f^{(0)} + f_1^{(1)} \, \psi_1 + f_2^{(1)} \, \psi_2 + f^{(2)} \, \psi_1 \, \psi_2 \,. \tag{5.7}$$

ordinary variables	Grassmann variables
$\frac{\partial}{\partial x_i} 1 = 0$	$\frac{\partial}{\partial \psi_i} 1 = 0$
$\frac{\partial x_j}{\partial x_i} = \delta_{ij}$	$\frac{\partial \psi_j}{\partial \psi_i} = \delta_{ij}$
$\frac{\partial}{\partial x_i}\Big(x_j\, x_k\Big) = \delta_{ij}\, x_k + \delta_{ik}\, x_j$	$\frac{\partial}{\partial \psi_i}\Big(\psi_j\, \psi_k\Big) = \delta_{ij}\, \psi_k - \delta_{ik}\, \psi_j$
$\frac{\partial}{\partial x_i}\Big[x_j\, f(x_1,\ldots,x_n)\Big] = \delta_{ij}\, f(x_1,\ldots,x_n) + x_j\, \frac{\partial f(x_1,\ldots,x_n)}{\partial x_i}$	$\frac{\partial}{\partial \psi_i}\Big[\psi_j\, f(\psi_1,\ldots,\psi_n)\Big] = \delta_{ij}\, f(\psi_1,\ldots,\psi_n) - \psi_j\, \frac{\partial f(\psi_1,\ldots,\psi_n)}{\partial \psi_i}$
$\left[\frac{\partial}{\partial x_i}, x_j\right]_- = \frac{\partial}{\partial x_i}\, x_j - x_j\, \frac{\partial}{\partial x_i} = \delta_{ij}$	$\left[\frac{\partial}{\partial \psi_i}, \psi_j\right]_+ = \frac{\partial}{\partial \psi_i}\, \psi_j + \psi_j\, \frac{\partial}{\partial \psi_i} = \delta_{ij}$
$\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right]_- = \frac{\partial}{\partial x_i}\, \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_j}\, \frac{\partial}{\partial x_i} = 0$	$\left[\frac{\partial}{\partial \psi_i}, \frac{\partial}{\partial \psi_j}\right]_+ = \frac{\partial}{\partial \psi_i}\, \frac{\partial}{\partial \psi_j} + \frac{\partial}{\partial \psi_j}\, \frac{\partial}{\partial \psi_i} = 0$

Figure 5.1: Comparison of calculation rules for differentiation with respect to ordinary and Grassmann variables.

With the help of the Taylor series and (5.1) we obtain, for instance, the following two Grassmann functions:

$$e^{\psi_1+\psi_2} = 1 + \psi_1 + \psi_2\,, \tag{5.8}$$

$$e^{\psi_1\,\psi_2} = 1 + \psi_1\,\psi_2\,. \tag{5.9}$$

Does a Grassmann number f only consist of an even number of generators, then it commutes with all Grassmann numbers and one assigns to it the parity $\pi(f) = 0$. In the opposite case that a Grassmann number f consists of an odd number of generators, then it anti-commutes with such Grassmann numbers, which also have an odd number of generators, and one assigns to it the parity $\pi(f) = 1$. Grassmann numbers, which contain both an even and an odd number of generators, do not have any parity. We have, for example, $\pi(\psi_1) = \pi(\psi_2) = 1$ and $\pi(e^{\psi_1\,\psi_2}) = 0$ due to (5.9), but we can assign to $e^{\psi_1+\psi_2}$ no parity due to (5.8).

5.1.3 Differentiation and Integration

Within a Grassmann algebra one can introduce the operations of differentiation and integration. But these are abstract constructions, which have properties differing considerably from the usual differentiation and integration calculus with real or complex numbers. In comparison of a differentiation with respect to an ordinary variable, the differentiation with respect to a Grassmann variable is defined via the rules in Fig. 5.1. As an example we consider again the Grassmann algebra of degree 2. For the respective derivatives of (5.7) we obtain

$$\frac{\partial f(\psi_1,\psi_2)}{\partial \psi_1} = f_1^{(1)} + f^{(2)}\,\psi_2\,, \qquad \frac{\partial f(\psi_1,\psi_2)}{\partial \psi_2} = f_2^{(1)} - f^{(2)}\,\psi_1\,, \tag{5.10}$$

$$\frac{\partial^2 f(\psi_1,\psi_2)}{\partial \psi_1^2} = \frac{\partial^2 f(\psi_1,\psi_2)}{\partial \psi_2^2} = 0\,, \qquad \frac{\partial^2 f(\psi_1,\psi_2)}{\partial \psi_1 \partial \psi_2} = -f^{(2)} = -\frac{\partial^2 f(\psi_1,\psi_2)}{\partial \psi_2 \partial \psi_1}\,. \tag{5.11}$$

ordinary variable	Grassmann variable
$\int_{-\infty}^{\infty} dx \Big[\alpha f(x) + \beta g(x)\Big] = \alpha \int_{-\infty}^{\infty} dx f(x) + \beta \int_{-\infty}^{\infty} dx g(x)$	$\int d\psi \Big[\alpha f(\psi) + \beta g(\psi)\Big] = \alpha \int d\psi f(\psi) + \beta \int d\psi g(\psi)$
$\int_{-\infty}^{\infty} dx f(x + x_0) = \int_{-\infty}^{\infty} dx f(x)$	$\int d\psi f(\psi + \psi_0) = \int d\psi f(\psi)$

Figure 5.2: Linearity and translational invariance as defining properties of integration with respect to Grassmann variables in comparison to integration with respect to ordinary variables.

Correspondingly the derivatives of (5.8) and (5.9) yield

$$\frac{\partial e^{\psi_1+\psi_2}}{\partial \psi_1} = 1\,, \qquad \frac{\partial e^{\psi_1+\psi_2}}{\partial \psi_2} = 1\,, \tag{5.12}$$

$$\frac{\partial e^{\psi_1\,\psi_2}}{\partial \psi_1} = \psi_2\,, \qquad \frac{\partial e^{\psi_1\,\psi_2}}{\partial \psi_2} = -\psi_1\,. \tag{5.13}$$

Introducing an integration with respect to a Grassmann variable one has to abstain from various usual properties. For instance, the integration with respect to Grassmann variables can not be defined via a Riemann sum as there does not exist any concrete interpretation for an area under a curve. But the integration can also not be defined by inverting the differentiation as integration boundaries do not make any sense. Let us consider at first the case of a single Grassmann variable ψ with the property

$$\Big[\psi, \psi\Big]_+ = 0\,. \tag{5.14}$$

Then a Grassmann function f of this Grassmann variable ψ is given by

$$f(\psi) = a + b\,\psi\,. \tag{5.15}$$

The integral $\int d\psi\, f(\psi)$ is determined according to Fig. 5.2 such that its properties are similar to those of a definite integral $\int_{-\infty}^{+\infty} dx\, f(x)$ of ordinary functions $f(x)$, which vanish at infinity. Demanding linearity and translational invariance of integration according to Fig. 5.2, we conclude

$$\int d\psi\, \Big[a + b\, (\psi + \psi_0)\Big] = \int d\psi\, \Big(a + b\,\psi\Big) + b \left(\int d\psi\, 1\right) \psi_0 = \int d\psi \Big(a + b\,\psi\Big)\,, \tag{5.16}$$

from which we can read off the following important integration rule:

$$\int d\psi\, 1 = 0\,. \tag{5.17}$$

This is integration rule is complemented by the arbitrary normalization

$$\int d\psi\, \psi = 1\,. \tag{5.18}$$

Those integration rules (5.17), (5.18) have to be compared with the corresponding differentiation rules, see Fig. 5.1:

$$\frac{d}{d\psi}\,1 = 0\,, \qquad \frac{d}{d\psi}\,\psi = 1\,. \tag{5.19}$$

Thus, one can conclude that in the space of Grassmann numbers integration and differentiation are surprisingly identical. For instance, we get for the function (5.15):

$$\int d\psi\, f(\psi) = b\,, \qquad \frac{d}{d\psi}\, f(\psi) = b\,. \tag{5.20}$$

The generalization of the integration rules (5.17), (5.18) to the case of higher-dimensional Grassmann algebra with generators ψ_i for $i = 1, \ldots, n$ is given by

$$\int d\psi_i\, 1 = 0\,, \qquad \int d\psi_i\, \psi_j = \delta_{ij}\,, \tag{5.21}$$

which corresponds to the differentiation rules in Fig. 5.1. Multiple integrals are calculated in the usual way by performing successively the respective one-dimensional integrals. As an example we determine the integral over the function (5.7):

$$\int d\psi_2 \int d\psi_1\, f(\psi_1, \psi_2) = f^{(2)}\,. \tag{5.22}$$

Thus, a multiple integration has the effect of projecting the corresponding coefficient in the expansion (5.3) of the Grassmann function

$$\int d\psi_n \ldots \int d\psi_1\, f(\psi_1, \ldots, \psi_n) = f^{(n)}\,. \tag{5.23}$$

5.1.4 Complex Grassmann Numbers

In view of dealing with a quantum many-body problem with an arbitrary number of fermions it is reasonable to also introduce complex Grassmann numbers. To this end one deals with two disjunct sets of Grassmann numbers $\psi_1, \ldots, \psi_n$ and $\psi_1^*, \ldots, \psi_n^*$, which anti-commute:

$$\Big[\psi_i, \psi_j\Big]_+ = \Big[\psi_i^*, \psi_j^*\Big]_+ = \Big[\psi_i, \psi_j^*\Big]_+ = 0\,. \tag{5.24}$$

Those generators constitute together a $2n$-dimensional Grassmann algebra. Both sets $\psi_1, \ldots, \psi_n$ and $\psi_1^*, \ldots, \psi_n^*$ are interconnected via the operation of conjugation:

$$(\psi_i)^* = \psi_i^*\,, \quad (\psi_i^*)^* = \psi_i\,, \quad (\psi_{i_1}\,\psi_{i_2}\,\cdots\,\psi_{i_n})^* = \psi_{i_n}^*\,\cdots\,\psi_{i_2}^*\,\psi_{i_1}^*\,, \quad (\lambda\,\psi_i)^* = \lambda^*\,\psi_i^*\,, \tag{5.25}$$

where λ denotes a complex number. Differentiation and integration are the defined in such a way that both sets $\psi_1, \ldots, \psi_n$ and $\psi_1^*, \ldots, \psi_n^*$ are treated as independent numbers.

5.1.5 Grassmann Fields

Finally, in order to apply complex Grassmann numbers in the realm of classical field theory for fermions, we have to introduce also anti-commuting fields, which amounts to the continuum limit $\psi_i \to \psi(x)$ and $\psi_i^* \to \psi^*(x)$. Thus, the anti-commutation relations (5.24) go over to

$$\Big[\psi(x), \psi(x')\Big]_+ = \Big[\psi^*(x), \psi^*(x')\Big]_+ = \Big[\psi(x), \psi^*(x')\Big]_+ = 0\,. \tag{5.26}$$

With this the anti-commuting fields $\psi(x)$ and $\psi^*(x)$ are the generators of an infinite-dimensional Grassmann algebra. An arbitrary element of this algebra then represents a functional $f[\psi^*, \psi]$, which can be expanded in generalization to (5.3) according to

$$f[\psi^*(\bullet), \psi(\bullet)] = f^{(0)} + \int dx_1 \, \Big\{ f_1^{(1)}(x_1)\psi^*(x_1) + f_2^{(1)}(x_1)\psi(x_1) \Big\} + \int dx_1 \int dx_2 \\ \times \Big\{ f_1^{(2)}(x_1, x_2)\psi^*(x_1)\psi^*(x_2) + f_2^{(2)}(x_1, x_2)\psi^*(x_1)\psi(x_2) + f_3^{(2)}(x_1, x_2)\psi(x_1)\psi(x_2) \Big\} + \ldots\,. \tag{5.27}$$

In this continuum limit the differentiation with respect to Grassmann variables becomes the functional derivative with respect to complex Grassmann fields, which obey the rules

$$\frac{\delta\psi(x)}{\delta\psi(x')} = \frac{\delta\psi^*(x)}{\delta\psi^*(x')} = \delta(x - x')\,, \qquad \frac{\delta\psi(x)}{\delta\psi^*(x')} = \frac{\delta\psi^*(x)}{\delta\psi(x')} = 0\,. \tag{5.28}$$

5.2 Lagrange Field Theory for Fermions

Now we develop a classical field theory for fermions and assume to this end that the Schrödinger fields $\psi^*(\mathbf{x}, t)$ and $\psi(\mathbf{x}, t)$ are anti-commuting complex Grassmann fields. As in the bosonic case (4.8)–(4.10) the action is a space-time integral

$$\mathcal{A} = \int dt\, L\left[\psi^*(\bullet, t), \frac{\partial\psi^*(\bullet, t)}{\partial t}; \psi(\bullet, t), \frac{\partial\psi(\bullet, t)}{\partial t}\right] \tag{5.29}$$

of the Lagrange function

$$L = \int d^3x\, \mathcal{L}\left(\psi^*(\mathbf{x}, t), \boldsymbol{\nabla}\psi^*(\mathbf{x}, t), \frac{\partial\psi^*(\mathbf{x}, t)}{\partial t}; \psi(\mathbf{x}, t), \boldsymbol{\nabla}\psi(\mathbf{x}, t), \frac{\partial\psi(\mathbf{x}, t)}{\partial t}\right), \tag{5.30}$$

where the Lagrange density is given by

$$\mathcal{L} = i\hbar\, \psi^*(\mathbf{x}, t)\, \frac{\partial\psi(\mathbf{x}, t)}{\partial t} - \frac{\hbar^2}{2M}\, \boldsymbol{\nabla}\psi^*(\mathbf{x}, t) \cdot \boldsymbol{\nabla}\psi(\mathbf{x}, t) - V_1(\mathbf{x})\psi^*(\mathbf{x}, t)\psi(\mathbf{x}, t)\,. \tag{5.31}$$

Instead of the bosonic Hamilton principle of the Lagrange field theory (4.37), (4.38) we obtain now the corresponding fermionic version:

$$\delta\mathcal{A} = \int dt \int d^3x\, \left\{ \delta\psi^*(\mathbf{x}, t)\, \frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x}, t)} + \delta\psi(\mathbf{x}, t)\, \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x}, t)} \right\} = 0\,. \tag{5.32}$$

As the variations of the complex Grassmann fields $\delta\psi^*(\mathbf{x},t)$ and $\delta\psi(\mathbf{x},t)$ are considered to be independent, we obtain from (5.32) like in (4.38):

$$\frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x},t)} = 0\,, \qquad \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} = 0\,. \tag{5.33}$$

Calculating the functional derivatives of the action (5.29) with respect to the complex Grassmann fields $\psi^*(\mathbf{x},t)$ and $\psi(\mathbf{x},t)$ yields the same Euler-Lagrange equations like in the bosonic case (4.42) and (4.43):

$$\frac{\delta\mathcal{A}}{\delta\psi^*(\mathbf{x},t)} = \frac{\delta L}{\delta\psi^*(\mathbf{x},t)} - \frac{\partial}{\partial t}\frac{\delta L}{\delta\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}}\,, \tag{5.34}$$

$$\frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} = \frac{\delta L}{\delta\psi(\mathbf{x},t)} - \frac{\partial}{\partial t}\frac{\delta L}{\delta\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}}\,. \tag{5.35}$$

Also the respective functional derivatives of the Lagrange function (5.30) with respect to the complex Grassmann fields $\psi^*(\mathbf{x},t)$ and $\psi(\mathbf{x},t)$ formally coincide with the bosonic calculation (4.44), (4.45):

$$\frac{\delta L}{\delta\psi^*(\mathbf{x},t)} = \frac{\partial\mathcal{L}}{\partial\psi^*(\mathbf{x},t)} - \boldsymbol{\nabla}\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi^*(\mathbf{x},t)}\,, \qquad \frac{\delta L}{\delta\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}} = \frac{\partial\mathcal{L}}{\partial\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}}\,, \tag{5.36}$$

$$\frac{\delta L}{\delta\psi(\mathbf{x},t)} = \frac{\partial\mathcal{L}}{\partial\psi(\mathbf{x},t)} - \boldsymbol{\nabla}\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi(\mathbf{x},t)}\,, \qquad \frac{\delta L}{\delta\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}} = \frac{\partial\mathcal{L}}{\partial\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}}\,. \tag{5.37}$$

Thus, also the Euler-Lagrange equations for the complex Grassmann fields have formally the same structure as in the bosonic case (4.46), (4.47)

$$\frac{\partial\mathcal{L}}{\partial\psi^*(\mathbf{x},t)} - \boldsymbol{\nabla}\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi^*(\mathbf{x},t)} - \frac{\partial\mathcal{L}}{\partial\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}} = 0\,, \tag{5.38}$$

$$\frac{\partial\mathcal{L}}{\partial\psi(\mathbf{x},t)} - \boldsymbol{\nabla}\frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi(\mathbf{x},t)} - \frac{\partial\mathcal{L}}{\partial\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}} = 0\,. \tag{5.39}$$

A difference between the Schrödinger field theory for bosons and fermions only occurs once the partial derivatives of the Lagrange density (5.31) are determined:

$$\frac{\partial\mathcal{L}}{\partial\psi^*(\mathbf{x},t)} = -V_1(\mathbf{x})\psi(\mathbf{x},t) + i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t}\,, \frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi^*(\mathbf{x},t)} = -\frac{\hbar^2}{2M}\boldsymbol{\nabla}\psi(\mathbf{x},t)\,, \frac{\partial\mathcal{L}}{\partial\dfrac{\partial\psi^*(\mathbf{x},t)}{\partial t}} = 0\,, \tag{5.40}$$

$$\frac{\partial\mathcal{L}}{\partial\psi(\mathbf{x},t)} = V_1(\mathbf{x})\psi(\mathbf{x},t)\,, \quad \frac{\partial\mathcal{L}}{\boldsymbol{\nabla}\psi(\mathbf{x},t)} = \frac{\hbar^2}{2M}\boldsymbol{\nabla}\psi^*(\mathbf{x},t)\,, \quad \frac{\partial\mathcal{L}}{\partial\dfrac{\partial\psi(\mathbf{x},t)}{\partial t}} = -i\hbar\psi^*(\mathbf{x},t)\,. \tag{5.41}$$

Namely, whereas (4.49) and (5.40) have the same signs, we observe different signs in (4.50) and (5.41). Despite of that we obtain in the fermionic case from (5.38)–(5.41) formally the same equations of motion for the Schrödinger Grassmann fields

$$i\hbar \frac{\partial \psi(\mathbf{x},t)}{\partial t} = \left\{ -\frac{\hbar^2}{2M} \Delta + V_1(\mathbf{x}) \right\} \psi(\mathbf{x},t) \,, \tag{5.42}$$

$$-i\hbar \frac{\partial \psi^*(\mathbf{x},t)}{\partial t} = \left\{ -\frac{\hbar^2}{2M} \Delta + V_1(\mathbf{x}) \right\} \psi^*(\mathbf{x},t) \,. \tag{5.43}$$

as in the bosonic case (4.1) and (4.2).

5.3 Hamilton Field Theory for Fermions

Going over from the Lagrange to the Hamilton formulation of field theory one needs the momentum fields, which are canonically conjugated to the anti-commuting Schrödinger fields $\psi^*(\mathbf{x},t)$ and $\psi(\mathbf{x},t)$. From (5.36), (5.37) as well as from (5.40), (5.41) we conclude:

$$\pi^*(\mathbf{x},t) = \frac{\delta L}{\delta \dfrac{\partial \psi^*(\mathbf{x},t)}{\partial t}} = \frac{\partial \mathcal{L}}{\partial \dfrac{\partial \psi^*(\mathbf{x},t)}{\partial t}} = 0 \,, \tag{5.44}$$

$$\pi(\mathbf{x},t) = \frac{\delta L}{\delta \dfrac{\partial \psi(\mathbf{x},t)}{\partial t}} = \frac{\partial \mathcal{L}}{\partial \dfrac{\partial \psi(\mathbf{x},t)}{\partial t}} = -i\hbar \psi^*(\mathbf{x},t) \,. \tag{5.45}$$

Thus, $\pi^*(\mathbf{x},t)$ vanishes and the momentum field $\pi(\mathbf{x},t)$, which is canonically conjugated to the Grassmann field $\psi(\mathbf{x},t)$, turns out be also a Grassmann field as it is given by $\psi^*(\mathbf{x},t)$. Furthermore, we remark that a comparison of (4.52) with (5.45) reveals a sign change. The Legendre transformation between the Lagrange function L and the Hamilton function H reads

$$L = \int d^3x \left\{ \frac{\partial \psi^*(\mathbf{x},t)}{\partial t} \, \pi^*(\mathbf{x},t) + \frac{\partial \psi(\mathbf{x},t)}{\partial t} \, \pi(\mathbf{x},t) \right\} - H\Big[\pi(\bullet,t); \psi(\bullet,t) \Big] \,. \tag{5.46}$$

Note that here the order of the Grassmann fields $\partial\psi(\mathbf{x},t)/\partial t$ and $\pi(\mathbf{x},t)$ and their complex conjugate is chosen in such a way that the Legendre transformation (5.46) is consistent with the definition of the canonical conjugated momentum fields in (5.44) and (5.45). Taking into account (5.30), (5.31) as well as (5.44) and (5.46), the Hamilton function turns out to be of the form

$$H = \int d^3x \, \mathcal{H} \left(\pi(\mathbf{x},t), \boldsymbol{\nabla}\pi(\mathbf{x},t); \psi(\mathbf{x},t), \boldsymbol{\nabla}\psi(\mathbf{x},t) \right) \,, \tag{5.47}$$

where the Hamilton density $\mathcal{H}$ is given by

$$\mathcal{H} = -\frac{\hbar}{2Mi} \, \boldsymbol{\nabla}\pi(\mathbf{x},t) \cdot \boldsymbol{\nabla}\psi(\mathbf{x},t) - \frac{V_1(\mathbf{x})}{i\hbar} \pi(\mathbf{x},t)\psi(\mathbf{x},t) \,. \tag{5.48}$$

Note that the fermionic Hamilton density (5.48) has the opposite sign of the bosonic Hamilton density (4.55). Furthermore, we remark that, in order to derive (5.48), we used the anti-commutativity of the Grassmann fields so that two terms proportional to $[\partial\psi(\mathbf{x},t)/\partial t]\pi(\mathbf{x},t)$

and $\pi(\mathbf{x},t)[\partial\psi(\mathbf{x},t)/\partial t]$ just cancel each other. Taking into account the relation (5.45) between $\pi(\mathbf{x},t)$ and $\psi^*(\mathbf{x},t)$ yields

$$H = \int d^3x \left\{ \frac{\hbar^2}{2M} \boldsymbol{\nabla}\psi^*(\mathbf{x},t) \cdot \boldsymbol{\nabla}\psi(\mathbf{x},t) + V_1(\mathbf{x})\psi^*(\mathbf{x},t)\psi(\mathbf{x},t) \right\}, \tag{5.49}$$

so a subsequent partial integration then converts (5.49) to the standard form

$$H = \int d^3x\, \psi^*(\mathbf{x},t) \left\{ -\frac{\hbar^2}{2M}\Delta + V_1(\mathbf{x}) \right\} \psi(\mathbf{x},t). \tag{5.50}$$

Thus, we conclude that the two sign changes in (5.45) and (5.48) in comparison to the bosonic case compensate each other and the Hamilton function of anti-commuting Schrödinger fields (5.49) finally coincides formally with the corresponding one of commuting Schrödinger fields in (4.56).

The Hamilton principle of classical field theory reads in the Hamilton formulation

$$\delta\mathcal{A}\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right] = \int dt \int d^3x \left\{ \delta\pi(\mathbf{x},t)\, \frac{\delta\mathcal{A}}{\delta\pi(\mathbf{x},t)} + \delta\psi(\mathbf{x},t)\, \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} \right\} = 0. \tag{5.51}$$

As the variations of the Grassmann fields $\delta\pi(\mathbf{x},t)$ and $\delta\psi(\mathbf{x},t)$ can be arbitrary, we obtain

$$\frac{\delta\mathcal{A}}{\delta\pi(\mathbf{x},t)} = 0, \qquad \frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} = 0. \tag{5.52}$$

Due to (5.29) and (5.46) the action $\mathcal{A}$ depends on the Hamilton function H as follows:

$$\mathcal{A} = \int dt \int d^3x\, \frac{\partial\psi(\mathbf{x},t)}{\partial t}\, \pi(\mathbf{x},t) - \int dt\, H\left[\pi(\bullet,t);\psi(\bullet,t)\right]. \tag{5.53}$$

Performing the functional derivatives (5.52) of the action (5.53) then leads to the Hamilton equations of the anti-commuting Schrödinger fields:

$$\frac{\delta\mathcal{A}}{\delta\pi(\mathbf{x},t)} = -\frac{\partial\psi(\mathbf{x},t)}{\partial t} - \frac{\delta H}{\delta\pi(\mathbf{x},t)} = 0, \tag{5.54}$$

$$\frac{\delta\mathcal{A}}{\delta\psi(\mathbf{x},t)} = -\frac{\partial\pi(\mathbf{x},t)}{\partial t} - \frac{\delta H}{\delta\psi(\mathbf{x},t)} = 0. \tag{5.55}$$

Note that the first term of the Hamilton equation equation (5.54) has an opposite sign in comparison with the corresponding bosonic case in (4.61). As the Hamilton function H is of the form (5.47), the respective functional derivatives in (5.54) and (5.55) yield

$$\frac{\delta H}{\delta\pi(\mathbf{x},t)} = \frac{\partial\mathcal{H}}{\partial\pi(\mathbf{x},t)} - \boldsymbol{\nabla}\, \frac{\partial\mathcal{H}}{\partial\boldsymbol{\nabla}\pi(\mathbf{x},t)}, \tag{5.56}$$

$$\frac{\delta H}{\delta\psi(\mathbf{x},t)} = \frac{\partial\mathcal{H}}{\partial\psi(\mathbf{x},t)} - \boldsymbol{\nabla}\, \frac{\partial\mathcal{H}}{\partial\boldsymbol{\nabla}\psi(\mathbf{x},t)}, \tag{5.57}$$

which formally agree with the corresponding formulas of the bosonic case (4.63) and (4.64). Thus, inserting (5.56), (5.57) into (5.54), (5.55) the Hamilton equations of the Grassmann field theory have the form

$$\frac{\partial \psi(\mathbf{x},t)}{\partial t} = -\frac{\partial \mathcal{H}}{\partial \pi(\mathbf{x},t)} + \boldsymbol{\nabla} \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \pi(\mathbf{x},t)}\,, \tag{5.58}$$

$$\frac{\partial \pi(\mathbf{x},t)}{\partial t} = -\frac{\partial \mathcal{H}}{\partial \psi(\mathbf{x},t)} + \boldsymbol{\nabla} \frac{\partial \mathcal{H}}{\partial \boldsymbol{\nabla} \psi(\mathbf{x},t)}\,. \tag{5.59}$$

Due to the Hamilton density of the Schrödinger theory (5.47) the respective partial derivatives read

$$\frac{\partial \mathcal{H}}{\partial \pi(\mathbf{x},t)} = -\frac{V_1(\mathbf{x})}{i\hbar}\psi(\mathbf{x},t)\,, \qquad \frac{\partial \mathcal{H}}{\boldsymbol{\nabla} \pi(\mathbf{x},t)} = -\frac{\hbar}{2Mi}\,\boldsymbol{\nabla}\psi(\mathbf{x},t)\,, \tag{5.60}$$

$$\frac{\partial \mathcal{H}}{\partial \psi(\mathbf{x},t)} = \frac{V_1(\mathbf{x})}{i\hbar}\pi(\mathbf{x},t)\,, \qquad \frac{\partial \mathcal{H}}{\boldsymbol{\nabla} \psi(\mathbf{x},t)} = \frac{\hbar}{2Mi}\,\boldsymbol{\nabla}\pi(\mathbf{x},t)\,. \tag{5.61}$$

Thus, we recover from (5.58)–(5.61) due to (5.45) the Schrödinger equations for the Grassmann fields (5.42) and (5.43), which formally agree with the Schrödinger equations of the bosonic case (4.1) and (4.2).

5.4 Poisson Brackets

Also in the classical field theory of anti-commuting Schrödinger fields one can introduce Poisson brackets. For two Grassmann functionals $F\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right]$ and $G\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right]$ their Poisson bracket is defined as

$$\left\{F,G\right\}_+ = (-1)^{\pi(F)} \int d^3x \left(\frac{\delta F}{\delta\psi(\mathbf{x},t)}\,\frac{\delta G}{\delta\pi(\mathbf{x},t)} + \frac{\delta F}{\delta\pi(\mathbf{x},t)}\,\frac{\delta G}{\delta\psi(\mathbf{x},t)} \right)\,, \tag{5.62}$$

where $\pi(F)$ denotes the parity of the Grassmann functional F. For instance, the anti-commuting Schrödinger fields $\psi(\mathbf{x},t)$ and $\pi(\mathbf{x},t)$ have an odd parity $\pi = 1$, whereas the Lagrange function (5.30), (5.31) or the Hamilton function (5.47), (5.48) have an even parity $\pi = 0$. Now we investigate the symmetry of the Poisson bracket (5.62), which leads to three cases:

1. case: $\pi(F) = \pi(G) = 0$

$$\begin{aligned} \left\{F,G\right\}_+ &= \int d^3x \left(\frac{\delta F}{\delta\psi(\mathbf{x},t)}\,\frac{\delta G}{\delta\pi(\mathbf{x},t)} + \frac{\delta F}{\delta\pi(\mathbf{x},t)}\,\frac{\delta G}{\delta\psi(\mathbf{x},t)} \right) \\ &= -\int d^3x \left(\frac{\delta G}{\delta\pi(\mathbf{x},t)}\,\frac{\delta F}{\delta\psi(\mathbf{x},t)} + \frac{\delta G}{\delta\psi(\mathbf{x},t)}\,\frac{\delta F}{\delta\pi(\mathbf{x},t)} \right) = -\left\{G,F\right\}_+ \end{aligned} \tag{5.63}$$

2. case: $\pi(F) = 0$, $\pi(G) = 1$

$$\begin{aligned} \left\{F,G\right\}_+ &= \int d^3x \left(\frac{\delta F}{\delta\psi(\mathbf{x},t)}\,\frac{\delta G}{\delta\pi(\mathbf{x},t)} + \frac{\delta F}{\delta\pi(\mathbf{x},t)}\,\frac{\delta G}{\delta\psi(\mathbf{x},t)} \right) \\ &= \int d^3x \left(\frac{\delta G}{\delta\pi(\mathbf{x},t)}\,\frac{\delta F}{\delta\psi(\mathbf{x},t)} + \frac{\delta G}{\delta\psi(\mathbf{x},t)}\,\frac{\delta F}{\delta\pi(\mathbf{x},t)} \right) = -\left\{G,F\right\}_+ \end{aligned} \tag{5.64}$$

   Note that the case $\pi(F) = 1$, $\pi(G) = 0$ follows from reading (5.64) in the opposite direction and exchanging F and G.

3. case: $\pi(F) = \pi(G) = 1$

$$\begin{aligned} \left\{F, G\right\}_+ &= -\int d^3x \left(\frac{\delta F}{\delta\psi(\mathbf{x},t)} \frac{\delta G}{\delta\pi(\mathbf{x},t)} + \frac{\delta F}{\delta\pi(\mathbf{x},t)} \frac{\delta G}{\delta\psi(\mathbf{x},t)} \right) \\ &= -\int d^3x \left(\frac{\delta G}{\delta\pi(\mathbf{x},t)} \frac{\delta F}{\delta\psi(\mathbf{x},t)} + \frac{\delta G}{\delta\psi(\mathbf{x},t)} \frac{\delta F}{\delta\pi(\mathbf{x},t)} \right) = \left\{G, F\right\}_+ \end{aligned} \tag{5.65}$$

Thus, the Poisson bracket is symmetric for two odd Grassmann fields, otherwise it is anti-symmetric provided that the involved functionals do have a specific parity.

With the help of the Poisson bracket (5.62), the Hamilton equations (5.54), (5.55) for Grassmann fields read

$$\left\{\psi(\mathbf{x},t), H\right\}_+ = -\frac{\delta H}{\delta\pi(\mathbf{x},t)} = \frac{\partial\psi(\mathbf{x},t)}{\partial t}\,, \tag{5.66}$$

$$\left\{\pi(\mathbf{x},t), H\right\}_+ = -\frac{\delta H}{\delta\psi(\mathbf{x},t)} = \frac{\partial\pi(\mathbf{x},t)}{\partial t}\,. \tag{5.67}$$

Thus, the Hamilton equations of the fermionic and bosonic case (5.66), (5.67) and (4.70), (4.71), respectively, have the same general structure and only differ by the used Poisson bracket. Also the temporal change of a Grassmann functional $F\left[\pi(\bullet,\bullet);\psi(\bullet,\bullet)\right]$ can be formulated with the help of a Poisson bracket. At first we obtain with the chain rule of functional differentiation:

$$\frac{\partial F}{\partial t} = \int d^3x \left(\frac{\partial\pi(\mathbf{x},t)}{\partial t} \frac{\delta F}{\delta\pi(\mathbf{x},t)} + \frac{\partial\psi(\mathbf{x},t)}{\partial t} \frac{\delta F}{\delta\psi(\mathbf{x},t)} \right)\,, \tag{5.68}$$

which reduces due to the Hamilton equations (5.54), (5.55) to

$$\begin{aligned} \frac{\partial F}{\partial t} &= -\int d^3x \left(\frac{\delta H}{\delta\psi(\mathbf{x},t)} \frac{\delta F}{\delta\pi(\mathbf{x},t)} + \frac{\delta H}{\delta\pi(\mathbf{x},t)} \frac{\delta F}{\delta\psi(\mathbf{x},t)} \right) \\ &= (-1)^{\pi(F)} \int d^3x \left(\frac{\delta F}{\delta\psi(\mathbf{x},t)} \frac{\delta H}{\delta\pi(\mathbf{x},t)} + \frac{\delta F}{\delta\pi(\mathbf{x},t)} \frac{\delta H}{\delta\psi(\mathbf{x},t)} \right) = \left\{F, H\right\}_+ . \end{aligned} \tag{5.69}$$

Thus, the formulation of the Hamilton equations in form of Poisson brackets according to (5.66), (5.67) represent a special case of (5.69). Furthermore, we obtain for the fundamental Poisson brackets:

$$\left\{\psi(\mathbf{x},t), \psi(\mathbf{x}',t)\right\}_+ = -\int d^3x'' \left(\frac{\delta\psi(\mathbf{x},t)}{\delta\psi(\mathbf{x}'',t)} \frac{\delta\psi(\mathbf{x}',t)}{\delta\pi(\mathbf{x}'',t)} - \frac{\delta\psi(\mathbf{x},t)}{\delta\pi(\mathbf{x}'',t)} \frac{\delta\psi(\mathbf{x}',t)}{\delta\psi(\mathbf{x}'',t)} \right) = 0\,. \tag{5.70}$$

$$\left\{\pi(\mathbf{x},t), \pi(\mathbf{x}',t)\right\}_+ = -\int d^3x'' \left(\frac{\delta\pi(\mathbf{x},t)}{\delta\psi(\mathbf{x}'',t)} \frac{\delta\pi(\mathbf{x}',t)}{\delta\pi(\mathbf{x}'',t)} - \frac{\delta\pi(\mathbf{x},t)}{\delta\pi(\mathbf{x}'',t)} \frac{\delta\pi(\mathbf{x}',t)}{\delta\psi(\mathbf{x}'',t)} \right) = 0\,. \tag{5.71}$$

$$\left\{\psi(\mathbf{x},t), \pi(\mathbf{x}',t)\right\}_+ = -\int d^3x'' \left(\frac{\delta\psi(\mathbf{x},t)}{\delta\psi(\mathbf{x}'',t)} \frac{\delta\pi(\mathbf{x}',t)}{\delta\pi(\mathbf{x}'',t)} - \frac{\delta\psi(\mathbf{x},t)}{\delta\pi(\mathbf{x}'',t)} \frac{\delta\pi(\mathbf{x}',t)}{\delta\psi(\mathbf{x}'',t)} \right) = -\delta(\mathbf{x}-\mathbf{x}')\,. \tag{5.72}$$

Note the additional minus sign in (5.72) in comparison with (4.76).

5.5 Canonical Field Quantization

No we implement the canonical field quantization for fermions in the Heisenberg picture by going over from the anti-commuting Schrödinger fields $\psi(\mathbf{x},t)$ and $\pi(\mathbf{x},t)$ to corresponding second quantized field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\pi}(\mathbf{x},t)$. Here the question arises whether the Poisson bracket (5.62) corresponds to a commutator or to an anti-commutator. We show now that this depends upon which parity the Grassmann functions F and G have.

In case of the fundamental Poisson brackets (5.70)–(5.72) we observe that they are all symmetric. Therefore we postulate in case of symmetric Poisson brackets (5.65) a transition to anti-commutators, which are also symmetric:

$$\pi(F)=\pi(G)=1: \qquad \Big\{F,G\Big\}_+ \qquad \Longrightarrow \qquad \frac{1}{i\hbar}\Big[\hat{F},\hat{G}\Big]_+ . \tag{5.73}$$

In this way, the fundamental Poisson brackets (5.70)–(5.72) go over into equal-time anti-commutation relations

$$\Big[\hat{\psi}(\mathbf{x},t),\hat{\psi}(\mathbf{x}',t)\Big]_- = \Big[\hat{\pi}(\mathbf{x},t),\hat{\pi}(\mathbf{x}',t)\Big]_- = 0\,, \quad \Big[\hat{\psi}(\mathbf{x},t),\hat{\pi}(\mathbf{x}',t)\Big]_- = -i\hbar\,\delta(\mathbf{x}-\mathbf{x}')\,. \tag{5.74}$$

As (5.45) implies that the momentum field operator $\hat{\pi}(\mathbf{x},t)$ is given by the adjoint field operator $\hat{\psi}^\dagger(\mathbf{x},t)$ via

$$\hat{\pi}(\mathbf{x},t) = -i\hbar\,\hat{\psi}^\dagger(\mathbf{x},t)\,, \tag{5.75}$$

we recognize that the previous equal-time anti-commutation relations (3.87) between the field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$ follow from (5.74).

Afterwards, we consider the Hamilton equations (5.66), (5.67), where the involved Poisson brackets are anti-symmetric. Therefore we postulate in case of (5.64))that the Poisson brackets go over to commutators, which are also anti-symmetric:

$$\pi(F)=1,\ \pi(G)=0: \qquad \Big\{F,G\Big\}_+ \qquad \Longrightarrow \qquad \frac{1}{i\hbar}\Big[\hat{F},\hat{G}\Big]_- . \tag{5.76}$$

Then we obtain from the Hamilton equations (5.66), (5.67) the corresponding Heisenberg equations

$$i\hbar\,\frac{\partial\hat{\psi}(\mathbf{x},t)}{\partial t} = \Big[\hat{\psi}(\mathbf{x},t),\hat{H}\Big]_-\,, \tag{5.77}$$

$$i\hbar\,\frac{\partial\hat{\pi}(\mathbf{x},t)}{\partial t} = \Big[\hat{\pi}(\mathbf{x},t),\hat{H}\Big]_-\,. \tag{5.78}$$

Due to (5.75) they turn out to agree with the Heisenberg equations of motion of the fields operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$ in (3.68) and (3.73). And the Hamilton function (5.50) is converted within the canonical field quantization to the Hamilton operator (3.67) in the Heisenberg picture without the 2-particle interaction.

Part II:

Free Relativistic Fields and Their Quantization

Chapter 6

Poincaré Group

According to special relativity the space-time in the absence of gravity has a flat Minkowskian structure. The group of symmetries, which leaves distances between events in this Minkowskian space-time invariant, is named after the mathematician Henri Poincaré as the Poincaré group. In the following we work out its properties as a Lie group, which unifies mathematical structures of a group and a manifold as its group elements depend continuously and differentiably on certain parameters. In fact, the Poincaré group turns out to be a ten-parametric, non-abelian Lie group, which contains rotations in space, boosts between inertial systems, and translations in space-time. Thus, the elements of the Poincaré group depend continuously and differentiably on the rotation angles, the boost velocities and the translations. Furthermore, we discuss the Poincaré algebra, which amounts to restricting the Poincaré group to the tangent plane at the identity element, yielding the generators of rotations, boosts, and translations. And, conversely, the Lie theorem turns out to allow to reconstruct the full Poincaré group by evaluating an exponential function involving both the generators, i.e. the elements of the Lie algebra, and the group parameters. And, finally, we determine the Casimir operators of the Poincaré group, i.e. those operators commuting with all elements of the Poincaré algebra. Their eigenvalues turn out to characterize all irreducible representations of the Poincaré group, to one of which each elementary particle of the standard model has to belong. In this way, the Poincaré group characterizes the underlying symmetry of relativistic quantum field theory and, thus, represents its very backbone.

6.1 Special Relativity

Albert Einstein formulated the special relativity in 1905, which has changed since then the basic concept of space and time in the absence of gravity. It is based on two basic postulates:

1. Postulate: The velocity of light is the same in all inertial systems.

2. Postulate: The fundamental laws of physics have the same form in all inertial systems.

On the one hand, this implies concrete physical consequences for fast moving particles, which are nowadays confirmed, for instance, in the Large Hadron Collider (LHC) at Cern on a daily basis. A prominent example is provided by the time dilatation, i.e. for an observer in an inertial frame of reference, a clock that is moving relative to it in another inertial frame of reference will be measured to tick slower than a clock that is at rest in its frame of reference. On the other hand, special relativity also unifies the fundamental description of space and time. In view of formalising the second postulate, a point in space-time, which is also called the Minskowski space, is characterized by the contravariant space-time four-vector

$$(x^\mu) = \left(x^0, x^1, x^2, x^3\right) = \left(ct, x^i\right) = (ct, \mathbf{x}) \ . \tag{6.1}$$

Here we use the convention that Greek (Latin) indices run from 0 to 3 (from 1 to 3). Furthermore, from the first postulate follows for a light ray in two different inertial systems:

$$(ct)^2 - \mathbf{x}^2 = (ct')^2 - \mathbf{x}'^2 \,. \tag{6.2}$$

This condition can be reformulated with the help of the covariant Minkowski metric

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.3}$$

as the invariance of the scalar product of the space-time four-vectors x^μ and x'^μ in the respective inertial systems:

$$g_{\mu\nu}\, x^\mu x^\nu = g_{\mu\nu}\, x'^\mu x'^\nu \ . \tag{6.4}$$

Note that we use here the Einstein summation convention that one has to sum over all indices, which appear twice, i.e. once in form of an upper or contravariant index and once in form of a lower or covariant index. Apart from the contravariant space-time four-vector (6.1) we also introduce the covariant space-time four-vector

$$x_\mu = g_{\mu\nu}\, x^\nu \ . \tag{6.5}$$

Thus, the contravariant space-time four-vector x^ν is transformed via contraction with the covariant metric $g_{\mu\nu}$ to the corresponding covariant space-time four-vector x_μ. Inserting (6.1) and (6.3) in (6.5) the respective components of the covariant space-time four-vector turn out to be

$$(x_\mu) = (x_0, x_1, x_2, x_3) = \left(ct, -x^i\right) = (ct, -\mathbf{x}) \ . \tag{6.6}$$

With this the invariance of the scalar product (6.4) reduces to

$$x^\mu x_\mu = x'^\mu x'_\mu \,. \tag{6.7}$$

Furthermore, the obvious identity

$$g_{\mu\nu}\,\delta^{\nu}{}_{\kappa} = g_{\mu\kappa} \tag{6.8}$$

with the Kronecker symbol $\delta^{\nu}{}_{\kappa}$ means that the latter can be identified with the Minkowski metric $g^{\nu}{}_{\kappa}$, which consists of both the contravariant index ν and the covariant index κ:

$$(g^{\nu}{}_{\kappa}) = (\delta^{\nu}{}_{\kappa}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{6.9}$$

In addition we also define

$$(g^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \tag{6.10}$$

for which we read off with (6.3) and (6.9) the obvious identity

$$g^{\mu\nu}\,g_{\nu\kappa} = \delta^{\mu}{}_{\kappa} = g^{\mu}{}_{\kappa} . \tag{6.11}$$

Thus, (6.10) represents the contravariant Minkoswki metric. Due to (6.5) and (6.11) the covariant space-time four-vector x_ν is transformed via contraction with the contravariant metric $g^{\mu\nu}$ to the corresponding contravariant space-time four-vector x^μ:

$$g^{\mu\nu}\,x_{\nu} = g^{\mu\nu}\,g_{\nu\kappa}\,x^{\kappa} = \delta^{\mu}{}_{\kappa}\,x^{\kappa} = x^{\mu} . \tag{6.12}$$

Therefore, we can summarize that the co- and contravariant Minkowski metrices allow to pull down and up indices according to (6.5) and (6.12).

But the concept of four-vectors is much more general than the mere description of space-time four-vectors. Namely, a four-vector represents objects whose scalar products coincide in all inertial systems. Let us consider in view of another example the seminal energy-momentum dispersion relation of a relativistic particle, see Fig. 6.1, in two different inertial systems:

$$E^2 = M^2 c^4 + \mathbf{p}^2 c^2 , \qquad E'^2 = M'^2 c'^4 + \mathbf{p}'^2 c^2 . \tag{6.13}$$

Due to the equality of the rest masses M and M' in both inertial systems

$$M = M' \tag{6.14}$$

the energy-momentum dispersion relations (6.13) reduce to the identity

$$\left(\frac{E}{c}\right)^2 - \mathbf{p}^2 = \left(\frac{E'}{c}\right)^2 - \mathbf{p}'^2 . \tag{6.15}$$

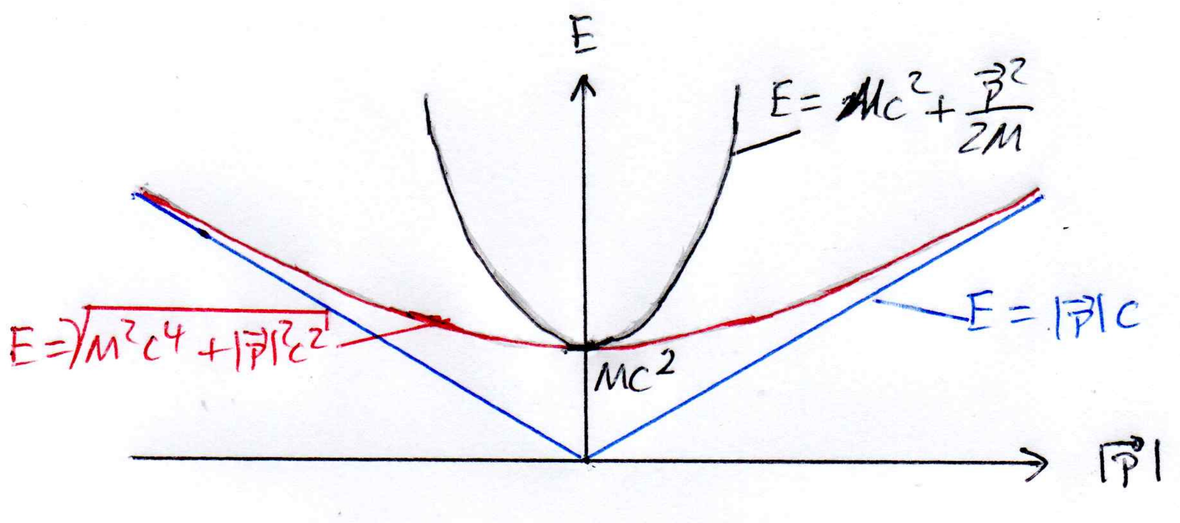


Figure 6.1: A relativistic energy-momentum dispersion (6.13) for massive (red) and massless (blue) particles in comparison with the non-relativistic limit (black).

Thus, introducing the contravariant momentum four-vector

$$(p^\mu) = \left(p^0, p^1, p^2, p^3\right) = \left(\frac{E}{c}, p^i\right) = \left(\frac{E}{c}, \mathbf{p}\right) \tag{6.16}$$

allows to formulate the identity (6.15) as the invariance of the scalar products of the contravariant momentum four-vectors p^μ and p'^μ:

$$g_{\mu\nu}\, p^\mu p^\nu = g_{\mu\nu}\, p'^\mu p'^\nu \,. \tag{6.17}$$

Defining in analogy to (6.5)

$$p_\mu = g_{\mu\nu}\, p^\nu \tag{6.18}$$

also the components of the covariant momentum four-vector

$$(p_\mu) = (p_0, p_1, p_2, p_3) = \left(\frac{E}{c}, -p^i\right) = \left(\frac{E}{c}, -\mathbf{p}\right) \,. \tag{6.19}$$

the invariance of the scalar product (6.17) can also be formulated as

$$p^\mu p_\mu = p'^\mu p'_\mu \,. \tag{6.20}$$

Furthermore, we conclude from (6.13), (6.14), and (6.19) that the scalar product of the four-momentum vector with itself is given by the rest mass M of the particle:

$$p^\mu p_\mu = M^2 c^2 \,. \tag{6.21}$$

6.2 Defining Representation of Lorentz Group

Now we study the consequences of the invariance of the scalar product of four-vectors with respect to a change from one inertial system to another. To this end we consider that the two inertial systems are connected via a linear coordinate transformation, which is mediated by a 4×4 matrix $\Lambda^\mu{}_\nu$:

$$x'^\mu = \Lambda^\mu{}_\nu \, x^\nu \,. \tag{6.22}$$

The invariance (6.4) then reads explicitly

$$g_{\mu\nu} \, x^\mu x^\nu = g_{\mu\nu} \, x'^\mu x'^\nu = g_{\mu\nu} \, \Lambda^\mu{}_\sigma \Lambda^\nu{}_\rho \, x^\sigma x^\rho = g_{\sigma\rho} \, \Lambda^\sigma{}_\mu \Lambda^\rho{}_\nu \, x^\mu x^\nu \,. \tag{6.23}$$

As (6.23) holds for arbitrary components x^μ of a space-time four-vector, we conclude the identity

$$g_{\mu\nu} = \Lambda^\sigma{}_\mu \, g_{\sigma\rho} \, \Lambda^\rho{}_\nu \,. \tag{6.24}$$

This represents the defining relation for Lorentz transformations Λ, which can be interpreted in two different ways. At first we write (6.24) in matrix notation

$$g = \Lambda^T \, g \, \Lambda \,, \tag{6.25}$$

where we have introduced the elements of the transposed matrix Λ^T according to

$$\left(\Lambda^T\right)_\mu{}^\sigma = g_{\mu\kappa} \left(\Lambda^T\right)^{\kappa\sigma} = g_{\mu\kappa} \Lambda^{\sigma\kappa} = \Lambda^{\sigma\kappa} g_{\kappa\mu} = \Lambda^\sigma{}_\mu \,. \tag{6.26}$$

Note that a left (right) index denotes the respective row (column) of the matrix, so we have concretely

$$\begin{pmatrix} \Lambda^T{}_0{}^0 & \Lambda^T{}_0{}^1 & \Lambda^T{}_0{}^2 & \Lambda^T{}_0{}^3 \\ \Lambda^T{}_1{}^0 & \Lambda^T{}_1{}^1 & \Lambda^T{}_1{}^2 & \Lambda^T{}_1{}^3 \\ \Lambda^T{}_2{}^0 & \Lambda^T{}_2{}^1 & \Lambda^T{}_2{}^2 & \Lambda^T{}_2{}^3 \\ \Lambda^T{}_3{}^0 & \Lambda^T{}_3{}^1 & \Lambda^T{}_3{}^2 & \Lambda^T{}_3{}^3 \end{pmatrix} = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^1{}_0 & \Lambda^2{}_0 & \Lambda^3{}_0 \\ \Lambda^0{}_1 & \Lambda^1{}_1 & \Lambda^2{}_1 & \Lambda^3{}_1 \\ \Lambda^0{}_2 & \Lambda^1{}_2 & \Lambda^2{}_2 & \Lambda^3{}_2 \\ \Lambda^0{}_3 & \Lambda^1{}_3 & \Lambda^2{}_3 & \Lambda^3{}_3 \end{pmatrix} . \tag{6.27}$$

The set $\mathcal{L}$ of all 4×4 matrices Λ, which transform the Minkowski matrix g according to (6.25) into the Minkowski metric g, defines the so-called Lorentz transformations. Note that another equivalent way to interpret the invariance (6.24) follows from contracting it with $g^{\nu\kappa}$. Taking into account (6.11) and (6.26) yields

$$\delta^\kappa{}_\mu = \delta_\mu{}^\kappa = \left(\Lambda^T\right)^\kappa{}_\sigma \, \Lambda^\sigma{}_\mu = \left(\Lambda^T \Lambda\right)^\kappa{}_\mu \,. \tag{6.28}$$

Thus, we conclude that Lorentz transformations Λ are also defined by the identity

$$\Lambda^T = \Lambda^{-1} \qquad \Longleftrightarrow \qquad \left(\Lambda^T\right)^\mu{}_\nu = \Lambda_\nu{}^\mu = \left(\Lambda^{-1}\right)^\mu{}_\nu \,. \tag{6.29}$$

By inspection we find that the set $\mathcal{L}$ fulfills all group axioms:

- At first we show that the closedness axiom is valid. Provided that Λ_1, Λ_2 belong to $\mathcal{L}$ we obtain from (6.25) that also $\Lambda_1\Lambda_2$ belongs to $\mathcal{L}$:

$$(\Lambda_1\Lambda_2)^T \, g \, (\Lambda_1\Lambda_2) = \Lambda_2^T \left(\Lambda_1^T \, g \, \Lambda_1\right) \Lambda_2 = \Lambda_2^T \, g \, \Lambda_2 = g \, . \tag{6.30}$$

- Then we take advantage of the associativity of matrix multiplication. For Λ_1, Λ_2, and Λ_3 belonging to $\mathcal{L}$ we conclude $(\Lambda_1\Lambda_2)\,\Lambda_3 = \Lambda_1\,(\Lambda_2\Lambda_3)$ from (6.25):

$$[(\Lambda_1\Lambda_2)\,\Lambda_3]^T \, g \, [(\Lambda_1\Lambda_2)\,\Lambda_3] = \Lambda_3^T \left[\Lambda_2^T \left(\Lambda_1^T \, g \, \Lambda_1\right) \Lambda_2\right] \Lambda_3 = g = [\Lambda_1\,(\Lambda_2\Lambda_3)]^T \, g \, [\Lambda_1\,(\Lambda_2\Lambda_3)] \, . \tag{6.31}$$

- The identity element is represented by the Kronecker symbol from (6.9):

$$\Lambda_e = I = \left(g^\nu{}_\kappa\right) \, . \tag{6.32}$$

  On the one hand we conclude that Λ_e belongs to $\mathcal{L}$ because of the identity

$$g = \Lambda_e^T \, g \, \Lambda_e \, . \tag{6.33}$$

  On the other hand we observe for any Λ belonging to $\mathcal{L}$:

$$\Lambda_e\Lambda = \Lambda\Lambda_e = \Lambda \, . \tag{6.34}$$

- And, finally, for each Λ from $\mathcal{L}$ we obtain for its determinant from (6.25):

$$\mathrm{Det}\, g = \mathrm{Det}\,\Lambda^T \cdot \mathrm{Det}\, g \cdot \mathrm{Det}\,\Lambda \qquad \Longrightarrow \qquad (\mathrm{Det}\,\Lambda)^2 = 1 \, . \tag{6.35}$$

  We conclude then that Λ from $\mathcal{L}$ has a non-vanishing determinant, i.e. $\mathrm{Det}\,\Lambda \neq 0$, so there exists an inverse transformation Λ^{-1}. Furthermore, from (6.25) we yield:

$$\left(\Lambda^T\right)^{-1} g \, \Lambda^{-1} = g \qquad \Longrightarrow \qquad \left(\Lambda^{-1}\right)^T g \, \Lambda^{-1} = g \, . \tag{6.36}$$

  Thus, there exists an inverse Λ^{-1} from $\mathcal{L}$.

One denotes the set $\mathcal{L}$ of all Lorentz transformations as the Lorentz group or, more concretely, as the pseudo-orthogonal group $O(1,3)$ due to the concrete form of the covariant Minkowski metric (6.3). The Lorentz group $\mathcal{L}$ can be classified with respect to the following two properties:

- Due to (6.35) we read off that $\mathrm{Det}\,\Lambda = \pm 1$. A Lorentz transformation with $\mathrm{Det}\,\Lambda = +1$ ($\mathrm{Det}\,\Lambda = -1$) is denoted to be special (non-special).

- From (6.24) we conclude for $\mu = \nu = 0$ due to (6.3):

$$1 = g_{00} = \Lambda^\sigma{}_0 \, g_{\sigma\rho} \, \Lambda^\rho{}_0 = \left(\Lambda^0{}_0\right)^2 - \left(\Lambda^i{}_0\right)^2 \qquad \Longrightarrow \qquad \left(\Lambda^0{}_0\right)^2 = 1 + \left(\Lambda^i{}_0\right)^2 \geq 1 \, . \tag{6.37}$$

  A Lorentz transformation Λ with $\Lambda^0{}_0 \geq 1$ ($\Lambda^0{}_0 \leq -1$) is called orthochronous (non-orthochronous).

branch	Det Λ	$\Lambda^0{}_0$	example
$\mathcal{L}_1$	+1	> 0	identity: diag (1,1,1,1)
$\mathcal{L}_2$	-1	> 0	space inversion: diag (1,-1,-1,-1)
$\mathcal{L}_3$	-1	< 0	time inversion: diag (-1,1,1,1)
$\mathcal{L}_4$	+1	< 0	space-time inversion: diag (-1,-1,-1,-1)

Table 6.1: Overview of the four branches of the Lorentz group.

Thus, we conclude that the Lorentz group consists of four different branches as indicated in Tab. 6.1. As the Lorentz transformations from the different branches can not be transformed into each other, the Lorentz group is not connected. Only the branch $\mathcal{L}_1$ of the special orthochronous Lorentz transformations represents a subgroup of the Lorentz group as performing consecutively two transformations from this branch does not allow to leave this branch. Therefore, in the following we deal with only this branch $\mathcal{L}_1$ and call these special orthochronous Lorentz transformations for the sake of simplicity as the Lorentz group.

6.3 Defining Representation of Lorentz Algebra

The set of all 4×4 matrices Λ is described in total by $4 \cdot 4 = 16$ degrees of freedom, where the invariance (6.24) leads to $4 \cdot 5/2 = 10$ restrictions. Therefore the dimension of the Lorentz group is

$$16 - 10 = 6\,. \tag{6.38}$$

Here we investigate, in particular, the elements of the Lorentz group in the vicinity of the unity element (6.32). All elements of the Lorentz group, which deviate infinitesimally from the unity element, can be represented as

$$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu\,. \tag{6.39}$$

Inserting (6.39) into the defining identity for Lorentz transformations (6.24), we obtain up to first order of the deviations $\omega^\mu{}_\nu$:

$$\begin{aligned}\Lambda^\sigma{}_\mu \Lambda^\rho{}_\nu\, g_{\sigma\rho} &= \left(g^\sigma{}_\mu + \omega^\sigma{}_\mu\right)\left(g^\rho{}_\nu + \omega^\rho{}_\nu\right) g_{\sigma\rho} \approx g^\sigma{}_\mu g^\rho{}_\nu g_{\sigma\rho} + \omega^\sigma{}_\mu g^\rho{}_\nu g_{\sigma\rho} + g^\sigma{}_\mu \omega^\rho{}_\nu g_{\sigma\rho} \\ &= g^\sigma{}_\mu g_{\sigma\nu} + \omega^\sigma{}_\mu g_{\sigma\nu} + \omega^\rho{}_\nu g_{\mu\rho} = g_{\mu\nu} + \omega_{\nu\mu} + \omega_{\mu\nu} = g_{\mu\nu}\,.\end{aligned} \tag{6.40}$$

Thus we conclude that the deviations of the Lorentz transformation from the unity element are represented by anti-symmetric 4×4 matrices:

$$\omega_{\mu\nu} + \omega_{\nu\mu} = 0\,. \tag{6.41}$$

The set of all anti-symmetric 4×4 matrices are called the Lorentz algebra of the Lorentz group. The dimension of the Lorentz algebra is 6, which coincides with the dimension of the Lorentz group determined in (6.38). Using the anti-symmetry (6.41) the elements $\omega^\mu{}_\nu$ of the Lorentz algebra can be represented as

$$\omega^\mu{}_\nu = g^{\alpha\mu} g^\beta{}_\nu \omega_{\alpha\beta} = \frac{1}{2}\left(g^{\alpha\mu} g^\beta{}_\nu - g^{\beta\mu} g^\alpha{}_\nu\right)\omega_{\alpha\beta}\,. \tag{6.42}$$

Thus, all elements $\omega^\mu{}_\nu$ of the Lorentz algebra can be expanded with respect to basis elements as follows:

$$\omega^\mu{}_\nu = -\frac{i}{2}\left(L^{\alpha\beta}\right)^\mu{}_\nu\,\omega_{\alpha\beta}\,. \tag{6.43}$$

Here $\omega_{\alpha\beta}$ represent expansion coefficients and the representation matrices of the basis elements $L^{\alpha\beta}$ read:

$$\left(L^{\alpha\beta}\right)^\mu{}_\nu = i\left(g^{\alpha\mu} g^\beta{}_\nu - g^{\beta\mu} g^\alpha{}_\nu\right)\,. \tag{6.44}$$

The indices α, β characterize the respective basis elements $L^{\alpha\beta}$, whereas the indices μ, ν indicate the components $\left(L^{\alpha\beta}\right)^\mu{}_\nu$ of their respective 4×4 representation matrices. One calls (6.44) the defining representation of the Lorentz algebra as it was derived via (6.39) and (6.43) from the elements Λ of the Lorentz group acting on space-time. Its representation matrices (6.44) have obviously the properties to be anti-symmetric with respect to both pairs of indices α, β and μ, ν:

$$\left(L^{\beta\alpha}\right)^\mu{}_\nu = -\left(L^{\alpha\beta}\right)^\mu{}_\nu\,, \tag{6.45}$$

$$\left(L^{\alpha\beta}\right)_\nu{}^\mu = -\left(L^{\alpha\beta}\right)^\mu{}_\nu\,. \tag{6.46}$$

And now we determine the commutator between two basis elements $L^{\alpha\beta}$ and $L^{\gamma\delta}$. After a lengthy but straight-forward calculation, which we have relegated to the exercises, one obtains

$$\left[L^{\alpha\beta}, L^{\gamma\delta}\right]_- = i\left(g^{\alpha\delta}L^{\beta\gamma} + g^{\beta\gamma}L^{\alpha\delta} - g^{\alpha\gamma}L^{\beta\delta} - g^{\beta\delta}L^{\alpha\gamma}\right)\,. \tag{6.47}$$

This means that the Lorentz algebra is closed with respect to performing the commutator between two of its basis elements. Furthermore, the result (6.47) can be summarized according to

$$\left[L^{\alpha\beta}, L^{\gamma\delta}\right]_- = iC^{\alpha\beta\gamma\delta}_{\epsilon\xi}\,L^{\epsilon\xi}\,, \tag{6.48}$$

where the structure constants of the Lorentz algebra are given by

$$C^{\alpha\beta\gamma\delta}_{\epsilon\xi} = g^{\alpha\delta}g^\beta{}_\epsilon g^\gamma{}_\xi + g^{\beta\gamma}g^\alpha{}_\epsilon g^\delta{}_\xi - g^{\alpha\gamma}g^\beta{}_\epsilon g^\delta{}_\xi - g^{\beta\delta}g^\alpha{}_\epsilon g^\gamma{}_\xi\,. \tag{6.49}$$

6.4 Classification of Basis Elements

The basis elements $L^{\alpha\beta}$ of the Lorentz algebra can be sorted into two classes by specializing the indices α, β into spatial and spatio-temporal indices, respectively:

$$L_k = \frac{1}{2}\,\epsilon_{klm}L^{lm}\,, \tag{6.50}$$

$$M_k = L^{0k}\,. \tag{6.51}$$

Here ϵ_{klm} denotes the three-dimensional Levi-Cività tensor, which has the value $\epsilon_{123} = 1$ and is anti-symmetric with respect to two of its three indices:

$$\epsilon_{klm} = -\epsilon_{lkm} = -\epsilon_{mlk} = -\epsilon_{kml} \,. \tag{6.52}$$

According to (6.44) we obtain by taking into account (6.9) and (6.10) the following explicit representations for the basis elements (6.50):

$$\begin{aligned}
L_1 &= L^{23} = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -g^{22}g^3{}_3 \\ 0 & 0 & g^{33}g^2{}_2 & 0 \end{pmatrix} = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & 0 & -1 & 0 \end{pmatrix} , \\
L_2 &= L^{31} = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g^{11}g^3{}_3 \\ 0 & 0 & 0 & 0 \\ 0 & -g^{33}g^1{}_1 & 0 & 0 \end{pmatrix} = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \end{pmatrix} , \\
L_3 &= L^{12} = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -g^{11}g^2{}_2 & 0 \\ 0 & g^{22}g^1{}_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .
\end{aligned} \tag{6.53}$$

Correspondingly, we yield for the basis elements (6.51):

$$\begin{aligned}
M_1 &= L^{01} = i \begin{pmatrix} 0 & g^{00}g^1{}_1 & 0 & 0 \\ -g^{11}g^0{}_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = i \begin{pmatrix} 0 & +1 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \\
M_2 &= L^{02} = i \begin{pmatrix} 0 & 0 & g^{00}g^2{}_2 & 0 \\ 0 & 0 & 0 & 0 \\ -g^{22}g^0{}_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = i \begin{pmatrix} 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \\
M_3 &= L^{03} = i \begin{pmatrix} 0 & 0 & 0 & g^{00}g^3{}_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -g^{33}g^0{}_0 & 0 & 0 & 0 \end{pmatrix} = i \begin{pmatrix} 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 \end{pmatrix} .
\end{aligned} \tag{6.54}$$

Specializing the commutator (6.47) to the respective spatial and temporal indices, we obtain corresponding commutator relations for the two classes of basis elements (6.50) and (6.51). To this end, however, one has to take into account the inversion of (6.50)

$$L^{ij} = \epsilon_{ijk} L_k \,, \tag{6.55}$$

which can be proven with the help of the contraction rule of the three-dimensional Levi-Cività symbol ϵ_{ijk}:

$$\epsilon_{ijk}\epsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \,. \tag{6.56}$$

With this we yield

$$
\begin{aligned}
[L_k, L_l]_- &= i\epsilon_{klm} L_m \,, & (6.57)\\
[L_k, M_l]_- &= i\epsilon_{klm} M_m \,, & (6.58)\\
[M_k, M_l]_- &= -i\epsilon_{klm} L_m \,. & (6.59)
\end{aligned}
$$

From the commutator (6.57) we read off that the basis elements (6.50) represent a subalgebra of the Lorentz algebra.

6.5 Lie Theorem

Considering the Lorentz group in the vicinity of the unity element (6.32), we recognize from (6.39) and (6.43) that there the basis elements $L^{\alpha\beta}$ appear:

$$
\Lambda^\mu{}_\nu = g^\mu{}_\nu - \frac{i}{2} \left(L^{\alpha\beta}\right)^\mu{}_\nu \, \omega_{\alpha\beta} \,. \tag{6.60}
$$

Conversely, the Lie theorem states that the knowledge of the basis elements $L^{\alpha\beta}$ of the Lorentz algebra allows to determine each element of the Lorentz group by evaluating a matrix exponential function:

$$
\Lambda = \exp\left\{ -\frac{i}{2} \, L^{\alpha\beta} \, \omega_{\alpha\beta} \right\} \,. \tag{6.61}
$$

The statement of the Lie theorem suggests that the (basis) elements of the Lorentz algebra are called to be the (basis) generators of the Lorentz group. Corresponding to the decomposition of the basis generators $L^{\alpha\beta}$ into the two classes (6.50) and (6.51) also the expansion coefficients $\omega_{\alpha\beta}$ are decomposed into

$$
\begin{aligned}
\varphi^k &= \frac{1}{2} \, \epsilon_{klm} \, \omega_{lm} \,, & (6.62)\\
\xi^k &= \omega_{0k} \,. & (6.63)
\end{aligned}
$$

By taking into account the anti-symmetric properties (6.41) and (6.45) as well as the definitions (6.51) and (6.55) the Lie theorem (6.61) reads

$$
\Lambda = \exp\left(-\frac{i}{2} \, L^{kl} \omega_{kl} - \frac{i}{2} \, L^{0k} \omega_{0k} \right) = \exp\Big(- i\boldsymbol{\varphi}\mathbf{L} - i\boldsymbol{\xi}\mathbf{M} \Big) \,. \tag{6.64}
$$

In the following we investigate further the Lie theorem (6.64) and show that $\boldsymbol{\xi} = \mathbf{0}$ corresponds to rotations and $\boldsymbol{\varphi} = \mathbf{0}$ to boosts, respectively. Thus, $\boldsymbol{\varphi}$ ($\boldsymbol{\xi}$) denote the vector of rotation angles (rapidities) and $\mathbf{L}$ ($\mathbf{M}$) represent the generators for the rotations (boosts).

6.6 Rotations

According to the Lie theorem (6.64) a general rotation with the vector of rotation angles $\boldsymbol{\varphi}$ is defined by the matrix exponential function

$$R(\boldsymbol{\varphi}) = \exp\Big\{ - i\boldsymbol{\varphi}\mathbf{L}\Big\}\,, \tag{6.65}$$

where the explicit representation matrices for the basis generators of rotations $\mathbf{L}$ are defined in (6.53). In the exercises Eq. (6.65) is evaluated by investigating the Taylor series for the matrix exponential function and using, for instance, the Cayleigh-Hamilton theorem. To this end one uses the fact that for any matrix A its characteristic polynomial

$$f(\lambda) = \text{Det}\ (A - \lambda E) = (-1)^n \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + \text{Det}\, A \tag{6.66}$$

yields an analogous polynomial $f(A)$ by substituting the scalar variable λ by the matrix A:

$$f(A) = (-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + \text{Det}\ A\,. \tag{6.67}$$

The Cayley-Hamilton theorem then states that this polynomial expression is equal to the zero matrix, i.e. $f(A) = 0$, implying that the matrix A fulfills the property

$$(-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + \text{Det}\, A = 0\,. \tag{6.68}$$

With this one obtains from (6.65) that the representation matrix of a rotation is of the form

$$R_{00} = 1\,,\ \ R_{0j} = R_{j0} = 0\,,\ \ R_{jk}(\boldsymbol{\varphi}) = \frac{\varphi_i}{|\boldsymbol{\varphi}|}\,\epsilon_{ikj}\ \sin|\boldsymbol{\varphi}| + \frac{\varphi_j\varphi_k}{|\boldsymbol{\varphi}|^2}\ (1 - \cos|\boldsymbol{\varphi}|) + \delta_{jk}\ \cos|\boldsymbol{\varphi}|\,. \tag{6.69}$$

Note that the 4×4 matrix defined by (6.69) fulfills two properties, which are characteristic for describing a rotation along the axis $\boldsymbol{\varphi}$ with the angle $|\boldsymbol{\varphi}|$. On the one hand the rotation axis $\boldsymbol{\varphi}$ is an eigenvalue of the rotation matrix $R(\boldsymbol{\varphi})$ with eigenvalue 1:

$$R(\boldsymbol{\varphi})\begin{pmatrix} 0 \\ \boldsymbol{\varphi} \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{\varphi} \end{pmatrix}. \tag{6.70}$$

On other hand the trace of the rotation matrix $R(\boldsymbol{\varphi})$ is related to the rotation angle $|\boldsymbol{\varphi}|$ via

$$\text{Tr}\, R(\boldsymbol{\varphi}) \ \ = \ \ 2 + 2\cos|\boldsymbol{\varphi}|\,. \tag{6.71}$$

Furthermore, we note that the spatial components of a representation matrix of a rotation obey the orthonormality relation

$$R_{kl}(\boldsymbol{\varphi})R_{km}(\boldsymbol{\varphi}) = \delta_{lm}\,, \tag{6.72}$$

which follows from (6.28) and (6.29) but can also be proven by using the explicit expression (6.69).

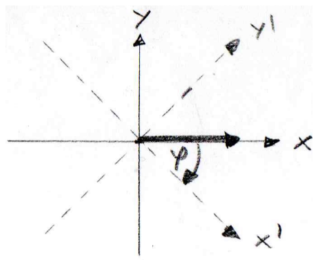
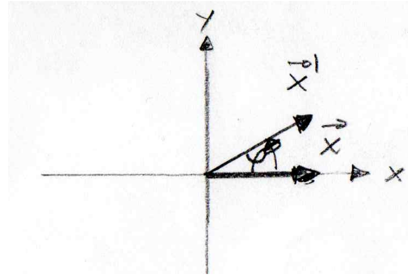
passive rotation	active rotation
vector is fixed	vector is rotated
coordinate system is rotated	coordinate system is fixed
	

Table 6.2: Passive and active rotations act in opposite directions.

Now we apply the rotation matrix (6.69) to a vector $\mathbf{x}$, which has a component parallel to the rotation axis

$$\mathbf{x}_{\parallel} = \frac{\boldsymbol{\varphi}\cdot\mathbf{x}}{|\boldsymbol{\varphi}|}\,\frac{\boldsymbol{\varphi}}{|\boldsymbol{\varphi}|} \tag{6.73}$$

and another one perpendicular to the rotation axis: $\mathbf{x}_{\perp} = \mathbf{x} - \mathbf{x}_{\parallel}$. For the rotated vector

$$x'_j = R_{jk}\, x_k \tag{6.74}$$

we then obtain the decomposition

$$\mathbf{x}' = \mathbf{x}_{\parallel} + \mathbf{x}_{\perp}\,\cos|\boldsymbol{\varphi}| + \frac{\boldsymbol{\varphi}}{|\boldsymbol{\varphi}|}\times\mathbf{x}_{\perp}\,\sin|\boldsymbol{\varphi}|\,. \tag{6.75}$$

Specializing (6.75) to a rotation around the axis $\boldsymbol{\varphi} = \varphi\,\mathbf{e}_z$ yields

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{6.76}$$

Note that a coordinate transformation like the rotation in (6.76) allows for both a passive and an active interpretation, see Tab. 6.2. For instance, the transformation

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \Longrightarrow \qquad \mathbf{x}' = \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix} \tag{6.77}$$

can be interpreted either as the description of a fixed vector under the clockwise rotation of the coordinate system or an anti-clockwise rotation of the vector for a fixed coordinate system.

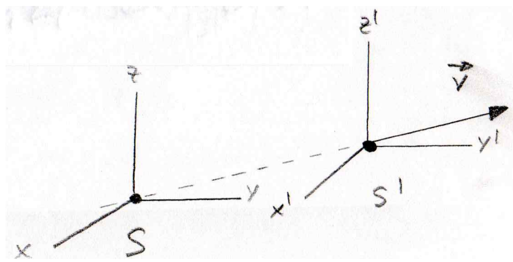


Figure 6.2: Inertial system S' moves with the velocity $\mathbf{v}$ relative to inertial system S.

6.7 Boosts

According to the Lie theorem (6.64) a general boost with the vector of rapidities $\boldsymbol{\xi}$ is defined by the matrix exponential function

$$B(\boldsymbol{\xi}) = \exp\left\{ -i\boldsymbol{\xi}\mathbf{M}\right\}, \tag{6.78}$$

where the explicit representation matrices for the basis generators of boosts $\mathbf{M}$ are defined in (6.54). In the exercises (6.78) is evaluated, yielding the representation matrix of a boost in the form

$$B(\boldsymbol{\xi}) = \begin{pmatrix} \cosh|\boldsymbol{\xi}| & \dfrac{\xi_j}{|\boldsymbol{\xi}|}\sinh|\boldsymbol{\xi}| \\ \dfrac{\xi_i}{|\boldsymbol{\xi}|}\sinh|\boldsymbol{\xi}| & \delta_{ij} + \dfrac{\xi_i\xi_j}{|\boldsymbol{\xi}|^2}\left(\cosh|\boldsymbol{\xi}| - 1\right) \end{pmatrix}. \tag{6.79}$$

We interpret the boost (6.79) passively in order to determine a relation between the rapidity $\boldsymbol{\xi}$ and the velocity $\mathbf{v}$, with which the inertial system S' is moving with respect to the inertial system S, see Fig. 6.2. To this end we observe that the coordinate origin of S' is described in both inertial systems S and S' with the following space-time four-vectors:

$$(x^\mu) = \begin{pmatrix} ct \\ \mathbf{v}t \end{pmatrix}, \qquad (x'^\mu) = \begin{pmatrix} ct' \\ \mathbf{0} \end{pmatrix}. \tag{6.80}$$

Thus, mapping the four-vector (x^μ) to (x'^μ) via the boost (6.80) according to

$$x'^\mu = B^\mu{}_\nu(\boldsymbol{\xi})\, x^\nu \tag{6.81}$$

we obtain from taking to account (6.79):

$$t' = t\cosh|\boldsymbol{\xi}| + \frac{\boldsymbol{\xi}\mathbf{v}t}{|\boldsymbol{\xi}|c}\sinh|\boldsymbol{\xi}|, \tag{6.82}$$

$$\mathbf{0} = \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\sinh|\boldsymbol{\xi}| + \frac{\mathbf{v}}{c} + \frac{\boldsymbol{\xi}\mathbf{v}}{|\boldsymbol{\xi}|c}\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\left(\cosh|\boldsymbol{\xi}| - 1\right). \tag{6.83}$$

At first, we conclude from (6.83) that rapidity $\boldsymbol{\xi}$ and velocity $\mathbf{v}$ are anti-parallel with respect to each other:

$$\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} = -\frac{\mathbf{v}}{|\mathbf{v}|}. \tag{6.84}$$

Inserting (6.84) into (6.83) we conclude how the amounts of both the rapidity vector and the velocity vector are related:

$$\frac{|\mathbf{v}|}{c} = \sinh|\boldsymbol{\xi}| - \frac{|\mathbf{v}|}{c}\left(\cosh|\boldsymbol{\xi}| - 1\right) \qquad \Longrightarrow \qquad \frac{|\mathbf{v}|}{c} = \tanh|\boldsymbol{\xi}|\,. \tag{6.85}$$

Thus, due to hyperbolic relations we obtain

$$\cosh|\boldsymbol{\xi}| = \frac{1}{\sqrt{1-\tanh^2|\boldsymbol{\xi}|}} = \gamma\,, \tag{6.86}$$

$$\sinh|\boldsymbol{\xi}| = \frac{\tanh|\boldsymbol{\xi}|}{\sqrt{1-\tanh^2|\boldsymbol{\xi}|}} = \frac{|\mathbf{v}|}{c}\,\gamma\,, \tag{6.87}$$

where we have introduced the Lorentz factor of special relativity as an abbreviation:

$$\gamma = \frac{1}{\sqrt{1-|\mathbf{v}|^2/c^2}}\,. \tag{6.88}$$

With (6.79) and (6.84)–(6.88) the representation matrix of a boost turns out to be

$$B(\mathbf{v}) = \begin{pmatrix} \gamma & -\frac{v_j}{c}\,\gamma \\ -\frac{v_i}{c}\,\gamma & \delta_{ij} + \frac{v_i v_j}{|\mathbf{v}|^2}\,(\gamma-1) \end{pmatrix}. \tag{6.89}$$

Note that the components of a representation matrix of a boost obey the relation

$$B^{\mu}{}_{\nu}(\mathbf{v})B_{\mu}{}^{\kappa}(\mathbf{v}) = \delta_{\nu}{}^{\kappa}\,, \tag{6.90}$$

which follows from (6.28) and (6.29) but can also be proven by using the explicit expression (6.79). And finally, as a concrete example, we read off from (6.82) and (6.84)–(6.88) the time dilatation

$$t' = t\,\gamma\left(1-\frac{\mathbf{v}^2}{c^2}\right) = t\sqrt{1-\frac{\mathbf{v}^2}{c^2}}\,, \tag{6.91}$$

i.e. an observer in the inertial system S detects that the clock in the moving inertial system S' goes slower than the clock in S.

6.8 Scalar Field Representation

Let us consider a scalar field $\phi(x^\mu)$, which represents a tensor field of rank $n = 0$ as it is invariant with respect to any Lorentz transformation Λ. Within a passive interpretation of the Lorentz transformation

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}\,x^{\nu} \qquad \Longleftrightarrow \qquad x^{\mu} = \left(\Lambda^{-1}\right)^{\mu}{}_{\nu}\,x'^{\nu} \tag{6.92}$$

the four-vectors x^μ and x'^μ denote one and the same space-time point at the original and the transformed coordinate system S and S', respectively. Due to the invariance of the scalar field the original scalar field $\phi(x^\mu)$ in S must coincide with the transformed scalar field $\phi'(x'^\mu)$ in S':

$$\phi'(x'^\mu) = \phi(x^\mu)\,. \tag{6.93}$$

Expressing the original scalar field ϕ via the transformed coordinate system S' we obtain from (6.92) and (6.93):

$$\phi'(x'^\mu) = \phi\left(\left(\Lambda^{-1}\right)^\mu{}_\nu\, x'^\nu\right)\,. \tag{6.94}$$

In order to simplify our notation in view of the following considerations, we omit from now on the prime $'$ at the respective four-vectors:

$$\phi'(x^\mu) = \phi\left(\left(\Lambda^{-1}\right)^\mu{}_\nu\, x^\nu\right)\,. \tag{6.95}$$

Specializing (6.95) with the help of (6.39) and (6.42) to infinitesimal Lorentz transformations, we obtain up to first order in the expansion coefficients $\omega_{\alpha\beta}$:

$$\phi'(x^\mu) = \phi\left(x^\mu + \frac{i}{2}\,\omega_{\alpha\beta}\left(L^{\alpha\beta}\right)^\mu{}_\nu\, x^\nu\right) = \left(1 - \frac{i}{2}\,\omega_{\alpha\beta}\,\hat{L}^{\alpha\beta}\right)\phi(x^\mu)\,, \tag{6.96}$$

where the differential operators $\hat{L}^{\alpha\beta}$ are given by

$$\hat{L}^{\alpha\beta} = -\left(L^{\alpha\beta}\right)^\mu{}_\nu\, x^\nu\,\partial_\mu\,. \tag{6.97}$$

Due to the representation matrices (6.44) the differential operators turn out to be of the form

$$\hat{L}^{\alpha\beta} = i\left(x^\alpha\partial^\beta - x^\beta\partial^\alpha\right)\,. \tag{6.98}$$

Taking into account the definition of the four-momentum operator in quantum mechanics

$$\hat{p}^\alpha = i\hbar\,\partial^\alpha \tag{6.99}$$

Eq. (6.98) reduces to dimensionless angular momentum operators

$$\hat{L}^{\alpha\beta} = \frac{1}{\hbar}\left(x^\alpha\hat{p}^\beta - x^\beta\hat{p}^\alpha\right)\,. \tag{6.100}$$

Note that the components of the space-time four-vector and the momentum four-vector operator fulfill

$$\left[\hat{p}^\alpha, x^\beta\right]_- = i\hbar\, g^{\alpha\gamma}\,\left[\partial_\gamma, x^\beta\right]_- = i\hbar\, g^{\alpha\beta}\,. \tag{6.101}$$

Here we have taken into account that differentiating with respect to the components of a contravariant four-vector yields the components of a covariant four-vector:

$$\partial_\alpha = \frac{\partial}{\partial x^\alpha}\,. \tag{6.102}$$

From (3.10) and (6.101) we get the following set of commutation relations:

$$\left[\hat{L}^{\alpha\beta}, x^\gamma\right]_- = -\left(L^{\alpha\beta}\right)^\gamma{}_\delta \, x^\delta \,, \tag{6.103}$$

$$\left[\hat{L}^{\alpha\beta}, \hat{p}^\gamma\right]_- = -\left(L^{\alpha\beta}\right)^\gamma{}_\delta \, \hat{p}^\delta \,. \tag{6.104}$$

Due to the commutation relations (6.103) and (6.104) one denotes the space-time four-vector x^λ and the momentum four-vector operator $\hat{p}^\lambda$ as vector operators. Correspondingly one considers an operator $\hat{O}^{\lambda_1 \dots \lambda_n}$ as a tensor operator of rank n if it transforms in each index $\lambda_1, \dots, \lambda_n$ as a vector:

$$\left[\hat{L}^{\mu\nu}, \hat{O}^{\lambda_1 \dots \lambda_n}\right]_- = -\sum_{k=1}^{n} \left(L^{\mu\nu}\right)^{\lambda_k}{}_\kappa \, \hat{O}^{\lambda_1 \dots \lambda_{k-1}\kappa\lambda_{k+1}\dots\lambda_n} \,. \tag{6.105}$$

The commutation relations (6.103) and (6.104) now allow to determine the commutation relations between the angular momentum operators (6.100) by taking into account (3.43):

$$\left[\hat{L}^{\alpha\beta}, \hat{L}^{\gamma\delta}\right]_- = i\left(g^{\alpha\delta}\,\hat{L}^{\beta\gamma} + g^{\beta\gamma}\,\hat{L}^{\alpha\delta} - g^{\alpha\gamma}\,\hat{L}^{\beta\delta} - g^{\beta\delta}\,\hat{L}^{\alpha\gamma}\right) \,. \tag{6.106}$$

Comparing (6.47) with (6.106) we conclude that also the angular momentum operators $\hat{L}^{\alpha\beta}$ fulfill the commutation relations of the Lorentz algebra. Therefore, the angular momentum operators $\hat{L}^{\alpha\beta}$ are considered as a representation of the Lorentz algebra in the Hilbert space of scalar fields. Furthermore, with the help of the representation matrices (6.43) we can rewrite (6.106) according to

$$\left[\hat{L}^{\alpha\beta}, \hat{L}^{\gamma\delta}\right]_- = -\left(L^{\alpha\beta}\right)^\gamma{}_\sigma \,\hat{L}^{\sigma\delta} - \left(L^{\alpha\beta}\right)^\delta{}_\sigma \,\hat{L}^{\gamma\sigma} \,. \tag{6.107}$$

Thus, the angular momentum operators $\hat{L}^{\alpha\beta}$ represent in the sense of (6.105) tensor operators of rank 2.

6.9 Tensor/Spinor Field Representation

Now we consider a tensor or a spinor field $\psi^\sigma(x^\mu)$, where the index σ stands for the respective tensor or spinor indices. Performing a Lorentz transformation one has to take into account that this affects both the space-time four-vector x^μ and the tensor or spinor components ψ^σ.

6.9.1 Four-Vector Example

Let us consider at first the concrete example of a four-vector $A^\sigma(x^\mu)$, which represents a tensor field of rank $n = 1$ as one Lorentz matrix Λ is involved in transforming the tensor or spinor components ψ^σ:

$$A'^\sigma(x'^\mu) = \Lambda^\sigma{}_\tau \, A^\tau(x^\mu) \,. \tag{6.108}$$

Reexpressing the space-time components of the original four-vector A^τ in S via the transformed coordinate system S' according to (6.92), one yields

$$A'^\sigma(x^\mu) = \Lambda^\sigma{}_\tau \, A^\tau \left(\left(\Lambda^{-1}\right)^\mu{}_\nu \, x^\nu \right) , \tag{6.109}$$

where again the prime $'$ at the space-time four-vector has been omitted in order to simplify the notation. Afterwards we specialize (6.109) with the help of (6.39) and (6.42) to infinitesimal Lorentz transformations and obtain up to first order in the expansion coefficients $\omega_{\alpha\beta}$:

$$\begin{aligned} A'^\sigma(x^\mu) &= \left\{ g^\sigma{}_\tau - \frac{i}{2} \left(L^{\alpha\beta}\right)^\sigma{}_\tau \, \omega_{\alpha\beta} \right\} \left\{ A^\tau(x^\mu) + \frac{i}{2} \left(L^{\alpha\beta}\right)^\mu{}_\nu \, \omega_{\alpha\beta} \, x^\nu \, \partial_\mu \, A^\tau(x^\mu) \right\} \\ &\Longrightarrow \; A'^\sigma(x^\mu) = \left\{ g^\sigma{}_\tau - \frac{i}{2} \left(\hat{M}^{\alpha\beta}\right)^\sigma{}_\tau \, \omega_{\alpha\beta} \right\} A^\tau(x^\mu) \, . \end{aligned} \tag{6.110}$$

Here the operator $\hat{M}^{\alpha\beta}$ turns out to be additive in the representation matrices (6.43) and the angular momentum operator (6.100):

$$\hat{M}^{\alpha\beta} = \hat{L}^{\alpha\beta} + L^{\alpha\beta} \, . \tag{6.111}$$

Thus, from (6.45) and (6.100) we read off the anti-symmetry

$$\hat{M}^{\alpha\beta} = -\hat{M}^{\beta\alpha} \, . \tag{6.112}$$

As both the representation matrices $L^{\alpha\beta}$ and the angular momentum operators $\hat{L}^{\alpha\beta}$ fulfill according to (6.47) and (6.106) the Lorentz algebra as well as they commute with each other

$$\left[L^{\alpha\beta}, \hat{L}^{\gamma\delta} \right]_- = 0 \, , \tag{6.113}$$

we conclude that also the operators $\hat{M}^{\alpha\beta}$ fulfill the Lorentz algebra:

$$\left[\hat{M}^{\alpha\beta}, \hat{M}^{\gamma\delta} \right]_- = i \left(g^{\alpha\delta} \, \hat{M}^{\beta\gamma} + g^{\beta\gamma} \, \hat{M}^{\alpha\delta} - g^{\alpha\gamma} \, \hat{M}^{\beta\delta} - g^{\beta\delta} \, \hat{M}^{\alpha\gamma} \right) . \tag{6.114}$$

6.9.2 General Case

Now we return back to the general case of a tensor or spinor field $\psi^\sigma(x^\mu)$. Performing an infinitesimal Lorentz transformation we have then in analogy to (6.110)

$$\psi'^\sigma(x^\mu) = \left\{ g^\sigma{}_\tau - \frac{i}{2} \left(\hat{M}^{\alpha\beta}\right)^\sigma{}_\tau \, \omega_{\alpha\beta} \right\} \psi^\tau(x^\mu) \, , \tag{6.115}$$

where the operator $\hat{M}^{\alpha\beta}$ has a decomposition similar to (6.111):

$$\hat{M}^{\alpha\beta} = \hat{L}^{\alpha\beta} + N^{\alpha\beta} \, . \tag{6.116}$$

Although we can not write down the explicit form of the matrices $N^{\alpha\beta}$ for a general tensor or spinor representation of the Lorentz algebra, we do know that they must fulfill the commutator relation

$$\left[N^{\alpha\beta}, N^{\gamma\delta} \right]_- = i \left(g^{\alpha\delta} \, N^{\beta\gamma} + g^{\beta\gamma} \, N^{\alpha\delta} - g^{\alpha\gamma} \, N^{\beta\delta} - g^{\beta\delta} \, N^{\alpha\gamma} \right) . \tag{6.117}$$

Furthermore, both representations $\hat{L}^{\alpha\beta}$ and $N^{\alpha\beta}$ of the Lorentz algebra in Minkowski space and in the space of the tensor or spinor components are independent from each other, implying

$$\left[N^{\alpha\beta}, \hat{L}^{\gamma\delta}\right]_- = 0 \,. \tag{6.118}$$

From this we then read off that also the operators $\hat{M}^{\alpha\beta}$ defined in (6.116) fulfill the commutation relation (6.114) of the Lorentz algebra. They are a representation of the Lorentz algebra in the Hilbert space of tensor or spinor fields. In addition, as $\hat{L}^{\alpha\beta}$ coincides with the orbital angular momentum (6.100), one can identify the representation $N^{\alpha\beta}$ of the Lorentz algebra in the space of the tensor or spinor components with the spin angular momentum and, thus, $\hat{M}^{\alpha\beta}$ with the total angular momentum.

6.10 Defining Representation of Poincaré Group

Poincaré transformations in Minkowski space are put together from a Lorentz transformation ${\Lambda^\mu}_\nu$ and a shift a^μ:

$$x'^\mu = {\Lambda^\mu}_\nu \, x^\nu + a^\mu \,. \tag{6.119}$$

Whereas Lorentz transformations do not change the scalar product of four-vectors due to (6.4) and (6.17), Poincaré transformations (6.119) only leave distances between four-vectors invariant:

$$g_{\mu\nu} \, (x^\mu - y^\mu) \, (x^\nu - y^\nu) = g_{\mu\nu} \, \left(x'^\mu - y'^\mu\right) \left(x'^\nu - y'^\nu\right) \,. \tag{6.120}$$

Therefore, Poincaré transformations are also called to be inhomogeneous Lorentz transformations.

We show now that the set $\mathcal{P}$ of all Poincaré transformations is a group. To this end we characterize an element from $\mathcal{P}$ with (Λ, a):

- At first we prove the closedness and assume, to this end, that both (Λ_1, a_1) and (Λ_2, a_2) belong to $\mathcal{P}$. Taking into account (6.119) we then conclude

$$\begin{aligned} x_2^\mu &= {\Lambda_2}^\mu{}_\nu \, x_1^\nu + a_2^\mu = {\Lambda_2}^\mu{}_\nu \, \left({\Lambda_1}^\nu{}_\kappa \, x^\kappa + a_1^\nu\right) + a_2^\mu = {\Lambda_2}^\mu{}_\nu \, {\Lambda_1}^\nu{}_\kappa \, x^\kappa + {\Lambda_2}^\mu{}_\nu \, a_1^\nu + a_2^\mu \\ \Longrightarrow & \qquad {\Lambda^\mu}_\nu = {\Lambda_2}^\mu{}_\nu \, {\Lambda_1}^\nu{}_\kappa \,, \qquad a^\mu = {\Lambda_2}^\mu{}_\nu \, a_1^\nu + a_2^\mu \,. \end{aligned} \tag{6.121}$$

  Thus, also

$$(\Lambda_2, a_2)(\Lambda_1, a_1) = (\Lambda, a) = (\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2) \tag{6.122}$$

  belongs to $\mathcal{P}$. One calls the multiplication rule (6.122) a semi-direct product of the Lorentz group $\mathcal{L}$ and the translation group $\mathcal{T}$. In case of a direct product one would have had the simpler multiplication rule:

$$(\Lambda_2, a_2)(\Lambda_1, a_1) = (\Lambda, a) = (\Lambda_2\Lambda_1, a_1 + a_2) \,. \tag{6.123}$$

- In the next step we consider the associativity, so we assume that (Λ_1, a_1), (Λ_2, a_2) and (Λ_3, a_3) belong to $\mathcal{P}$. Thus, we obtain from (6.122)

$$(\Lambda_1, a_1)\left((\Lambda_2, a_2)(\Lambda_3, a_3)\right) = (\Lambda_1, a_1)(\Lambda_2\Lambda_3, \Lambda_2 a_3 + a_2) = (\Lambda_1\Lambda_2\Lambda_3, \Lambda_1\Lambda_2 a_3 + \Lambda_1 a_2 + a_1) \tag{6.124}$$

$$\left((\Lambda_1, a_1)(\Lambda_2, a_2)\right)(\Lambda_3, a_3) = (\Lambda_1\Lambda_2, \Lambda_1 a_2 + a_1)(\Lambda_3, a_3) = (\Lambda_1\Lambda_2\Lambda_3, \Lambda_1\Lambda_2 a_3 + \Lambda_1 a_2 + a_1) \tag{6.125}$$

  and deduce with this the associativity

$$(\Lambda_1, a_1)\left((\Lambda_2, a_2)(\Lambda_3, a_3)\right) = \left((\Lambda_1, a_1)(\Lambda_2, a_2)\right)(\Lambda_3, a_3)\,. \tag{6.126}$$

- Then we identify the unity element of $\mathcal{P}$ with $(\Lambda_e, a_e) = (I, 0)$ due to (6.32). Namely, with (Λ, a) from $\mathcal{P}$ we read off from (6.122)

$$(I, 0)(\Lambda, a) = (\Lambda, a) = (\Lambda, a)(I, 0)\,. \tag{6.127}$$

- And, finally, the inverse element of some element (Λ, a) belonging to $\mathcal{P}$ is given by $(\Lambda, a)^{-1} = (\Lambda^{-1}, -\Lambda^{-1}a)$ from $\mathcal{P}$ as taking into account (6.122) leads to

$$(\Lambda, a)^{-1}(\Lambda, a) = (\Lambda^{-1}, -\Lambda^{-1}a)(\Lambda, a) = (\Lambda^{-1}\Lambda, \Lambda^{-1}a - \Lambda^{-1}a) = (I, 0)\,. \tag{6.128}$$

Similar to the Lorentz group also the Poincaré group is divided with the help of the values $\mathrm{Det}\,\Lambda$ and ${\Lambda^0}_0$ into the four branches $\mathcal{P}_i$ with $i = 1, 2, 3, 4$, see Tab. 6.1. In the following we restrict ourselves to consider the subgroup $\mathcal{P}_1$ of the Poincaré group $\mathcal{P}$, which is characterized by $\mathrm{Det}\,\Lambda > 0$ and ${\Lambda^0}_0 > 0$.

6.11 Tensor/Spinor Representation of Poincaré Algebra

Let us analyse a tensor or spinor field $\psi^\sigma(x^\mu)$, which is invariant with respect to a translation with an arbitrary four-vector a^μ. Within a passive interpretation of the translation

$$x'^\mu = x^\mu + a^\mu \qquad \Longleftrightarrow \qquad x^\mu = x'^\mu - a^\mu \tag{6.129}$$

both x^μ and x'^μ denote one and the same space-time point with respect to the original and the translated coordinate system S and S'. Due to the invariance of the tensor or spinor field its descriptions $\psi^\sigma(x^\mu)$ and $\psi'^\sigma(x'^\mu)$ in S and S' must coincide:

$$\psi'^\sigma(x'^\mu) = \psi^\sigma(x^\mu)\,. \tag{6.130}$$

Considering in (6.130) the original tensor or spinor field ψ^σ with respect to the transformed coordinate system S', we obtain from (6.129) and (6.130)

$$\psi'^\sigma(x^\mu) = \psi^\sigma(x^\mu - a^\mu)\,, \tag{6.131}$$

where we have omitted again the prime $'$ at the four-vectors in order to simplify the notation. For an infinitesimal translation $a^\mu = \epsilon^\mu$ we then have

$$\psi'^\sigma(x^\mu) = \psi^\sigma(x^\mu) - \epsilon^\alpha \, \partial_\alpha \, \psi^\sigma(x^\mu) \,. \tag{6.132}$$

Taking into account the momentum operator (6.99) this reduces to

$$\psi'^\sigma(x^\mu) = \left(1 + \frac{i}{\hbar} \, \epsilon_\alpha \, \hat{p}^\alpha \right) \psi^\sigma(x^\mu) \,. \tag{6.133}$$

Thus, the basis generators of the translations can be identified with the components of the momentum operator (6.99). Together with the basis generators of the Lorentz transformations, which are given by the total momentum operators (6.116), they span the Poincaré algebra. In order to characterize the Poincaré algebra completely, it remains to deduce the commutation relations between its basis generators $\hat{p}^\alpha$ and $\hat{M}^{\alpha\beta}$, which can be accomplished straight-forwardly. To this end we read off from (6.99) that the commutator between two basis generators of translations vanishes:

$$\left[\hat{p}^\alpha, \hat{p}^\beta\right]_- \;=\; 0 \,. \tag{6.134}$$

Thus, the momentum operators $\hat{p}^\alpha$ represent a commutative subalgebra of the Poincaré algebra, which implies via the Lie theorem that the translations form an abelian subgroup of the Poincaré group. Afterwards, we consider the commutator between the generators $\hat{p}^\alpha$ and $\hat{M}^{\alpha\beta}$ themselves. Here we use that the representation of the basis generators of translations (6.99) and the representation $N^{\alpha\beta}$ of the Lorentz algebra in the space of the tensor or spinor components are independent from each other, implying

$$\left[\hat{p}^\alpha, N^{\beta\gamma}\right]_- \;=\; 0 \,. \tag{6.135}$$

With this as well as (6.44), (6.104), and (6.116) we then obtain

$$\left[\hat{M}^{\alpha\beta}, \hat{p}^\gamma\right]_- \;=\; i\left(g^{\beta\gamma}\hat{p}^\alpha - g^{\alpha\gamma}\hat{p}^\beta\right) \,. \tag{6.136}$$

And we remark that the commutator relations between the total momentum (6.116) were already obtained in (6.114) and are characteristic of the Lorentz algebra. From (6.116) we read off due to the Lie theorem that the Lorentz group is a non-abelian subgroup of the Poincaré group.

Finally, the definition (6.105) of a tensor operators $\hat{O}^{\lambda_1,\ldots,\lambda_n}$ of rank n for the Lorentz algebra is straight-forwardly extended to the Poincaré algebra according to

$$\left[\hat{M}^{\mu\nu}, \hat{O}^{\lambda_1\ldots\lambda_n}\right]_- = -\sum_{k=1}^{n} \left(L^{\mu\nu}\right)^{\lambda_k}{}_{\kappa} \; \hat{O}^{\lambda_1\ldots\lambda_{k-1}\kappa\lambda_{k+1}\ldots\lambda_n} \,. \tag{6.137}$$

With the help of the representation matrices (6.44) the commutator relations (6.116) and (6.136) can then be rewritten as

$$\left[\hat{M}^{\alpha\beta}, \hat{p}^\gamma\right]_- \;=\; -\left(L^{\alpha\beta}\right)^\gamma{}_\delta \; \hat{p}^\delta \,, \tag{6.138}$$

$$\left[\hat{M}^{\alpha\beta}, \hat{M}^{\gamma\delta}\right]_- \;=\; -\left(L^{\alpha\beta}\right)^\gamma{}_\sigma \; \hat{M}^{\sigma\delta} - \left(L^{\alpha\beta}\right)^\delta{}_\sigma \; \hat{M}^{\gamma\sigma} \,. \tag{6.139}$$

Thus, according to (6.137), $\hat{p}^\alpha$ and $\hat{M}^{\alpha\beta}$ represent tensor operators of rank $n = 1$ and $n = 2$, respectively.

6.12 Casimir Operators of Poincaré Algebra

Those operators, which commute with all basis generators of a Lie algebra, are called Casimir operators. The first Casimir operator of the Poincaré algebra is given by the scalar product of the momentum operator with itself:

$$\hat{p}^2 = g_{\alpha\beta}\, \hat{p}^\alpha \hat{p}^\beta \, . \tag{6.140}$$

Taking into account (3.10) and (6.134) one can directly show that $\hat{p}^2$ commutes with all momentum operators:

$$\left[\hat{p}^2, \hat{p}^\alpha\right]_- = g_{\beta\gamma}\, \left[\hat{p}^\beta \hat{p}^\gamma, \hat{p}^\alpha\right]_- = g_{\beta\gamma}\, \left\{\hat{p}^\beta \left[\hat{p}^\gamma, \hat{p}^\alpha\right]_- + \left[\hat{p}^\beta, \hat{p}^\alpha\right]_- \hat{p}^\gamma\right\} = 0\, . \tag{6.141}$$

Furthermore, $\hat{p}^2$ is per construction a Lorentz scalar and, thus, commutes with all generators of the Lorentz algebra $\hat{M}^{\alpha\beta}$ due to (3.10), (6.136), and (6.140):

$$\begin{aligned} \left[\hat{p}^2, \hat{M}^{\alpha\beta}\right]_- &= g_{\gamma\delta}\, \left[\hat{p}^\gamma \hat{p}^\delta, \hat{M}^{\alpha\beta}\right]_- = g_{\gamma\delta}\, \left\{\hat{p}^\gamma \left[\hat{p}^\delta, \hat{M}^{\alpha\beta}\right]_- + \left[\hat{p}^\gamma, \hat{M}^{\alpha\beta}\right]_- \hat{p}^\delta\right\} \\ &= i g_{\gamma\delta} \left\{\hat{p}^\gamma \left(g^{\alpha\delta}\hat{p}^\beta - g^{\beta\delta}\hat{p}^\alpha\right) + \left(g^{\alpha\gamma}\hat{p}^\beta - g^{\beta\gamma}\hat{p}^\beta\right)\hat{p}^\delta\right\} = 0\, . \end{aligned} \tag{6.142}$$

In order to construct a second Casimir operator, we define now the Pauli-Lubanski operator

$$\hat{W}_\alpha = \frac{1}{2}\, \epsilon_{\alpha\beta\gamma\delta}\, \hat{p}^\beta \hat{M}^{\gamma\delta}\, . \tag{6.143}$$

Here $\epsilon_{\alpha\beta\gamma\delta}$ denotes the four-dimensional, totally anti-symmetric unity tensor, which is a relativistic extension of the three-dimensional Levi-Cività symbol used in (6.50). It has the value $\epsilon_{1234} = 1$ and is anti-symmetric with respect to two of its four indices:

$$\epsilon_{\alpha\beta\gamma\delta} = -\epsilon_{\alpha\beta\delta\gamma} = -\epsilon_{\alpha\delta\gamma\beta} = -\epsilon_{\alpha\gamma\beta\delta} = -\epsilon_{\delta\beta\gamma\alpha} = -\epsilon_{\gamma\beta\alpha\delta} = -\epsilon_{\beta\alpha\gamma\delta}\, . \tag{6.144}$$

The scalar product of the Pauli-Lubanski operator $\hat{W}_\alpha$ with the four-momentum operator $\hat{p}^\alpha$ vanishes due to (6.143) and (6.144):

$$\hat{W}_\alpha\, \hat{p}^\alpha = \frac{1}{2}\, \epsilon_{\alpha\beta\gamma\delta}\, \hat{p}^\beta \hat{M}^{\gamma\delta} \hat{p}^\alpha = 0\, . \tag{6.145}$$

Furthermore, we read off from (3.102), (6.134), (6.136), (6.143), and (6.144) that the Pauli-Lubanski vector commutes with the four-momentum operator:

$$\begin{aligned} \left[\hat{W}^\alpha, \hat{p}^\sigma\right]_- &= g^{\alpha\alpha'} \left[\hat{W}_{\alpha'}, \hat{p}^\sigma\right]_- = \frac{1}{2}\, g^{\alpha\alpha'} \epsilon_{\alpha'\beta\gamma\delta} \left[\hat{p}^\beta \hat{M}^{\gamma\delta}, \hat{p}^\sigma\right]_- = \frac{1}{2}\, g^{\alpha\alpha'} \epsilon_{\alpha'\beta\gamma\delta} \left\{\hat{p}^\beta \left[\hat{M}^{\gamma\delta}, \hat{p}^\sigma\right]_- \right. \\ &\left. + \left[\hat{p}^\beta, \hat{p}^\sigma\right]_- \hat{M}^{\gamma\delta}\right\} = \frac{i}{2}\, g^{\alpha\alpha'} \epsilon_{\alpha'\beta\gamma\delta} \left(g^{\delta\sigma}\hat{p}^\beta \hat{p}^\gamma - g^{\gamma\sigma}\hat{p}^\beta \hat{p}^\delta\right) = 0\, . \end{aligned} \tag{6.146}$$

Now we determine the commutator of the Pauli-Lubanski operator with the basis generators of the Lorentz algebra. To this end we use (3.43), (6.116), (6.136), (6.143), and (6.144) and obtain at first:

$$\begin{aligned}\left[\hat{M}^{\alpha\beta},\hat{W}^{\gamma}\right]_- &= g^{\gamma\delta}\left[\hat{M}^{\alpha\beta},\hat{W}_{\delta}\right]_- = \frac{1}{2}\,g^{\gamma\delta}\epsilon_{\delta\rho\sigma\tau}\left[\hat{M}^{\alpha\beta},\hat{p}^{\rho}\hat{M}^{\sigma\tau}\right]_- \\ &= \frac{1}{2}\,g^{\gamma\delta}\epsilon_{\delta\rho\sigma\tau}\left\{\left[\hat{M}^{\alpha\beta},\hat{p}^{\rho}\right]_-\hat{M}^{\sigma\tau}+\hat{p}^{\rho}\left[\hat{M}^{\alpha\beta},\hat{M}^{\sigma\tau}\right]_-\right\} \\ &= \frac{i}{2}\,g^{\gamma\delta}\left\{g^{\beta\rho}\epsilon_{\delta\rho\sigma\tau}\left(\hat{p}^{\alpha}\hat{M}^{\sigma\tau}-2\hat{p}^{\sigma}M^{\alpha\tau}\right)-g^{\alpha\rho}\epsilon_{\delta\rho\sigma\tau}\left(\hat{p}^{\beta}\hat{M}^{\sigma\tau}-2\hat{p}^{\sigma}M^{\beta\tau}\right)\right\}. \end{aligned} \tag{6.147}$$

In order to identify the right-hand side of (6.147) with known operators, several additional calculations are necessary. At first we apply the contraction rule for the ϵ-tensor

$$\begin{aligned}\epsilon_{\alpha\beta\gamma\delta}\,\epsilon^{\alpha'\beta'\gamma'\delta} &= \delta_{\alpha}{}^{\alpha'}\delta_{\beta}{}^{\beta'}\delta_{\gamma}{}^{\gamma'}+\delta_{\alpha}{}^{\beta'}\delta_{\beta}{}^{\gamma'}\delta_{\gamma}{}^{\alpha'}+\delta_{\alpha}{}^{\gamma'}\delta_{\beta}{}^{\alpha'}\delta_{\gamma}{}^{\beta'} \\ &\quad -\delta_{\alpha}{}^{\beta'}\delta_{\beta}{}^{\alpha'}\delta_{\gamma}{}^{\gamma'}-\delta_{\alpha}{}^{\alpha'}\delta_{\beta}{}^{\gamma'}\delta_{\gamma}{}^{\beta'}-\delta_{\alpha}{}^{\gamma'}\delta_{\beta}{}^{\beta'}\delta_{\gamma}{}^{\alpha'}\,,\end{aligned} \tag{6.148}$$

which is similar to (6.56), so that the relation (6.143) can be inverted in analogy to (6.50) and (6.55) due to the anti-symmetry (6.112):

$$\hat{W}_{\alpha}\,\epsilon^{\alpha\beta\gamma\delta}=\hat{p}^{\beta}\hat{M}^{\gamma\delta}+\hat{p}^{\gamma}\hat{M}^{\delta\beta}+\hat{p}^{\delta}\hat{M}^{\beta\gamma}\,. \tag{6.149}$$

Furthermore, we conclude from the contraction rule (6.148) the special case

$$\epsilon_{\alpha\beta\gamma\delta}\,\epsilon^{\alpha'\beta'\gamma\delta}=2\left(\delta_{\alpha}{}^{\alpha'}\delta_{\beta}{}^{\beta'}-\delta_{\alpha}{}^{\beta'}\delta_{\beta}{}^{\alpha'}\right)\,, \tag{6.150}$$

so that (6.149) can be contracted with the ϵ-tensor. On the one hand we then obtain

$$\hat{W}_{\alpha}\,\epsilon^{\alpha\beta\gamma\delta}\,\epsilon_{\sigma\tau\gamma\delta}=2\left(\hat{W}_{\sigma}\,\delta_{\tau}{}^{\beta}-\hat{W}_{\tau}\,\delta_{\sigma}{}^{\beta}\right)\,, \tag{6.151}$$

whereas we read off from (6.149)

$$\hat{W}_{\alpha}\,\epsilon^{\alpha\beta\gamma\delta}\,\epsilon_{\sigma\tau\gamma\delta}=\hat{p}^{\beta}\hat{M}^{\gamma\delta}\,\epsilon_{\sigma\tau\gamma\delta}+\hat{p}^{\gamma}\hat{M}^{\delta\beta}\,\epsilon_{\sigma\tau\gamma\delta}+\hat{p}^{\delta}\hat{M}^{\beta\gamma}\,\epsilon_{\sigma\tau\gamma\delta}\,. \tag{6.152}$$

Thus, taking into account (6.112) and (6.144) we result in

$$\epsilon_{\sigma\tau\gamma\delta}\left(\hat{p}^{\beta}\hat{M}^{\gamma\delta}-2\hat{p}^{\gamma}\hat{M}^{\beta\delta}\right)=2\left(\hat{W}_{\sigma}\,\delta_{\tau}{}^{\beta}-\hat{W}_{\tau}\,\delta_{\sigma}{}^{\beta}\right)\,. \tag{6.153}$$

Inserting then (6.153) into (6.147) determines the commutator of the Lubanski operator with the basis generators of the Lorentz algebra in the following form:

$$\left[\hat{M}^{\alpha\beta},\hat{W}^{\gamma}\right]_-=i\left(g^{\beta\gamma}W^{\alpha}-g^{\alpha\gamma}W^{\beta}\right)\,. \tag{6.154}$$

With the help of the representation matrices (6.44) one recognizes that the Pauli-Lubaski operator represents a tensor operator of rank $n=1$:

$$\left[\hat{M}^{\alpha\beta},\hat{W}^{\gamma}\right]_-=-\left(L^{\alpha\beta}\right)^{\gamma}{}_{\delta}\,\hat{W}^{\delta}\,. \tag{6.155}$$

We consider now the scalar product of the Pauli-Lubanski operator with itself

$$\hat{W}^2 = g_{\alpha\beta}\hat{W}^\alpha \hat{W}^\beta \tag{6.156}$$

and show that it represents the second Casimir operator of the Poincaré algebra. At first, we yield for the commutator of $\hat{W}^2$ and $\hat{p}^\alpha$ due to (3.10) and (6.146)

$$\left[\hat{W}^2, \hat{p}^\alpha\right]_- = 0\,. \tag{6.157}$$

In addition, we obtain that $\hat{W}^2$ also commutes with $\hat{M}^{\alpha\beta}$ by taking into account (3.43) and (6.154)

$$\left[\hat{W}^2, \hat{M}^{\alpha\beta}\right]_- = 0\,. \tag{6.158}$$

Finally, the question arises how to physically interpret both Casimir operators of the Poincaré group. To this end we describe a particle with fixed four-momentum $p = (p^\mu)$ via a tensor or spinor field $\psi^\sigma(x)$ and the eigenvalue problem

$$\hat{p}^\mu \psi^\sigma(x) = p^\mu \psi^\sigma(x)\,. \tag{6.159}$$

Then the first Casimir operator (6.140) has an eigenvalue, which is determined by the rest mass M due to (6.21). Thus, in view of the second Casmir operator $\hat{W}^2$ it remains to interpret physically also the Pauli-Lubanski operator $\hat{W}_\alpha$. To this end we insert the decomposition (6.116) of the representation $\hat{M}^{\alpha\beta}$ of the Lorentz algebra in the Hilbert space of the tensor or spinor fields in the representation $\hat{L}^{\alpha\beta}$ of the Lorentz algebra in Minskowski space and the representation $N^{\alpha\beta}$ of the Lorentz algebra in the space of the tensor or spinor components into (6.143). Due to the anti-symmetry of the ϵ-tensor (6.144) this yields:

$$\hat{W}_\alpha = \frac{1}{6}\,\epsilon_{\alpha\beta\gamma\delta}\left(\hat{p}^\beta \hat{L}^{\gamma\delta} + \hat{p}^\gamma \hat{L}^{\delta\beta} + \hat{p}^\delta \hat{L}^{\beta\gamma}\right) + \frac{1}{2}\,\epsilon_{\alpha\beta\gamma\delta} N^{\gamma\delta}\,. \tag{6.160}$$

Taking into account the definition of the orbital angular momentum operators (6.100) as well as the commutation relations (6.101) and (6.134) we observe that (6.160) reduces to

$$\hat{W}_\alpha = \frac{1}{2}\,\epsilon_{\alpha\beta\gamma\delta}\,\hat{p}^\beta\, N^{\gamma\delta}\,. \tag{6.161}$$

Thus, it turns out that the orbital angular momentum operator $\hat{L}^{\alpha\beta}$ does not contribute to the Pauli-Lubanski operator W_α. Describing again a particle with fixed four-momentum $p = (p^\mu)$ via a tensor or spinor field $\psi^\sigma(x)$, the eigenvalue problem with respect to the Pauli-Lubanski operator reads

$$\hat{W}_\alpha \psi^\sigma(x) = W_\alpha \psi^\sigma(x)\,, \tag{6.162}$$

where the eigenvector is given by the Pauli-Lubanski four-vector

$$W_\alpha = \frac{1}{2}\,\epsilon_{\alpha\beta\gamma\delta}\, p^\beta N^{\gamma\delta}\,. \tag{6.163}$$

Decomposing the basis generators $N^{\alpha\beta}$ of the Lorentz algebra in the space of tensor or spinor components in analogy to (6.50), (6.51) into two classes

$$S_k = \frac{1}{2}\epsilon_{klm}N^{lm}, \tag{6.164}$$
$$K_k = N^{0k}, \tag{6.165}$$

we introduce the two vectors

$$\mathbf{S} = (S_1, S_2, S_3) = \left(N^{23}, N^{31}, N^{12}\right), \tag{6.166}$$
$$\mathbf{K} = (K_1, K_2, K_3) = \left(N^{01}, N^{02}, N^{03}\right). \tag{6.167}$$

With this the covariant components of the Pauli-Lubanski four-vector (6.163) are defined similar to (6.6) and (6.19)

$$(W_\alpha) = (W_0, W_1, W_2, W_3) = \left(W_0, -W^i\right) = (W_0, -\mathbf{W}), \tag{6.168}$$

where the temporal and spatial components read

$$W_0 = \mathbf{p}\cdot\mathbf{S}, \tag{6.169}$$
$$\mathbf{W} = p^0\,\mathbf{S} + \mathbf{S}\times\mathbf{K}, \tag{6.170}$$

respectively. In the rest frame of the particle we have $p^0 = Mc$ and $\mathbf{p} = \mathbf{0}$, so that the temporal and spatial components of the Pauli-Lubanski vector (6.169) and (6.170) reduce to

$$W_0 = 0, \tag{6.171}$$
$$\mathbf{W} = Mc\,\mathbf{S}. \tag{6.172}$$

Analogously to the calculation of (6.57) we obtain from the commutation relation (6.117) a corresponding commutation relation for the vector components S_k:

$$[S_k, S_l]_- = i\epsilon_{klm}S_m. \tag{6.173}$$

Thus, we conclude that in the rest frame of the particle the Pauli-Lubanski four-vector represents the spin angular momentum of the particle. Therefore, $\hat{W}_\alpha$ in (6.143) is a relativistic generalization of the spin angular momentum in any inertial frame.

6.13 Irreducible Representations of Poincaré Group

With the help of the eigenvalues of the Casimir operators (6.140) and (6.143) of the Poincaré algebra one can classify the irreducible representations of the Poincaré group. Note that they are infinite dimensional as they describe particles with an unbounded momentum. In contrast to that the defining representation of the Lorentz group was finite dimensional. The eigenvalue of the first Casimir operator (6.140) is characterized due to (6.21) by the rest mass M of the particle:

$$p^2 = M^2 c^2 \,. \tag{6.174}$$

Depending whether the rest mass M is non-zero or vanishes one distinguishes two different classes of representations.

6.13.1 Massive Representations

Let us consider first the case that the rest mass is non-zero, i.e. $M > 0$, which defines the massive representations. Then we remark that the second Casimir operator (6.143) has an eigenvalue, which is a Lorentz scalar, so it has in each inertial system the same value. In particular in the rest frame the eigenvalue of (6.143) reduces due to (6.171) to

$$W^2 = -\mathbf{W}^2 = -M^2 c^2 \, \mathbf{S}^2 \,. \tag{6.175}$$

As the components of the vector $\mathbf{S}$ obey the commutation relations (6.173) of the angular momentum algebra, the eigenvalues of (6.175) are given by

$$W^2 = -M^2 c^2 \, S(S+1) \,; \qquad S = 0, 1/2, 1, 3/2, \ldots \,. \tag{6.176}$$

Such a massive representation is, thus, characterized by both the mass M and the spin S. As these are the fundamental properties of elementary particles, we have obtained the result that the elementary particles themselves can be identified with the irreducible representations of the Poincaré group. States within such a representation only differ in the third component of the spin vector, where $2S + 1$ different eigenvalues can occur.

6.13.2 Massless Representations

For a particle with a vanishing rest mass, i.e. $M = 0$, it is not possible to reach its rest frame by applying any Lorentz transformation. If this was possible, then this would have the unphysical consequence that the energy of the particle would vanish due to $p^0 = 0$. Therefore, massless particles need as a basic principle a different treatment.

Within a massless representation both four-vectors p^α and W^α have a vanishing scalar product with respect to each other due to (6.145):

$$p_\alpha \, W^\alpha = 0 \,. \tag{6.177}$$

Furthermore, due to (6.174) and (6.175), they represent light-like four-vectors, i.e. they obey

$$p_\alpha p^\alpha = 0 \,, \qquad W_\alpha W^\alpha = 0 \,, \tag{6.178}$$

Decomposing (6.178) into its temporal and spatial components

$$\left(p^0\right)^2 = \mathbf{p}^2 \,, \qquad \left(W^0\right)^2 = \mathbf{W}^2 \,, \tag{6.179}$$

then we directly conclude from $p^\alpha \neq 0$ and $W^\alpha \neq 0$:

$$p^0 \neq 0 \,, \qquad W^0 \neq 0 \,. \tag{6.180}$$

Let us consider now the linear combination

$$A p^\alpha + B W^\alpha = 0 \,. \tag{6.181}$$

Obviously, (6.181) does not only have the trivial solution $A = B = 0$ as we obtain from $\alpha = 0$ and from taking into account (6.180)

$$B = -\frac{p^0}{W^0} A \,. \tag{6.182}$$

Thus, both light-like four-vectors p^α and W^α are linear dependent. Therefore, for their respective operators $\hat{p}^\alpha$ and $\hat{W}^\alpha$ there must exist a proportionality factor operator $\hat{h}$ with the property

$$\hat{W}^\alpha = \hat{h}\, \hat{p}^\alpha \,. \tag{6.183}$$

Now we determine for this proportionality factor $\hat{h}$ the commutator relations with the generators of the Poincaré algebra. At first we get from (3.10), (6.134), (6.146), and (6.183)

$$\left[\hat{W}^\alpha, \hat{p}^\beta\right]_- = \left[\hat{h}\, \hat{p}^\alpha, \hat{p}^\beta\right]_- = \hat{h}\left[\hat{p}^\alpha, \hat{p}^\beta\right]_- + \left[\hat{h}, \hat{p}^\beta\right]_- \hat{p}^\alpha \quad \Longrightarrow \quad \left[\hat{h}, \hat{p}^\alpha\right]_- = 0 \,. \tag{6.184}$$

In a similar way we determine from (3.43), (6.136), (6.154), and (6.183):

$$\left[\hat{M}^{\alpha\beta}, \hat{W}^\gamma\right]_- = \left[\hat{M}^{\alpha\beta}, \hat{h}\hat{p}^\gamma\right]_- = \left[\hat{M}^{\alpha\beta}, \hat{h}\right]_- \hat{p}^\gamma + \hat{h}\left[\hat{M}^{\alpha\beta}, \hat{p}^\gamma\right]_- \quad \Longrightarrow \quad \left[\hat{M}^{\alpha\beta}, \hat{h}\right]_- = 0 \,. \tag{6.185}$$

This means that the proportionality factor $\hat{h}$ represents an additional Casimir operator. For the corresponding eigenvalues of $\hat{W}^\alpha$, $\hat{h}$, and $\hat{p}$ we then obtain from (6.183)

$$W^\alpha = h\, p^\alpha \,, \tag{6.186}$$

so we read off for the zeroth component $\alpha = 0$

$$h = \frac{W^0}{p^0} \,. \tag{6.187}$$

Thus, taking into account (6.169) and (6.179) the eigenvalue h of this additional Casimir operator $\hat{h}$ is given by

$$h = \frac{\mathbf{pS}}{|\mathbf{p}|} \,, \tag{6.188}$$

which is intuitively accessible as the projection of the particle spin upon the direction of motion. Therefore, one calls $\hat{h}$ as the helicity operator. For a given spin S and momentum $\mathbf{p}$ the eigenvalue (6.188) of the helicity operator $\hat{h}$ has a fixed sign, i.e. either positive or negative, which is the same in all inertial systems.

One can define the helicity operator $\hat{h}$ also for massive particles, but then it does not represent a Casimir operator. This means, for instance, that then an appropriate Lorentz transformation can convert a state of positive helicity into another state with negative helicity. Thus, the helicity describes for a massive particle its state but not the massive particle itself. The latter is only possible for massless particles as they always move with light velocity.

6.13.3 Other Representations

From a mathematical point of view the Poincaré group does allow also for other classes of unitary representations. Among them is one with the constraint $p_\mu p^\mu = 0$ and a continuous spin. Another one obeys the constraint $p_\mu p^\mu < 0$ for particles moving with a velocity larger than the light velocity, which are known hypothetically as tachyons. But so far there is no experimental indication that these other representations of the Poincaré group are realised in nature by any elementary particle. But, although this is purely speculative, one of these other representations of the Poincaré group might indicate a solution for the virulent problem of our time that the physical nature of dark matter is yet unknown.

Chapter 7

Noether Theorem

In 1918 the mathematician Emmy Noether published a theorem, which had far-reaching consequences for different branches of theoretical physics. It states that every differentiable symmetry of the action of a physical system is relaed to a corresponding conservation law. Although Noether's theorem has implications also in classical mechanics, we focus here on classical field theory, where it provides a fundamental connection between continuous symmetries and conserved quantities. Namely, each continuous symmetry, which leaves the action invariant, leads inevitably to a corresponding conserved quantity. For instance, translations in time and space are related with the energy and the momentum conservation. In a similar way spatial rotations and boosts imply the conservation of angular momentum and the center of mass, respectively. And, finally, an invariance with respect to a translation of the phase in a wave function turns out to be connected with the charge conservation. In the following we derive the Noether theorem in its most general form in the realm of classical field theory and then specialize it to these important applications.

7.1 Invariance

The action $\mathcal{A}$ represents a functional of the underlying tensor or spinor field $\Psi^\sigma(x^\lambda)$, i.e. we have

$$\mathcal{A} = \mathcal{A}\left[\Psi^\sigma(\bullet)\right] , \tag{7.1}$$

which is defined as a spatio-temporal integral over the Lagrange density $\mathcal{L}$:

$$\mathcal{A} = \frac{1}{c}\int_\Omega d^4x\, \mathcal{L}\left(\Psi^\sigma(x^\lambda), \partial_\mu \Psi^\sigma(x^\lambda)\right) . \tag{7.2}$$

Here we restrict ourselves to a local field theory, where the Lagrange density $\mathcal{L}$ can only depend on the tensor or spinor field itself and its first partial derivatives but not from higher partial derivatives with respect to space and time. Now we consider a transformation, which involves both the space-time coordinates and the tensor or spinor field:

$$x'^{\lambda} = x'^{\lambda}(x^{\kappa}) \,, \tag{7.3}$$

$$\Psi'^{\sigma}(x'^{\lambda}) = \Psi'^{\sigma}(\Psi^{\tau}(x^{\kappa})) \,. \tag{7.4}$$

Here we use the convention that the unprimed (primed) quantities denote the ones before (after) the transformation. The transformation (7.3), (7.4) then changes the action (7.1), (7.2) to

$$\mathcal{A}'\left[\Psi'^{\sigma}(\bullet)\right] = \frac{1}{c}\int_{\Omega'} d^4x'\,\mathcal{L}\left(\Psi'^{\sigma}(x'^{\lambda}), \partial'_{\mu}\Psi'^{\sigma}(x'^{\lambda})\right) \,. \tag{7.5}$$

The transformation (7.4), (7.5) is exactly then a symmetry transformation, when it leaves the action invariant:

$$\mathcal{A}\left[\Psi^{\sigma}(\bullet)\right] = \mathcal{A}'\left[\Psi'^{\sigma}(\bullet)\right] \,. \tag{7.6}$$

In the following we analyze the physical consequences of this invariance of the action.

7.2 Infinitesimal Transformation

For the proof of the Noether theorem it turns out to be sufficient to consider an infinitesimal symmetry transformation. For such an infinitesimal symmetry transformation Eqs. (7.3), (7.4) reduce to

$$x'^{\lambda} = x^{\lambda} + \delta x^{\lambda} \,, \tag{7.7}$$

$$\Psi'^{\sigma}(x'^{\lambda}) = \Psi^{\sigma}(x^{\lambda}) + \delta\Psi^{\sigma}(x^{\lambda}) \,. \tag{7.8}$$

Here $\delta\Psi^{\sigma}(x^{\lambda})$ denotes the total variation of the tensor or spinor field, which generically contains two contributions. On the one hand it involves a change due to transforming the space-time coordinates from x^{λ} to x'^{λ}, on the other hand it includes a change of the tensor or spinor field from Ψ^{σ} to Ψ'^{σ}. For technical reasons it is, therefore, advantageous to introduce the local variation $\tilde{\delta}\Psi^{\sigma}(x^{\lambda})$ of the tensor or spinor field $\Psi^{\sigma}(x^{\lambda})$ as the infinitesimal transformation of the tensor or spinor field for fixed space-time coordinates:

$$\tilde{\delta}\Psi^{\sigma}(x^{\lambda}) = \Psi'^{\sigma}(x^{\lambda}) - \Psi^{\sigma}(x^{\lambda}) \,. \tag{7.9}$$

Combining (7.8) and (7.9) we recognize the following connection between the total variation $\delta\Psi^{\sigma}(x^{\lambda})$ and the local variation $\tilde{\delta}\Psi^{\sigma}(x^{\lambda})$:

$$\tilde{\delta}\Psi^{\sigma}(x^{\lambda}) = \delta\Psi^{\sigma}(x^{\lambda}) - \left[\Psi'^{\sigma}(x'^{\lambda}) - \Psi'^{\sigma}(x^{\lambda})\right] \,. \tag{7.10}$$

Inserting (7.7) into (7.10) and taking into account only the first order of the variations yields

$$\tilde{\delta}\Psi^{\sigma}(x^{\lambda}) \approx \delta\Psi^{\sigma}(x^{\lambda}) - \partial_{\mu}\Psi'^{\sigma}(x^{\lambda})\delta x^{\mu} \approx \delta\Psi^{\sigma}(x^{\lambda}) - \partial_{\mu}\Psi^{\sigma}(x^{\lambda})\delta x^{\mu} \,. \tag{7.11}$$

The result (7.11) means that one has to subtract from the the total variation $\delta\Psi^\sigma(x^\lambda)$ that contribution, which stems from the change of the space-time coordinates, in order to obtain the local variation $\tilde{\delta}\Psi^\sigma(x^\lambda)$. Note that using the local variation $\tilde{\delta}\Psi^\sigma(x^\lambda)$ has the formal advantage that it commutes with the differentiation

$$\partial_\mu\left[\tilde{\delta}\Psi^\sigma(x^\lambda)\right] = \tilde{\delta}\left[\partial_\mu\Psi^\sigma(x^\lambda)\right], \tag{7.12}$$

whereas this is not true for the total variation $\delta\Psi^\sigma(x^\lambda)$ due to (7.11):

$$\partial_\mu\left[\tilde{\delta}\Psi^\sigma(x^\lambda)\right] \neq \delta\left[\partial_\mu\Psi^\sigma(x^\lambda)\right]. \tag{7.13}$$

7.3 Total and Local Variation of Action

According to (7.2) the action $\mathcal{A}$ before the transformation is integrated with respect to Ω, whereas the action $\mathcal{A}'$ after the transformation is integrated with respect to Ω' due to (7.5). Within a passive interpretation of the space-time transformation (7.3) both Ω and Ω' denote one and the same four-dimensional integration volume, which is described by different coordinate systems. Transforming Ω' back into Ω with the help of the infinitesimal transformation (7.7), the respective differential volumes transform with the Jacobi determinant. Up to first order in the variation we have

$$\frac{d^4x'}{d^4x} = \frac{\partial(x'^\lambda)}{\partial(x^\mu)} = \left|g^\lambda{}_\mu + \frac{\partial\delta x^\lambda}{\partial x^\mu}\right| = \begin{vmatrix} 1+\dfrac{\partial\delta x^0}{\partial x^0} & \dfrac{\partial\delta x^0}{\partial x^1} & \dots \\ \dfrac{\partial\delta x^1}{\partial x^0} & 1+\dfrac{\partial\delta x^1}{\partial x^1} & \dots \\ \vdots & \vdots & \ddots \end{vmatrix} \approx 1 + \frac{\partial\delta x^\mu}{\partial x^\mu}. \tag{7.14}$$

The result (7.14) states that the relative change of the differential volumes is given by the four-divergence of the variation of the space-time coordinates. The Lagrange density transforms accordingly via

$$\mathcal{L}' = \mathcal{L} + \delta\mathcal{L}. \tag{7.15}$$

Taking into account (7.2) and (7.4) the total variation of the action

$$\delta\mathcal{A} = \mathcal{A}'\left[\Psi'^\sigma(\bullet)\right] - \mathcal{A}\left[\Psi^\sigma(\bullet)\right] = \frac{1}{c}\int_{\Omega'} d^4x'\,\mathcal{L}' - \frac{1}{c}\int_\Omega d^4x\,\mathcal{L} \tag{7.16}$$

can be evaluated with the help of (7.14) and (7.15) up to first order:

$$\delta\mathcal{A} = \frac{1}{c}\int_\Omega d^4x\left[\left(1+\frac{\partial\delta x^\mu}{\partial x^\mu}\right)(\mathcal{L}+\delta\mathcal{L}) - \mathcal{L}\right] \approx \frac{1}{c}\int_\Omega d^4x\left(\delta\mathcal{L} + \frac{\partial\delta x^\mu}{\partial x^\mu}\mathcal{L}\right). \tag{7.17}$$

Similar to (7.11) the following relation holds between the total variation $\delta\mathcal{L}$ and the local variation $\tilde{\delta}\mathcal{L}$ of the Lagrange density:

$$\delta\mathcal{L} = \tilde{\delta}\mathcal{L} + \partial_\mu\mathcal{L}\,\delta x^\mu. \tag{7.18}$$

For the Lagrange density $\mathcal{L}$ of a local field theory as it appears in (7.2) the local variation $\tilde{\delta}\mathcal{L}$ is given by

$$\tilde{\delta}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \Psi^\sigma(x^\lambda)}\,\tilde{\delta}\Psi^\sigma(x^\lambda) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\tilde{\delta}\Big[\partial_\mu \Psi^\sigma(x^\lambda)\Big]\,. \tag{7.19}$$

As the local variation $\tilde{\delta}$ has according to (7.12) the property that it commutates with the differentiation, we conclude from (7.19)

$$\tilde{\delta}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \Psi^\sigma(x^\lambda)}\,\tilde{\delta}\Psi^\sigma(x^\lambda) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\partial_\mu\Big[\tilde{\delta}\Psi^\sigma(x^\lambda)\Big]\,, \tag{7.20}$$

which can be rewritten as

$$\tilde{\delta}\mathcal{L} = \left[\frac{\partial \mathcal{L}}{\partial \Psi^\sigma(x^\lambda)} - \partial_\mu\,\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\right]\tilde{\delta}\Psi^\sigma(x^\lambda) + \partial_\mu\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\tilde{\delta}\Psi^\sigma(x^\lambda)\right]. \tag{7.21}$$

7.4 Continuity Equation

Inserting both the total and the local variation of the Lagrange density from (7.18) and (7.21) into (7.17), the total variation of the action results in

$$\begin{aligned}\delta\mathcal{A} \;=\; & \frac{1}{c}\int_\Omega d^4x\,\left\{\left[\frac{\partial \mathcal{L}}{\partial \Psi^\sigma(x^\lambda)} - \partial_\mu\,\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\right]\tilde{\delta}\Psi^\sigma(x^\lambda)\right.\\ & \left.+\partial_\mu\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\tilde{\delta}\Psi^\sigma(x^\lambda)\right] + \mathcal{L}\,\delta x^\mu\right\}.\end{aligned} \tag{7.22}$$

The first term in (7.22) turns out to vanish as the tensor or spinor field $\Psi^\sigma(x^\lambda)$ has to fulfill the Euler-Lagrange equations corresponding to the action $\mathcal{A}$ due to the Hamilton principle of classical field theory:

$$\frac{\delta\mathcal{A}}{\delta\Psi^\sigma(x^\lambda)} \;=\; \frac{\partial \mathcal{L}}{\partial \Psi^\sigma(x^\lambda)} - \partial_\mu\,\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))} = 0\,. \tag{7.23}$$

From (7.22), (7.23) as well as the connection (7.10) between the local variation $\tilde{\delta}\Psi^\sigma(x^\lambda)$ and the total variation $\delta\Psi^\sigma(x^\lambda)$ of the tensor or spinor field $\Psi^\sigma(x^\lambda)$ we then conclude

$$\delta\mathcal{A} = \frac{1}{c}\int_\Omega d^4x\,\partial_\mu\left\{\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\delta\Psi^\sigma(x^\lambda) - \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\partial_\nu\,\Psi^\sigma(x^\lambda) - \delta^\mu{}_\nu\,\mathcal{L}\right]\delta x^\nu\right\}. \tag{7.24}$$

As the infinitesimal symmetry transformation leaves the action invariant according to (7.6), the total variation of the action must vanish:

$$\delta\mathcal{A} = 0\,. \tag{7.25}$$

Furthermore, we note that the four-dimensional integration volume Ω in (7.24) can be chosen arbitrarily, so we read off that also the integrand of (7.24) must vanish due to (7.25). In this way one obtains a continuity equation

$$\partial_\mu\,f^\mu(x^\lambda) = 0\,, \tag{7.26}$$

where the current density $f^\mu(x^\lambda)$ is additive in the variations of both the tensor or spinor fields $\delta\Psi^\sigma(x^\lambda)$ and the space-time coordinates δx^ν:

$$f^\mu(x^\lambda) = \frac{1}{c}\,\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\delta\Psi^\sigma(x^\lambda) - \frac{1}{c}\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi^\sigma(x^\lambda))}\,\partial_\nu\,\Psi^\sigma(x^\lambda) - \delta^\mu{}_\nu\,\mathcal{L}\right]\delta x^\nu\,. \tag{7.27}$$

7.5 Conserved Quantities

A continuity equation of the form (7.26) represents a differential formulation of a conservation law. Integrating (7.26) over the total three-dimensional volume, we obtain

$$0 = \int d^3x\,\partial_\mu\, f^\mu(x^\lambda) \;=\frac{1}{c}\int d^3x\,\frac{\partial f^0(x^\lambda)}{\partial t} + \int d^3x\,\mathrm{div}\,\mathbf{f}(\mathbf{x},t)\,. \tag{7.28}$$

Thus, applying the theorem of Gauß and assuming that the tensor or spinor field as well as its first derivatives vanish fast enough at infinity, we get

$$\frac{\partial}{\partial t}\int d^3x\,f^0(x^\lambda) \;= -c\int d^3x\,\mathrm{div}\,\mathbf{f}(\mathbf{x},t) \;=\; -c\oint d\mathbf{o}\cdot\mathbf{f}(\mathbf{x},t) = 0\,. \tag{7.29}$$

With this we conclude that the spatial integral of the temporal component of the current density

$$F(t) = \int d^3x\,f^0(x^\lambda) \tag{7.30}$$

represents the conserved quantity of the Noether theorem:

$$\frac{\partial}{\partial t}F(t) = 0\,. \tag{7.31}$$

After having derived with this the most general form of the Noether theorem, we discuss now case by case important applications.

7.6 Canonical Energy-Momentum Tensor

Due to the Poincaré symmetry of the flat Minkowskian space-time structure the action must be invariant with respect to translations of both time and space. According to (6.129) and (6.130) this leads to the following infinitesimal variations of the space-time coordinates x^λ and the tensor or spinor field $\Psi'^\sigma(x'^\lambda)$:

$$\begin{aligned} \delta x^\lambda &= x'^\lambda - x^\lambda = \epsilon^\lambda\,, & (7.32)\\ \delta\partial\Psi^\sigma(x^\lambda) &= \Psi'^\sigma(x'^\lambda) - \Psi^\sigma(x^\lambda) = 0\,. & (7.33)\end{aligned}$$

As the infinitesimal translation four-vector ϵ^λ can be chosen arbitrarily, we read off from (7.26), (7.27) and (7.32), (7.33) the differential continuity equation

$$\partial_\mu\,\Theta^{\mu\nu}(x^\lambda) = 0\,, \tag{7.34}$$

where the canonical energy-momentum tensor $\Theta^{\mu\nu}(x^\lambda) = g^{\nu\kappa}\,\Theta^\mu{}_\kappa(x^\lambda)$ is given by

$$\Theta^\mu{}_\kappa(x^\lambda) = \frac{1}{c}\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi^\sigma(x^\lambda))}\,\partial_\kappa\Psi^\sigma(x^\lambda) - \delta^\mu{}_\kappa\,\mathcal{L}\right]. \tag{7.35}$$

Evaluating (7.34) for $\nu = 0, 1, 2, 3$ we obtain four conserved quantities, namely the energy E and the momentum $\mathbf{p}$ of the particle:

$$(p^\nu) = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix} = \int d^3x\,\left(\Theta^{0\nu}(x^\lambda)\right), \qquad \frac{\partial p^\nu}{\partial t} = 0\,. \tag{7.36}$$

The energy turns out to be of the form

$$E = \int d^3x\,\mathcal{H}\,, \tag{7.37}$$

where the Hamilton density $\mathcal{H}$ is given by

$$\mathcal{H} = c\Theta^{00} = \frac{\partial\mathcal{L}}{\partial\left(\dfrac{\partial\Psi^\sigma}{\partial t}\right)}\,\frac{\partial\Psi^\sigma}{\partial t} - \mathcal{L} \tag{7.38}$$

and, thus, corresponds to a Legendre transformation. And the momentum results in

$$p^i = \int d^3x\,\mathcal{P}^i\,, \tag{7.39}$$

where the momentum density $\mathcal{P}^i$ turns out to be

$$\mathcal{P}^i = \Theta^{0i} = -\frac{\partial\mathcal{L}}{\partial\left(\dfrac{\partial\Psi^\sigma}{\partial t}\right)}\,\frac{\partial\Psi^\sigma}{\partial x^i}\,. \tag{7.40}$$

Later on we will specialize (7.37)–(7.40) for concrete field theories as, for instance, the Maxwell and the Dirac theory. Then it will also become transparent that the conserved quantities defined according to the Noether theorem via (7.37)–(7.40) have, indeed, the proper physical SI units. Furthermore, we remark that, using the definition (7.35) of the canonical energy-momentum tensor, the current density (7.27) of the Noether theorem can also be written as

$$f^\mu(x^\lambda) = \frac{1}{c}\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi^\sigma(x^\lambda))}\,\delta\Psi^\sigma(x^\lambda) - \Theta^\mu{}_\nu(x^\lambda)\delta x^\nu\,. \tag{7.41}$$

7.7 Angular Momentum Tensor

The action must also be invariant with respect to Lorentz transformations. According to (6.22), (6.39), (6.43) and Section 6.9 this involves the following infinitesimal transformations of the space-time coordinates x^λ in (7.7) and of the tensor or spinor field $\Psi^\sigma(x^\lambda)$ in (7.8):

$$\delta x^\lambda = x'^\lambda - x^\lambda = -\frac{i}{2}\,\omega_{\nu\kappa}\,(L^{\nu\kappa})^\lambda{}_\mu\,x^\mu\,, \tag{7.42}$$

$$\delta\Psi^\sigma(x^\lambda) = \Psi'^\sigma(x'^\lambda) - \Psi^\sigma(x^\lambda) = -\frac{i}{2}\,\omega_{\nu\kappa}\,(N^{\nu\kappa})^\sigma{}_\tau\,\Psi^\tau(x^\lambda)\,. \tag{7.43}$$

As the infinitesimal rotation angles and rapidities can be chosen arbitrarily, we obtain from (7.26), (7.27) and (7.42), (7.43) the differential conservation law

$$\partial_\mu J^{\mu\nu\kappa}(x^\lambda) = 0 \,, \tag{7.44}$$

where the angular momentum tensor $J^{\mu\nu\kappa}$ consists of two contributions:

$$J^{\mu\nu\kappa}(x^\lambda) = L^{\mu\nu\kappa}(x^\lambda) + S^{\mu\nu\kappa}(x^\lambda) \,. \tag{7.45}$$

The first term in (7.45) depends on the representation matrices of the Lorentz algebra $L^{\nu\kappa}$ in the space-time

$$L^{\mu\nu\kappa}(x^\lambda) = i\Theta^\mu{}_\rho(x^\lambda)\,(L^{\nu\kappa})^\rho{}_{\rho'}\; x^{\rho'} \tag{7.46}$$

and is, therefore, identified with the orbital angular momentum tensor. Inserting therein the respective representation matrices (6.44), the orbital angular tensor reduces to

$$L^{\mu\nu\kappa}(x^\lambda) = \Theta^{\mu\kappa}(x^\lambda)\,x^\nu - \Theta^{\mu\nu}(x^\lambda)\,x^\kappa \,. \tag{7.47}$$

Analogously, the second term in (7.45) stems from the representation matrices of the Lorentz algebra $N^{\nu\kappa}$ in the tensor or spinor field space

$$S^{\mu\nu\kappa}(x^\lambda) = \frac{-i}{c}\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi^\sigma(x^\lambda))}\,(N^{\nu\kappa})^\sigma{}_\tau\;\Psi^\tau(x^\lambda) \tag{7.48}$$

and, therefore, corresponds to the spin angular momentum tensor. As $N^{\nu\kappa}$ is anti-symmetric with respect to its indices ν, κ, the spin angular momentum tensor (7.48) fulfills the property

$$S^{\mu\nu\kappa}(x^\lambda) = -S^{\mu\kappa\nu}(x^\lambda) \,. \tag{7.49}$$

Furthermore, we read off from (7.34), (7.44), (7.45), and (7.47) that the four-divergence of the spin angular momentum tensor coincides with the anti-symmetric contribution of the canonical energy-momentum tensor:

$$\partial_\mu S^{\mu\nu\kappa}(x^\lambda) = \Theta^{\kappa\nu}(x^\lambda) - \Theta^{\nu\kappa}(x^\lambda) \,. \tag{7.50}$$

Note that in addition to the differential conservation law (7.44) also an integral version exists, which states that an anti-symmetric tensor of second rank represents a constant of motion:

$$M^{\nu\kappa} = \int d^3x\; J^{0\nu\kappa}(x^\lambda) \,, \qquad \frac{\partial M^{\nu\kappa}}{\partial t} = 0 \,. \tag{7.51}$$

Specializing ν, κ to the values $j,k = 1,2,3$ one can interpret M^{jk} as the total angular momentum. According to (7.45) and (7.51) it decomposes into

$$M^{\nu\kappa} = L^{jk} + S^{jk} \,. \tag{7.52}$$

Here the angular momentum L^{jk} follows from (7.47)

$$L^{jk} = \int d^3x \left[\Theta^{0k}(\mathbf{x},t)\,x^j - \Theta^{0j}(\mathbf{x},t)\,x^k\right] , \tag{7.53}$$

which reduces with the help of (7.35) to

$$L^{jk} = \frac{1}{c} \int d^3x \, \frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi^\sigma(\mathbf{x},t))} \left(x^k \partial_j - x^j \partial_k \right) \Psi^\sigma(\mathbf{x},t) \,. \tag{7.54}$$

Correspondingly the spin angular momentum S^{jk} reads due to (7.48)

$$S^{jk} = \frac{-i}{c} \int d^3x \, \frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi^\sigma(\mathbf{x},t))} \, \left(N^{jk} \right)^\sigma{}_\tau \, \Psi^\tau(x^\lambda) \,, \tag{7.55}$$

so that the spin angular momentum vector

$$S_i = \frac{1}{2} \, \epsilon_{jk} S^{jk} \tag{7.56}$$

can be expressed with the help of (6.164) according to

$$S_i = \frac{-i}{c} \int d^3x \, \frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi^\sigma(\mathbf{x},t))} \, \left(S_i \right)^\sigma{}_\tau \, \Psi^\tau(x^\lambda) \,. \tag{7.57}$$

7.8 Symmetrizing Canonical Energy-Momentum Tensor

In general the canonical energy-momentum tensor $\Theta^{\nu\kappa}(x^\lambda)$ following from (7.35) due to the Noether theorem turns out to be not symmetric with respect to its indices ν, κ. This represents at a first glance a quite fundamental theoretical problem as in Albert Einstein's general relativity a symmetric energy-momentum tensor $T^{\nu\kappa}(x^\lambda)$ appears as an inhomogeneity of the field equations defining the underlying metric of space-time. In the following we show how this problem can generically be solved in a constructive way. To this end we work out the so-called Belifante construction, which allows for any underlying field theory to determine for each canonical energy-momentum tensor $\Theta^{\nu\kappa}(x^\lambda)$ a symmetrized energy-momentum tensor $T^{\nu\kappa}(x^\lambda)$ by adding an additional tensor of second rank $t^{\nu\kappa}$:

$$T^{\nu\kappa}(x^\lambda) = \Theta^{\nu\kappa}(x^\lambda) + t^{\nu\kappa}(x^\lambda) \,. \tag{7.58}$$

Demanding that the modified energy-momentum tensor $T^{\nu\kappa}(x^\lambda)$ is symmetric, i.e. that

$$T^{\nu\kappa}(x^\lambda) - T^{\kappa\nu}(x^\lambda) = 0 \tag{7.59}$$

holds, we obtain from (7.50) and (7.58) a relation between the tensor of second rank $t^{\nu\kappa}$ and the spin angular momentum tensor $S^{\mu\nu\kappa}$

$$t^{\nu\kappa}(x^\lambda) - t^{\kappa\nu}(x^\lambda) = \partial_\mu S^{\mu\nu\kappa}(x^\lambda) \,. \tag{7.60}$$

In order to solve (7.60) for the tensor of second rank $t^{\nu\kappa}$ we perform the ansatz that it follows from the four-divergence of a tensor of third rank $\chi^{\mu\nu\kappa}$

$$t^{\nu\kappa}(x^\lambda) = \partial_\mu \chi^{\mu\nu\kappa}(x^\lambda) \,, \tag{7.61}$$

where the tensor of third rank $\chi^{\mu\nu\kappa}$ is anti-symmetric with respect to its first two indices, i.e.

$$\chi^{\mu\nu\kappa}(x^\lambda) = -\chi^{\nu\mu\kappa}(x^\lambda) . \tag{7.62}$$

In case that such a tensor of third rank $\chi^{\mu\nu\kappa}$ would exist, the symmetric energy-momentum tensor $T^{\nu\kappa}(x^\lambda)$ would be physically equivalent to the canonical energy-momentum tensor $\Theta^{\nu\kappa}(x^\lambda)$. On the one hand we read off from (7.58), (7.61), and (7.62) that the differential energy-momentum conservation law (7.34) for the canonical energy-momentum tensor $\Theta^{\nu\kappa}(x^\lambda)$ implies a similar one for the symmetrized energy-momentum tensor $T^{\nu\kappa}(x^\lambda)$:

$$\partial_\nu T^{\nu\kappa}(x^\lambda) = 0 . \tag{7.63}$$

On the other hand it also follows from (7.58), (7.61), and (7.62) that the conserved quantities of energy and momentum following from the canonical energy-momentum tensor $\Theta^{\nu\kappa}(x^\lambda)$ agree with the ones from the symmetrized energy-momentum tensor $T^{\nu\kappa}(x^\lambda)$

$$\int d^3x\, T^{0\kappa}(x^\lambda) = \int d^3x\, \Theta^{0\kappa}(x^\lambda) , \tag{7.64}$$

as the Gauß law implies

$$\int d^3x\, \partial_j \chi^{j0\kappa}(x^\lambda) = 0 . \tag{7.65}$$

Inserting the ansatz (7.61) into (7.60) one obtains the following relation between the tensor $\chi^{\mu\nu\kappa}$ and the spin angular momentum tensor $S^{\mu\nu\kappa}$:

$$\chi^{\mu\nu\kappa}(x^\lambda) - \chi^{\mu\kappa\nu}(x^\lambda) = S^{\mu\nu\kappa}(x^\lambda) . \tag{7.66}$$

Now we uniquely determine the tensor $\chi^{\mu\nu\kappa}$ by taking into account (7.62) and (7.66). To this end we decompose the tensor $\chi^{\mu\nu\kappa}$ via

$$\chi^{\mu\nu\kappa}(x^\lambda) = \chi_{\rm s}^{\mu\nu\kappa}(x^\lambda) + \chi_{\rm a}^{\mu\nu\kappa}(x^\lambda) \tag{7.67}$$

into the two tensors $\chi_{\rm s}^{\mu\nu\kappa}$ and $\chi_{\rm a}^{\mu\nu\kappa}$, which are symmetric and anti-symmetric with respect to the indices ν, κ, respectively:

$$\chi_{\rm s}^{\mu\nu\kappa}(x^\lambda) = \chi_{\rm s}^{\mu\kappa\nu}(x^\lambda) , \tag{7.68}$$

$$\chi_{\rm a}^{\mu\nu\kappa}(x^\lambda) = -\chi_{\rm a}^{\mu\kappa\nu}(x^\lambda) . \tag{7.69}$$

Inserting (7.67)–(7.69) into (7.66) the tensor $\chi_{\rm s}^{\mu\nu\kappa}$ drops out and the tensor $\chi_{\rm a}^{\mu\nu\kappa}$ follows to be

$$\chi_{\rm a}^{\mu\nu\kappa}(x^\lambda) = \frac{1}{2}\, S^{\mu\nu\kappa}(x^\lambda) . \tag{7.70}$$

Here the anti-symmetry (7.69) of the tensor $\chi_{\rm a}^{\mu\nu\kappa}$ is guaranteed due to the anti-symmetry (7.49) of the spin angular momentum tensor $S^{\mu\nu\kappa}$. Taking into account (7.67)–(7.70) one deduces from (7.62)

$$\chi_{\rm s}^{\mu\nu\kappa}(x^\lambda) + \chi_{\rm s}^{\nu\mu\kappa}(x^\lambda) = -\frac{1}{2}\, S^{\mu\nu\kappa}(x^\lambda) - \frac{1}{2}\, S^{\nu\mu\kappa}(x^\lambda) . \tag{7.71}$$

Due to (7.49) we find that (7.71) is straightforwardly solved by

$$\chi_{\mathrm{s}}^{\mu\nu\kappa}(x^\lambda) = -\frac{1}{2}\,S^{\nu\mu\kappa}(x^\lambda) - \frac{1}{2}\,S^{\kappa\mu\nu}(x^\lambda)\,. \tag{7.72}$$

And the tensor of third rank $\chi^{\mu\nu\kappa}$ follows, finally, from combining (7.49), (7.67), (7.70), and (7.72):

$$\chi^{\mu\nu\kappa}(x^\lambda) = \frac{1}{2}\left[S^{\mu\nu\kappa}(x^\lambda) + S^{\nu\kappa\mu}(x^\lambda) - S^{\kappa\mu\nu}(x^\lambda)\right]. \tag{7.73}$$

7.9 Modified Angular Momentum Tensor

Finally, we show for the sake of completeness that the symmetrization of the energy-momentum tensor also leads to a simplied angular momentum tensor. To this end we consider a modified angular momentum tensor $I^{\mu\nu\kappa}$, which follows from the symmetrized energy-momentum tensor $T^{\nu\kappa}$ in the same way as the orbital angular momentum tensor (7.47):

$$I^{\mu\nu\kappa}(x^\lambda) = T^{\mu\kappa}(x^\lambda)\,x^\nu - T^{\mu\nu}(x^\lambda)\,x^\kappa\,. \tag{7.74}$$

Combining (7.45), (7.49), (7.58), (7.61), (7.73), and (7.74) it turns out that the canonical angular momentum tensor $J^{\mu\nu\kappa}$ and the modified angular momentum temsor $I^{\mu\nu\kappa}$ differ by the four-divergence of a tensor of fourth rank $\eta^{\rho\mu\nu\kappa}$:

$$I^{\mu\nu\kappa}(x^\lambda) = J^{\mu\nu\kappa}(x^\lambda) + \partial_\rho\eta^{\rho\mu\nu\kappa}(x^\lambda)\,. \tag{7.75}$$

Here the tensor of fourth rank $\eta^{\rho\mu\nu\kappa}$ turns out to be

$$\eta^{\rho\mu\nu\kappa}(x^\lambda) = \chi^{\rho\mu\kappa}(x^\lambda)x^\nu - \chi^{\rho\mu\nu}(x^\lambda)x^\kappa\,, \tag{7.76}$$

which is anti-symmetric with respect to its first two indices due to (7.62):

$$\eta^{\rho\mu\nu\kappa}(x^\lambda) = -\eta^{\mu\rho\nu\kappa}(x^\lambda)\,. \tag{7.77}$$

Therefore, both angular momentum tensors $I^{\mu\nu\kappa}$ and $J^{\mu\nu\kappa}$ are physically equivalent. On the one hand we read off from (7.75) and (7.77) that the differential angular momentum conservation law (7.44) for the canonical angular momentum tensor $J^{\mu\nu\kappa}(x^\lambda)$ implies a similar one for the modified angular momentum tensor $I^{\mu\nu\kappa}(x^\lambda)$:

$$\partial_\mu I^{\mu\nu\kappa}(x^\lambda) = 0\,. \tag{7.78}$$

On the other hand it also follows from (7.75), (7.77), and the Gauß law that the conserved angular momenta (7.51) following from the canonical angular momentum tensor $J^{\mu\nu\kappa}(x^\lambda)$ agree with the ones from the modified angular momentum tensor $I^{\mu\nu\kappa}(x^\lambda)$:

$$\int d^3x\, J^{0\nu\kappa}(x^\lambda) = \int d^3x\, I^{0\nu\kappa}(x^\lambda)\,. \tag{7.79}$$

Chapter 8

Klein-Gordon Field

The first relativistic quantum field, which we deal with here, is the Klein-Gordon field. It represents a free scalar field and describes in its second-quantized form particles with spin 0. One example for such particles within the realm of the standard model of elementary particles is the Higgs particle H, which is electrically neutral and gives all particles their mass due to its interaction with them. Another example is provided by the pions, which were originally introduced by Hideki Yukawa as the exchange particles giving rise to the nuclear force. There exists a neutral pion π^0 and two charged pions, namely π^+ and its antiparticle π^-. Note that the pions turned out to be the lightest mesons, i.e. they consist of two quarks. Therefore, they are unstable, decay via weak or electromagnetic interaction, and are considered nowadays no longer as elementary particles.

Coupling the charged pions minimally to the electromagnetic field yields a theory, which is called scalar electrodynamics. In its second quantized form it microscopically describes the interaction between charged pions due to the exchange of photons. From a pedagogical point of view it would be reasonable to introduce scalar QED before QED as the description of matter by the Klein-Gordon theory is much simpler than the Dirac theory. Therefore, starting with scalar QED would make it easier to understand several technical issues as, for instance, the Feynman diagrams of QED without having to deal with the intricate spinor algebra of the Dirac theory. Another motivation to study scalar electrodynamics would be that it represents the relativistic generalization of the Ginzburg-Landau theory of superconductivity. However, due to time constraints, we will not be able to work out scalar electrodynamics, so here we can only refer the interested reader to the relevant literature.

8.1 Action and Equations of Motions

The action of the Schrödinger fields $\psi(\mathbf{x}, t)$ and $\psi^*(\mathbf{x}, t)$ in (4.8)–(4.10) is not invariant with respect to Lorentz transformations as it contains partial derivatives of first (second) order with respect to the time (space) coordinate(s). In contrast to that a relativistic action must treat

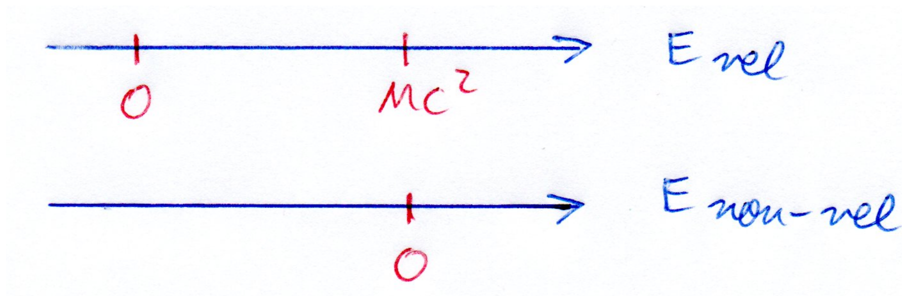


Figure 8.1: Relativistic and non-relativistic energy scales differ according to (8.5) by the rest energy Mc^2.

temporal and spatial partial derivatives on an equal footing. Depending on the respective internal spin degrees of freedom there are different ways how to convert the non-relativistic Schrödinger action (4.8)–(4.10) into a relativistic one.

In the following we deal with charged relativistic particles like the pions π^+ and π^-, which do not have any internal spin degree of freedom. Such particles are described by scalar fields $\Psi(x^\lambda)$ and $\Psi^*(x^\lambda)$. The corresponding action

$$\mathcal{A} = \mathcal{A}\left[\Psi^*(\bullet);\Psi(\bullet)\right] \tag{8.1}$$

is defined by a spatio-temporal integral over the Lagrange density according to

$$\mathcal{A} = \frac{1}{c}\int d^4x\,\mathcal{L}\Big(\Psi^*(x^\lambda),\partial_\mu\Psi^*(x^\lambda);\Psi(x^\lambda),\partial_\mu\Psi(x^\lambda)\Big)\,, \tag{8.2}$$

where we have $d^4x = cdt\,d^3x$. The Lagrange density of the Klein-Gordon fields is given by the real-valued Lorentz invariant

$$\mathcal{L} = A\,g^{\mu\nu}\,\partial_\mu\Psi^*(x^\lambda)\,\partial_\nu\Psi(x^\lambda) + B\,\Psi^*(x^\lambda)\Psi(x^\lambda)\,. \tag{8.3}$$

In the following we choose the yet unknown constants A and B in such a way that the Lagrange density of the Klein-Gordon fields (8.3) goes over in the non-relativistic limit into the Lagrange density (4.10) of the Schrödinger fields. To this end we decompose at first the derivatives in (8.3) into their respective temporal and spatial contributions:

$$\mathcal{L} = A\left[\frac{1}{c^2}\frac{\partial\Psi^*(\mathbf{x},t)}{\partial t}\frac{\partial\Psi(\mathbf{x},t)}{\partial t} - \boldsymbol{\nabla}\,\Psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi(\mathbf{x},t)\right] + B\,\Psi^*(\mathbf{x},t)\,\Psi(\mathbf{x},t)\,. \tag{8.4}$$

Performing the transition from a relativistic to the corresponding non-relativistic theory one has to take into account that the corresponding energy scales are shifted by the rest energy Mc^2 of the particles with mass M with respect to each other as is illustrated in Fig. 8.1:

$$E_{\rm rel} = E_{\rm non-rel} + Mc^2\,. \tag{8.5}$$

This becomes apparent from Fig. 6.1, where the relativistic dispersion relation is compared with its non-relativistic limit. As a quantum mechanical wave function depends exponentially via $e^{-iEt/\hbar}$ from the energy E, (8.5) suggests to perform the separation ansatz

$$\Psi(\mathbf{x},t) = e^{-iMc^2t/\hbar}\,\psi(\mathbf{x},t)\,, \tag{8.6}$$
$$\Psi^*(\mathbf{x},t) = e^{iMc^2t/\hbar}\,\psi^*(\mathbf{x},t)\,. \tag{8.7}$$

Inserting (8.6), (8.7) into the Lagrange density of the Klein-Gordon fields (8.4), we obtain

$$\begin{aligned}\mathcal{L} &= \frac{A}{c^2}\left\{\frac{\partial\psi^*(\mathbf{x},t)}{\partial t}\,\frac{\partial\psi(\mathbf{x},t)}{\partial t} + \frac{i}{\hbar}\,Mc^2\left[\psi^*(\mathbf{x},t)\,\frac{\partial\psi(\mathbf{x},t)}{\partial t} - \frac{\partial\psi^*(\mathbf{x},t)}{\partial t}\,\psi(\mathbf{x},t)\right]\right\}\\ &\quad -A\,\boldsymbol{\nabla}\psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\psi(\mathbf{x},t) + \left(B + \frac{M^2c^2}{\hbar^2}\,A\right)\,\psi^*(\mathbf{x},t)\,\psi(\mathbf{x},t)\,.\end{aligned} \tag{8.8}$$

In the non-relativistic limit $c \to \infty$ we have now to guarantee that (8.8) reduces term by term to (4.10):

- Due to a partial integration in time the second and third term in (8.8) can be merged. A comparison with (4.10) then fixes the constant A:

$$\frac{2Mi}{\hbar}\,A = i\hbar \qquad \Longrightarrow \qquad A = \frac{\hbar^2}{2M}\,. \tag{8.9}$$

- With this choice of A the first term in (8.8) vanishes in the non-relativistic limit $c \to \infty$ and the fourth term turns out to yield the correct kinetic energy of the Schrödinger field.

- The last term in (8.8) must vanish as the Schrödinger field does not have such a mass term, so also the constant B is determined by taking into account (8.9):

$$B = -\frac{M^2c^2}{\hbar^2}\,A \qquad \Longrightarrow \qquad B = -\frac{1}{2}\,Mc^2\,. \tag{8.10}$$

Inserting (8.9) and (8.10) into (8.4) the action of the Klein-Gordon field

$$\mathcal{A} = \mathcal{A}\left[\Psi^*(\bullet,\bullet);\Psi(\bullet,\bullet)\right] \tag{8.11}$$

is given by a spatio-temporal integral

$$\mathcal{A} = \int dt \int d^3x\,\mathcal{L}\Big(\Psi^*(\mathbf{x},t),\boldsymbol{\nabla}\Psi^*(\mathbf{x},t),\frac{\partial\Psi^*(\mathbf{x},t),}{\partial t};\Psi(\mathbf{x},t),\boldsymbol{\nabla}\Psi(\mathbf{x},t),\frac{\partial\Psi(\mathbf{x},t)}{\partial t}\Big) \tag{8.12}$$

with the Lagrange density

$$\mathcal{L} = \frac{\hbar^2}{2Mc^2}\,\frac{\partial\Psi^*(\mathbf{x},t)}{\partial t}\,\frac{\partial\Psi(\mathbf{x},t)}{\partial t} - \frac{\hbar^2}{2M}\,\boldsymbol{\nabla}\,\Psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi(\mathbf{x},t) - \frac{Mc^2}{2}\,\Psi^*(\mathbf{x},t)\,\Psi(\mathbf{x},t)\,. \tag{8.13}$$

Similar to the discussion of the Schrödinger fields in Section 4.4 the Hamilton principle of classical field theory

$$\frac{\delta\mathcal{A}}{\delta\Psi^*(\mathbf{x},t)} = 0\,, \qquad \frac{\delta\mathcal{A}}{\delta\Psi(\mathbf{x},t)} = 0 \tag{8.14}$$

leads to the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \Psi^*(\mathbf{x},t)} - \boldsymbol{\nabla} \frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\Psi^*(\mathbf{x},t)} - \frac{\partial}{\partial t} \frac{\partial \mathcal{L}}{\partial \frac{\partial \Psi^*(\mathbf{x},t)}{\partial t}} = 0\,, \tag{8.15}$$

$$\frac{\partial \mathcal{L}}{\partial \Psi(\mathbf{x},t)} - \boldsymbol{\nabla} \frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\Psi(\mathbf{x},t)} - \frac{\partial}{\partial t} \frac{\partial \mathcal{L}}{\partial \frac{\partial \Psi(\mathbf{x},t)}{\partial t}} = 0\,. \tag{8.16}$$

In order to evaluate (8.15), (8.16) we need the following partial derivatives from the Lagrange density (8.13):

$$\frac{\partial \mathcal{L}}{\partial \Psi^*(\mathbf{x},t)} = -\frac{1}{2}\, Mc^2 \Psi(\mathbf{x},t)\,, \frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\Psi^*(\mathbf{x},t)} = -\frac{\hbar^2}{2M}\, \boldsymbol{\nabla}\Psi(\mathbf{x},t)\,, \frac{\partial \mathcal{L}}{\partial \frac{\partial \Psi^*(\mathbf{x},t)}{\partial t}} = \frac{\hbar^2}{2Mc^2} \frac{\partial \Psi(\mathbf{x},t)}{\partial t}\,, \tag{8.17}$$

$$\frac{\partial \mathcal{L}}{\partial \Psi(\mathbf{x},t)} = -\frac{1}{2}\, Mc^2 \Psi^*(\mathbf{x},t)\,, \frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\Psi(\mathbf{x},t)} = -\frac{\hbar^2}{2M}\, \boldsymbol{\nabla}\Psi^*(\mathbf{x},t)\,, \frac{\partial \mathcal{L}}{\partial \frac{\partial \Psi(\mathbf{x},t)}{\partial t}} = \frac{\hbar^2}{2Mc^2} \frac{\partial \Psi^*(\mathbf{x},t)}{\partial t}\,. \tag{8.18}$$

Inserting the additional calculation (8.17) and (8.18) into the Euler-Lagrange equations (8.15), (8.16), we obtain the Klein-Gordon equations for the fields $\Psi(\mathbf{x},t)$ and $\Psi^*(\mathbf{x},t)$:

$$\frac{1}{c^2} \frac{\partial^2 \Psi(\mathbf{x},t)}{\partial t^2} - \boldsymbol{\nabla}^2 \Psi(\mathbf{x},t) + \frac{M^2c^2}{\hbar^2}\, \Psi(\mathbf{x},t) = 0\,, \tag{8.19}$$

$$\frac{1}{c^2} \frac{\partial^2 \Psi^*(\mathbf{x},t)}{\partial t^2} - \boldsymbol{\nabla}^2 \Psi^*(\mathbf{x},t) + \frac{M^2c^2}{\hbar^2}\, \Psi^*(\mathbf{x},t) = 0\,. \tag{8.20}$$

They represent wave equations, which contain an additional term due to the finiteness of the Compton wave length of the particles

$$\lambda_{\mathrm{C}} = 2\pi\, \frac{\hbar}{Mc}\,. \tag{8.21}$$

For a pion π^+ or π^- with the rest energy $Mc^2 = 139.6$ MeV the Compton wave length (8.21) amounts to $\lambda_{\mathrm{C}} \approx 9$ fm, which is of the order of magnitude of the size of the atomic nucleus.

The appearance of the Compton wave length (8.21) can be physically understood as follows. A relativistic particle with the momentum uncertainty $\Delta p = Mc$ yields via the Heisenberg uncertainty relation a corresponding spatial uncertainty

$$\Delta x = \frac{\hbar}{Mc}\,, \tag{8.22}$$

which is of the order of the Compton wave length (8.21). Wherever a relativistic particle is confined to a region, which is of the order of the Compton wave length, the resulting energy uncertainty becomes so large that particle-antiparticle pairs are generated out of the vacuum. This peculiar phenomenon is best illustrated by the Klein paradox, which arises for a pion π^+ running against a potential step of height V, see Fig. 8.2. Provided that the potential height V reaches the order of the rest energy $2Mc^2$ of two pions, the wave function falls off exponentially in the region of the potential threshold. This then leads to the generation of particle-antiparticle pairs, which have to move due to momentum conservation in opposite directions. As a consequence, one observes within the potential threshold a negative charge

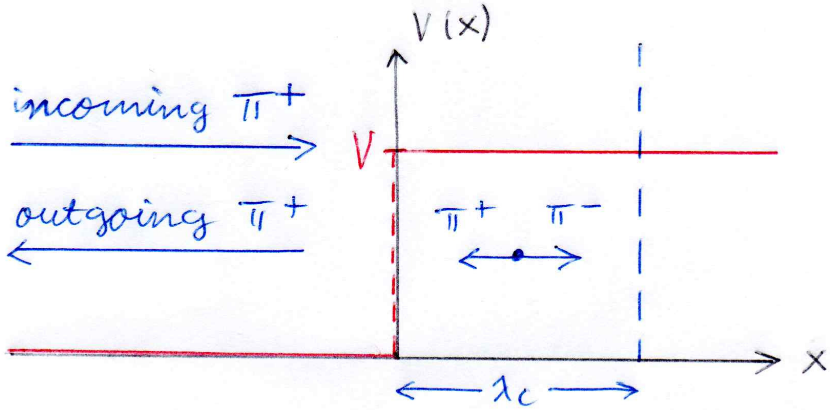


Figure 8.2: The scattering of a pion π^+ at a potential threshold with height $V \sim 2Mc^2$ leads to the Klein paradox that the reflection coefficient becomes larger than one. This is due to the creation of particle-antiparticle pairs within a region, which has the extension of the Compton wave length (8.22).

density, so that the situation emerges as depicted in Fig. 8.2. Surprisingly, this leads to a reflection coefficient of this one-particle scattering problem, which is larger than one. The Klein paradox has, therefore, the consequence that a relativistic quantum theory can never be restricted to a one-particle theory. Instead, it has to be extended to a relativistic quantum field theory in order to incorporate adequately the inherent many-body phenomena. Inserting the ansatz (8.6), (8.7) in the Klein-Gordon equations (8.19), (8.20) for the wave functions $\Psi(\mathbf{x}, t)$, $\Psi^*(\mathbf{x}, t)$, we obtain

$$\frac{1}{c^2}\frac{\partial^2\psi(\mathbf{x},t)}{\partial t^2} - \frac{2iM}{\hbar}\frac{\partial\psi(\mathbf{x},t)}{\partial t} - \boldsymbol{\nabla}^2\psi(\mathbf{x},t) = 0\,, \tag{8.23}$$

$$\frac{1}{c^2}\frac{\partial^2\psi^*(\mathbf{x},t)}{\partial t^2} + \frac{2iM}{\hbar}\frac{\partial\psi^*(\mathbf{x},t)}{\partial t} - \boldsymbol{\nabla}^2\psi^*(\mathbf{x},t) = 0\,. \tag{8.24}$$

In the non-relativistic limit $c \to \infty$ both (8.23) and (8.24) go over into the corresponding Schrödinger equations for the wave functions $\psi(\mathbf{x}, t)$ and $\psi^*(\mathbf{x}, t)$, as expected:

$$i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} = -\frac{\hbar^2}{2M}\boldsymbol{\nabla}^2\psi(\mathbf{x},t)\,, \tag{8.25}$$

$$-i\hbar\frac{\partial\psi^*(\mathbf{x},t)}{\partial t} = -\frac{\hbar^2}{2M}\boldsymbol{\nabla}^2\,\psi^*(\mathbf{x},t)\,. \tag{8.26}$$

Note that, historically, Erwin Schrödinger discovered on his quest for a quantum mechanical wave equation in 1926 at first the Klein-Gordon equation. But solving this relativistic wave equation for the example of the Coulomb potential he found that the resulting energy eigenvalues disagreed with the measured spectral lines of the hydrogen atom. In retrospect we know that this is due to the fact that the Klein-Gordon equation does not take into account the spin 1/2 degree of freedom of the electron in the hydrogen atom. Due to this discrepancy he abandoned the Klein-Gordon equation and derived instead in the non-relativistic limit the

Schrödinger equation, where he obtained a much better agreement between the corresponding solution of the Coulomb problem and the measured spectral lines of the hydrogen atom.

8.2 Continuity Equation

Now we multiply (8.19) with $\Psi^*(\mathbf{x},t)$ and (8.20) with $\Psi(\mathbf{x},t)$ and subtract both from each other, yielding at first

$$\frac{1}{c^2}\Psi^*(\mathbf{x},t)\frac{\partial^2\Psi(\mathbf{x},t)}{\partial t^2} - \frac{1}{c^2}\Psi(\mathbf{x},t)\frac{\partial^2\Psi^*(\mathbf{x},t)}{\partial t^2} - \Psi^*(\mathbf{x},t)\boldsymbol{\nabla}^2\Psi(\mathbf{x},t) + \Psi(\mathbf{x},t)\boldsymbol{\nabla}^2\Psi^*(\mathbf{x},t) = 0\,, \tag{8.27}$$

where the mass terms have dropped out. This can be recast into the form

$$\frac{\partial}{\partial t}\left[\Psi^*(\mathbf{x},t)\,\frac{\partial\Psi(\mathbf{x},t)}{\partial t} - \frac{\partial\Psi^*(\mathbf{x},t)}{\partial t}\,\Psi(\mathbf{x},t)\right] + \boldsymbol{\nabla}\Big[\Psi(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi^*(\mathbf{x},t) - \Psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi(\mathbf{x},t)\Big] = 0\,, \tag{8.28}$$

which corresponds to a continuity equation:

$$\frac{\partial\rho(\mathbf{x},t)}{\partial t} + \boldsymbol{\nabla}\mathbf{j}(\mathbf{x},t) = 0\,. \tag{8.29}$$

Here both density $\rho(\mathbf{x},t)$ and current density $\mathbf{j}(\mathbf{x},t)$ are only determined up to a yet unknown constant K:

$$\rho(\mathbf{x},t) = \frac{K}{c^2}\left[\Psi^*(\mathbf{x},t)\,\frac{\partial\Psi(\mathbf{x},t)}{\partial t} - \frac{\partial\Psi^*(\mathbf{x},t)}{\partial t}\,\Psi(\mathbf{x},t)\right], \tag{8.30}$$

$$\mathbf{j}(\mathbf{x},t) = K\Big[\Psi(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi^*(\mathbf{x},t) - \Psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi(\mathbf{x},t)\Big]\,. \tag{8.31}$$

The constant K can now be fixed uniquely by considering the non-relativistic limit $c\to\infty$. To this end one inserts the ansatz (8.19), (8.20) into (8.30), (8.31) and gets

$$\rho(\mathbf{x},t) = \frac{K}{c^2}\left[\psi^*(\mathbf{x},t)\,\frac{\partial\psi(\mathbf{x},t)}{\partial t} - \frac{\partial\psi^*(\mathbf{x},t)}{\partial t}\,\psi(\mathbf{x},t) - \frac{2iMc^2}{\hbar}\,\psi^*(\mathbf{x},t)\psi(\mathbf{x},t)\right], \tag{8.32}$$

$$\mathbf{j}(\mathbf{x},t) = K\Big[\psi(\mathbf{x},t)\,\boldsymbol{\nabla}\psi^*(\mathbf{x},t) - \psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\psi(\mathbf{x},t)\Big]\,. \tag{8.33}$$

We have then to demand that (8.32), (8.33) go over in the non-relativistic limit $c\to\infty$ to the corresponding non-relativistic expressions:

$$\rho(\mathbf{x},t) = \psi^*(\mathbf{x},t)\psi(\mathbf{x},t)\,, \tag{8.34}$$

$$\mathbf{j}(\mathbf{x},t) = \frac{i\hbar}{2M}\Big[\psi(\mathbf{x},t)\,\boldsymbol{\nabla}\psi^*(\mathbf{x},t) - \psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\psi(\mathbf{x},t)\Big]\,. \tag{8.35}$$

This fixes the constant K to the value

$$K = \frac{i\hbar}{2M}\,. \tag{8.36}$$

Thus, we obtain from (8.30), (8.31), and (8.36) for the density $\rho(\mathbf{x},t)$ and the current density $\mathbf{j}(\mathbf{x},t)$ of the Klein-Gordon fields

$$\rho(\mathbf{x},t) = \frac{i\hbar}{2Mc^2}\left[\Psi^*(\mathbf{x},t)\,\frac{\partial\Psi(\mathbf{x},t)}{\partial t} - \frac{\partial\Psi^*(\mathbf{x},t)}{\partial t}\,\Psi(\mathbf{x},t)\right], \tag{8.37}$$

$$\mathbf{j}(\mathbf{x},t) = \frac{i\hbar}{2M}\left[\Psi(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi^*(\mathbf{x},t) - \Psi^*(\mathbf{x},t)\,\boldsymbol{\nabla}\Psi(\mathbf{x},t)\right]. \tag{8.38}$$

From the continuity equation (8.29) follows the existence of the conserved quantity. Namely, considering the time derivative of the quantity

$$Q = \int d^3x\,\rho(\mathbf{x},t)\,, \tag{8.39}$$

we obtain from (8.29) and applying the theorem of Gauß

$$\frac{\partial Q}{\partial t} = -\oint d\mathbf{f}\cdot\mathbf{j}(\mathbf{x},t)\,. \tag{8.40}$$

Here the surface integral at infinity vanishes as the fields $\Psi^*(\mathbf{x},t)$, $\Psi(\mathbf{x},t)$ as well as the current density $\mathbf{j}(\mathbf{x},t)$ in (8.38) are assumed to vanish fast enough at infinity, yielding

$$\frac{\partial Q}{\partial t} = 0\,. \tag{8.41}$$

Now it turns out to be useful to define a scalar product between two arbitrary fields $\Psi_1(\mathbf{x},t)$ and $\Psi_2(\mathbf{x},t)$ according to

$$\langle\Psi_1,\Psi_2\rangle = \frac{i\hbar}{2Mc^2}\int d^3x\left[\Psi_1^*(\mathbf{x},t)\,\frac{\partial\Psi_2(\mathbf{x},t)}{\partial t} - \frac{\partial\Psi_1^*(\mathbf{x},t)}{\partial t}\,\Psi_2(\mathbf{x},t)\right]. \tag{8.42}$$

But note that this scalar product is not positive definite. For instance, choosing the ansatz

$$\Psi_1(\mathbf{x},t) = \Psi_2(\mathbf{x},t) = N\,e^{iMc^2t/\hbar} \tag{8.43}$$

we obtain

$$\langle\Psi_1,\Psi_2\rangle = -N^2 < 0\,. \tag{8.44}$$

In order to investigate the non-relativistic limit of this scalar product, we insert (8.6), (8.7) into (8.42):

$$\langle\Psi_1,\Psi_2\rangle = \frac{i\hbar}{2Mc^2}\int d^3x\left[\psi_1^*(\mathbf{x},t)\frac{\partial\psi_2(\mathbf{x},t)}{\partial t} - \frac{\partial\psi_1(\mathbf{x},t)}{\partial t}\psi_2(\mathbf{x},t) - \frac{2iMc^2}{\hbar}\psi_1^*(\mathbf{x},t)\psi_2(\mathbf{x},t)\right]. \tag{8.45}$$

Performing the limit $c\to\infty$, we conclude

$$\langle\Psi_1,\Psi_2\rangle = \lim_{c\to\infty}\langle\Psi_1,\Psi_2\rangle = \int d^3x\,\psi_1^*(\mathbf{x},t)\psi_2(\mathbf{x},t)\,, \tag{8.46}$$

which is just the positive definite scalar product used in the Schrödinger theory. Thus, from the fact, that the scalar products of the Klein-Gordon and the Schrödinger theory differ, we read off that each quantum field theory has its own natural scalar product. It turns out that this conclusion has far-reaching consequences, as the natural scalar product of a quantum field theory represents a central technical tool. For instance, in the present case of the Klein-Gordon theory, taking into account (8.37) we finally obtain a useful relation between the conserved quantity (8.42) and the scalar product (8.39):

$$Q = \langle \Psi, \Psi \rangle \,. \tag{8.47}$$

As the scalar product is not positive definite, the conserved quantity can have both positive and negative values. This makes it possible to identify Q, or more precisely eQ with the electric charge of a complex-valued Klein-Gordon field, where e denotes the elementary charge. Furthermore, we conclude that a real-valued Klein-Gordon field, where $\Psi^*(\mathbf{x},t) = \Psi(\mathbf{x},t)$ holds, leads to a vanishing charge Q due to (8.42) and (8.47).

8.3 Canonical Field Quantization

The two independent Klein-Gordon fields $\Psi^*(\mathbf{x},t)$ and $\Psi(\mathbf{x},t)$ have the two following two canonically conjugated momentum fields:

$$\Pi^*(\mathbf{x},t) = \frac{\partial \mathcal{L}}{\frac{\partial \Psi^*(\mathbf{x},t)}{\partial t}} = \frac{\hbar^2}{2Mc^2}\,\frac{\partial \Psi(\mathbf{x},t)}{\partial t}\,, \tag{8.48}$$

$$\Pi(\mathbf{x},t) = \frac{\partial \mathcal{L}}{\frac{\partial \Psi(\mathbf{x},t)}{\partial t}} = \frac{\hbar^2}{2Mc^2}\,\frac{\partial \Psi^*(\mathbf{x},t)}{\partial t}\,, \tag{8.49}$$

where $\mathcal{L}$ denotes the Lagrange density of the Klein-Gordon field from (8.13). With the help of a Legendre transformation we then obtain the Hamilton density from the Lagrange density:

$$\mathcal{H} = \Pi^*(\mathbf{x},t)\,\frac{\partial \Psi^*(\mathbf{x},t)}{\partial t} + \Pi(\mathbf{x},t)\,\frac{\partial \Psi(\mathbf{x},t)}{\partial t} - \mathcal{L}\,. \tag{8.50}$$

Inserting therein (8.13) together with (8.48), (8.49) this yields

$$\mathcal{H} = \frac{2Mc^2}{\hbar^2}\,\Pi^*(\mathbf{x},t)\Pi(\mathbf{x},t) + \frac{\hbar^2}{2M}\,\boldsymbol{\nabla}\Psi^*(\mathbf{x},t)\boldsymbol{\nabla}\Psi(\mathbf{x},t) + \frac{Mc^2}{2}\,\Psi^*(\mathbf{x},t)\Psi(\mathbf{x},t)\,. \tag{8.51}$$

The Hamilton function H of the charged Klein-Gordon field then follows from spatially integrating this Hamilton density $\mathcal{H}$:

$$H = \int d^3x\,\mathcal{H}\,. \tag{8.52}$$

With this one can perform a canonical field quantization along the lines outlined in Chapter 4. For the sake of brevity we do not work this out in detail for the Klein-Gordon field but mention instead the result. At first, one assigns to the classical fields $\Psi^*(\mathbf{x},t)$, $\Psi(\mathbf{x},t)$, $\Pi^*(\mathbf{x},t)$, and $\Pi(\mathbf{x},t)$ corresponding second-quantized operators $\hat{\Psi}^\dagger(\mathbf{x},t)$, $\hat{\Psi}(\mathbf{x},t)$, $\hat{\Pi}^\dagger(\mathbf{x},t)$, and $\hat{\Pi}(\mathbf{x},t)$.

Due to the spin-statistic theorem of Pauli one performs for the Klein-Gordon field a bosonic field quantization and obtains between both $\hat{\Psi}(\mathbf{x},t)$, $\hat{\Pi}(\mathbf{x},t)$ and $\hat{\Psi}^\dagger(\mathbf{x},t)$, $\hat{\Pi}^\dagger(\mathbf{x},t)$ equal-time canonical commutation relations:

$$\left[\hat{\Psi}(\mathbf{x},t),\hat{\Psi}(\mathbf{x}',t)\right]_- = \left[\hat{\Pi}(\mathbf{x},t),\hat{\Pi}(\mathbf{x}',t)\right]_- = 0, \quad \left[\hat{\Psi}(\mathbf{x},t),\hat{\Pi}(\mathbf{x}',t)\right]_- = i\hbar\,\delta(\mathbf{x}-\mathbf{x}')\,, \tag{8.53}$$

$$\left[\hat{\Psi}^\dagger(\mathbf{x},t),\hat{\Psi}^\dagger(\mathbf{x}',t)\right]_- = \left[\hat{\Pi}^\dagger(\mathbf{x},t),\hat{\Pi}^\dagger(\mathbf{x}',t)\right]_- = 0, \quad \left[\hat{\Psi}^\dagger(\mathbf{x},t),\hat{\Pi}^\dagger(\mathbf{x}',t)\right]_- = i\hbar\,\delta(\mathbf{x}-\mathbf{x}')\,. \tag{8.54}$$

Due to the independence of the quantized degrees of freedom all mixed equal-time commutator relations vanish:

$$\begin{aligned}\left[\hat{\Psi}(\mathbf{x},t),\hat{\Psi}^\dagger(\mathbf{x}',t)\right]_- = \left[\hat{\Psi}(\mathbf{x},t),\hat{\Pi}^\dagger(\mathbf{x}',t)\right]_- &= 0\,,\\ \left[\hat{\Pi}(\mathbf{x},t),\hat{\Psi}^\dagger(\mathbf{x}',t)\right]_- = \left[\hat{\Pi}(\mathbf{x},t),\hat{\Pi}^\dagger(\mathbf{x}',t)\right]_- &= 0.\end{aligned} \tag{8.55}$$

Furthermore, the canonical field quantization converts the classical Hamilton function (8.51), (8.52) to the Hamilton operator:

$$\hat{H} = \int d^3x \left[\frac{2Mc^2}{\hbar^2}\,\hat{\Pi}^\dagger(\mathbf{x},t)\hat{\Pi}(\mathbf{x},\mathbf{t}) + \frac{\hbar^2}{2M}\,\boldsymbol{\nabla}\hat{\Psi}^\dagger(\mathbf{x},t)\boldsymbol{\nabla}\hat{\Psi}(\mathbf{x},t) + \frac{Mc^2}{2}\,\hat{\Psi}^\dagger(\mathbf{x},\mathbf{t})\hat{\Psi}(\mathbf{x},t)\right]. \tag{8.56}$$

Note that the respective order of the operators in (8.56) does not play a role due to (8.55). With the Hamilton operator we then obtain the following Heisenberg equations:

$$i\hbar\,\frac{\partial\hat{\Psi}(\mathbf{x},t)}{\partial t} = \left[\hat{\Psi}(\mathbf{x},t),\hat{H}\right]_- \quad\Longrightarrow\quad \frac{\partial\hat{\Psi}(\mathbf{x},t)}{\partial t} = \frac{2Mc^2}{\hbar^2}\,\hat{\Pi}^\dagger(\mathbf{x},t)\,, \tag{8.57}$$

$$i\hbar\,\frac{\partial\hat{\Psi}^\dagger(\mathbf{x},t)}{\partial t} = \left[\hat{\Psi}^\dagger(\mathbf{x},t),\hat{H}\right]_- \quad\Longrightarrow\quad \frac{\partial\hat{\Psi}^\dagger(\mathbf{x},t)}{\partial t} = \frac{2Mc^2}{\hbar^2}\,\hat{\Pi}(\mathbf{x},t)\,, \tag{8.58}$$

$$i\hbar\,\frac{\partial\hat{\Pi}(\mathbf{x},t)}{\partial t} = \left[\hat{\Pi}(\mathbf{x},t),\hat{H}\right]_- \quad\Longrightarrow\quad \frac{\partial\hat{\Pi}(\mathbf{x},t)}{\partial t} = \frac{\hbar^2}{2M}\,\Delta\,\hat{\Psi}^\dagger(\mathbf{x},t) - \frac{Mc^2}{2}\,\hat{\Psi}^\dagger(\mathbf{x},t)\,, \tag{8.59}$$

$$i\hbar\,\frac{\partial\hat{\Pi}^\dagger(\mathbf{x},t)}{\partial t} = \left[\hat{\Pi}^\dagger(\mathbf{x},t),\hat{H}\right]_- \quad\Longrightarrow\quad \frac{\partial\hat{\Pi}^\dagger(\mathbf{x},t)}{\partial t} = \frac{\hbar^2}{2M}\,\Delta\,\hat{\Psi}(\mathbf{x},t) - \frac{Mc^2}{2}\,\hat{\Psi}(\mathbf{x},t)\,. \tag{8.60}$$

Note that the respective commutators are evaluated either with the operator identity (3.43) or with functional derivatives similar to Section 4.3. Furthermore, combining (8.57) and (8.60) as well as (8.58) and (8.59), we read off that both field operators $\hat{\Psi}^\dagger(\mathbf{x},t)$ and $\hat{\Psi}(\mathbf{x},t)$ obey the Klein-Gordon equation:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta + \frac{M^2c^2}{\hbar^2}\right)\hat{\Psi}(\mathbf{x},t) = 0\,, \tag{8.61}$$

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta + \frac{M^2c^2}{\hbar^2}\right)\hat{\Psi}^\dagger(\mathbf{x},t) = 0\,. \tag{8.62}$$

In the following we determine the solutions of the operator-valued partial differential equations (8.61), (8.62) and work out their corresponding physical interpretation.

8.4 Plane Waves

The field operator $\hat{\Psi}(\mathbf{x},t)$ as a function of its spatial degree of freedom $\mathbf{x}$ is now expanded into plane waves:

$$\hat{\Psi}(\mathbf{x},t) = \int d^3p\, \hat{a}_{\mathbf{p}}(t) N_{\mathbf{p}} \exp\left(\frac{i}{\hbar}\,\mathbf{p}\,\mathbf{x}\right). \tag{8.63}$$

Here $N_{\mathbf{p}}$ denotes a normalization constant, which is fixed later on appropriately. Inserting the decomposition (8.63) into the Klein-Gordon equation (8.61) of the field operator, yields an ordinary differential equation of second order for the respective Fourier operators $\hat{a}_{\mathbf{p}}(t)$:

$$\frac{\partial}{\partial t^2}\,\hat{a}_{\mathbf{p}}(t) + \frac{\mathbf{p}^2c^2 + M^2c^4}{\hbar^2}\,\hat{a}_{\mathbf{p}}(t) = 0\,. \tag{8.64}$$

The general solution of (8.64) reads

$$\hat{a}_{\mathbf{p}}(t) = \hat{a}_{\mathbf{p}}^{(1)}\, \exp\left(-\frac{i}{\hbar}\,E_{\mathbf{p}}t\right) + \hat{a}_{\mathbf{p}}^{(2)} \exp\left(\frac{i}{\hbar}\,E_{\mathbf{p}}t\right). \tag{8.65}$$

Here we have introduced as an abbreviation the relativistic energy-momentum dispersion

$$E_{\mathbf{p}} = \sqrt{\mathbf{p}^2c^2 + M^2c^4}\,, \tag{8.66}$$

which obeys the symmetry

$$E_{\mathbf{p}} = E_{-\mathbf{p}}\,. \tag{8.67}$$

Inserting (8.65) into the plane wave expansion (8.63), we obtain at first

$$\hat{\Psi}(\mathbf{x},t) = \int d^3p\, N_{\mathbf{p}} \left\{\hat{a}_{\mathbf{p}}^{(1)}\, \exp\left[\frac{i}{\hbar}\,(\mathbf{p}\mathbf{x} - E_{\mathbf{p}}t)\right] + \hat{a}_{\mathbf{p}}^{(2)}\, \exp\left[\frac{i}{\hbar}\,(\mathbf{p}\mathbf{x} + E_{\mathbf{p}}t)\right]\right\}. \tag{8.68}$$

Performing in the second term the substitution $\mathbf{p} \to -\mathbf{p}$, taking into account (8.67), and assuming

$$N_{\mathbf{p}} = N_{-\mathbf{p}} \tag{8.69}$$

converts (8.68) into

$$\hat{\Psi}(\mathbf{x},t) = \int d^3p\, N_{\mathbf{p}} \left\{\hat{a}_{\mathbf{p}}^{(1)}\, \exp\left[\frac{i}{\hbar}\,(\mathbf{p}\mathbf{x} - E_{\mathbf{p}}t)\right] + \hat{a}_{-\mathbf{p}}^{(2)}\, \exp\left[-\frac{i}{\hbar}\,(\mathbf{p}\mathbf{x} - E_{\mathbf{p}}t)\right]\right\}. \tag{8.70}$$

Thus, redefining $\hat{a}_{-\mathbf{p}}^{(2)}$ as $\hat{a}_{\mathbf{p}}^{(2)}$ allows to compactly summarize (8.70) as

$$\hat{\Psi}(\mathbf{x},t) = \sum_{r=1}^{2} \int d^3p\, \hat{a}_{\mathbf{p}}^{(r)}\, u_{\mathbf{p}}^{(r)}(\mathbf{x},t)\,. \tag{8.71}$$

Here we have introduced $u_{\mathbf{p}}^{(r)}(\mathbf{x},t)$ as an abbreviation for the plane waves

$$u_{\mathbf{p}}^{(r)}(\mathbf{x},t) = N_{\mathbf{p}}\, \exp\left[\varepsilon_r\, \frac{i}{\hbar}\,(\mathbf{p}\mathbf{x} - E_{\mathbf{p}}t)\right] \tag{8.72}$$

with the notation

$$\varepsilon_r = \begin{cases} +1; & r=1 \\ -1; & r=2 \end{cases} . \tag{8.73}$$

The normalization constant $N_{\mathbf{p}}$ is now fixed by demanding for the scalar product between the plane waves $u_{\mathbf{p}}^{(r)}(\mathbf{x},t)$ and $u_{\mathbf{p}'}^{(r')}(\mathbf{x},t)$:

$$\langle u_{\mathbf{p}}^{(r)}, u_{\mathbf{p}'}^{(r')} \rangle = \varepsilon_r \, \delta_{r,r'} \, \delta(\mathbf{p}-\mathbf{p}') \, . \tag{8.74}$$

Thus, this condition amounts to demanding that the plane waves (8.72) with $r=1$ and $r=2$ correspond to the charge $+1$ and -1, respectively, as follows from (8.47) and (8.73). Taking into account the scalar product of the Klein-Gordon theory defined in (8.42) as well as (8.72), we get at first

$$\begin{aligned} \langle u_{\mathbf{p}}^{(r)}, u_{\mathbf{p}'}^{(r')} \rangle &= \frac{\varepsilon_r E_{\mathbf{p}} + \varepsilon_{r'} E_{\mathbf{p}'}}{2Mc^2} \, N_{\mathbf{p}} N_{\mathbf{p}'} \, \exp\left[\frac{i}{\hbar} \left(\varepsilon_r E_{\mathbf{p}} - \varepsilon_{r'} E_{\mathbf{p}'}\right) t\right] \\ &\quad \times \int d^3x \, \exp\left[\frac{i}{\hbar} \left(\varepsilon_{r'} \mathbf{p}' - \varepsilon_r \mathbf{p}\right) \mathbf{x}\right] . \end{aligned} \tag{8.75}$$

Performing the spatial integral yields $\delta(\varepsilon_{r'}\mathbf{p}' - \varepsilon_r \mathbf{p}) = \delta(\mathbf{p}' - \varepsilon_r \varepsilon_{r'} \mathbf{p})$, so we conclude from the symmetries (8.67) and (8.69):

$$\langle u_{\mathbf{p}}^{(r)}, u_{\mathbf{p}'}^{(r')} \rangle = \frac{(2\pi\hbar)^3 E_{\mathbf{p}}}{Mc^2} \, N_{\mathbf{p}}^2 \, \frac{\varepsilon_r + \varepsilon_{r'}}{2} \, \exp\left[\frac{i}{\hbar} \left(\varepsilon_r - \varepsilon_{r'}\right) E_{\mathbf{p}} t\right] \delta(\mathbf{p}' - \varepsilon_r \varepsilon_{r'} \mathbf{p}) \, . \tag{8.76}$$

Due to the observation

$$\frac{\varepsilon_r + \varepsilon_{r'}}{2} = \begin{cases} \varepsilon_r; & r = r' \\ 0; & r \neq r' \end{cases} = \varepsilon_r \, \delta_{r,r'} \; , \tag{8.77}$$

which follows from (8.73), Eq. (8.76) reduces to (8.74) provided the normalization is fixed by

$$N_{\mathbf{p}} = \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 \, E_{\mathbf{p}}}} \, . \tag{8.78}$$

Indeed, the normalization (8.78) obeys the imposed symmetry (8.69) due to (8.67).

For the following calculations we need another technical result. Namely, considering the complex conjugated plane wave $u_{\mathbf{p}}^{(r)*}(\mathbf{x},t)$, this just corresponds to exchanging the indices $r=1$ and $r=2$ according to (8.72):

$$u_{\mathbf{p}}^{(1)*}(\mathbf{x},t) = u_{\mathbf{p}}^{(2)}(\mathbf{x},t) \, , \qquad u_{\mathbf{p}}^{(2)*}(\mathbf{x},t) = u_{\mathbf{p}}^{(1)}(\mathbf{x},t) \, . \tag{8.79}$$

Therefore, we read off from (8.74) and (8.79) the scalar product between two complex conjugated plane waves:

$$\langle u_{\mathbf{p}}^{(r)*}, u_{\mathbf{p}'}^{(r')*} \rangle = -\varepsilon_r \, \delta_{r,r'} \, \delta(\mathbf{p}-\mathbf{p}') \, . \tag{8.80}$$

8.5 Fourier Operators

According to (8.71) and its adjoint

$$\hat{\Psi}^\dagger(\mathbf{x},t) = \sum_{r=1}^{2} \int d^3p \, \hat{a}_{\mathbf{p}}^{(r)\dagger} \, u_{\mathbf{p}}^{(r)*}(\mathbf{x},t) \, . \tag{8.81}$$

both field operator $\hat{\Psi}(\mathbf{x},t)$ and $\hat{\Psi}^\dagger(\mathbf{x},t)$ are expanded in plane waves with time-independent Fourier operators $\hat{a}_{\mathbf{p}}^{(r)}$ and $\hat{a}_{\mathbf{p}}^{(r)\dagger}$. With the help of the scalar product of the Klein-Gordon field both relations can be inverted so that, conversely, the Fourier operators $\hat{a}_{\mathbf{p}}^{(r)}$ and $\hat{a}_{\mathbf{p}}^{(r)\dagger}$ are expressed in terms of the field operator $\hat{\Psi}(\mathbf{x},t)$ and its adjoint $\hat{\Psi}^\dagger(\mathbf{x},t)$. Taking into account (8.74) and (8.80) we get at first

$$\hat{a}_{\mathbf{p}}^{(r)} = \varepsilon_r \, \langle u_{\mathbf{p}}^{(r)}, \hat{\Psi} \rangle \, , \tag{8.82}$$

$$\hat{a}_{\mathbf{p}}^{(r)\dagger} = -\varepsilon_r \, \langle u_{\mathbf{p}}^{(r)*}, \hat{\Psi}^\dagger \rangle \, , \tag{8.83}$$

which reduces due to (8.42) to

$$\hat{a}_{\mathbf{p}}^{(r)} = \frac{i\hbar\varepsilon_r}{2Mc^2} \int d^3x \left[u^{(r)*}(\mathbf{x},t) \, \frac{\partial \hat{\Psi}(\mathbf{x},t)}{\partial t} - \frac{\partial u^{(r)*}(\mathbf{x},t)}{\partial t} \, \hat{\Psi}(\mathbf{x},t) \right] , \tag{8.84}$$

$$\hat{a}_{\mathbf{p}}^{(r)\dagger} = \frac{-i\hbar\varepsilon_r}{2Mc^2} \int d^3x \left[u^{(r)}(\mathbf{x},t) \, \frac{\partial \hat{\Psi}^\dagger(\mathbf{x},t)}{\partial t} - \frac{\partial u^{(r)}(\mathbf{x},t)}{\partial t} \, \hat{\Psi}^\dagger(\mathbf{x},t) \right] . \tag{8.85}$$

Applying the Heisenberg equations of motion (8.57) and (8.58) we arrive at the following representation for the Fourier operators:

$$\hat{a}_{\mathbf{p}}^{(r)} = \frac{i\hbar\varepsilon_r}{2Mc^2} \int d^3x \left[\frac{2Mc^2}{\hbar^2} \, u^{(r)*}(\mathbf{x},t) \, \hat{\Pi}^\dagger(\mathbf{x},t) - \frac{\partial u^{(r)*}(\mathbf{x},t)}{\partial t} \, \hat{\Psi}(\mathbf{x},t) \right] , \tag{8.86}$$

$$\hat{a}_{\mathbf{p}}^{(r)\dagger} = \frac{-i\hbar\varepsilon_r}{2Mc^2} \int d^3x \left[\frac{2Mc^2}{\hbar^2} \, u^{(r)}(\mathbf{x},t) \, \hat{\Pi}(\mathbf{x},t) - \frac{\partial u^{(r)}(\mathbf{x},t)}{\partial t} \, \hat{\Psi}^\dagger(\mathbf{x},t) \right] . \tag{8.87}$$

With this and the canonical equal-time commutator relations between the field operators and the momentum operators (8.53)–(8.55) we determine the commutation relations between the Fourier operators $\hat{a}_{\mathbf{p}}^{(r)}$ and $\hat{a}_{\mathbf{p}}^{(r)\dagger}$. At first we get straight-forwardly the trivial commutators

$$\left[\hat{a}_{\mathbf{p}}^{(r)}, \hat{a}_{\mathbf{p}'}^{(r')} \right]_- = \left[\hat{a}_{\mathbf{p}}^{(r)\dagger}, \hat{a}_{\mathbf{p}'}^{(r')\dagger} \right]_- = 0 \, . \tag{8.88}$$

And for the non-trivial commutator we obtain at first

$$\left[\hat{a}_{\mathbf{p}}^{(r)}, \hat{a}_{\mathbf{p}'}^{(r')\dagger} \right]_- = \varepsilon_r \varepsilon_{r'} \, \frac{i\hbar}{2Mc^2} \int d^3x \left[u_{\mathbf{p}}^{(r)}(\mathbf{x},t) \, \frac{\partial u_{\mathbf{p}'}^{(r')}(\mathbf{x},t)}{\partial t} - \frac{\partial u_{\mathbf{p}}^{(r)}(\mathbf{x},t)}{\partial t} \, u_{\mathbf{p}'}^{(r')}(\mathbf{x},t) \right] , \tag{8.89}$$

so taking into account (8.42), $\varepsilon_r^2 = 1$ due to (8.73), and (8.74) finally yields

$$\left[\hat{a}_{\mathbf{p}}^{(r)}, \hat{a}_{\mathbf{p}'}^{(r')\dagger} \right]_- = \varepsilon_r \, \delta_{r,r'} \, \delta(\mathbf{p} - \mathbf{p}') \, . \tag{8.90}$$

Here the appearance of the factor ε_r indicates due to (8.73) that $\hat{a}_{\mathbf{p}}^{(2)}$ and $\hat{a}_{\mathbf{p}}^{(2)\dagger}$ do not represent a creation and annihilation operator, respectively. We come back to this observation in due course, but before we determine how both the Hamilton operator and the charge operator are decomposed in terms of the Fourier operators $\hat{a}_{\mathbf{p}}^{(r)}$ and $\hat{a}_{\mathbf{p}}^{(r)\dagger}$.

8.6 Hamilton Operator

The plane wave expansions (8.71) and (8.81) of the field operators $\hat{\Psi}(\mathbf{x},t)$ and $\hat{\Psi}^\dagger(\mathbf{x},t)$ have together with (8.57), (8.58), and (8.72), the following consequences:

$$\nabla\hat{\Psi}(\mathbf{x},t) = \frac{i}{\hbar}\sum_{r=1}^{2}\int d^3p\,\varepsilon_r\mathbf{p}\,\hat{a}_{\mathbf{p}}^{(r)}\,u_{\mathbf{p}}^{(r)}(\mathbf{x},t)\,, \tag{8.91}$$

$$\nabla\hat{\Psi}^\dagger(\mathbf{x},t) = -\frac{i}{\hbar}\sum_{r=1}^{2}\int d^3p\,\varepsilon_r\mathbf{p}\,\hat{a}_{\mathbf{p}}^{(r)\dagger}\,u_{\mathbf{p}}^{(r)*}(\mathbf{x},t)\,, \tag{8.92}$$

$$\hat{\Pi}(\mathbf{x},t) = \frac{i\hbar}{2Mc^2}\sum_{r=1}^{2}\int d^3p\,\varepsilon_r E_{\mathbf{p}}\,\hat{a}_{\mathbf{p}}^{(r)\dagger}\,u_{\mathbf{p}}^{(r)*}(\mathbf{x},t)\,, \tag{8.93}$$

$$\hat{\Pi}^\dagger(\mathbf{x},t) = \frac{-i\hbar}{2Mc^2}\sum_{r=1}^{2}\int d^3p\,\varepsilon_r E_{\mathbf{p}}\,\hat{a}_{\mathbf{p}}^{(r)}\,u_{\mathbf{p}}^{(r)}(\mathbf{x},t)\,. \tag{8.94}$$

Using now all plane wave expansions (8.71), (8.81) and (8.91)–(8.94) in the Hamilton operator of the Klein-Gordon field (8.56) we get at first

$$\begin{aligned}\hat{H} &= \sum_{r=1}^{2}\sum_{r'=1}^{2}\int d^3p\int d^3p'\left(\frac{\varepsilon_r\varepsilon_{r'}E_{\mathbf{p}}E'_{\mathbf{p}}}{2Mc^2}+\frac{\varepsilon_r\varepsilon_{r'}\mathbf{p}\mathbf{p'}}{2M}+\frac{Mc^2}{2}\right)\hat{a}_{\mathbf{p}}^{(r)\dagger}\hat{a}_{\mathbf{p'}}^{(r')}\\ &\quad\times\int d^3x\,u_{\mathbf{p}}^{(r)*}(\mathbf{x},t)\,u_{\mathbf{p'}}^{(r')}(\mathbf{x},t)\,.\end{aligned} \tag{8.95}$$

The remaining spatial integral is evaluated with (8.67), (8.72), and (8.78), yielding

$$\int d^3x\,u_{\mathbf{p}}^{(r)*}(\mathbf{x},t)\,u_{\mathbf{p'}}^{(r')}(\mathbf{x},t) = \frac{Mc^2}{E_{\mathbf{p}}}\,\exp\left[\frac{i}{\hbar}\left(\varepsilon_r-\varepsilon_{r'}\right)E_{\mathbf{p}}t\right]\,\delta(\mathbf{p'}-\varepsilon_r\varepsilon_{r'}\mathbf{p})\,. \tag{8.96}$$

Inserting (8.96) into (8.95) the integration with respect to the momenta $\mathbf{p'}$ can be evaluated by taking into account the symmetry (8.67)

$$\hat{H} = \sum_{r=1}^{2}\sum_{r'=1}^{2}\int d^3p\left(\frac{\varepsilon_r\varepsilon_{r'}E_{\mathbf{p}}^2}{2Mc^2}+\frac{\mathbf{p^2}}{2M}+\frac{Mc^2}{2}\right)\frac{Mc^2}{E_{\mathbf{p}}}\,\exp\left[\frac{i}{\hbar}\left(\varepsilon_r-\varepsilon_{r'}\right)E_{\mathbf{p}}t\right]\,\hat{a}_{\mathbf{p}}^{(r)\dagger}\hat{a}_{\varepsilon_r\varepsilon_{r'}\mathbf{p}}^{(r')}\,. \tag{8.97}$$

With the relativistic energy-momentum dispersion (8.66) this simplifies to

$$\hat{H} = \sum_{r=1}^{2}\sum_{r'=1}^{2}\int d^3p\,\frac{\varepsilon_r\varepsilon_{r'}+1}{2}\,E_{\mathbf{p}}\,\exp\left[\frac{i}{\hbar}\left(\varepsilon_r-\varepsilon_{r'}\right)E_{\mathbf{p}}t\right]\hat{a}_{\mathbf{p}}^{(r)\dagger}\hat{a}_{\varepsilon_r\varepsilon_{r'}\mathbf{p}}^{(r')}\,. \tag{8.98}$$

As Eq. (8.73) implies the auxiliary calculation

$$\frac{\varepsilon_r\varepsilon_{r'}+1}{2} = \begin{cases}1; & r=r'\\ 0; & r\neq r'\end{cases} = \delta_{r,r'}\;, \tag{8.99}$$

the Hamilton operator of the Klein-Gordon field (8.98) finally reduces to

$$\hat{H} = \sum_{r=1}^{2}\int d^3p\,E_{\mathbf{p}}\,\hat{a}_{\mathbf{p}}^{(r)\dagger}\hat{a}_{\mathbf{p}}^{(r)}\,. \tag{8.100}$$

Thus, whereas the intermediate results (8.97) and (8.98) suggest that the second-quantized Hamilton operator $\hat{H}$ of the Klein-Gordon theory may explicitly depend on time, the result (8.100) reveals that it turns out to be time-independent. This is consistent with the fact that the energy of the Klein-Gordon theory is a conserved quantity due its time translational invariance.

8.7 Charge Operator

According to (8.37), (8.39) and (8.42), (8.47), respectively, the charge of the Klein-Gordon field is defined by

$$Q = \frac{i\hbar}{2Mc^2} \int d^3x \left[\Psi^*(\mathbf{x},t) \, \frac{\partial \Psi(\mathbf{x},t)}{\partial t} - \frac{\partial \Psi^*(\mathbf{x},t)}{\partial t} \, \Psi(\mathbf{x},t) \right] . \tag{8.101}$$

Due to (8.48) and (8.49) the charge (8.101) can be reexpressed as follows:

$$Q = \frac{i}{\hbar} \int d^3x \left[\Psi^*(\mathbf{x},t) \, \Pi^*(\mathbf{x},t) - \Pi(\mathbf{x},t) \, \Psi(\mathbf{x},t) \right] . \tag{8.102}$$

Within the second quantization we assign to the charge (8.102) a corresponding operator:

$$\hat{Q} = \frac{i}{\hbar} \int d^3x \left[\hat{\Psi}^\dagger(\mathbf{x},t) \, \hat{\Pi}^\dagger(\mathbf{x},t) - \hat{\Pi}(\mathbf{x},t) \, \hat{\Psi}(\mathbf{x},t) \right] . \tag{8.103}$$

Note that also here the respective order of the operators does play a role due to (8.53) and (8.54). The particular operator order chosen in (8.103) guarantees that the charge operator $\hat{Q}$ commutes with the Hamilton operator (8.56) due to applying (3.10) and (3.43):

$$\left[\hat{Q}, \hat{H} \right]_- = 0 \, . \tag{8.104}$$

Thus energy and charge remain to be both conserved quantities also in the second quantized Klein-Gordon theory. Inserting in (8.103) the plane wave expansions (8.71), (8.81) and (8.93), (8.94) we get at first

$$\hat{Q} = \sum_{r=1}^{2} \sum_{r'=1}^{2} \int d^3p \int d^3p' \, \frac{\varepsilon_r E_{\mathbf{p}} + \varepsilon_{r'} E_{\mathbf{p}'}}{2Mc^2} \, \hat{a}_{\mathbf{p}}^{(r)\dagger} \hat{a}_{\mathbf{p}'}^{(r')} \int d^3x \, u_{\mathbf{p}}^{(r)*}(\mathbf{x},t) \, u_{\mathbf{p}'}^{(r')}(\mathbf{x},t) \, . \tag{8.105}$$

Taking into account the symmetry (8.67), the integral (8.96), and the auxiliary calculation (8.77), the charge operator (8.105) reduces finally to the form

$$\hat{Q} = \sum_{r=1}^{2} \int d^3p \, \varepsilon_r \, \hat{a}_{\mathbf{p}}^{(r)\dagger} \hat{a}_{\mathbf{p}}^{(r)} \, . \tag{8.106}$$

Thus, also the charge operator $\hat{Q}$ turns out to be time independent, which confirms that the charge is a conserved quantity for the Klein-Gordon field.

8.8 Redefinition of Fourier Operators

Now we aim for a consistent physical interpretation of the results obtained so far within the second quantization of the Klein-Gordon field. From the commutation relations (8.88) and (8.90) we read off that the Fourier operators $\hat{a}_{\mathbf{p}}^{(1)}$, $\hat{a}_{\mathbf{p}}^{(2)\dagger}$ and $\hat{a}_{\mathbf{p}}^{(1)\dagger}$, $\hat{a}_{\mathbf{p}}^{(2)}$ have to be interpreted as annihilation and creation operators, respectively. This observation suggests to reinterpret the Fourier operators as follows:

$$\hat{a}_{\mathbf{p}} = \hat{a}_{\mathbf{p}}^{(1)} , \qquad \hat{a}_{\mathbf{p}}^{\dagger} = \hat{a}_{\mathbf{p}}^{(1)\dagger} , \qquad \hat{b}_{\mathbf{p}} = \hat{a}_{\mathbf{p}}^{(2)\dagger} , \qquad \hat{b}_{\mathbf{p}}^{\dagger} = \hat{a}_{\mathbf{p}}^{(2)} . \tag{8.107}$$

By using different letters a and b we express that the corresponding operators $\hat{a}_{\mathbf{p}}$, $\hat{b}_{\mathbf{p}}$ and $\hat{a}_{\mathbf{p}}^{\dagger}$, $\hat{b}_{\mathbf{p}}^{\dagger}$ describe the annihilation and the creation of different kinds of particles. Furthermore, this redefinition leaves the trivial commutation relations (8.88) invariant:

$$\begin{aligned} \left[\hat{a}_{\mathbf{p}}, \hat{a}_{\mathbf{p}'}\right]_{-} = \left[\hat{a}_{\mathbf{p}}^{\dagger}, \hat{a}_{\mathbf{p}'}^{\dagger}\right]_{-} = \left[\hat{b}_{\mathbf{p}}, \hat{b}_{\mathbf{p}'}\right]_{-} = \left[\hat{b}_{\mathbf{p}}^{\dagger}, \hat{b}_{\mathbf{p}'}^{\dagger}\right]_{-} = 0 , \\ \left[\hat{a}_{\mathbf{p}}, \hat{b}_{\mathbf{p}'}\right]_{-} = \left[\hat{a}_{\mathbf{p}}, \hat{b}_{\mathbf{p}'}^{\dagger}\right]_{-} = \left[\hat{a}_{\mathbf{p}}^{\dagger}, \hat{b}_{\mathbf{p}'}\right]_{-} = \left[\hat{a}_{\mathbf{p}}^{\dagger}, \hat{b}_{\mathbf{p}'}^{\dagger}\right]_{-} = 0 . \end{aligned} \tag{8.108}$$

But the non-trivial commutation relations (8.90) are converted to

$$\left[\hat{a}_{\mathbf{p}}, \hat{a}_{\mathbf{p}'}^{\dagger}\right]_{-} = \left[\hat{b}_{\mathbf{p}}, \hat{b}_{\mathbf{p}'}^{\dagger}\right]_{-} = \delta(\mathbf{p} - \mathbf{p}') . \tag{8.109}$$

And the plane wave expansions (8.71) and (8.81) of the field operators $\hat{\Psi}(\mathbf{x}, t)$ and $\hat{\Psi}^{\dagger}(\mathbf{x}, t)$ then read due to (8.79):

$$\hat{\Psi}(\mathbf{x}, t) = \int d^3p \left[\hat{a}_{\mathbf{p}}\, u_{\mathbf{p}}(\mathbf{x}, t) + \hat{b}_{\mathbf{p}}^{\dagger}\, u_{\mathbf{p}}^{*}(\mathbf{x}, t)\right] , \tag{8.110}$$

$$\hat{\Psi}^{\dagger}(\mathbf{x}, t) = \int d^3p \left[\hat{a}_{\mathbf{p}}^{\dagger}\, u_{\mathbf{p}}^{*}(\mathbf{x}, t) + \hat{b}_{\mathbf{p}}\, u_{\mathbf{p}}(\mathbf{x}, t)\right] . \tag{8.111}$$

Here we have introduced according to (8.72) and (8.78)

$$u_{\mathbf{p}}(\mathbf{x}, t) = u_{\mathbf{p}}^{(1)}(\mathbf{x}, t) = \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}}}}\, \exp\left[\frac{i}{\hbar}\, (\mathbf{p}\mathbf{x} - E_{\mathbf{p}} t)\right] . \tag{8.112}$$

In addition, the Hamilton operator (8.100) and the charge operator (8.106) read due to the redefinition (8.107)

$$\hat{H} = \int d^3p\, E_{\mathbf{p}} \left(\hat{a}_{\mathbf{p}}^{\dagger}\hat{a}_{\mathbf{p}} + \hat{b}_{\mathbf{p}}\hat{b}_{\mathbf{p}}^{\dagger}\right) , \tag{8.113}$$

$$\hat{Q} = \int d^3p \left(\hat{a}_{\mathbf{p}}^{\dagger}\hat{a}_{\mathbf{p}} - \hat{b}_{\mathbf{p}}\hat{b}_{\mathbf{p}}^{\dagger}\right) . \tag{8.114}$$

In order to obtain a normal ordering of the operators we have to use the commutator (8.109), yielding

$$\hat{H} = \int d^3p\, E_{\mathbf{p}} \left(\hat{a}_{\mathbf{p}}^{\dagger}\hat{a}_{\mathbf{p}} + \hat{b}_{\mathbf{p}}^{\dagger}\hat{b}_{\mathbf{p}}\right) + \delta(\mathbf{0}) \int d^3p\, E_{\mathbf{p}} , \tag{8.115}$$

$$\hat{Q} = \int d^3p \left(\hat{a}_{\mathbf{p}}^{\dagger}\hat{a}_{\mathbf{p}} - \hat{b}_{\mathbf{p}}^{\dagger}\hat{b}_{\mathbf{p}}\right) - \delta(\mathbf{0}) \int d^3p . \tag{8.116}$$

The vacuum state is defined here as usual

$$\hat{a}_{\mathbf{p}}|0\rangle = 0\,, \qquad \hat{b}_{\mathbf{p}}|0\rangle = 0\,. \tag{8.117}$$

With this the vacuum expectation values of both the Hamilton operator and the charge operator result to

$$\langle 0|\hat{H}|0\rangle = \delta(\mathbf{0}) \int d^3p\, E_{\mathbf{p}}\,, \tag{8.118}$$

$$\langle 0|\hat{Q}|0\rangle = -\delta(\mathbf{0}) \int d^3p\,, \tag{8.119}$$

which are divergent due to two reasons. On the one hand, the factor $\delta(\mathbf{0})$ is divergent and, on the other hand, the respective momentum integrals are divergent as well. Therefore, one considers instead of the Hamilton operator and the charge operator the respective renormalized quantities:

$$:\hat{H}: \;=\; \hat{H} - \langle 0|\hat{H}|0\rangle = \int d^3p\, E_{\mathbf{p}} \left(\hat{a}^\dagger_{\mathbf{p}} \hat{a}_{\mathbf{p}} + \hat{b}^\dagger_{\mathbf{p}} \hat{b}_{\mathbf{p}} \right), \tag{8.120}$$

$$:\hat{Q}: \;=\; \hat{Q} - \langle 0|\hat{Q}|0\rangle = \int d^3p \left(\hat{a}^\dagger_{\mathbf{p}} \hat{a}_{\mathbf{p}} - \hat{b}^\dagger_{\mathbf{p}} \hat{b}_{\mathbf{p}} \right). \tag{8.121}$$

We recognize that both renormalized operators $:\langle 0|\hat{H}|0\rangle:$ and $:\langle 0|\hat{Q}|0\rangle:$ are normal ordered, i.e. the creation (annihilation) operators stand on the left-hand (right-hand) side.

The results (8.120) and (8.121) allow now for the following physical interpretation. The operators $\hat{a}^\dagger_{\mathbf{p}}$, $\hat{a}_{\mathbf{p}}$ ($\hat{b}^\dagger_{\mathbf{p}}$, $\hat{b}_{\mathbf{p}}$) describe particles of charge 1 (-1) and energy $E_{\mathbf{p}}$. As the two particle types only differ by their charge, they represent particles and their respective antiparticles. The particle type a (b) can be identified with the pion π^+ (π^-). On the basis of this insight, we recognize in (8.110) that the field operator $\hat{\Psi}(\mathbf{x}, t)$ contains both the annihilation of particles a with charge 1 and the creation of antiparticles b with charge -1. These microscopic processes act together such that the field operator $\hat{\Psi}(\mathbf{x}, t)$ describes the annihilation of a charge 1 and, correspondingly, the adjoint field operator $\hat{\Psi}^\dagger(\mathbf{x}, t)$ represents the creation of a charge 1 at the space-point $(\mathbf{x}, t)$. This physical interpretation of the second-quantized operators $\hat{\Psi}(\mathbf{x}, t)$ and $\hat{\Psi}^\dagger(\mathbf{x}, t)$ turns out to be crucial for the corresponding propagator of the Klein-Gordon theory.

8.9 Definition of Propagator

In the following we investigate in more detail the Klein-Gordon propagator, which is an important ingredient of quantum field theory when the interaction of the Klein-Gordon field with other quantum fields is treated perturbatively. For instance, the Klein-Gordon propagator is an essential building block of scalar quantum electrodynamics, where the photon exchange between charged pions is described graphically in terms of corresponding Feynman diagrams. But the Klein-Gordon propagator turns out to be also central for this lecture from a technical point of view. On the one hand, its non-relativistic limit leads to the Schrödinger propagator, which is used in the context of non-relativistic quantum many-body theory. On the other hand, we will see later on that the propagator of the Dirac theory is determined by partial derivatives from the Klein-Gordon propagator. Thus, having a profound understanding of the Klein-Gordon propagator represents a prerequisite for the Dirac propagator, which is a key element of the Feynman diagrams of quantum electrodynamics.

Let us start with defining the Klein-Gordon propagator as the vacuum expectation value of the product of two field operators:

$$G(\mathbf{x},t;\mathbf{x}',t') = \left\langle 0 \left| \hat{T}\left(\hat{\Psi}(\mathbf{x},t)\,\hat{\Psi}^{\dagger}(\mathbf{x}',t')\right)\right| 0\right\rangle . \tag{8.122}$$

The symbol $\hat{T}$ denotes the time-ordered product of the field operators $\hat{\Psi}(\mathbf{x},t)$ and $\hat{\Psi}^{\dagger}(\mathbf{x}',t')$. Given two time-dependent bosonic operators $\hat{A}(t)$ and $\hat{B}(t')$, their time-ordered product reads

$$\hat{T}\left(\hat{A}(t)\,\hat{B}(t')\right) = \Theta(t-t')\,\hat{A}(t)\,\hat{B}(t') + \Theta(t'-t)\,\hat{B}(t')\,\hat{A}(t)\,, \tag{8.123}$$

where we have used the Heaviside function

$$\Theta(t) = \begin{cases} 1\,; & t>0 \\ 0\,; & t<0 \end{cases} . \tag{8.124}$$

Thus, the operator-valued factors in (8.123) are put into chronological order so that the operator having the later time argument is put first, i.e. to the left. If the two time arguments happen to be equal, problems might arise since the operator ordering is then not well defined. In the present case (8.122), however, this is not the case since the operators $\hat{\Psi}(\mathbf{x},t)$ and $\hat{\Psi}^{\dagger}(\mathbf{x}',t)$ at equal time commute due to (8.55). Taking into account (8.123) in (8.122) leads to

$$G(\mathbf{x},t;\mathbf{x}',t') = \Theta(t-t')\left\langle 0\left|\hat{\Psi}(\mathbf{x},t)\,\hat{\Psi}^{\dagger}(\mathbf{x}',t')\right|0\right\rangle + \Theta(t'-t)\left\langle 0\left|\hat{\Psi}^{\dagger}(\mathbf{x}',t')\,\hat{\Psi}(\mathbf{x},t)\right|0\right\rangle . \tag{8.125}$$

Note that this introduction of the Klein-Gordon propagator with a time-ordered product of field operators appears admittedly to be quite unmotivated at this stage of the lecture. But it will be justified a posteriori when dealing perturbatively with interacting quantum fields. Namely, such a perturbative treatment is performed systematically in the so-called Dirac interaction picture, where the unperturbed Hamiltonian determines the time dependence of the field operators, so that their interpretation of representing creation and annihilation operators is preserved, and the perturbative Hamiltonian affects the quantum states. And the latter turns out to lead to the time evolution operator in the Dirac interaction picture, whose perturbative expansion naturally involves the time-ordered product of field operators. Thus, in conclusion, any perturbative treatment in quantum field theory is based on the time-ordered product of field operators.

In order to determine the equation of motion for the Klein-Gordon propagator we calculate the first time derivative:

$$\frac{\partial G(\mathbf{x},t;\mathbf{x}',t')}{\partial t} = \delta(t-t')\left\langle 0\left|\left[\hat{\Psi}(\mathbf{x},t),\hat{\Psi}^{\dagger}(\mathbf{x}',t')\right]_{-}\right|0\right\rangle \tag{8.126}$$

$$+\Theta(t-t')\left\langle 0\left|\frac{\partial\hat{\Psi}(\mathbf{x},t)}{\partial t}\,\hat{\Psi}^{\dagger}(\mathbf{x}',t')\right|0\right\rangle + \Theta(t'-t)\left\langle 0\left|\hat{\Psi}^{\dagger}(\mathbf{x}',t')\,\frac{\partial\hat{\Psi}(\mathbf{x},t)}{\partial t}\right|0\right\rangle .$$

Here we have used the fact that the time derivative of the Heaviside function yields the delta function:

$$\frac{\partial \Theta(t)}{\partial t} = \delta(t) \, . \tag{8.127}$$

As the commutator of the field operators $\hat{\Psi}(\mathbf{x}, t)$ and $\hat{\Psi}^\dagger(\mathbf{x}', t)$ at the same time t vanish according to (8.55), the first term in (8.126) vanishes. Another time derivative leads then with (8.127) to

$$\frac{\partial^2 G(\mathbf{x}, t; \mathbf{x}', t')}{\partial t^2} = \delta(t - t') \left\langle 0 \right| \left[\frac{\partial \hat{\Psi}(\mathbf{x}, t)}{\partial t}, \hat{\Psi}^\dagger(\mathbf{x}', t') \right]_- \left| 0 \right\rangle \tag{8.128}$$
$$+ \Theta(t - t') \left\langle 0 \left| \frac{\partial^2 \hat{\Psi}(\mathbf{x}, t)}{\partial t^2} \, \hat{\Psi}^\dagger(\mathbf{x}', t') \right| 0 \right\rangle + \Theta(t' - t) \left\langle 0 \left| \hat{\Psi}^\dagger(\mathbf{x}', t') \, \frac{\partial^2 \hat{\Psi}(\mathbf{x}, t)}{\partial t^2} \right| 0 \right\rangle .$$

Taking into account (8.54), (8.57), (8.61), and (8.125) we finally obtain

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta + \frac{M^2 c^2}{\hbar^2} \right) G(\mathbf{x}, t; \mathbf{x}', t') = -i \, \frac{2M}{\hbar} \, \delta(t - t') \, \delta(\mathbf{x} - \mathbf{x}') \, . \tag{8.129}$$

Thus, we recognize that the Klein-Gordon propagator represents the Green function of the Klein-Gordon equation. As a coupling of the Klein-Gordon field to other quantum fields yields as a Heisenberg equation an inhomogeneous Klein-Gordon equation, its perturbative solution is based on the knowledge of the corresponding Green function, i.e. the Klein-Gordon propagator.

In view of the non-relativistic limit $c \to \infty$ we have to separate the rest energy from the Klein-Gordon propagator due to (8.5):

$$G(\mathbf{x}, t; \mathbf{x}', t') = g(\mathbf{x}, t; \mathbf{x}', t') \, \exp\left(-\frac{i}{\hbar} \, M c^2 t \right) . \tag{8.130}$$

Inserting the ansatz (8.130) in the equation of motion (8.129) we get

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{2iM}{\hbar} \frac{\partial}{\partial t} - \Delta \right) g(\mathbf{x}, t; \mathbf{x}', t') = -i \, \frac{2M}{\hbar} \, \delta(t - t') \, \delta(\mathbf{x} - \mathbf{x}') \, . \tag{8.131}$$

Performing then the non-relativistic limit $c \to \infty$ Eq. (8.131) reduces to

$$\left(i\hbar \, \frac{\partial}{\partial t} + \frac{\hbar^2}{2M} \, \Delta \right) g(\mathbf{x}, t; \mathbf{x}', t') = i\hbar \, \delta(t - t') \, \delta(\mathbf{x} - \mathbf{x}') \, . \tag{8.132}$$

Thus, $g(\mathbf{x}, t; \mathbf{x}', t')$ coincides with the Green function of the Schrödinger equation and can be identified with the Schrödinger propagator.

8.10 Interpretation of Propagator

Now we deal with the physical interpretation of the Klein-Gordon propagator (8.125). To this end we state two commutation relations for the charge operator (8.103):

$$\left[\hat{Q}, \hat{\Psi}(\mathbf{x}, t) \right]_- = -\hat{\Psi}(\mathbf{x}, t) \, , \tag{8.133}$$

$$\left[\hat{Q}, \hat{\Psi}^\dagger(\mathbf{x}, t) \right]_- = \hat{\Psi}^\dagger(\mathbf{x}, t) \, . \tag{8.134}$$

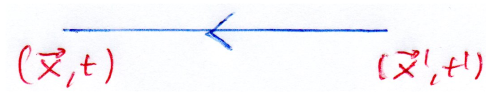


Figure 8.3: Graphical representation of the Klein-Gordon propagator (8.125) describing the propagation of the charge 1 from $(\mathbf{x}', t')$ to $(\mathbf{x}, t)$.

Thus, the field operators $\hat{\Psi}(\mathbf{x}, t)$ and $\hat{\Psi}^\dagger(\mathbf{x}, t)$ decrease and increase the charge by one unit, respectively, as was already anticipated at the end of Section 8.8. Namely, denoting with $|q\rangle$ an eigenstate of the charge operator $\hat{Q}$ with eigenvalue q, i.e.

$$\hat{Q}|q\rangle = q\,|q\rangle\,, \tag{8.135}$$

we conclude with the help of the commutator relations (8.133), (8.134):

$$\hat{Q}\hat{\Psi}(\mathbf{x}, t)|q\rangle = \hat{\Psi}(\mathbf{x}, t)\big(\hat{Q} - 1\big)|q\rangle = (q-1)\hat{\Psi}(\mathbf{x}, t)|q\rangle \quad\Longrightarrow\quad |q-1\rangle \sim \hat{\Psi}(\mathbf{x}, t)|q\rangle\,, \tag{8.136}$$

$$\hat{Q}\hat{\Psi}^\dagger(\mathbf{x}, t)|q\rangle = \hat{\Psi}^\dagger(\mathbf{x}, t)\big(\hat{Q} + 1\big)|q\rangle = (q+1)\hat{\Psi}^\dagger(\mathbf{x}, t)|q\rangle \quad\Longrightarrow\quad |q+1\rangle \sim \hat{\Psi}^\dagger(\mathbf{x}, t)|q\rangle\,. \tag{8.137}$$

Against this background the Klein-Gordon propagator (8.125) describes the propagation of the charge 1 from $(\mathbf{x}', t')$ to $(\mathbf{x}, t)$, see Fig. 8.3, via two microscopic processes. Taking into account the plane wave decompositions (8.110), (8.111) the first term in (8.125) describes the propagation of a particle of charge +1 from $(\mathbf{x}', t')$ to $(\mathbf{x}, t)$, whereas the second term considers the propagation of an antiparticle of charge −1 from $(\mathbf{x}, t)$ to $(\mathbf{x}', t')$. Thus, the Klein-Gordon propagator (8.125) takes both processes of particle and antiparticle propagation into account. But, according to the intuitive physical picture of Richard Feynman, particles with positive energy propagate forward in time, whereas antiparticles are considered to have negative energy, which move backwards in time.

8.11 Calculation of Propagator

Now we insert the plane wave decompositions (8.110), (8.111) of the field operators $\hat{\Psi}(\mathbf{x}, t)$, $\hat{\Psi}^\dagger(\mathbf{x}, t)$ into the definition of the Klein-Gordon propagator (8.125). Due to the commutation relations (8.108)–(8.109) and the definition of the vacuum state (8.117) we obtain the plane wave representation

$$G(\mathbf{x}, t; \mathbf{x}', t') = \int d^3p\,\Big[\Theta(t-t')u_{\mathbf{p}}(\mathbf{x}, t)u^*_{\mathbf{p}}(\mathbf{x}', t') + \Theta(t'-t)u_{\mathbf{p}}(\mathbf{x}', t')u^*_{\mathbf{p}}(\mathbf{x}, t)\Big]\,. \tag{8.138}$$

Inserting the plane wave (8.112) together with the relativistic energy-momentum dispersion (8.66), one obtains the following Fourier representation of the Klein-Gordon propagator:

$$
\begin{aligned}
G(\mathbf{x},t;\mathbf{x}',t') &= \frac{Mc^2}{(2\pi\hbar)^3}\int d^3p\,\frac{1}{\sqrt{\mathbf{p}^2c^2+M^2c^4}} \\
&\times\left(\Theta(t-t')\,\exp\left\{\frac{i}{\hbar}\left[\mathbf{p}(\mathbf{x}-\mathbf{x}')-\sqrt{\mathbf{p}^2c^2+M^2c^4}\,(t-t')\right]\right\}\right. \\
&\left.+\Theta(t'-t)\,\exp\left\{\frac{i}{\hbar}\left[\mathbf{p}(\mathbf{x}'-\mathbf{x})-\sqrt{\mathbf{p}^2c^2+M^2c^4}\,(t'-t)\right]\right\}\right)\,. \qquad (8.139)
\end{aligned}
$$

In the following we evaluate this momentum integral analytically. At first, substituting in the second term $\mathbf{p}\to-\mathbf{p}$, both terms are combined as follows:

$$
\begin{aligned}
G(\mathbf{x},t;\mathbf{x}',t') &= \frac{Mc^2}{(2\pi\hbar)^3}\int d^3p\,\frac{1}{\sqrt{\mathbf{p}^2c^2+M^2c^4}} \\
&\times\exp\left[\frac{i}{\hbar}\,\mathbf{p}(\mathbf{x}-\mathbf{x}')-\frac{i}{\hbar}\sqrt{\mathbf{p}^2c^2+M^2c^4}\,|t-t'|\right]\,. \qquad (8.140)
\end{aligned}
$$

Introducing subsequently spherical coordinates for the momentum integral, we obtain at first

$$
\begin{aligned}
G(\mathbf{x},t;\mathbf{x}',t') &= \frac{Mc^2}{(2\pi\hbar)^3}\int_0^{2\pi}d\varphi\int_0^{\pi}d\theta\sin\theta\int_0^{\infty}dp\,p^2\,\frac{1}{\sqrt{p^2c^2+M^2c^4}} \\
&\times\exp\left[\frac{i}{\hbar}\,p|\mathbf{x}-\mathbf{x}'|\,\cos\theta-\frac{i}{\hbar}\,\sqrt{p^2c^2+M^2c^4}\,|t-t'|\right]\,. \qquad (8.141)
\end{aligned}
$$

Evaluating the angle integrals explicitly, one gets two remaining integrals over the absolute value of the momentum. Performing the substitution $p\to-p$ in the second integral, both integrals over half axis can be combined into a single one over the whole real axis, yielding

$$
\begin{aligned}
G(\mathbf{x},t;\mathbf{x}',t') &= \frac{-iMc^2}{4\pi^2\hbar^2|\mathbf{x}-\mathbf{x}'|}\int_{-\infty}^{\infty}dp\,\frac{p}{\sqrt{p^2c^2+M^2c^4}} \\
&\times\exp\left\{\frac{i}{\hbar}\left[p|\mathbf{x}-\mathbf{x}'|-\sqrt{p^2c^2+M^2c^4}\,|t-t'|\right]\right\}\,. \qquad (8.142)
\end{aligned}
$$

Here the factor p in the integrand can be represented in terms of a partial derivative with respect to the distance $|\mathbf{x}-\mathbf{x}'|$:

$$
\begin{aligned}
G(\mathbf{x},t;\mathbf{x}',t') &= \frac{-Mc^2}{4\pi^2\hbar|\mathbf{x}-\mathbf{x}'|}\,\frac{\partial}{\partial|\mathbf{x}-\mathbf{x}'|}\int_{-\infty}^{\infty}dp\,\frac{1}{\sqrt{p^2c^2+M^2c^4}} \\
&\times\exp\left\{\frac{i}{\hbar}\left[p|\mathbf{x}-\mathbf{x}'|-\sqrt{p^2c^2+M^2c^4}\,|t-t'|\right]\right\}\,. \qquad (8.143)
\end{aligned}
$$

Due to the substitution

$$
p\,(z)=Mc\sinh z\,, \qquad (8.144)
$$

where we have

$$
\frac{dp\,(z)}{dz}=Mc\cosh z=Mc\sqrt{1+\sinh^2z}=\frac{1}{c}\,\sqrt{p^2c^2+M^2c^4}\,, \qquad (8.145)
$$

Eq. (8.143) is converted to

$$G(\mathbf{x},t;\mathbf{x}',t') = \frac{-Mc}{4\pi^2\hbar|\mathbf{x}-\mathbf{x}'|}\frac{\partial}{\partial|\mathbf{x}-\mathbf{x}'|} \times \int_{-\infty}^{\infty} dz\, \exp\left\{\frac{iMc}{\hbar}\Big[|\mathbf{x}-\mathbf{x}'|\sinh z - c|t-t'|\cosh z\Big]\right\}. \tag{8.146}$$

We now aim at simplifying the integral (8.146) by combining the two terms in the argument of the exponential function into a single one. This is accomplished by the trick to perform the substitution $z = z' + z_0$, which introduces a new variable z_0 into the calculation:

$$G(\mathbf{x},t;\mathbf{x}',t') = \frac{-Mc}{4\pi^2\hbar|\mathbf{x}-\mathbf{x}'|}\frac{\partial}{\partial|\mathbf{x}-\mathbf{x}'|} \times \int_{-\infty}^{\infty} dz\, \exp\left\{\frac{iMc}{\hbar}\Big[|\mathbf{x}-\mathbf{x}'|\sinh(z+z_0) - c|t-t'|\cosh(z+z_0)\Big]\right\}. \tag{8.147}$$

Taking into account the addition theorems of hyperbolic functions

$$\sinh(z+z_0) = \sinh z\,\cosh z_0 + \cosh z\,\sinh z_0\,, \tag{8.148}$$

$$\cosh(z+z_0) = \cosh z\,\cosh z_0 + \sinh z\,\sinh z_0\,, \tag{8.149}$$

the integral (8.147) gets at first more involved:

$$G(\mathbf{x},t;\mathbf{x}',t') = \frac{-Mc}{4\pi^2\hbar|\mathbf{x}-\mathbf{x}'|}\frac{\partial}{\partial|\mathbf{x}-\mathbf{x}'|}\int_{-\infty}^{\infty} dz \times \exp\left\{\frac{iMc}{\hbar}\Big[\Big(|\mathbf{x}-\mathbf{x}'|\cosh z_0 - c|t-t'|\sinh z_0\Big)\sinh z + \Big(|\mathbf{x}-\mathbf{x}'|\sinh z_0 - c|t-t'|\cosh z_0\Big)\cosh z\Big]\right\}. \tag{8.150}$$

But a closer inspection then reveals that the yet undetermined parameter z_0 can be chosen in such a way that the argument of the exponential function in (8.150) does only depend on one term, for instance on the $\cosh z$ function:

$$\tanh z_0 = \frac{\sinh z_0}{\cosh z_0} = \frac{|\mathbf{x}-\mathbf{x}'|}{c|t-t'|}\,. \tag{8.151}$$

The subsequent hyperbolic side calculations

$$\sinh z_0 = \frac{\tanh z_0}{\sqrt{1-\tanh^2 z_0}} = \frac{|\mathbf{x}-\mathbf{x}'|}{\sqrt{c^2\,(t-t')^2 - (\mathbf{x}-\mathbf{x}')^2}}\,, \tag{8.152}$$

$$\cosh z_0 = \frac{1}{\sqrt{1-\tanh^2 z_0}} = \frac{c|t-t'|}{\sqrt{c^2\,(t-t')^2 - (\mathbf{x}-\mathbf{x}')^2}} \tag{8.153}$$

together with (8.151) then simplify the integral in (8.150) to

$$G(\mathbf{x},t;\mathbf{x}',t') = \frac{-Mc}{4\pi^2\hbar|\mathbf{x}-\mathbf{x}'|}\frac{\partial}{\partial|\mathbf{x}-\mathbf{x}'|}\int_{-\infty}^{\infty} dz \times \exp\left[-i\,\frac{Mc}{\hbar}\sqrt{c^2\,(t-t')^2 - (\mathbf{x}-\mathbf{x}')^2}\,\cosh z\right]. \tag{8.154}$$

Here we can use the Hankel function of second kind [(8.405.2), Gradshteyn/Ryzhik]

$$H_\nu^{(2)}(x) = J_\nu(x) - iN_\nu(x)\,, \tag{8.155}$$

which consists of the Bessel function $J_\nu(x)$ and the von Neumann function $N_\nu(x)$, due to its integral representation [(8.421.2), Gradshteyn/Ryzhik]

$$H_\nu^{(2)}(x) = -\frac{e^{i\nu\pi/2}}{\pi\, i} \int_{-\infty}^{\infty} dt\, e^{-ix\cosh t - \nu t}\,. \tag{8.156}$$

With this we obtain from (8.154)

$$G(\mathbf{x},t;\mathbf{x}',t') = \frac{iMc}{4\pi\hbar|\mathbf{x}-\mathbf{x}'|}\,\frac{\partial}{\partial|\mathbf{x}-\mathbf{x}'|}\,H_0^{(2)}\left(\frac{Mc}{\hbar}\sqrt{c^2\,(t-t')^2 - (\mathbf{x}-\mathbf{x}')^2}\right). \tag{8.157}$$

Thus, it remains to evaluate the derivative, where we have to take into account [(8.473.6), Gradshteyn/Ryzhik]

$$\frac{d}{dx}\,H_0^{(2)}(x) = -H_1^{(2)}(x)\,. \tag{8.158}$$

Thus we get for the Klein-Gordon propagator the following explicit result:

$$G(\mathbf{x},t;\mathbf{x}',t') = \frac{i(Mc/\hbar)^2}{4\pi\sqrt{c^2\,(t-t')^2 - (\mathbf{x}-\mathbf{x}')^2}}\,H_1^{(2)}\left(\frac{Mc}{\hbar}\sqrt{c^2\,(t-t')^2 - (\mathbf{x}-\mathbf{x}')^2}\right). \tag{8.159}$$

We note that the particle M enters here only in form of the Compton wave length (8.21).

In the non-relativistic limit $c \to \infty$ the argument of the Hankel function becomes arbitrarily large, so we use [(8.451.4), Gradshtey/Ryzhik]:

$$H_\nu^{(2)}\,(x) \approx \sqrt{\frac{2}{\pi x}}\,e^{-i\left(x - \frac{\pi}{2}\nu - \frac{\pi}{4}\right)}\,, \qquad x \gg 1\,. \tag{8.160}$$

With this the non-relativistic limit of the Klein-Gordon propagator (8.159) is for $t > t'$ of the form (8.130) with

$$g(\mathbf{x},t;\mathbf{x}',t') = \sqrt{\left(\frac{M}{2\pi i\hbar\,(t-t')}\right)^3}\,\exp\left[\frac{iM\,(\mathbf{x}-\mathbf{x}')^2}{2\hbar\,(t-t')}\right]. \tag{8.161}$$

One can show that Eq. (8.161) represents the solution of the inhomogeneous Schrödinger equation (8.132). Thus, indeed, the Klein-Gordon propagator reduces in the non-relativistic limit to the Schrödinger propagator.

8.12 Covariant Form of Propagator

In view of obtaining a manifestly covariant form of the Klein-Gordon propagator, we extend now its three-dimensional Fourier representation (8.140) to a four-dimensional one. To this end we consider the auxiliary integral

$$I\,(t-t') \;=\; \lim_{\eta\downarrow 0}\int_{-\infty}^{\infty}\frac{dE}{2\pi\hbar}\,\frac{e^{-\frac{i}{\hbar}E(t-t')}}{E^2 - E_{\mathbf{p}}^2 + i\eta}\,. \tag{8.162}$$

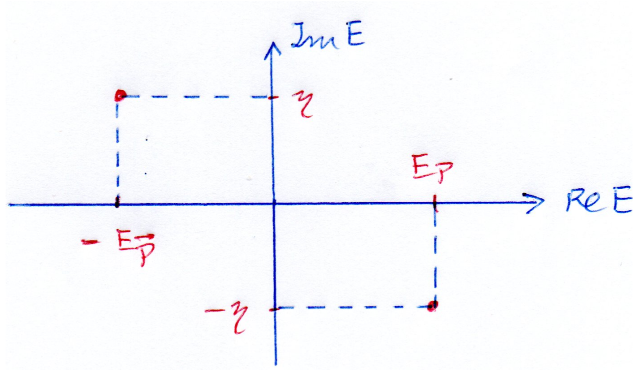


Figure 8.4: Shift of energy poles according to the $i\eta$ prescription of Richard Feynman.

Here the term $i\eta$ with $\eta > 0$ shifts infinitesimally the poles of the integrand on the real axis into the complex plane in a particular way. According to this $i\eta$ prescription, which was introduced by Richard Feynman, the pole at $E = E_{\mathbf{p}}$ is shifted below the real axis, whereas the pole at $E = -E_{\mathbf{p}}$ is shifted above the real axis, see Fig. 8.4. As we see in due course this guarantees that particles (antiparticles) move forward (backward) in time. To this end we evaluate the integral (8.162) with the help of the residue theorem. In order to guarantee the convergence of the integral one has to close the integration contour along the real axis for $t > t'$ ($t < t'$) in the lower (upper) part of the complex plane, yielding

$$t > t' : \quad I\,(t-t') = \frac{-2\pi i}{2\pi\hbar} \lim_{\eta \downarrow 0} \underset{E=\sqrt{E_{\mathbf{p}}^2 - i\eta}}{\mathrm{Res}} \frac{e^{-\frac{i}{\hbar}E(t-t')}}{E^2 - E_{\mathbf{p}}^2 + i\eta} = -\frac{i}{2\hbar E_{\mathbf{p}}}\, e^{-\frac{i}{\hbar}E_{\mathbf{p}}(t-t')}\,, \tag{8.163}$$

$$t < t' : \quad I\,(t-t') = \frac{2\pi i}{2\pi\hbar} \lim_{\eta \downarrow 0} \underset{E=-\sqrt{E_{\mathbf{p}}^2 - i\eta}}{\mathrm{Res}} \frac{e^{-\frac{i}{\hbar}E(t-t')}}{E^2 - E_{\mathbf{p}}^2 + i\eta} = -\frac{i}{2\hbar E_{\mathbf{p}}}\, e^{\frac{i}{\hbar}E_{\mathbf{p}}(t-t')}\,. \tag{8.164}$$

Here we have used the fact that the residue of a function $f(z)$ with a simple pole at $z = z_0$ is determined via

$$\underset{z\,=\,z_0}{\mathrm{Res}}\; f(z) = \lim_{z \to z_0} (z - z_0)\, f(z)\,. \tag{8.165}$$

Both results (8.163), (8.164) can be summarized as follows:

$$I\,(t-t') = -\frac{i}{2\hbar E_{\mathbf{p}}} \left[\Theta(t-t')\, e^{-\frac{i}{\hbar}E_{\mathbf{p}}(t-t')} + \Theta(t'-t)\, e^{\frac{i}{\hbar}E_{\mathbf{p}}(t-t')}\right] = -\frac{i}{2\hbar E_{\mathbf{p}}}\, e^{-\frac{i}{\hbar}E_{\mathbf{p}}|t-t'|}\,. \tag{8.166}$$

Inserting (8.162) and (8.166) into (8.140) leads at first to

$$\begin{aligned} G(\mathbf{x}, t; \mathbf{x}', t') \;&=\; 2i\hbar M c^2 \lim_{\eta \downarrow 0} \int \frac{d^3p}{(2\pi\hbar)^3} \int \frac{dE}{2\pi\hbar} \frac{1}{E^2 - \mathbf{p}^2c^2 - M^2c^4 + i\eta} \\ &\quad \times \exp\left\{-\frac{i}{\hbar}\Big[E\,(t-t') - \mathbf{p}\,(\mathbf{x} - \mathbf{x}')\Big]\right\}. \end{aligned} \tag{8.167}$$

This can be rewritten in a manifestly Lorentz covariant form as follows:

$$G(x^\lambda; x'^\lambda) = 2i\hbar M c \lim_{\eta \downarrow 0} \int \frac{d^4p}{(2\pi\hbar)^4} \frac{1}{g_{\mu\nu}p^\mu p^\nu - M^2c^2 + i\eta}\, \exp\left[-\frac{i}{\hbar}\, g_{\mu\nu}\, p^\mu\, (x^\nu - x'^\nu)\right]. \tag{8.168}$$

In this form the equation of motion of the Klein-Gordon propagator (8.129) is obviously fulfilled:

$$\left(g_{\mu\nu}\hat{p}^\mu\hat{p}^\nu + \frac{M^2c^2}{\hbar^2}\right) G(x^\lambda; x'^\lambda) = 2i\hbar Mc \lim_{\eta\downarrow 0} \int \frac{d^4p}{(2\pi\hbar)^4} \frac{g_{\mu\nu}\, \frac{-i}{\hbar}\, p^\mu\, \frac{-i}{\hbar}\, p^\nu + \frac{M^2c^2}{\hbar^2}}{g_{\mu\nu}p^\mu p^\nu - M^2c^2 + i\eta}$$
$$\times \exp\left[-\frac{i}{\hbar}\, g_{\mu\nu}\, p^\mu\, (x^\nu - x'^\nu)\right] = -\frac{2iMc}{\hbar} \int \frac{d^4p}{(2\pi\hbar)^4} \exp\left[-\frac{i}{\hbar}\, g_{\mu\nu}\, p^\mu\, (x^\nu - x'^\nu)\right]$$
$$= -\frac{2iMc}{\hbar}\, \delta^{(4)}(x - x') = -\frac{2iM}{\hbar}\, \delta(t - t')\, \delta\left(\mathbf{x} - \mathbf{x}'\right) . \tag{8.169}$$

Comparing (8.168) with the four-dimensional Fourier transformation the Klein-Gordon propagator

$$G(x^\lambda; x'^\lambda) = \int \frac{d^4p}{(2\pi\hbar)^4}\, G(p^\lambda)\, \exp\left[-\frac{i}{\hbar}\, g_{\mu\nu}\, p^\mu\, (x^\nu - x'^\nu)\right], \tag{8.170}$$

we read off

$$G(p^\lambda) = G(\mathbf{p}, E) = \lim_{\eta\downarrow 0} \frac{2i\hbar Mc}{E^2 - \mathbf{p}^2c^2 - M^2c^4 + i\eta} . \tag{8.171}$$

Here a singularity appears when the energy variable E coincides with the physical energy of a relativistic massive particle, which is given by the energy-momentum dispersion (8.66). In the non-relativistic limit $c \to \infty$ the Fourier transformed of the Klein-Gordon propagator (8.171) goes over into the Fourier transformed of the Schrödinger propagator:

$$\begin{aligned} g(\mathbf{p}, E) &= \lim_{c\to\infty} \frac{1}{c}\, G(\mathbf{p}, E + Mc^2) = \lim_{\eta\downarrow 0} \lim_{c\to\infty} \frac{2i\hbar M}{(E/c + Mc)^2 - \mathbf{p}^2 - M^2c^2 + i\eta} \\ &= \lim_{\eta\downarrow 0} \lim_{c\to\infty} \frac{i\hbar}{E - \frac{\mathbf{p}^2}{2M} + \frac{E^2}{2Mc^2} + i\eta} = \lim_{\eta\downarrow 0} \frac{i\hbar}{E - \frac{\mathbf{p}^2}{2M} + i\eta} . \end{aligned} \tag{8.172}$$

Indeed, solving the inhomogeneous Schrödinger equation (8.132) via a four-dimensional Fourier transformation

$$g(\mathbf{x}, t; \mathbf{x}', t') = \int_{-\infty}^{\infty} \frac{dE}{2\pi\hbar} \int \frac{d^3p}{(2\pi\hbar)^3}\, g(\mathbf{p}, E)\, \exp\left\{\frac{i}{\hbar}\left[\mathbf{p}\,(\mathbf{x} - \mathbf{x}') - E(t - t')\right]\right\} \tag{8.173}$$

yields straight-forwardly (8.173).

Chapter 9

Maxwell Field

All electrodynamic processes are described by the Maxwell equations. Surprisingly they represent the equations of motion of a first-quantized theory, although the Planck constant $\hbar$ does not appear explicitly. This apparent contradiction is resolved by the following consideration. If the quanta of the Maxwell field, i.e. the photons, had a finite rest mass M, then it would appear due to dimensional reasons together with spatio-temporal derivatives as a mass term in the equations of motion in form of the inverse Compton wave length (8.21). Thus, performing the limit of a vanishing rest mass, i.e. $M \to 0$, also the Planck constant $\hbar$ vanishes automatically from the respective equations of motion.

In this chapter we first review the relativistic covariant formulation of this first-quantized Maxwell theory. Afterwards, we invoke the canonical field quantization formalism and work out systematically the second quantization of the Maxwell theory. In particular, we have to deal with the intricate consequences of the underlying local gauge symmetry, which occur due to the vanishing rest mass of the quanta of the Maxwell field. In this way we determine step by step the respective properties of a single photon as, for instance, its energy, its momentum, and its spin. Finally, we discuss the photon propagator, which represents an important building block in the Feynman diagrams of quantum electrodynamics describing the interaction between light and matter.

9.1 Maxwell Equations

Forces of an electromagnetic field upon electric charges, which are at rest or move, are mediated by both the electric field strength $\mathbf{E}$ and the magnetic induction $\mathbf{B}$. Physically both vector fields are generated by the charge density ρ and the current density $\mathbf{j}$. Mathematically they are determined by partial differential equations, which were first formulated by James Clerk Maxwell. The general structure of the Maxwell equations is prescribed by the Helmholtz vector decomposition theorem, which states that any vector field is uniquely determined by its respective divergence and rotation in combination with appropriate boundary conditions. With this the electric field strength $\mathbf{E}$ follows from

$$\mathrm{div}\,\mathbf{E} = \frac{\rho}{\varepsilon_0}\,, \tag{9.1}$$

$$\mathrm{rot}\,\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}\,, \tag{9.2}$$

whereas the magnetic induction $\mathbf{B}$ is defined by

$$\mathrm{div}\,\mathbf{B} = 0\,, \tag{9.3}$$

$$\mathrm{rot}\,\mathbf{B} = \mu_0\,\mathbf{j} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}\,. \tag{9.4}$$

Here the vacuum dielectric constant ε_0, the vacuum permeability μ_0, and the vacuum light velocity c are related via

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}\,. \tag{9.5}$$

We remark that (9.1), (9.4) and (9.2), (9.3) are denoted as the inhomogeneous and homogeneous Maxwell equations, respectively. Furthermore, we read off from the inhomogeneous Maxwell equations (9.1) and (9.4) the consistency equation that charge density ρ and current density $\mathbf{j}$ are not independent from each other but must fulfill the continuity equation

$$\frac{\partial \rho}{\partial t} + \mathrm{div}\,\mathbf{j} = 0\,, \tag{9.6}$$

which corresponds to the charge conservation similar to the discussion in (8.39)–(8.41). Note that we formulate the Maxwell equations (9.1)–(9.4) according to the International System of Units, which is abbreviated by SI from the French Système International d'Unités. Instead, in quantum field theory quite often the rational Lorentz-Heaviside unit system is used, where one assumes $\varepsilon_0 = \mu_0 = c = 1$ in order to simplify the notation. But we stick consistently to the SI unit system, although this might be considered to be more cumbersome, as this has the advantage that at each stage of the calculation one obtains results, which are, at least in principle, directly accessible in an experiment.

9.2 Local Gauge Symmetry

From the homogeneous Maxwell equations (9.2) and (9.3) we conclude straight-forwardly that both the electric field strength $\mathbf{E}$ and the magnetic induction $\mathbf{B}$ follow from differentiation of a scalar field φ and a vector potential $\mathbf{A}$:

$$\mathbf{B} = \mathrm{rot}\,\mathbf{A}\,, \tag{9.7}$$

$$\mathbf{E} = -\mathrm{grad}\,\varphi - \frac{\partial \mathbf{A}}{\partial t}\,. \tag{9.8}$$

From the inhomogeneous Maxwell equations (9.1) and (9.4) as well as from (9.7) and (9.8) we then determine coupled partial differential equations for the scalar field φ and the vector potential $\mathbf{A}$:

$$-\Delta\,\varphi - \frac{\partial}{\partial t}\,\mathrm{div}\,\mathbf{A} = \frac{\rho}{\varepsilon_0}\,, \tag{9.9}$$

$$\frac{1}{c^2}\,\frac{\partial^2\mathbf{A}}{\partial t^2} - \Delta\mathbf{A} + \mathrm{grad}\left(\frac{1}{c^2}\,\frac{\partial\varphi}{\partial t} + \mathrm{div}\,\mathbf{A}\right) = \mu_0\,\mathbf{j}\,. \tag{9.10}$$

The equations (9.7)–(9.10) turn out to be invariant with respect to a local gauge transformation with an arbitrary gauge function Λ:

$$\varphi' = \varphi + \frac{\partial\Lambda}{\partial t}\,, \tag{9.11}$$

$$\mathbf{A}' = \mathbf{A} - \mathrm{grad}\,\Lambda\,. \tag{9.12}$$

Thus, a local gauge transformation does not have any physical consequences, but it changes the mathematical description of the electromagnetic field. For instance, choosing a particular gauge allows to decouple the coupled equations of motion (9.9) and (9.10). In the following we briefly discuss the two most prominent gauges.

The *Coulomb gauge* assumes that the longitudinal part of the vector potential $\mathbf{A}$ vanishes, i.e.

$$\mathrm{div}\,\mathbf{A} = 0\,. \tag{9.13}$$

With this (9.9) and (9.10) reduce to

$$\Delta\,\varphi = -\frac{\rho}{\varepsilon_0}\,, \tag{9.14}$$

$$\frac{1}{c^2}\,\frac{\partial^2\mathbf{A}}{\partial t^2} - \Delta\,\mathbf{A} = \mu_0\,\mathbf{j} - \frac{1}{c^2}\,\frac{\partial}{\partial t}\,\mathrm{grad}\varphi\,. \tag{9.15}$$

As the scalar potential $\varphi(\mathbf{x},t)$ obeys the Poisson equation (9.14), it is determined at each time instant t by the corresponding value of the charge density $\rho(\mathbf{x},t)$ according to

$$\varphi(\mathbf{x},t) = \int d^3x'\,\frac{\rho(\mathbf{x}',t)}{4\pi\varepsilon_0|\mathbf{x}-\mathbf{x}'|}\,. \tag{9.16}$$

Due to (9.13) and (9.16) we conclude that from the original four fields φ and $\mathbf{A}$ only two of them represent dynamical degrees of freedom. As a consequence, the quantization of the electromagnetic field thus yields later on two types of photons. The advantage of the Coulomb gauge is that the remaining two dynamical degrees of freedom of the electromagnetic field can be physically identified with the two transversal degrees of freedom of the vector potential $\mathbf{A}$. The disadvantage of the Coulomb gauge is that it is not manifestly Lorentz invariant. Thus, the Coulomb gauge is only valid in a particular inertial system.

The *Lorentz gauge* is defined via

$$\frac{1}{c^2}\,\frac{\partial\varphi}{\partial t} + \mathrm{div}\,\mathbf{A} = 0\,. \tag{9.17}$$

With this the coupled equations of motions (9.9) and (9.10) yield uncoupled wave equations:

$$\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} - \Delta\,\varphi = \frac{\rho}{\varepsilon_0}\,, \tag{9.18}$$

$$\frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} - \Delta\,\mathbf{A} = \mu_0\,\mathbf{j}\,. \tag{9.19}$$

The advantage is here that the Lorentz gauge (9.17) as well as the decoupled equations of motion (9.18), (9.19) are Lorentz invariant. On the other hand, the quantization of the electromagnetic field on the basis of the Lorentz gauge, as worked out by Suraj Gupta and Konrad Bleuler, turns out to have an essential disadvantage. Namely, apart from the two physical transversal degrees of freedom also an unphysical longitudinal degree of freedom of the electromagnetic field emerges, which has to be eliminated afterwards with some effort.

9.3 Field Strength Tensors

In view of a manifestly Lorentz invariant formulation of the Maxwell theory both the electric field strength **E** and the magnetic induction **B** are considered as elements of an anti-symmetric 4×4 matrix F, which is called the electromagnetic field strength tensor. Its contravariant components read

$$(F^{\mu\nu}) = \Big(F^{\mu\nu}(\mathbf{E},\mathbf{B})\Big) = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}, \tag{9.20}$$

which fulfill, indeed, the anti-symmetry condition:

$$F^{\mu\nu} = -F^{\nu\mu}\,. \tag{9.21}$$

Its corresponding covariant components

$$F_{\mu\nu} = g_{\mu\lambda}g_{\nu\kappa}\,F^{\lambda\kappa} \tag{9.22}$$

are given by

$$(F_{\mu\nu}) = \Big(F^{\mu\nu}(-\mathbf{E},\mathbf{B})\Big) = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}. \tag{9.23}$$

Furthermore, it turns out to be useful to introduce in addition the dual electromagnetic field strength tensor ${}^*F$ by contracting the electromagnetic field strength tensor F with the totally anti-symmetric unity tensor ϵ, which was already used in (6.143):

$${}^*F^{\mu\nu} = \frac{1}{2}\,\epsilon^{\mu\nu\lambda\kappa}\,F_{\lambda\kappa}\,. \tag{9.24}$$

Thus, its contravariant components turn out to be

$$\left({}^*F^{\mu\nu}\right) = \Big(F^{\mu\nu}(c\mathbf{B}, -\mathbf{E}/c)\Big) = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix} \tag{9.25}$$

and the covariant components

$${}^*F_{\mu\nu} = g_{\mu\lambda} g_{\nu\kappa} \, {}^*F^{\lambda\kappa} \tag{9.26}$$

result in

$$\left({}^*F_{\mu\nu}\right) = \Big(F^{\mu\nu}(-c\mathbf{B}, -\mathbf{E}/c)\Big) = \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & E_z/c & -E_y/c \\ -B_y & -E_z/c & 0 & E_x/c \\ -B_z & E_y/c & -E_x/c & 0 \end{pmatrix} . \tag{9.27}$$

With these definitions we can now concisely summarize the homogeneous Maxwell equations (9.2), (9.3) with the help of the dual electromagnetic field strength tensor ${}^*F$

$$\partial_\mu {}^*F^{\mu\nu} = 0 \, , \tag{9.28}$$

whereas, correspondingly, the inhomogeneous Maxwell equations (9.1), (9.4) can be united with the help of the electrodynamic field strength tensor F:

$$\partial_\mu F^{\mu\nu} = \mu_0 \, j^\nu \, . \tag{9.29}$$

Here the contravariant current density four-vector j^λ consists of both the charge density ρ in the temporal component and the current density $\mathbf{j}$ in the spatial components:

$$\left(j^\lambda\right) = (c\rho, \mathbf{j}) \, . \tag{9.30}$$

Indeed, taking into account (6.102), an explicit calculation reproduces the homogeneous Maxwell equations

$$\begin{aligned} \left(\partial_\mu {}^*F^{\mu\nu}\right) &= \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix} \\ &= \left(\operatorname{div}\mathbf{B}, -\frac{1}{c}\operatorname{rot}\mathbf{E} - \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}\right) = (0, \mathbf{0}) \end{aligned} \tag{9.31}$$

as well as also the inhomogeneous Maxwell equations

$$\begin{aligned} \left(\partial_\mu F^{\mu\nu}\right) &= \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \\ &= \left(\frac{1}{c}\operatorname{div}\mathbf{E}, \operatorname{rot}\mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}\right) = \mu_0 (c\rho, \mathbf{j}) \, . \end{aligned} \tag{9.32}$$

Evaluating the four-divergence of (9.29) one obtains due to the anti-symmetry (9.21) a consistency condition, which is the continuity equation for the contravariant current density

$$\partial_\nu \partial_\mu F^{\mu\nu} = \mu_0 \, \partial_\nu j^\nu \qquad \Longrightarrow \qquad \partial_\nu j^\nu = 0 \,. \tag{9.33}$$

Note that (9.33) represents the manifest Lorentz invariant formulation of (9.6).

9.4 Four-Vector Potential

We now combine both the scalar potential φ and the vector potential $\mathbf{A}$ to the contravariant four-vector potential

$$\left(A^\lambda\right) = \left(\frac{\varphi}{c}, \mathbf{A}\right) . \tag{9.34}$$

With this the relations (9.7) and (9.8) between the electric field strength $\mathbf{E}$ and the magnetic induction $\mathbf{B}$ as well as the scalar potential φ and the vector potential $\mathbf{A}$ are combined into one single relation between the electromagnetic field strength tensor $F^{\mu\nu}$ and the four-vector potential A^λ:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \,. \tag{9.35}$$

Here the contravariant nabla four-vector is defined via

$$(\partial^\mu) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\boldsymbol{\nabla}\right) . \tag{9.36}$$

For instance, we obtain from (9.34)–(9.36):

$$F^{01} = \partial^0 A^1 - \partial^1 A^0 = \frac{1}{c}\frac{\partial A_x}{\partial t} + \frac{1}{c}\frac{\partial \varphi}{\partial x} = -\frac{1}{c}\, E_x \,, \tag{9.37}$$

$$F^{12} = \partial^1 A^2 - \partial^2 A^1 = \frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x} = -B_z \,. \tag{9.38}$$

We remark that the definitions (9.24) and (9.35) have the consequence that the homogeneous Maxwell equations (9.28) are automatically fulfilled:

$$\partial_\mu {}^*F^{\mu\nu} = \frac{1}{2}\,\epsilon^{\mu\nu\lambda\kappa}\,\partial_\mu F_{\lambda\kappa} = \frac{1}{2}\,\epsilon^{\mu\nu\lambda\kappa}\left(\partial_\mu \partial_\lambda A_\kappa - \partial_\mu \partial_\kappa A_\lambda\right) = 0 \,. \tag{9.39}$$

Note that we have used here the anti-symmetry of the ϵ tensor and that we have assumed that the covariant four-vector potential fulfills the theorem of Schwarz, i.e. partial derivatives commute:

$$\left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu\right) A_\kappa = 0 \,. \tag{9.40}$$

Furthermore, due to the definition (9.35) the inhomogeneous Maxwell equations (9.29) go over into the manifest Lorentz invariant formulation of the coupled equations of motion (9.9) and (9.10):

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = \mu_0 \, j^\nu \,. \tag{9.41}$$

And, finally, the manifest Lorentz invariant formulation of the local gauge transformation (9.11), (9.12) reads

$$A'^{\mu} = A^{\mu} + \partial^{\mu}\Lambda\,. \tag{9.42}$$

Due to such local gauge transformations (9.42) the electromagnetic field strength tensor F defined via (9.35) does not change

$$F'^{\mu\nu} = \partial^{\mu}A'^{\nu} - \partial^{\nu}A'^{\mu} = \partial^{\mu}A^{\nu} + \partial^{\mu}\partial^{\nu}\Lambda - \partial^{\nu}A^{\mu} - \partial_{\nu}\partial_{\mu}\Lambda = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} = F^{\mu\nu}\,, \tag{9.43}$$

provided that the gauge function Λ also fulfills the theorem of Schwarz:

$$\left(\partial_{\mu}\partial_{\nu} - \partial_{\nu}\partial_{\mu}\right)\Lambda = 0\,. \tag{9.44}$$

Furthermore, we conclude from (9.24) and (9.43) that then also the dual electromagnetic field strength tensor ${}^{*}F$ is gauge invariant:

$${}^{*}F'^{\mu\nu} = {}^{*}F^{\mu\nu}\,. \tag{9.45}$$

Thus, finally, we conclude that the local gauge transformation (9.42) leaves both the homogeneous and the inhomogeneous Maxwell equations (9.28) and (9.29) invariant.

9.5 Euler-Lagrange Equations

Now we set up a covariant variational principle, whose Euler-Lagrange equations are equivalent to the Maxwell equations. According to (9.35) the electromagnetic field strength tensor is completely determined from the knowledge of the four-vector potential. Therefore, we take here the point of view that the primary dynamical degree of freedom is provided by the four-vector potential. As the homogeneous Maxwell equations (9.28) are already automatically fulfilled by defining (9.35), the covariant variational principle must only reproduce the inhomogeneous Maxwell equations (9.29) or (9.41).

The action $\mathcal{A}$ as a functional of the covariant components A_{ν} of the four-vector potential is defined as an integral of a Lagrange density $\mathcal{L}$ over a volume Ω of the four-dimensional space-time:

$$\mathcal{A}\left[A_{\nu}(\bullet)\right] = \frac{1}{c}\int_{\Omega} d^4x\,\mathcal{L}\,. \tag{9.46}$$

As the inhomogeneous Maxwell equations (9.29) or (9.41) are of second order in the derivatives of the four-vector potential, the Lagrange density can only contain derivatives up to first order:

$$\mathcal{L} = \mathcal{L}\left(A_{\nu}\left(x^{\lambda}\right); \partial_{\mu}A_{\nu}\left(x^{\lambda}\right)\right)\,. \tag{9.47}$$

The corresponding Hamilton principle states that the functional derivative of the action with respect to the covariant components of the four-vector potential vanishes:

$$\frac{\delta\mathcal{A}}{\delta A_{\nu}(x^{\lambda})} = 0\,. \tag{9.48}$$

The resulting Euler-Lagrange equations of this classical field theory then read

$$\frac{\partial \mathcal{L}}{\partial A_\nu(x^\lambda)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu(x^\lambda))} = 0 \,. \tag{9.49}$$

Thus, it remains to find a Lagrange density, whose Euler-Lagrange equations (9.49) coincide with the inhomogeneous Maxwell equations (9.29) or (9.41). As the Maxwell equations are Lorentz invariant, the same must also hold for the Lagrange density. To this end we perform the following covariant ansatz for the Lagrange density of the electrodynamic field:

$$\mathcal{L} = \alpha F^{\lambda\kappa} F_{\lambda\kappa} + \beta j^\lambda A_\lambda \,. \tag{9.50}$$

Here α and β denote some constants, which are fixed below. Taking into account (9.35) the ansatz (9.50) reduces after some straight-forward algebraic calculations to the expression

$$\mathcal{L} = 2\alpha g^{\lambda\rho} g^{\kappa\sigma} (\partial_\rho A_\sigma - \partial_\sigma A_\rho) \partial_\lambda A_\kappa + \beta j^\lambda A_\lambda \,. \tag{9.51}$$

With this we obtain the partial derivative

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = \beta j^\nu \tag{9.52}$$

and, correspondingly, due to (9.35) also

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = 4\alpha F^{\mu\nu} \,. \tag{9.53}$$

Thus, with (9.52) and (9.53) the Euler-Lagrange equations (9.49) turn out to be of the form

$$\partial_\mu F^{\mu\nu} = \frac{\beta}{4\alpha} j^\nu \,. \tag{9.54}$$

A comparison of (9.54) with the inhomogeneous Maxwell equations (9.29) allows to fix the constant β according to

$$\frac{\beta}{4\alpha} = \mu_0 \qquad \Longrightarrow \qquad \beta = 4\alpha\mu_0 \,. \tag{9.55}$$

Due to (9.55) the Lagrange density (9.50) is then given by

$$\mathcal{L} = \alpha F^{\mu\nu} F_{\mu\nu} + 4\alpha\mu_0 \, j^\nu A_\nu \,, \tag{9.56}$$

where the constant α is still not yet determined.

9.6 Hamilton Function

We consider now the free electrodynamic field, where neither electric charges nor currents are present:

$$\rho(\mathbf{x}, t) = 0 \,, \qquad \mathbf{j}(\mathbf{x}, t) = \mathbf{0} \,. \tag{9.57}$$

Furthermore, we restrict ourselves from now on to the Coulomb gauge (9.13) as it represents the basis of the standard formulation for the second quantization of the Maxwell theory and is commonly used in quantum optics. From (9.13), (9.16), and (9.57) we then conclude that the scalar potential vanishes:

$$\varphi(\mathbf{x}, t) = 0\,. \tag{9.58}$$

Note that (9.13) and (9.58) together is also known as the radiation gauge. From (9.14), (9.57), and (9.58) we then read off that the vector potential obeys the wave equation:

$$\frac{1}{c^2}\frac{\partial^2 \mathbf{A}(\mathbf{x}, t)}{\partial t^2} - \Delta \mathbf{A}(\mathbf{x}, t) = \mathbf{0}\,. \tag{9.59}$$

Thus, in radiation gauge the vector potential $\mathbf{A}(\mathbf{x}, t)$ is determined from solving the wave equation (9.59) by taking into account the Coulomb gauge (9.13). Once the vector potential is known, one obtains from (9.7) the magnetic induction, whereas the electric field (9.8) reduces due to the radiation gauge (9.13) and (9.58) to

$$\mathbf{E}(\mathbf{x}, t) = -\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t}\,. \tag{9.60}$$

Furthermore, the Lagrange density of the free electrodynamic field reads due to (9.20), (9.23), (9.56), and (9.57)

$$\mathcal{L} = 2\alpha \left(\mathbf{B}^2 - \frac{\mathbf{E}^2}{c^2} \right)\,. \tag{9.61}$$

Due to (9.7) and (9.60) the Lagrange density (9.61) can be expressed in terms of the vector potential:

$$\mathcal{L} = 2\alpha \left\{ \left[\boldsymbol{\nabla} \times \mathbf{A}(\mathbf{x}, t) \right]^2 - \frac{1}{c^2} \left[\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t} \right]^2 \right\}\,. \tag{9.62}$$

With this the momentum field $\boldsymbol{\pi}$, which is canonically conjugated to the vector potential $\mathbf{A}$, follows as

$$\boldsymbol{\pi}(\mathbf{x}, t) = \frac{\delta \mathcal{A}\left[\mathbf{A}(\bullet, \bullet)\right]}{\delta \frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t}} = \frac{\partial \mathcal{L}}{\partial \frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t}} = -\frac{4\alpha}{c^2}\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t}\,. \tag{9.63}$$

A subsequent Legendre transformation

$$\mathcal{H} = \boldsymbol{\pi}(\mathbf{x}, \mathbf{t})\,\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t} - \mathcal{L} \tag{9.64}$$

converts then the Lagrange density (9.62) to the Hamilton density

$$\mathcal{H} = -\frac{c^2}{8\alpha}\,\boldsymbol{\pi}(\mathbf{x}, t)^2 - 2\alpha \left[\boldsymbol{\nabla} \times \mathbf{A}(\mathbf{x}, t) \right]^2, \tag{9.65}$$

which should coincide with the well-known energy density of the free electromagnetic field in SI units

$$\mathcal{H} = \frac{\varepsilon_0}{2} \left[\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t} \right]^2 + \frac{1}{2\mu_0} \left[\boldsymbol{\nabla} \times \mathbf{A}(\mathbf{x}, t) \right]^2\,. \tag{9.66}$$

Here the first term corresponds to the electric field energy density due to (9.60), where the second terms stands for the magnetic field energy density due to (9.7). By taking into account (9.5) a comparison of (9.65) and (9.66) fixes the parameter α according to

$$\alpha = -\frac{1}{4\mu_0} \,. \tag{9.67}$$

Thus, we obtain from (9.5), (9.63), and (9.67) the following result for the momentum field:

$$\boldsymbol{\pi}(\mathbf{x},t) = \varepsilon_0 \, \frac{\partial \mathbf{A}(\mathbf{x},t)}{\partial t} \,. \tag{9.68}$$

This corresponds to the classical expression for the momentum $\mathbf{p} = m\dot{\mathbf{x}}$, provided we identify the coordinate $\mathbf{x}$ with the vector potential $\mathbf{A}$ and the mass m with the vacuum dielectric constant ε_0. Furthermore, a spatial integral over the Hamilton density yields the Hamilton function

$$H = \int d^3x \, \mathcal{H} \,, \tag{9.69}$$

which follows from (9.66) to be

$$H = \frac{1}{2} \int d^3x \left\{ \frac{1}{\varepsilon_0} \, \boldsymbol{\pi}(\mathbf{x},t)^2 + \frac{1}{\mu_0} \left[\boldsymbol{\nabla} \times \mathbf{A}(\mathbf{x},t) \right]^2 \right\} . \tag{9.70}$$

Note that the first (second) term represents the kinetic (potential) energy of the electromagnetic field. With an additional calculation the Hamilton function (9.70) can be simplified. To this end we consider

$$(\boldsymbol{\nabla} \times \mathbf{A})^2 = \epsilon_{jkl} \, \partial_k A_l \, \epsilon_{jmn} \, \partial_m A_n \,, \tag{9.71}$$

which reduces with the help of (6.56) to

$$(\boldsymbol{\nabla} \times \mathbf{A})^2 = \partial_k A_l \partial_k A_l - \partial_k \left(A_l \partial_l A_k \right) + A_l \partial_l \partial_k A_k \,. \tag{9.72}$$

Inserting (9.72) into (9.70), the second term vanishes due to applying the theorem of Gauß and the third term is zero in the Coulomb gauge (9.13), so we end up with

$$H = \frac{1}{2} \int d^3x \left[\frac{1}{\varepsilon_0} \, \pi_k(\mathbf{x},t) \pi_k(\mathbf{x},t) + \frac{1}{\mu_0} \, \partial_k A_l(\mathbf{x},t) \partial_k A_l(\mathbf{x},t) \right] . \tag{9.73}$$

9.7 Canonical Field Quantization

The electrodynamic field is now quantized by exchanging the fields $A_j(\mathbf{x},t)$ and $\pi_j(\mathbf{x},t)$ with their corresponding field operators $\hat{A}_j(\mathbf{x},t)$ and $\hat{\pi}_j(\mathbf{x},t)$. To this end we perform a bosonic field quantization and demand equal-time commutation relations. At first, we demand that the field operators $\hat{A}_j(\mathbf{x},t)$ and $\hat{\pi}_j(\mathbf{x},t)$ commute, as usual, among themselves, respectively:

$$\left[\hat{A}_k(\mathbf{x},t), \hat{A}_l(\mathbf{x}',t) \right]_- = 0, \tag{9.74}$$

$$\left[\hat{\pi}_k(\mathbf{x},t), \hat{\pi}_l(\mathbf{x}',t) \right]_- = 0. \tag{9.75}$$

But when it comes to the equal-time commutation relations between the field operators $\hat{A}_j(\mathbf{x}, t)$ and their canonical conjugated momentum field operators $\hat{\pi}_j(\mathbf{x}, t)$, the situation turns out to be more intriguing. Let us investigate whether naive equal-time commutation relations of the form

$$\left[\hat{A}_k(\mathbf{x}, t), \hat{\pi}_l(\mathbf{x}', t)\right]_- = i\hbar\, \delta_{kl}\delta(\mathbf{x} - \mathbf{x}') \tag{9.76}$$

are possible. On the one hand, a derivative with respect to x_k then yields at the left-hand side of (9.76) to

$$\partial_k \left[\hat{A}_k(\mathbf{x}, t), \hat{\pi}_l(\mathbf{x}', t)\right]_- = \left[\partial_k \hat{A}_k(\mathbf{x}, t), \hat{\pi}_l(\mathbf{x}', t)\right]_- = 0\,, \tag{9.77}$$

as we have to demand the quantized version of the Coulomb gauge (9.13):

$$\partial_j \hat{A}_j(\mathbf{x}, t) = 0\,. \tag{9.78}$$

On the other, a derivative with respect to x_k at the right-hand side of (9.76) leads to

$$i\hbar\, \delta_{kl}\partial_k \delta(\mathbf{x} - \mathbf{x}') = i\hbar\, \partial_l \delta(\mathbf{x} - \mathbf{x}') \neq 0\,, \tag{9.79}$$

i.e. to an expression, which is non-zero in obvious contradiction to (9.77). Therefore, we are forced to modify the naive equal-time commutation relations (9.76) in such a way that it becomes compatible with the quantized version of the Coulomb gauge (9.78). To this end we consider the Fourier transformed of the right-hand side of (9.76)

$$i\hbar\, \delta_{kl}\delta(\mathbf{x} - \mathbf{x}') = i\hbar \int \frac{d^3k}{(2\pi)^3}\, \delta_{kl}\, e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')} \tag{9.80}$$

and substitute this expression by a yet to be determined transversal delta function

$$i\hbar\, \delta^T_{kl}(\mathbf{x} - \mathbf{x}') = i\hbar \int \frac{d^3k}{(2\pi)^3}\, \delta^T_{kl}(\mathbf{k})\, e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')}\,. \tag{9.81}$$

The Fourier transformed of the transversal delta function is then fixed from demanding that the derivative of (9.81) with respect to x_k vanishes, i.e.

$$i\hbar\, \partial_k \delta^T_{kl}(\mathbf{x} - \mathbf{x}') = i\hbar \int \frac{d^3k}{(2\pi)^3}\, ik_k\, \delta^T_{kl}(\mathbf{k})\, e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')} = 0\,. \tag{9.82}$$

For this to be valid it is sufficient that the transversality condition

$$k_k\, \delta^T_{kl}(\mathbf{k}) = 0 \tag{9.83}$$

is fulfilled. By comparing (9.80) and (9.81) a suitable ansatz for the Fourier transformed of the transversal delta function reads

$$\delta^T_{kl}(\mathbf{k}) = \delta_{kl} + k_k k_l\, f(\mathbf{k})\,. \tag{9.84}$$

The yet unknown function $f(\mathbf{k})$ follows then from inserting (9.84) into (9.83):

$$f(\mathbf{k}) = -\frac{1}{\mathbf{k}^2}\,. \tag{9.85}$$

Thus, from (9.81), (9.84), and (9.85) we then conclude for the transversal delta function

$$\delta_{kl}^T(\mathbf{x}-\mathbf{x}') = \delta_{kl}\delta(\mathbf{x}-\mathbf{x}') + \partial_k'\partial_l' \int \frac{d^3k}{(2\pi)^3}\,\frac{1}{\mathbf{k}^2}\,e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')}\,. \tag{9.86}$$

The remaining integral is known, for instance, within the realm of electrostatics from determining the Green function of the Poisson equation and yields the Coulomb potential. Thus, we obtain for the transversal delta function

$$\delta_{kl}^T(\mathbf{x}-\mathbf{x}') = \delta_{kl}\,\delta(\mathbf{x}-\mathbf{x}') + \frac{1}{4\pi}\,\partial_k'\partial_l'\,\frac{1}{|\mathbf{x}-\mathbf{x}'|}\,. \tag{9.87}$$

And, finally, we summarize our derivation by stating that the naive equal-time commutation relations (9.76) have to be modified by

$$\left[\hat{A}_k(\mathbf{x},t), \hat{\pi}_l(\mathbf{x}',t)\right]_- \;=\; i\hbar\,\delta_{kl}^T(\mathbf{x}-\mathbf{x}') \tag{9.88}$$

in order to be compatible with the quantized version of the Coulomb gauge (9.78).

However, one should be aware that such a derivation of commutation relations has an essential caveat. As hard as one tries to consistently determine such basic principles, they are always attached with heuristic elements. Whether commutation relations are at the end correct or not can only be verified by checking any prediction following from them against experimental measurements. In this spirit we will show later on that demanding the bosonic equal-time commutation relations (9.74), (9.75), and (9.88) leads, indeed, to a consistent description of the electromagnetic field with the help of usual annihilation and creation operators for photons, i.e. the quanta of light.

9.8 Heisenberg Equations

Furthermore, proceeding with the second-quantized formalism, we obtain from the Hamilton function (9.73) the Hamilton operator

$$\hat{H} = \frac{1}{2}\int d^3x' \left[\frac{1}{\varepsilon_0}\,\hat{\pi}_k(\mathbf{x}',t)\hat{\pi}_k(\mathbf{x}',t) + \frac{1}{\mu_0}\,\partial_k'\hat{A}_l(\mathbf{x}',t)\partial_k'\hat{A}_l(\mathbf{x}',t)\right]\,. \tag{9.89}$$

Note that the order of the operators in (9.89) does not play a role due to the commutation relations (9.74) and (9.75). Let us now evaluate the Heisenberg equation (3.62) for the field operator

$$i\hbar\,\frac{\partial \hat{A}_j(\mathbf{x},t)}{\partial t} = \left[\hat{A}_j(\mathbf{x},t), \hat{H}\right]_- \tag{9.90}$$

by inserting therein the Hamilton operator (9.89). After applying (3.10) as well as the equal-time commutation relations (9.74), (9.75), and (9.88) we get at first

$$i\hbar \frac{\partial \hat{A}_j(\mathbf{x},t)}{\partial t} = \frac{i\hbar}{\varepsilon_0} \int d^3x' \, \delta^T_{jk}(\mathbf{x}-\mathbf{x}') \, \hat{\pi}_k(\mathbf{x}',t) \, . \tag{9.91}$$

Taking into account the transversal delta function (9.87), a partial integration yields

$$i\hbar \frac{\partial \hat{A}_j(\mathbf{x},t)}{\partial t} = \frac{i\hbar}{\varepsilon_0} \left[\hat{\pi}_j(\mathbf{x},t) - \frac{1}{4\pi} \int d^3x' \left(\partial'_j \frac{1}{|\mathbf{x}-\mathbf{x}'|} \right) \partial'_k \hat{\pi}_k(\mathbf{x}',t) \right] \, . \tag{9.92}$$

With this we reproduce the quantized version of (9.68), as the last term in (9.92) vanishes due to the quantized version of the Coulomb gauge (9.78):

$$\frac{\partial \hat{A}_j(\mathbf{x},t)}{\partial t} = \frac{1}{\varepsilon_0} \, \hat{\pi}_j(\mathbf{x},t) \, . \tag{9.93}$$

Correspondingly, the Heisenberg equation (3.62) for the momentum field operator reads

$$i\hbar \frac{\partial \hat{\pi}_j(\mathbf{x},t)}{\partial t} = \left[\hat{\pi}_j(\mathbf{x},t), \hat{H} \right]_- \, . \tag{9.94}$$

Using (3.10) as well as the equal-time commutation relations (9.74), (9.75), and (9.88) we get at first

$$i\hbar \frac{\partial \hat{\pi}_j(\mathbf{x},t)}{\partial t} = \frac{-i\hbar}{\mu_0} \int d^3x' \, \partial'_k \delta^T_{jl}(\mathbf{x}-\mathbf{x}') \, \partial'_k \hat{A}_l(\mathbf{x}',t) \, , \tag{9.95}$$

so a partial integration yields

$$i\hbar \frac{\partial \hat{\pi}_j(\mathbf{x},t)}{\partial t} = \frac{i\hbar}{\mu_0} \int d^3x' \, \delta^T_{jl}(\mathbf{x}-\mathbf{x}') \, \Delta' \hat{A}_l(\mathbf{x}',t) \, , \tag{9.96}$$

Due to the explicit form of the transversal delta function (9.87) and a partial integration we then get

$$i\hbar \frac{\partial \hat{\pi}_j(\mathbf{x},t)}{\partial t} = \frac{i\hbar}{\mu_0} \left[\partial_k \partial_k \hat{A}_j(\mathbf{x},t) - \frac{1}{4\pi} \int d^3x' \left(\partial'_j \frac{1}{|\mathbf{x}-\mathbf{x}'|} \right) \Delta' \, \partial'_l \hat{A}_l(\mathbf{x}',t) \right] \, . \tag{9.97}$$

With the quantized version of the Coulomb gauge (9.78) this reduces finally to

$$\frac{\partial \hat{\pi}_j(\mathbf{x},t)}{\partial t} = \frac{1}{\mu_0} \, \Delta \hat{A}_j(\mathbf{x},t) \, . \tag{9.98}$$

Thus, we conclude from (9.5), (9.93), and (9.98) that the field operator $\hat{\mathbf{A}}(\mathbf{x},t)$ obeys like the classical field $\mathbf{A}(\mathbf{x},t)$ in (9.59) the wave equation:

$$\frac{1}{c^2} \frac{\partial^2 \hat{\mathbf{A}}(\mathbf{x},t)}{\partial t^2} - \Delta \hat{\mathbf{A}}(\mathbf{x},t) = \mathbf{0} \, . \tag{9.99}$$

9.9 Decomposition in Plane Waves

The wave equation (9.99) can be solved with a Fourier decomposition into plane waves:

$$\hat{\mathbf{A}}(\mathbf{x},t) = \int d^3k\, \hat{\mathbf{A}}(\mathbf{k},t)\, e^{i\mathbf{kx}} . \tag{9.100}$$

Inserting (9.100) into (9.99) one obtains for the expansion operators $\hat{\mathbf{A}}(\mathbf{k},t)$ the differential equation of a harmonic oscillator:

$$\frac{\partial^2 \hat{\mathbf{A}}(\mathbf{k},t)}{\partial t^2} + \omega_{\mathbf{k}}^2\, \hat{\mathbf{A}}(\mathbf{k},t) = \mathbf{0} , \tag{9.101}$$

where the dispersion relation is given by

$$\omega_{\mathbf{k}} = c|\mathbf{k}| . \tag{9.102}$$

The general solution of (9.101) reads

$$\hat{\mathbf{A}}(\mathbf{k},t) = \hat{\mathbf{A}}^{(1)}(\mathbf{k})\, e^{-i\omega_{\mathbf{k}}t} + \hat{\mathbf{A}}^{(2)}(\mathbf{k})\, e^{i\omega_{\mathbf{k}}t} . \tag{9.103}$$

so that the field operator (9.100) results in

$$\hat{\mathbf{A}}(\mathbf{x},t) = \int d^3k \left[\hat{\mathbf{A}}^{(1)}(\mathbf{k}) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{A}}^{(2)}(\mathbf{k}) e^{i(\mathbf{kx}+\omega_{\mathbf{k}}t)}\right] . \tag{9.104}$$

Performing in the second integral the substitution $\mathbf{k} \to -\mathbf{k}$ and taking into account the symmetry of the dispersion relation (9.102), i.e.

$$\omega_{\mathbf{k}} = \omega_{-\mathbf{k}} , \tag{9.105}$$

the Fourier decomposition (9.104) is converted into

$$\hat{\mathbf{A}}(\mathbf{x},t) = \int d^3k \left[\hat{\mathbf{A}}^{(1)}(\mathbf{k}) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{A}}^{(2)}(-\mathbf{k}) e^{-i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\right] . \tag{9.106}$$

Thus, the adjoint field operator reads

$$\hat{\mathbf{A}}^{\dagger}(\mathbf{x},t) = \int d^3k \left[\hat{\mathbf{A}}^{(1)\dagger}(\mathbf{k}) e^{-i(\mathbf{kx}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{A}}^{(2)\dagger}(-\mathbf{k}) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\right] . \tag{9.107}$$

As the vector potential of electrodynamics is real, we demand that the field operator as its second-quantized counterpart is self-adjoint, i.e.

$$\hat{\mathbf{A}}(\mathbf{x},t) = \hat{\mathbf{A}}^{\dagger}(\mathbf{x},t) , \tag{9.108}$$

and conclude from (9.106) and (9.107):

$$\hat{\mathbf{A}}(\mathbf{k}) = \hat{\mathbf{A}}^{(1)}(\mathbf{k}) \quad , \quad \hat{\mathbf{A}}^{\dagger}(\mathbf{k}) = \hat{\mathbf{A}}^{(2)}(-\mathbf{k}) , \tag{9.109}$$

Inserting the finding (9.109) into the Fourier decomposition (9.106), we finally obtain

$$\hat{\mathbf{A}}(\mathbf{x},t) = \int d^3k \left[\hat{\mathbf{A}}(\mathbf{k}) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{A}}^{\dagger}(\mathbf{k}) e^{-i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\right] . \tag{9.110}$$

9.10 Construction of Polarization Vectors

Before we can continue with working out the second quantization of the Maxwell theory we have to acquire beforehand a more detailed understanding of the description of plane waves. To this end we define two linearly polarized plane waves with the wave vector $\mathbf{k}$ and the dispersion (9.102) via

$$\mathbf{A}_1(\mathbf{x},t) = A_1 \boldsymbol{\epsilon}_1 e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\,, \qquad \mathbf{A}_2(\mathbf{x},t) = A_2 \boldsymbol{\epsilon}_2 e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\,. \tag{9.111}$$

Here A_1, A_2 represent the respective complex-valued amplitudes and $\boldsymbol{\epsilon}_1$, $\boldsymbol{\epsilon}_2$ denote two complex-valued polarization vectors, which are orthonormal according to

$$\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^* = \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^* = 1\,, \qquad \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^* = 0\,. \tag{9.112}$$

Let us consider now the sum of those two linearly polarized plane waves:

$$\mathbf{A}(\mathbf{x},t) = \mathbf{A}_1(\mathbf{x},t) + \mathbf{A}_2(\mathbf{x},t) = \left(A_1\boldsymbol{\epsilon}_1 + A_2\boldsymbol{\epsilon}_2\right) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\,. \tag{9.113}$$

Provided that both complex amplitudes $A_1 = |A_1|e^{i\varphi}$ and $A_2 = |A_2|e^{i\varphi}$ have the same phase φ, also their sum (9.113) is linearly polarized and we get

$$\mathbf{A}(\mathbf{x},t) = A\boldsymbol{\epsilon} e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\,. \tag{9.114}$$

Here the resulting amplitude A is given by

$$A = \sqrt{|A_1|^2 + |A_2|^2}\, e^{i\varphi} \tag{9.115}$$

and the resulting polarization vector $\boldsymbol{\epsilon}$ has the angle

$$\vartheta = \arctan\frac{|A_2|}{|A_1|} \tag{9.116}$$

with respect to $\boldsymbol{\epsilon}_1$, see Fig. 9.1. However, in the more general case that both complex amplitudes $A_1 = |A_1|e^{i\varphi_1}$ and $A_2 = |A_2|e^{i\varphi_2}$ have different phases $\varphi_1 \neq \varphi_2$, the sum (9.113) represents an ellipticly polarized plane wave. Let us illustrate this for the simpler situation of a circularly polarized plane wave, which occurs provided that both complex amplitudes A_1 and A_2 have the same absolute value and their phases differ by 90°:

$$A_1 = \frac{A_0}{\sqrt{2}}\,, \qquad A_2 = \pm i\,\frac{A_0}{\sqrt{2}}\,. \tag{9.117}$$

Inserting (9.117) into (9.113) we obtain for the sum of the two linearly polarized plane waves

$$\mathbf{A}(\mathbf{x},t) = \frac{A_0}{\sqrt{2}}\,\left(\boldsymbol{\epsilon}_1 \pm i\boldsymbol{\epsilon}_2\right) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)}\,. \tag{9.118}$$

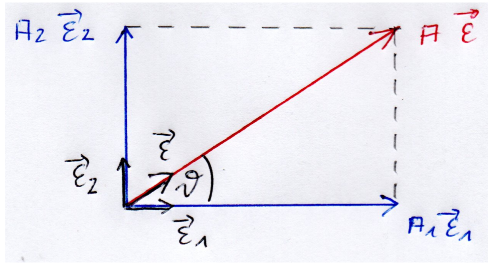


Figure 9.1: Adding two linearly polarized plane waves according to (9.113) with complex amplitudes A_1 and A_2, which have the same phase.

In order to be concrete we choose now the coordinate axes in such a way that the plane wave propagates in z-direction, whereas the two polarization vectors $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$, which are orthonormal according to (9.112), point in x- and y-direction:

$$\mathbf{k} = k \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} , \qquad \boldsymbol{\epsilon}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \qquad \boldsymbol{\epsilon}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} . \tag{9.119}$$

With this Eq. (9.118) reduces to

$$\mathbf{A}(\mathbf{x}, t) = \frac{A_0}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \\ 0 \end{pmatrix} e^{i(k\mathbf{e}_z\mathbf{x} - \omega_{k\mathbf{e}_z} t)} . \tag{9.120}$$

Considering the real part of the vector potential $\mathbf{A}(\mathbf{x}, t)$ at a fixed space point $\mathbf{x}$, it represents a vector in the xy-plane with constant absolute value A_0, which rotates on a circle with the frequency $\omega_{k\mathbf{e}_z}$:

$$\mathrm{Re}\, A_x(\mathbf{x}, t) = \frac{A_0}{\sqrt{2}} \cos\left(kz - \omega_{k\mathbf{e}_z} t\right) , \mathrm{Re}\, A_y(\mathbf{x}, t) = \mp \frac{A_0}{\sqrt{2}} \sin\left(kz - \omega_{k\mathbf{e}_z} t\right) , \mathrm{Re}\, A_z(\mathbf{x}, t) = 0 \,. \tag{9.121}$$

For the upper (lower) sign the rotation is performed anti-clockwise (clockwise) for an observer looking in the direction of the oncoming light beam. Such a plane wave is called in optics left-(right-) circularly polarized light, whereas in elementary particle physics one says that such a plane wave has positive (negative) helicity, see Fig. 9.2.

In view of a more detailed discussion of the helicity we remind us upon its definition in Eq. (6.188). Here the spin vector (6.166) of the electromagnetic field is given by the representation matrices $N^{\alpha\beta}$ of the Lorentz algebra in the space of the four-vector potential, which coincide with the representation matrices $L^{\alpha\beta}$ of the Lorentz algebra in the Minkowskian space-time according to (6.111) and (6.116). Thus, taking into account (6.53) and restricting us upon the spatial components, the helicity operator

$$\hat{h}(\mathbf{k}) = \frac{\mathbf{k}}{k} \mathbf{L} \tag{9.122}$$

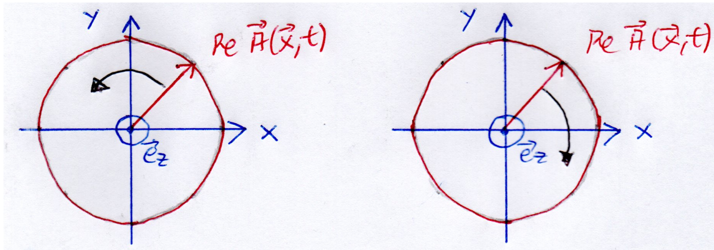


Figure 9.2: Adding two linearly polarized plane waves according to (9.113) with complex amplitudes A_1 and A_2 with the same absolute value and phases, which differ by 90°.

turns out to be defined by

$$\hat{h}(\mathbf{k}) = \frac{i}{k} \begin{pmatrix} 0 & -k_z & k_y \\ k_z & 0 & -k_x \\ -k_y & k_x & 0 \end{pmatrix} . \tag{9.123}$$

Now we introduce the polariation vectors $\boldsymbol{\epsilon}(\mathbf{k}, \lambda)$ for plane waves which propagate with the wave vector $\mathbf{k}$ and the helicity $\lambda = \pm 1$:

$$\mathbf{A}(\mathbf{x}, t) = A\boldsymbol{\epsilon}(\mathbf{k}, \lambda) e^{i(\mathbf{kx} - \omega_{\mathbf{k}} t)} . \tag{9.124}$$

Here the polarization vectors $\boldsymbol{\epsilon}(\mathbf{k}, \lambda)$ represent the eigenvectors of the helicity operator (9.122) with the eigenvalues λ:

$$\hat{h}(\mathbf{k})\boldsymbol{\epsilon}(\mathbf{k}, \lambda) = \lambda \boldsymbol{\epsilon}(\mathbf{k}, \lambda) . \tag{9.125}$$

From (9.120) and (9.124) we read off the polarization vectors $\boldsymbol{\epsilon}(k\mathbf{e}_z, \lambda)$ for a propagation in z-direction:

$$\boldsymbol{\epsilon}(k\mathbf{e}_z, \lambda) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \lambda i \\ 0 \end{pmatrix} . \tag{9.126}$$

Indeed, the polarization vectors (9.126) fulfill due to (9.123) the eigenvalue problem

$$\hat{h}(k\mathbf{e}_z)\boldsymbol{\epsilon}(k\mathbf{e}_z, \lambda) = \lambda \boldsymbol{\epsilon}(k\mathbf{e}_z, \lambda) . \tag{9.127}$$

Now we construct the polarization vectors $\boldsymbol{\epsilon}(\mathbf{k}, \lambda)$ with a general wave vector $\mathbf{k}$ by rotating the polarization vectors $\boldsymbol{\epsilon}(k\mathbf{e}_z, \lambda)$ in the same way as the original wave vector $k\mathbf{e}_z$. To this end we need the rotation matrix $R(\theta, \phi)$, which rotates the original wave vector $k\mathbf{e}_z$ to the general wave vector $\mathbf{k}$, where the latter is described in terms of spherical coordinates k, θ, and ϕ:

$$\mathbf{k} = k \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\phi \end{pmatrix} . \tag{9.128}$$

Here the rotation matrix $R(\theta, \phi)$ is constructed such that first the rotation $R_y(\theta)$ around the y-axis with angle θ and then the rotation $R_z(\phi)$ around the z-axis with angle ϕ is applied:

$$R(\theta, \phi) = R_z(\phi)\, R_y(\theta)\,. \tag{9.129}$$

The individual rotation matrices follow from evaluating matrix exponential functions

$$R_z(\phi) = e^{-iL_3\phi} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{9.130}$$

$$R_y(\theta) = e^{-iL_2\theta} = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \tag{9.131}$$

where the respective generators stem from (6.53). As a result we obtain for the rotation matrix (9.129)

$$R(\theta, \phi) = \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{9.132}$$

Indeed, the rotation matrix $R(\theta, \phi)$ maps the original wave vector $k\mathbf{e}_z$ to the general wave vector (9.128) as follows from the third column of (9.132):

$$R(\theta, \phi)k\mathbf{e}_z = \mathbf{k}\,. \tag{9.133}$$

Transforming correspondingly also the polarization vectors $\boldsymbol{\epsilon}(k\mathbf{e}_z, \lambda)$ from (9.126) with the rotation matrix $R(\theta, \phi)$, i.e.

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda) = R(\theta, \phi)\boldsymbol{\epsilon}(k\mathbf{e}_z, \lambda)\,, \tag{9.134}$$

we obtain the explicit result

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda) = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\theta\cos\phi - \lambda i\sin\phi \\ \cos\theta\sin\phi + \lambda i\cos\phi \\ -\sin\theta \end{pmatrix}. \tag{9.135}$$

Indeed, taking into account (9.123) and (9.128) one can show that the polarization vectors (9.135) fulfill the eigenvalue problem of the helicity operator (9.125). Furthermore, as expected, the polarization vectors (9.135) reduce for the special case $\theta = \phi = 0$ to the original polarization vectors (9.126).

9.11 Properties of Polarization Vectors

Due to the second-quantized formulation of the Coulomb gauge (9.78) the Fourier operators $\hat{\mathbf{A}}(\mathbf{k})$ in the decomposition (9.110) must obey the transversality condition

$$\mathbf{k}\hat{\mathbf{A}}(\mathbf{k}) = 0\,. \tag{9.136}$$

This means that the Fourier operators $\hat{\mathbf{A}}(\mathbf{k})$ have two transversal dynamical degrees of freedom. Performing the ansatz

$$\hat{\mathbf{A}}(\mathbf{k}) = N_{\mathbf{k}} \sum_{\lambda=\pm 1} \boldsymbol{\epsilon}(\mathbf{k}, \lambda) \hat{a}_{\mathbf{k},\lambda} \tag{9.137}$$

with some normalization constants $N_{\mathbf{k}}$ the transversality condition (9.136) is fulfilled provided that the polarization vectors $\boldsymbol{\epsilon}(\mathbf{k}, \lambda)$ are perpendicular to the propagation direction, which is defined by the wave vector $\mathbf{k}$:

$$\mathbf{k}\boldsymbol{\epsilon}(\mathbf{k}, \lambda) = 0\,. \tag{9.138}$$

Due to (9.128) it is straight-forward to show that the polarization vectors determined in (9.135) obey (9.138).

As another property of the polarization vectors (9.135) we investigate whether they obey orthonormality relations. Showing separately

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda)\boldsymbol{\epsilon}(\mathbf{k}, \lambda)^* = 1\,, \tag{9.139}$$

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda)\boldsymbol{\epsilon}(\mathbf{k}, -\lambda)^* = 0\,, \tag{9.140}$$

we arrive, indeed, due to $\lambda = \pm 1$ at the orthonormality relations

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda)\boldsymbol{\epsilon}(\mathbf{k}, \lambda')^* = \delta_{\lambda,\lambda'}\,. \tag{9.141}$$

Another property of the polarization vectors (9.135), which will turn out to be quite useful for later calculations, is their behaviour concerning the inversion $\mathbf{k} \to -\mathbf{k}$. Obviously, such an inversion is obtained in spherical coordinates (9.126) via

$$\phi \to \phi + \pi : \qquad \sin\phi \to -\sin\phi\,, \qquad \cos\phi \to -\cos\phi\,, \tag{9.142}$$

$$\theta \to \theta - \pi : \qquad \sin\theta \to \sin\theta\,, \qquad \cos\theta \to -\cos\theta\,. \tag{9.143}$$

With this we then conclude from (9.135

$$\boldsymbol{\epsilon}(-\mathbf{k}, \lambda) = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\theta\cos\phi + \lambda i \sin\phi \\ \cos\theta\sin\phi - \lambda i \cos\phi \\ -\sin\theta \end{pmatrix}\,. \tag{9.144}$$

Thus, from (9.135) and (9.144) we read off

$$\boldsymbol{\epsilon}(-\mathbf{k}, \lambda) = \boldsymbol{\epsilon}(\mathbf{k}, -\lambda) = \boldsymbol{\epsilon}(\mathbf{k}, \lambda)^*\,. \tag{9.145}$$

And, inserting the decomposition (9.137) into (9.110) by taking into account (9.145), we finally get

$$\hat{\mathbf{A}}(\mathbf{x}, t) = \sum_{\lambda=\pm 1} \int d^3k\, N_{\mathbf{k}} \left[\boldsymbol{\epsilon}(\mathbf{k}, \lambda) e^{i(\mathbf{k}\mathbf{x} - \omega_{\mathbf{k}} t)} \hat{a}_{\mathbf{k},\lambda} + \boldsymbol{\epsilon}(\mathbf{k}, \lambda)^* e^{-i(\mathbf{k}\mathbf{x} - \omega_{\mathbf{k}} t)} \hat{a}^\dagger_{\mathbf{k},\lambda}\right]. \tag{9.146}$$

Note that this plane wave decomposition fulfills, indeed, the Coulomb gauge (9.78) due to the transversality condition (9.138). In the following we aim at unravelling the physical interpretation of the Fourier operators $\hat{a}_{\mathbf{k},\lambda}$ and $\hat{a}^{\dagger}_{\mathbf{k},\lambda}$ in the plane wave decomposition (9.146), which leads to straight-forward but quite lengthy calculations. Therefore, we relegate the respective technical details to the exercises and restrict ourselves in the subsequent four sections to present a concise summary of the corresponding derivations.

9.12 Fourier Operators

We start with noting the plane wave decomposition for the momentum field operator, which follows from (9.93) and (9.146):

$$\hat{\boldsymbol{\pi}}(\mathbf{x},t) = \sum_{\lambda=\pm 1} \int d^3k\, \epsilon_0\, N_{\mathbf{k}} \Big[- i\omega_{\mathbf{k}} \boldsymbol{\epsilon}(\mathbf{k},\lambda) e^{i(\mathbf{k}\mathbf{x}-\omega_{\mathbf{k}}t)} \hat{a}_{\mathbf{k},\lambda} + i\omega_{\mathbf{k}} \boldsymbol{\epsilon}(\mathbf{k},\lambda)^* e^{-i(\mathbf{k}\mathbf{x}-\omega_{\mathbf{k}}t)} \hat{a}^{\dagger}_{\mathbf{k},\lambda} \Big] . \tag{9.147}$$

The plane wave decompositions (9.146) and (9.147) for both the field operator $\hat{\mathbf{A}}(\mathbf{x},t)$ and the momentum field operator $\hat{\boldsymbol{\pi}}(\mathbf{x},t)$ can now be solved for the Fourier operators $\hat{a}_{\mathbf{k},\lambda}$ and $\hat{a}^{\dagger}_{\mathbf{k},\lambda}$:

$$\hat{a}_{\mathbf{k},\lambda} = \frac{1}{2(2\pi)^3 N_{\mathbf{k}}} \int d^3x\, \boldsymbol{\epsilon}(\mathbf{k},\lambda)^* e^{-i(\mathbf{k}\mathbf{x}-\omega_{\mathbf{k}}t)} \left[\hat{\mathbf{A}}(\mathbf{x},t) + i\, \frac{\hat{\boldsymbol{\pi}}(\mathbf{x},t)}{\epsilon_0 \omega_{\mathbf{k}}} \right] , \tag{9.148}$$

$$\hat{a}^{\dagger}_{\mathbf{k},\lambda} = \frac{1}{2(2\pi)^3 N_{\mathbf{k}}} \int d^3x\, \boldsymbol{\epsilon}(\mathbf{k},\lambda) e^{i(\mathbf{k}\mathbf{x}-\omega_{\mathbf{k}}t)} \left[\hat{\mathbf{A}}(\mathbf{x},t) - i\, \frac{\hat{\boldsymbol{\pi}}(\mathbf{x},t)}{\epsilon_0 \omega_{\mathbf{k}}} \right] . \tag{9.149}$$

This allows us now to determine the commutator relations between the Fourier operators $\hat{a}_{\mathbf{k},\lambda}$ and $\hat{a}^{\dagger}_{\mathbf{k},\lambda}$ from the equal-time commutator relations (9.74), (9.75), and (9.88) for the field operator $\hat{\mathbf{A}}(\mathbf{x},t)$ and the momentum field operator $\hat{\boldsymbol{\pi}}(\mathbf{x},t)$:

$$\left[\hat{a}_{\mathbf{k},\lambda}, \hat{a}_{\mathbf{k}',\lambda'} \right]_- = 0 \,, \tag{9.150}$$

$$\left[\hat{a}^{\dagger}_{\mathbf{k},\lambda}, \hat{a}^{\dagger}_{\mathbf{k}',\lambda'} \right]_- = 0 \,, \tag{9.151}$$

$$\left[\hat{a}_{\mathbf{k},\lambda}, \hat{a}^{\dagger}_{\mathbf{k}',\lambda'} \right]_- = \frac{\hbar}{2(2\pi)^3 \epsilon_0 \omega_{\mathbf{k}} N_{\mathbf{k}}^2}\, \delta_{\lambda\lambda'}\, \delta(\mathbf{k}-\mathbf{k}') \,. \tag{9.152}$$

Thus, fixing the yet undetermined normalization constant according to

$$N_{\mathbf{k}} = \sqrt{\frac{\hbar}{2(2\pi)^3 \epsilon_0 \omega_{\mathbf{k}}}} \,, \tag{9.153}$$

we end up with the bosonic canonical commutation relation

$$\left[\hat{a}_{\mathbf{k},\lambda}, \hat{a}^{\dagger}_{\mathbf{k}',\lambda'} \right]_- = \delta_{\lambda\lambda'}\, \delta(\mathbf{k}-\mathbf{k}') \,. \tag{9.154}$$

This means that the Fourier operators $\hat{a}_{\mathbf{k},\lambda}$ and $\hat{a}^{\dagger}_{\mathbf{k},\lambda}$ can be interpreted as the annihilation and creation operators of bosonic particles, which are characterized by the wave vector $\mathbf{k}$ and the polarization λ. In order to determine the respective properties of these particles we investigate in the subsequent three sections their contribution to the energy, the momentum, and the spin angular momentum of the electromagnetic field in second quantization.

9.13 Energy

Taking into account the normalization constant (9.153) in the plane wave decompositions (9.146) and (9.147) for both the field operator $\hat{\mathbf{A}}(\mathbf{x},t)$ and the momentum field operator $\hat{\boldsymbol{\pi}}(\mathbf{x},t)$ we get

$$\hat{\mathbf{A}}(\mathbf{x},t) = \sum_{\lambda=\pm1} \int d^3k \sqrt{\frac{\hbar}{2(2\pi)^3\epsilon_0\omega_{\mathbf{k}}}} \left[\boldsymbol{\epsilon}(\mathbf{k},\lambda) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)} \hat{a}_{\mathbf{k},\lambda} + \boldsymbol{\epsilon}(\mathbf{k},\lambda)^* e^{-i(\mathbf{kx}-\omega_{\mathbf{k}}t)} \hat{a}^\dagger_{\mathbf{k},\lambda} \right], \quad (9.155)$$

$$\hat{\boldsymbol{\pi}}(\mathbf{x},t) = \sum_{\lambda=\pm1} \int d^3k \sqrt{\frac{\hbar\epsilon_0\omega_{\mathbf{k}}}{2(2\pi)^3}} \left[-i\boldsymbol{\epsilon}(\mathbf{k},\lambda) e^{i(\mathbf{kx}-\omega_{\mathbf{k}}t)} \hat{a}_{\mathbf{k},\lambda} + i\boldsymbol{\epsilon}_k(\mathbf{k},\lambda)^* e^{-i(\mathbf{kx}-\omega_{\mathbf{k}}t)} \hat{a}^\dagger_{\mathbf{k},\lambda} \right]. \quad (9.156)$$

Inserting (9.155) and (9.156) in the expression for the Hamilton operator (9.89) and using (9.5), (9.102), and (9.105) we yield

$$\hat{H} = \frac{1}{2} \sum_{\lambda=\pm1} \int d^3k \, \hbar\omega_{\mathbf{k}} \left(\hat{a}^\dagger_{\mathbf{k},\lambda} \hat{a}_{\mathbf{k},\lambda} + \hat{a}_{\mathbf{k},\lambda} \hat{a}^\dagger_{\mathbf{k},\lambda} \right). \quad (9.157)$$

Thus, comparing (3.6) with (9.157) we recognize that the second quantized electromagnetic field consists of independent harmonic oscillators, where each energy $\hbar\omega_{\mathbf{k}}$ is doubly degenerate due to the polarization degree of freedome λ. Defining the vacuum state as usual

$$\hat{a}_{\mathbf{k},\lambda}|0\rangle = 0 \qquad \Longleftrightarrow \qquad \langle 0|\hat{a}^\dagger_{\mathbf{k},\lambda} = 0\,, \quad (9.158)$$

we find that the vacuum energy of the electrodynamic field is given by a sum of the zero-point energy of all independent harmonic oscillators

$$\langle 0|\hat{H}|0\rangle = \int d^3k \, \hbar\omega_{\mathbf{k}}\,, \quad (9.159)$$

which turns out to be divergent due to the linear dispersion (9.102). Therefore, using the commutator relation (9.154) we obtain for the renormalized Hamilton operator

$$\hat{H} = \hat{H} - \langle 0|\hat{H}|0\rangle \quad (9.160)$$

the normal ordered result

$$\hat{H} = \sum_{\lambda=\pm1} \int d^3k \, \hbar\omega_{\mathbf{k}} \, \hat{a}^\dagger_{\mathbf{k},\lambda} \hat{a}_{\mathbf{k},\lambda}\,. \quad (9.161)$$

Here $\hat{a}^\dagger_{\mathbf{k},\lambda}\hat{a}_{\mathbf{k},\lambda}$ represents the occupation number operator, which counts the number of photons with wave vector $\mathbf{k}$ and polarization λ once it is applied to a photon state.

9.14 Momentum

Applying the Noether theorem from Chapter 7 to the Maxwell field yields according to the exercises the momentum of the electromagnetic field:

$$\mathbf{P} = \int d^3x \, \frac{\mathbf{S}(\mathbf{x},t)}{c^2} \quad (9.162)$$

with the Poynting vector

$$\mathbf{S}(\mathbf{x},t) = \frac{1}{\mu_0}\,\mathbf{E}(\mathbf{x},t) \times \mathbf{B}(\mathbf{x},t)\,. \tag{9.163}$$

Taking into account (9.5), (9.7), and (9.60), the momentum (9.163) is expressed in terms of the vector potential and the canonically conjugated momentum field via

$$\mathbf{P} = \int d^3x\;[\boldsymbol{\nabla} \times \mathbf{A}(\mathbf{x},t)] \times \boldsymbol{\pi}(\mathbf{x},t)\,. \tag{9.164}$$

Thus, in second quantization, the momentum operator of the electromagnetic field reads

$$\hat{\mathbf{P}} = \int d^3x\left(\boldsymbol{\nabla} \times \hat{\mathbf{A}}(\mathbf{x},t)\right) \times \hat{\boldsymbol{\pi}}(\mathbf{x},t)\,. \tag{9.165}$$

The further evaluation is based on taking into account the plane wave decompositions (9.155) and (9.156) for both the field operator $\hat{\mathbf{A}}(\mathbf{x},t)$ and the momentum field operator $\hat{\boldsymbol{\pi}}(\mathbf{x},t)$. Furthermore, the symmetry of the dispersion relation (9.105), the vector identity

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{ac})\mathbf{b} - (\mathbf{bc})\mathbf{a}\,, \tag{9.166}$$

the transversality condition (9.138), the orthonormality relation (9.141), and (9.145) are needed. Subsequently, performing the substitution $\mathbf{k} \to -\mathbf{k}$ and applying (9.105), (9.150), (9.151), we get the expression

$$\hat{\mathbf{P}} = \sum_{\lambda=\pm 1} \int d^3k\,\frac{\hbar\mathbf{k}}{2}\left(\hat{a}^\dagger_{\mathbf{k},\lambda}\hat{a}_{\mathbf{k},\lambda} + \hat{a}_{\mathbf{k},\lambda}\hat{a}^\dagger_{\mathbf{k},\lambda}\right)\,. \tag{9.167}$$

Note that the vacuum state has a vanishing momentum

$$\langle 0|\hat{\mathbf{P}}|0\rangle = \int d^3k\,\hbar\mathbf{k} = \mathbf{0} \tag{9.168}$$

due to the odd symmetry of the integrand. Thus, taking into account the commutator relation (9.154) we recognize that (9.167) coincides with the renormalized momentum operator

$$\hat{\mathbf{P}} = \hat{\mathbf{P}} - \langle 0|\hat{\mathbf{P}}|0\rangle\;, \tag{9.169}$$

which finally yields the normal ordered result

$$\hat{\mathbf{P}} = \sum_{\lambda=\pm 1} \int d^3k\,\hbar\mathbf{k}\,\hat{a}^\dagger_{\mathbf{k},\lambda}\hat{a}_{\mathbf{k},\lambda}\,. \tag{9.170}$$

9.15 Spin Angular Momentum

According to the Noether theorem from Chapter 7, which is applied to the electromagnetic field in the exercises, the spin angular momentum of the electromagnetic is given by

$$\mathbf{S} = \int d^3x\,\mathbf{A}(\mathbf{x},t) \times \boldsymbol{\pi}(\mathbf{x},t)\,. \tag{9.171}$$

Thus, the corresponding second quantized spin angular momentum operator reads

$$\hat{\mathbf{S}} = \int d^3x \, \hat{\mathbf{A}}(\mathbf{x}, t) \times \hat{\boldsymbol{\pi}}(\mathbf{x}, t) \,. \tag{9.172}$$

Inserting the plane wave decompositions (9.155) and (9.156) for both the field operator $\hat{\mathbf{A}}(\mathbf{x}, t)$ and the momentum field operator $\hat{\boldsymbol{\pi}}(\mathbf{x}, t)$ and performing the substitution $\mathbf{k} \to -\mathbf{k}$ then yields the intermediate result

$$\hat{\mathbf{S}} = \sum_{\lambda=\pm 1} \sum_{\lambda'=\pm 1} \int d^3k \, \frac{i\hbar}{2} \left[\boldsymbol{\epsilon}(\mathbf{k}, \lambda) \times \boldsymbol{\epsilon}(\mathbf{k}, \lambda')^* \hat{a}_{\mathbf{k},\lambda} \hat{a}^\dagger_{\mathbf{k},\lambda'} + \boldsymbol{\epsilon}(\mathbf{k}, \lambda') \times \boldsymbol{\epsilon}(\mathbf{k}, \lambda)^* \hat{a}^\dagger_{\mathbf{k},\lambda} \hat{a}_{\mathbf{k},\lambda'} \right] . \tag{9.173}$$

Now we evaluate the vector product between two polarization vectors. At first we obtain from (9.145)

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda) \times \boldsymbol{\epsilon}(\mathbf{k}, -\lambda)^* = \mathbf{0} \,, \tag{9.174}$$

whereas we get from (9.128) and (9.135)

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda) \times \epsilon(\mathbf{k}, \lambda) = -i\lambda \, \frac{\mathbf{k}}{k} \,. \tag{9.175}$$

Thus, both (9.174) and (9.175) can be summarized by

$$\boldsymbol{\epsilon}(\mathbf{k}, \lambda) \times \boldsymbol{\epsilon}(\mathbf{k}, \lambda')^* = -i\lambda \, \frac{\mathbf{k}}{k} \, \delta_{\lambda\lambda'} \,. \tag{9.176}$$

With this the intermediate result (9.173) for the spin angular momentum operator of the electromagnetic field reduces to

$$\hat{\mathbf{S}} = \sum_{\lambda=\pm 1} \int d^3k \, \lambda \, \frac{\hbar}{2} \, \frac{\mathbf{k}}{k} \left(\hat{a}^\dagger_{\mathbf{k},\lambda} \hat{a}_{\mathbf{k},\lambda} + \hat{a}_{\mathbf{k},\lambda} \hat{a}^\dagger_{\mathbf{k},\lambda} \right) . \tag{9.177}$$

Thus, the vacuum state has a vanishing spin angular momentum

$$\langle 0 | \hat{\mathbf{S}} | 0 \rangle = \hbar \left(\sum_{\lambda=\pm 1} \lambda \right) \left(\int d^3k \, \frac{\mathbf{k}}{k} \right) = \mathbf{0} \tag{9.178}$$

due to the odd symmetry in both the summand and the integrand. Using the commutator relation (9.154) we read off that (9.177) coincides with the renormalized spin angular momentum operator

$$\hat{\mathbf{S}} = \hat{\mathbf{S}} - \langle 0 | \hat{\mathbf{S}} | 0 \rangle \,, \tag{9.179}$$

leading to the normal ordered result

$$\hat{\mathbf{S}} = \sum_{\lambda=\pm 1} \int d^3k \, \lambda\hbar \, \frac{\mathbf{k}}{k} \, \hat{a}^\dagger_{\mathbf{k},\lambda} \hat{a}_{\mathbf{k},\lambda} \,. \tag{9.180}$$

We observe that the decompositions of the second quantized expressions for the energy (9.161), the momentum (9.170), and the spin angular momentum (9.180) of the electromagnetic field turn out to be time independent and, thus, represent conserved quantities. Together with the commutator relations (9.150), (9.151), and (9.154) we furthermore conclude that the Fourier operators $\hat{a}_{\mathbf{k},\lambda}$ and $\hat{a}^\dagger_{\mathbf{k},\lambda}$ represent the annihilation and creation operators of photons with the energy $\hbar\omega_{\mathbf{k}}$, the momentum $\hbar\mathbf{k}$, and the spin angular momentum $\lambda\hbar\mathbf{k}/k$, where the latter amounts to the helicity $\lambda\hbar$.

9.16 Definition of Maxwell Propagator

In close analogy to the Klein-Gordon propagator (8.122) we now define also the Maxwell propagator as the vacuum expectation value of the time-ordered product of two field operators $\hat{A}^\mu(\mathbf{x},t)$ and $\hat{A}^\nu(\mathbf{x}',t')$:

$$D^{\mu\nu}(\mathbf{x},t;\mathbf{x}',t') = \left\langle 0 \left| \hat{T}\left(\hat{A}^\mu(\mathbf{x},t)\hat{A}^\nu(\mathbf{x}',t')\right)\right| 0 \right\rangle . \tag{9.181}$$

Taking into account the definition of the time ordering operator (8.123) the Maxwell propagator reads explicitly

$$D^{\mu\nu}(\mathbf{x},t;\mathbf{x}',t') = \Theta(t-t')\langle 0|\hat{A}^\mu(\mathbf{x},t)\hat{A}^\nu(\mathbf{x}',t')|0\rangle + \Theta(t'-t)\langle 0|\hat{A}^\nu(\mathbf{x}',t')\hat{A}^\mu(\mathbf{x},t)|0\rangle . \tag{9.182}$$

Due to the radiation gauge (9.13) and (9.58) the zeroth component of the field operator $\hat{A}^\mu(\mathbf{x},t)$ vanishes, so only the spatial components of the Maxwell propagator can be non-zero:

$$D^{\mu\nu}(\mathbf{x},t;\mathbf{x}',t') = 0 \qquad \text{if either } \mu = 0 \text{ or } \nu = 0 . \tag{9.183}$$

In order to determine the equation of motion for the spatial components of the Maxwell propagator we evaluate initially their first temporal partial derivative. To this end we take into account (8.124) as well as (9.74) and get from (9.182)

$$\begin{aligned}\frac{\partial D^{jk}(\mathbf{x},t;\mathbf{x}',t')}{\partial t} &= \Theta(t-t')\left\langle 0\left|\frac{\partial \hat{A}_j(\mathbf{x},t)}{\partial t}\hat{A}_k(\mathbf{x}',t')\right|0\right\rangle \\ &\quad +\Theta(t'-t)\left\langle 0\left|\hat{A}_k(\mathbf{x}',t')\frac{\partial \hat{A}_j(\mathbf{x},t)}{\partial t}\right|0\right\rangle .\end{aligned} \tag{9.184}$$

A subsequent time derivative then yields by applying (8.124), (9.88), (9.93), and (9.99)

$$\begin{aligned}\frac{\partial^2 D^{jk}(\mathbf{x},t;\mathbf{x}',t')}{\partial t^2} &= \frac{-i\hbar}{\epsilon_0}\,\delta(t-t')\,\delta^T_{jk}(\mathbf{x}-\mathbf{x}') \\ &+c^2\Delta\Big[\Theta(t-t')\,\langle 0|\hat{A}_j(\mathbf{x},t)\hat{A}_k(\mathbf{x}',t')|0\rangle + \Theta(t'-t)\,\langle 0|\hat{A}_k(\mathbf{x}',t')\hat{A}_j(\mathbf{x},t)|0\rangle\Big] .\end{aligned} \tag{9.185}$$

From (9.182) and (9.185) we then obtain the result that the Maxwell propagator represents the Green function of the wave equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta\right) D^{jk}(\mathbf{x},t;\mathbf{x}',t') = -i\hbar\mu_0\,\delta^T_{jk}(\mathbf{x}-\mathbf{x}')\delta(t-t') . \tag{9.186}$$

We remark that not the delta function but the transversal delta function appears at the right-hand side of the inhomogeneous wave equation (9.186) due to the chosen Coulomb gauge. Therefore, one calls $D^{jk}(\mathbf{x},t;\mathbf{x}',t')$ more specifically to be the transveral Maxwell propagator.

9.17 Calculation of Maxwell Propagator

In order to further evaluate the spatial components of the Maxwell propagator (9.182), we insert the plane wave decomposition (9.155) for the field operator $\hat{\mathbf{A}}(\mathbf{x}, t)$ and use the commutation relation (9.154):

$$\begin{aligned} D^{jk}(\mathbf{x}, t; \mathbf{x}', t') &= \sum_{\lambda=\pm 1} \int d^3k \, \frac{\hbar}{2(2\pi)^3 \epsilon_0 \omega_{\mathbf{k}}} \Big[\Theta(t - t') \epsilon_j(\mathbf{k}, \lambda) \epsilon_k(\mathbf{k}, \lambda)^* \, e^{i[\mathbf{k}(\mathbf{x}-\mathbf{x}') - \omega_{\mathbf{k}}(t-t')]} \\ &\quad + \Theta(t' - t) \epsilon_k(\mathbf{k}, \lambda) \epsilon_j(\mathbf{k}, \lambda)^* \, e^{-i[\mathbf{k}(\mathbf{x}-\mathbf{x}') - \omega_{\mathbf{k}}(t-t')]} \Big] . \end{aligned} \tag{9.187}$$

Performing in the second term the substitution $\lambda \to -\lambda$, this reduces due to (9.145) to

$$\begin{aligned} D^{jk}(\mathbf{x}, t; \mathbf{x}', t') &= \int d^3k \, \frac{\hbar}{2(2\pi)^3 \epsilon_0 \omega_{\mathbf{k}}} \, P^{jk}(\mathbf{k}) \, \Big[\Theta(t - t') e^{i[\mathbf{k}(\mathbf{x}-\mathbf{x}') - \omega_{\mathbf{k}}(t-t')]} \\ &\quad + \Theta(t' - t) e^{-i[\mathbf{k}(\mathbf{x}-\mathbf{x}') - \omega_{\mathbf{k}}(t-t')]} \Big] . \end{aligned} \tag{9.188}$$

Here we have introduced the polarization sum

$$P^{jk}(\mathbf{k}) = \sum_{\lambda=\pm 1} \epsilon_j(\mathbf{k}, \lambda) \epsilon_k(\mathbf{k}, \lambda)^* , \tag{9.189}$$

which is symmetric with respect to the wave vector according to (9.145)

$$P^{jk}(-\mathbf{k}) = P^{jk}(\mathbf{k}) . \tag{9.190}$$

The latter symmetry property allows to simplify (9.188) further by performing the substitution $\mathbf{k} \to -\mathbf{k}$ in the second term, yielding

$$D^{jk}(\mathbf{x}, t; \mathbf{x}', t') = \int d^3k \, \frac{\hbar P^{jk}(\mathbf{k})}{2(2\pi)^3 \epsilon_0 \omega_{\mathbf{k}}} \, e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')} \Big[\Theta(t - t') e^{-i\omega_{\mathbf{k}}(t-t')} + \Theta(t' - t) e^{i\omega_{\mathbf{k}}(t-t')} \Big] . \tag{9.191}$$

Now we evaluate the polarization sum (9.189) explicitly by taking into account the polar coordinate representations for both the wave vector in (9.128) and the polarization vectors in (9.135). This yields

$$\big(P^{jk}(\mathbf{k})\big) = \begin{pmatrix} 1 - k_x^2/\mathbf{k}^2 & -k_x k_y/\mathbf{k}^2 & -k_x k_z/\mathbf{k}^2 \\ -k_x k_y/\mathbf{k}^2 & 1 - k_y^2/\mathbf{k}^2 & -k_y k_z/\mathbf{k}^2 \\ -k_x k_z/\mathbf{k}^2 & k_y k_z/\mathbf{k}^2 & 1 - k_z^2/\mathbf{k}^2 \end{pmatrix} , \tag{9.192}$$

which is concisely summarized by

$$P^{jk}(\mathbf{k}) = \delta_{jk} - \frac{k_j k_k}{|\mathbf{k}|^2} . \tag{9.193}$$

With this we read off the transversality property of the polarization sum

$$k_j P^{jk}(\mathbf{k}) = 0 , \tag{9.194}$$

which implies the corresponding transversality property of the Maxwell propagator (9.191)

$$\partial_j D^{jk}(\mathbf{x}, t; \mathbf{x}', t') = 0\,. \tag{9.195}$$

Due to this transversality property, which originally stems from having chosen the Coulomb gauge, the transversal Maxwell propagator (9.191) is not Lorentz invariant. Therefore, we aim now for decomposing the transversal Maxwell propagator into a Lorentz invariant and a Lorentz non-invariant contribution.

9.18 Four-Dimensional Fourier Representation

To this end we rewrite at first the three-dimensional Fourier representation of the Maxwell propagator in (9.191) in terms of a four-dimensional Fourier representation by using an integral identity, which is analogous to one obtained in (8.162) and (8.166):

$$\lim_{\eta\downarrow 0}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\frac{e^{-i\omega(t-t')}}{\omega^2-\omega_{\mathbf{k}}^2+i\eta} = \frac{-i}{2\omega_{\mathbf{k}}}\left[\Theta(t-t')e^{-i\omega_{\mathbf{k}}(t-t')}+\Theta(t'-t)e^{i\omega_{\mathbf{k}}(t-t')}\right]. \tag{9.196}$$

With this we obtain

$$D^{jk}(\mathbf{x}, t; \mathbf{x}', t') = \lim_{\eta\downarrow 0}\int\frac{d^3k}{(2\pi)^3}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\frac{i\hbar}{\epsilon_0}\,P^{jk}(\mathbf{k})\,\frac{e^{i[\mathbf{k}(\mathbf{x}-\mathbf{x}')-\omega(t-t')]}}{\omega^2-\omega_{\mathbf{k}}^2+i\eta}\,. \tag{9.197}$$

Note that the four-dimensional Fourier representation of the Maxwell propagator (9.197) solves evidently the equation of motion (9.186) by taking into account (9.5), (9.87), and (9.193). Introducing the contravariant four-wave vector

$$\left(k^\lambda\right) = \left(k^0, \mathbf{k}\right) = (\omega/c, \mathbf{k}) \tag{9.198}$$

and edging the spatial components of the Maxwell propagator with zeros, we deduce from (9.197) by taking into account the dispersion (9.102)

$$D^{\mu\nu}(x^\lambda; x'^\lambda) = \lim_{\eta\downarrow 0}\int\frac{d^4k}{(2\pi)^4}\,\frac{i\hbar}{c\epsilon_0}\,\frac{e^{-ik_\lambda(x^\lambda-x'^\lambda)}}{k_\lambda k^\lambda+i\eta}\,P^{\mu\nu}(k^\lambda)\,, \tag{9.199}$$

where the polarization sum does not explicitly depend on k^0:

$$P^{\mu\nu}(k^\lambda) = -g^{\mu\nu} + \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & -k_j k_k/\mathbf{k}^2 \end{pmatrix}^{\mu\nu}. \tag{9.200}$$

This polarization sum projects due to the transversality property (9.194) into the two-dimensional subspace perpendicular to $(0, \mathbf{k})$. But this projection is not covariant as the zeroth component of the four-vector potential vanishes due to the radiation gauge (9.13) and (9.58). In order to investigate the non-covariance of the polarization sum and, thus, of the transversal Maxwell propagator, in more detail we introduce the time-like vector

$$(\xi^\lambda) = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} \tag{9.201}$$

and a space-like vector perpendicular to it

$$(\bar{k}^\lambda) = \begin{pmatrix} 0 \\ \mathbf{k}/|\mathbf{k}| \end{pmatrix} . \tag{9.202}$$

An explicit calculation then yields the following decomposition:

$$\bar{k}^\lambda = \frac{k^\lambda - (k\xi)\xi^\lambda}{\sqrt{(k\xi)^2 - k^2}} . \tag{9.203}$$

From (9.200)–(9.203) we then obtain for the polarization sum

$$P^{\mu\nu}(k^\lambda) = -g^{\mu\nu} - k^2 \frac{\xi^\mu \xi^\nu}{(k\xi)^2 - k^2} - \frac{k^\mu k^\nu - (k\xi)(k^\mu \xi^\nu + k^\nu \xi^\mu)}{(k\xi)^2 - k^2} . \tag{9.204}$$

All terms, which contain the time-like vector ξ, are not covariant. Inserting the polarization sum (9.204) into (9.199) the transversal Maxwell propagator decomposes into three terms:

$$D^{\mu\nu}(x;x') = D_{\mathrm{F}}^{\mu\nu}(x;x') - D_{\mathrm{C}}^{\mu\nu}(x;x') - D_{\mathrm{R}}^{\mu\nu}(x;x') . \tag{9.205}$$

The first term is the covariant Maxwell propagator of Feynman

$$D_{\mathrm{F}}^{\mu\nu}(x;x') = \lim_{\eta \downarrow 0} \int \frac{d^4k}{(2\pi)^4} \frac{\hbar}{c\epsilon_0} \frac{i g^{\mu\nu}}{k^2 + i\eta} e^{-ik(x-x')} , \tag{9.206}$$

which also follows from the Gupta-Bleuler quantization of the electromagnetic field. Later on, when we discuss the perturbative calculation of quantum electrodynamic processes, it turns out that the Maxwell propagator in Feynman diagrams can be identified without loss of generality with (9.206). The other two non-covariant terms in the transversal Maxwell propagator (9.205) turn out to not contribute to any physical result. The second term reads

$$D_{\mathrm{C}}^{\mu\nu}(\mathbf{x},t;\mathbf{x}',t') = \lim_{\eta \downarrow 0} \int \frac{d^4k}{(2\pi)^4} \frac{i\hbar}{c\epsilon_0} \frac{k^2 \xi^\mu \xi^\nu}{(k\xi)^2 - k^2} \frac{e^{-ik(x-x')}}{k^2 + i\eta} . \tag{9.207}$$

Note that (9.207) reduces due to (9.6), (9.198), and (9.201) to

$$D_{\mathrm{C}}^{\mu\nu}(\mathbf{x},t;\mathbf{x}',t') = \frac{i\hbar\mu_0}{4\pi} \delta^{\mu 0} \delta^{\nu 0} \frac{\delta(t-t')}{|\mathbf{x}-\mathbf{x}'|} . \tag{9.208}$$

With this we conclude that this contribution of the transversal Maxwell propagator is instantaneous and couples exclusively to the zeroth component of the four-current density, i.e. the charge density. And the third residual term in (9.205) reads

$$D_{\mathrm{R}}^{\mu\nu}(x;x') = \lim_{\eta \downarrow 0} \int \frac{d^4k}{(2\pi)^4} \frac{i\hbar}{c\epsilon_0} \frac{k^\mu k^\nu - (k\xi)(k^\mu \xi^\nu + \xi^\mu k^\nu)}{(k\xi)^2 - k^2} \frac{e^{-ikx}}{k^2 + i\eta} . \tag{9.209}$$

It contains contributions, which are proportional to either k^μ or k^ν. As the electromagnetic field couples to four-current densities, which fulfill the continuity equation (9.33), we have in Fourier space

$$j_\mu(k) k^\mu = 0 . \tag{9.210}$$

Therefore the integral over $D_{\mathrm{R}}^{\mu\nu}(x;x')$ contracted with conserved currents $j_\mu^{(1)}(x)$ and $j_\nu^{(2)}(x')$ produces a vanishing result:

$$\int d^4x \int d^4x' \, j_\mu^{(1)}(x) D_{\mathrm{R}}^{\mu\nu}(x^\lambda; x'^\lambda) j_\nu^{(2)}(x') = \int \frac{d^4k}{(2\pi)^4} \, j_\mu(-k) D_{\mathrm{R}}^{\mu\nu}(k) j_\nu(k) = 0 \, . \qquad (9.211)$$

Later on we demonstrate explicitly by discussing the concrete example of a scattering process that both contributions (9.207) and (9.209) of the transversal Maxwell propagator do, indeed, not contribute to any observable quantity like the cross section.

Chapter 10

Dirac Field

In particle physics, the Dirac equation is a relativistic wave equation, which was derived by the British physicist Paul Dirac in 1928 by unifying the principles of both the quantum theory and the theory of special relativity. It describes massive spin-1/2 particles such as electrons and quarks. Historically, it was validated by accounting for the fine details of the hydrogen spectrum in a rigorous way. The equation also implies the existence of a new form of matter, the so-called antimatter, previously unsuspected as well as unobserved. In 1932 the positron as the antiparticle of the electron was the first antimatter to be detected in the cosmic radiation by Carl David Anderson.

The wave function in the Dirac theory consists of four complex fields, which are called a spinor as it transforms differently with respect to Lorentz transformations than a vector. For instance, one needs a rotation around a fixed axis by 720° in order to recover the original spinor instead of 360° for a vector. In the non-relativistic limit one obtains the Pauli two-component wave function, whereas the Schrödinger equation deals only with a wave function of one complex field. Moreover, in the limit of zero mass, the Dirac equation reduces to the Weyl equation, which was supposed to describe massless neutrinos for decades.

In the following we derive at first the Dirac theory group theoretically by systematically working out the spinor representation of the Lorentz group. Although this derivation does not correspond to the historic one of Paul Dirac and is technically more involved, it has several advantages. On the one hand it emphasizes the Lorentz invariance as one of the fundamental building blocks of any quantum field theory and explains as a side effect why a four-component Dirac spinor is needed to describe a massive spin 1/2 particle. On the other hand it enables to construct plane wave solutions by boosting trivial plane wave solutions in the rest frame to a uniformly moving reference frame as an elegant alternative to plainly solving the underlying Dirac equation.

Then we show the invariance of the Dirac theory with respect to discrete symmetries like charge conjugation, parity transformation, and time inversion. With this we prove exemplarily the seminal CPT theorem, which represents a fundamental property of physical laws. It states

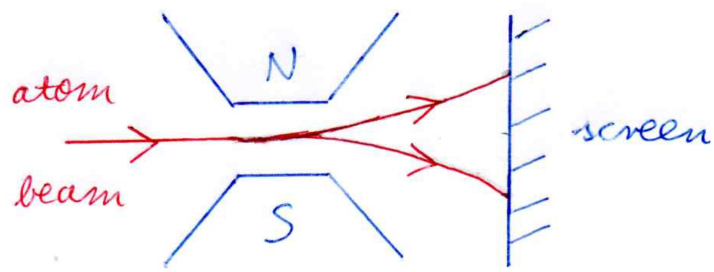


Figure 10.1: Set-up of the Stern-Gerlach experiment: a beam of silver or hydrogen atoms is split into two parts due to an inhomogeneous magnetic field.

that a mirror universe, where also all matter is replaced by antimatter, would evolve under exactly the same physical laws. As a consequence the masses and life-times of particles and antiparticles coincide.

Afterwards, we discuss how to quantize the Dirac theory within the realm of the canonical field quantization. With this we are able to deal with many massive spin 1/2 particles, whose description naturally also contains their respective antiparticles. And, finally, we determine the Dirac propagator, which describes the free motion of massive spin 1/2 particles and becomes important for a perturbative treatment of the light-matter interaction in terms of Feynman diagrams.

10.1 Pauli Matrices

The Stern-Gerlach experiment from 1922 involves sending a beam of silver or hydrogen atoms through an inhomogeneous magnetic field and observing their deflection. As each silver or hydrogen atom is in the ground state, its valence electron is in the $5s^1$ or the $1s^1$ state. Although the atoms should then not have any angular momentum, the beam is split into two parts, see Fig. 10.1. The reason for this is the spin angular momentum $s = 1/2$ of the valence electron, which leads to a residual magnetic moment of the atom and, thus, to a deflection in the applied inhomogeneous magnetic field. In order to mathematically describe the multiplicity of $2s+1 = 2$ spin degrees of freedom, Wolfgang Pauli introduced three complex 2×2 matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{10.1}$$

It is straight-forward to prove that the three Pauli matrices fulfill the following anti-commutators:

$$\left[\sigma^k, \sigma^l\right]_+ = 2\, \delta_{kl}\, I \,, \tag{10.2}$$

where I denotes the 2×2 unit matrix:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{10.3}$$

Here (10.2) means that the Pauli matrices represent a Clifford algebra with $N = 3$. Namely, a Clifford algebra with N generators $\xi^1, \ldots, \xi^N$ is defined by the anti-commutators

$$\left[\xi^k, \xi^l\right]_+ = 2\,\delta_{kl}\,. \tag{10.4}$$

But one can also convince oneself that the Pauli matrices additionally obey the commutators

$$\left[\sigma^k, \sigma^l\right]_- = 2i\,\epsilon_{klm}\,\sigma^m\,. \tag{10.5}$$

Here (10.5) means that the Pauli matrices also represent a Lie algebra with $N = 3$ generators. Namely, a Lie algebra with N generators $\xi^1, \ldots, \xi^N$ is defined by the commutators

$$\left[\xi^k, \xi^l\right]_- = i\,C_{klm}\,\xi^m\,, \tag{10.6}$$

where C_{klm} denote the structure constants of the Lie algebra. By adding (10.2) and (10.5) we result in the important calculation rule

$$\sigma^k \sigma^l = \delta_{kl} I + i\epsilon_{klm}\,\sigma^m\,, \tag{10.7}$$

which allows to simplify products of Pauli matrices.

10.2 Spinor Representation of Lorentz Algebra

With the help of the Pauli matrices one can construct two different representations of the Lorentz algebra. At first, we remark that the matrices

$$L_k = \frac{1}{2}\,\sigma^k \tag{10.8}$$

obey the commutator relations (6.57) of the generators of rotations. Furthermore, one can identify the generators of boosts via

$$M_k = \pm\,\frac{i}{2}\,\sigma^k\,, \tag{10.9}$$

where both signs are possible. In fact with the identifications (10.8), (10.9) also both commutator relations (6.58), (6.59) are valid. With this we define the following two representations of the Lorentz algebra:

$$D^{(1/2,0)}: \qquad (L_k, M_k) = \left(\frac{1}{2}\,\sigma^k, -\frac{i}{2}\,\sigma^k\right)\,, \tag{10.10}$$

$$D^{(0,1/2)}: \qquad (L_k, M_k) = \left(\frac{1}{2}\,\sigma^k, \frac{i}{2}\,\sigma^k\right)\,. \tag{10.11}$$

A general representation of the Lorentz algebra is characterized by $D^{(s_1,s_2)}$, where both quantum numbers s_1, s_2 can have all possible half-integer or integer values $0, 1/2, 1, 3/2, 2, \ldots$. It turns out that the space corresponding to the representation $D^{(s_1,s_2)}$ contains particles, whose spin lies in the interval $[|s_1 - s_2|, s_1 + s_2]$. In particular, particles with a single fixed spin s therefore belong to the representation $D^{(s,0)}$ or $D^{(0,s)}$. The trivial representation $D^{(0,0)}$ for a spinless particle assigns to each generator of the Lorentz algebra the number 1.

According to the Lie theorem of Section 6.5 the evaluation of the matrix-valued exponential function

$$D(\Lambda) = e^{-i\mathbf{L}\boldsymbol{\varphi} - i\mathbf{M}\boldsymbol{\xi}} \tag{10.12}$$

yields a representation of the Lorentz group, which corresponds to the representation of the Lorentz algebra. In both cases (10.10) and (10.11) we obtain from (10.12):

$$D^{(1/2,0)}(\Lambda) = \exp\left(-\frac{i}{2}\boldsymbol{\sigma}\boldsymbol{\varphi} - \frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\xi}\right) , \tag{10.13}$$

$$D^{(0,1/2)}(\Lambda) = \exp\left(-\frac{i}{2}\boldsymbol{\sigma}\boldsymbol{\varphi} + \frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\xi}\right) . \tag{10.14}$$

In the following we evaluate the respective matrix-valued exponential functions (10.13), (10.14) both for rotations $\boldsymbol{\xi} = \mathbf{0}$ and for boosts $\boldsymbol{\varphi} = \mathbf{0}$.

10.3 Spinor Representation of Rotations

According to (10.13) and (10.14) the spinor representation of rotations is given in both cases by

$$D\left(R(\boldsymbol{\varphi})\right) = \exp\left(-\frac{i}{2}\boldsymbol{\sigma}\boldsymbol{\varphi}\right) . \tag{10.15}$$

Due to the hermiticity of the Pauli matrices (10.1)

$$\left(\sigma^k\right)^\dagger = \sigma^k \tag{10.16}$$

the representation matrices of the rotations are unitary:

$$D\left(R(\boldsymbol{\varphi})\right)^\dagger = D\left(R(\boldsymbol{\varphi})\right)^{-1} . \tag{10.17}$$

Considering the Taylor series of the exponential function in (10.15) we evaluate separately the even and the odd terms:

$$D\left(R(\boldsymbol{\varphi})\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\frac{(\boldsymbol{\sigma}\boldsymbol{\varphi})^{2n}}{2^{2n}} - i\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\frac{(\boldsymbol{\sigma}\boldsymbol{\varphi})^{2n+1}}{2^{2n+1}} . \tag{10.18}$$

Applying the calculational rule (10.7) we obtain

$$\left(\boldsymbol{\sigma}\boldsymbol{\varphi}\right)^2 = \varphi_k\varphi_l\sigma^k\sigma^l = \varphi_k\varphi_l\left(\delta_{kl}\, I + i\epsilon_{klm}\,\sigma^m\right) = \boldsymbol{\varphi}^2\, I \,, \tag{10.19}$$

so that (10.18) leads to

$$D\left(R(\boldsymbol{\varphi})\right) = \left[\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \left(\frac{|\boldsymbol{\varphi}|}{2}\right)^{2n}\right] I - i \left[\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \left(\frac{|\boldsymbol{\varphi}|}{2}\right)^{2n+1}\right] \frac{\boldsymbol{\sigma\varphi}}{|\boldsymbol{\varphi}|} . \tag{10.20}$$

Taking into account the Taylor series of the trigonometric functions, one finally yields the spinor representation matrices for rotations

$$D\left(R(\boldsymbol{\varphi})\right) = I \cos\left(\frac{|\boldsymbol{\varphi}|}{2}\right) - i \frac{\boldsymbol{\sigma\varphi}}{|\boldsymbol{\varphi}|} \sin\left(\frac{|\boldsymbol{\varphi}|}{2}\right) , \tag{10.21}$$

which are, indeed, unitary (10.17) due to (10.16). Note that both representations $D^{(1/2,0)}$ and $D^{(0,1/2)}$ yield the same representation matrices for rotations. Furthermore, we observe that one needs in (10.21) a rotation of 4π instead of 2π in order to recover the identity. This is a consequence of the underlying spin 1/2 and represents a characteristic property for a spinor representation.

10.4 Spinor Representation of Boosts

According to (10.13) and (10.14) the representation of the boosts reads

$$D\left(B(\boldsymbol{\xi})\right) = \exp\left(\mp \frac{1}{2} \boldsymbol{\sigma\xi}\right) . \tag{10.22}$$

Due to the hermiticity of the Pauli matrices in (10.16) also the representation matrices of the boosts are hermitian:

$$D\left(B(\boldsymbol{\xi})\right)^{\dagger} = D\left(B(\boldsymbol{\xi})\right) . \tag{10.23}$$

The Taylor series of the matrix exponential function (10.22) is evaluated separately for even and odd terms:

$$D\left(B(\boldsymbol{\xi})\right) = \sum_{n=0}^{\infty} \frac{1}{(2n)!} \frac{(\boldsymbol{\sigma\xi})^{2n}}{2^{2n}} \mp \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \frac{(\boldsymbol{\sigma\xi})^{2n+1}}{2^{2n+1}} . \tag{10.24}$$

With the help of (10.19) this changes to

$$D\left(B(\boldsymbol{\xi})\right) = \left[\sum_{n=0}^{\infty} \frac{1}{(2n)!} \left(\frac{|\boldsymbol{\xi}|}{2}\right)^{2n}\right] I \mp \left[\sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \left(\frac{|\boldsymbol{\xi}|}{2}\right)^{2n+1}\right] \frac{\boldsymbol{\sigma\xi}}{|\boldsymbol{\xi}|} . \tag{10.25}$$

Taking into account the Taylor series of hyperbolic functions, one gets from (10.25) for the representation matrices (10.22) of the boosts

$$D\left(B(\boldsymbol{\xi})\right) = \exp\left(\mp\frac{1}{2} \boldsymbol{\sigma\xi}\right) = I \cosh\left(\frac{|\boldsymbol{\xi}|}{2}\right) \mp \frac{\boldsymbol{\sigma\xi}}{|\boldsymbol{\xi}|} \sinh\left(\frac{|\boldsymbol{\xi}|}{2}\right) . \tag{10.26}$$

As a reminder we note again that the upper and the lower sign stands for the representation $D^{(1/2,0)}$ and $D^{(0,1/2)}$, respectively. Furthermore, we remark that the representation matrices (10.26) are, indeed, hermitian (10.23) due to (10.16).

In order to simplify (10.26) further we consider now a particle of mass M in the rest frame, so that its contravariant four-momentum vector is given by

$$(p_{\mathrm{R}}^{\mu}) = (Mc, \mathbf{0}) \ . \tag{10.27}$$

Performing an active boost into the inertial frame the contravariant four-momentum vector (10.27) changes to

$$p^{\mu} = B^{\mu}{}_{\nu}(\boldsymbol{\xi}) p_{\mathrm{R}}^{\nu} \, , \tag{10.28}$$

where the respective matrix elements of the boost $B^{\mu}{}_{\nu}(\boldsymbol{\xi})$ were already determined in Section 6.7 in terms of the the rapidity $\boldsymbol{\xi}$. Using (6.79) we thus obtain

$$(p^{\mu}) = \left(p^0, \mathbf{p}\right) = \left(Mc \cosh|\boldsymbol{\xi}|, \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} Mc \sinh|\boldsymbol{\xi}|\right) \ . \tag{10.29}$$

Combining (10.29) with the hyperbolic Pythagoras

$$\cosh^2\alpha - \sinh^2\alpha = 1 \tag{10.30}$$

and the hyperbolic addition theorems

$$\cosh(\alpha+\beta) = \cosh\alpha\cosh\beta + \sinh(\alpha)\sinh\beta \, , \tag{10.31}$$

$$\sinh(\alpha+\beta) = \sinh(\alpha)\cosh\beta + \cosh\alpha\sinh\beta \, , \tag{10.32}$$

the following relations are derived:

$$\cosh\left(\frac{|\boldsymbol{\xi}|}{2}\right) = \sqrt{\frac{\cosh|\boldsymbol{\xi}|+1}{2}} = \sqrt{\frac{p^0+Mc}{2Mc}} \, , \tag{10.33}$$

$$\sinh\left(\frac{|\boldsymbol{\xi}|}{2}\right) = \sqrt{\frac{\cosh|\boldsymbol{\xi}|-1}{2}} = \sqrt{\frac{p^0-Mc}{2Mc}} \, , \tag{10.34}$$

$$\sinh(|\boldsymbol{\xi}|) = 2\sinh\left(\frac{|\boldsymbol{\xi}|}{2}\right)\cosh\left(\frac{|\boldsymbol{\xi}|}{2}\right) = \frac{\sqrt{(p^0-Mc)(p^0+Mc)}}{Mc} \ . \tag{10.35}$$

Using (10.33)–(10.35), the representation matrix (10.26) of the boost can be expressed by the components of the contravariant four-momentum vector (10.29)

$$D\left(B(\boldsymbol{\xi})\right) = \exp\left(\mp\frac{1}{2}\boldsymbol{\sigma\xi}\right) = I\sqrt{\frac{p^0+Mc}{2Mc}} \mp \frac{\boldsymbol{\sigma}\mathbf{p}}{Mc}\sqrt{\frac{p^0-Mc}{2Mc}}\frac{Mc}{\sqrt{(p^0-Mc)(p^0+Mc)}} \, , \tag{10.36}$$

yielding finally

$$D\left(B(\boldsymbol{\xi})\right) = \frac{(p^0+Mc)I \mp \boldsymbol{\sigma}\mathbf{p}}{\sqrt{2Mc(p^0+Mc)}} \ . \tag{10.37}$$

In the following it turns out to be technically advantageous to extend the three Pauli matrices σ^k by the unit matrix

$$\sigma^0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10.38}$$

to a four-vector of Pauli matrices:

$$(\sigma^\mu) = (\sigma^0, \sigma^k)\,. \tag{10.39}$$

Then (10.37) implies that the boost of the representation $D^{(1/2,0)}$ can be concisely written as

$$D^{(1/2,0)}\,(B(\boldsymbol{\xi})) = \exp\left(-\frac{1}{2}\,\boldsymbol{\sigma\xi}\right) = \frac{p\sigma + Mc}{\sqrt{2Mc(p^0 + Mc)}}\,, \tag{10.40}$$

where the scalar product between the four-vector of Pauli matrices (10.39) and the four-momentum vector is used:

$$p\sigma = p_\mu \sigma^\mu = p^0\sigma^0 - \mathbf{p}\boldsymbol{\sigma}\,. \tag{10.41}$$

Furthermore, we introduce the spatially inverted four-vector

$$\tilde{x} = (\tilde{x}^0, \tilde{x}^k) = (x^0, -x^k) \tag{10.42}$$

and, correspondingly, also the spatially inverted four-vector of Pauli matrices

$$\tilde{\sigma} = (\tilde{\sigma}^0, \tilde{\sigma}^k) = (\sigma^0, -\sigma^k)\,. \tag{10.43}$$

With this we read off from (10.37) that the boost of the representation $D^{(0,1/2}$ is given by

$$D^{(0,1/2)}\,(B(\boldsymbol{\xi})) = \exp\left(+\frac{1}{2}\,\boldsymbol{\sigma\xi}\right) = \frac{p\tilde{\sigma} + Mc}{\sqrt{2Mc(p^0 + Mc)}} \tag{10.44}$$

due to the scalar product

$$p\tilde{\sigma} = p_\mu \tilde{\sigma}^\mu = p^0\sigma^0 + \mathbf{p}\boldsymbol{\sigma}\,. \tag{10.45}$$

For various later calculations it turns out to be also useful to express the boost representations (10.40) and (10.44) as the square root of the same expression with a doubled rapidity. Indeed, taking into account (10.26) and (10.29) we obtain

$$\exp\left(\mp\frac{1}{2}\,\boldsymbol{\sigma\xi}\right) = \sqrt{\exp\left(\mp\boldsymbol{\sigma\xi}\right)} = \sqrt{\cosh|\boldsymbol{\xi}| \mp \frac{\boldsymbol{\sigma\xi}}{|\boldsymbol{\xi}|}\,\sinh|\boldsymbol{\xi}|} = \sqrt{\frac{p^0}{Mc} \mp \frac{\mathbf{p}\boldsymbol{\sigma}}{Mc}}\,. \tag{10.46}$$

Thus, together (10.41) and (10.45) we conclude

$$\exp\left(-\frac{1}{2}\,\boldsymbol{\sigma\xi}\right) = \sqrt{\frac{p\sigma}{Mc}}\,, \tag{10.47}$$

$$\exp\left(+\frac{1}{2}\,\boldsymbol{\sigma\xi}\right) = \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,. \tag{10.48}$$

Whenever we will use later on the spinor representations for boosts (10.47) and (10.48) we have to keep in mind that they present efficient shortcut notations for the more involved concrete expressions (10.40) and (10.44).

10.5 Lorentz Invariant Combinations of Weyl Spinors

So far we have constructed with $D^{(1/2,0)}$ and $D^{(0,1/2)}$ the smallest non-trivial representations of the Lorentz group. Now we define the corresponding Weyl spinors $\xi_\alpha(x)$ and $\eta^{\dot\alpha}(x)$ of type $(1/2,0)$ and $(0,1/2)$ upon which the representation matrices of the Lorentz group act. The different transformation properties of the Weyl spinors $\xi_\alpha(x)$ and $\eta^{\dot\alpha}(x)$ under a Lorentz transformation are expressed by using lower non-dotted and upper dotted indices, respectively:

$$\xi_\alpha(x) \qquad \longrightarrow \qquad \xi'_\alpha(x') = D^{(1/2,0)}(\Lambda)_\alpha{}^\beta\,\xi_\beta(x)\,, \tag{10.49}$$

$$\eta^{\dot\alpha}(x) \qquad \longrightarrow \qquad \eta'^{\dot\alpha}(x') = D^{(0,1/2)}(\Lambda)^{\dot\alpha}{}_{\dot\beta}\,\eta^{\dot\beta}(x)\,. \tag{10.50}$$

In the following we aim for constructing a Lorentz invariant action on the basis of using these Weyl spinors. To this end we restrict ourselves to consider quadratic terms in the Weyl spinors and their first partial derivatives.

At first, we only deal with quadratic terms in the Weyl spinors without any first partial derivative, which are needed to describe massive particles. In this case there are in total four different combinations of two Weyl spinors

$$\xi^\dagger\xi\,,\quad \eta^\dagger\eta\,,\quad \eta^\dagger\xi\,,\quad \xi^\dagger\eta\,, \tag{10.51}$$

which are converted by a Lorentz transformation Λ into

$$\begin{aligned} &\xi^\dagger D^{(1/2,0)}(\Lambda)^\dagger D^{(1/2,0)}(\Lambda)\xi\,,\quad \eta^\dagger D^{(0,1/2)}(\Lambda)^\dagger D^{(0,1/2)}(\Lambda)\eta\,,\\ &\eta^\dagger D^{(0,1/2)}(\Lambda)^\dagger D^{(1/2,0)}(\Lambda)\xi\,,\quad \xi^\dagger D^{(1/2,0)}(\Lambda)^\dagger D^{(0,1/2)}(\Lambda)\eta\,, \end{aligned} \tag{10.52}$$

respectively. In case of a rotation $\Lambda = R$ the representation matrices $D^{(1/2,0)}(R)$ and $D^{(0,1/2)}(R)$ coincide according to (10.13) and (10.14). Furthermore, we conclude from the unitarity (10.17) of these representation matrices that all four transformed combinations (10.52) are identical to the original combinations (10.51). But in case of a boost $\Lambda = B$ we read off from (10.13) and (10.14) that the representation matrices $D^{(1/2,0)}(B)$ and $D^{(0,1/2)}(B)$ are just inverse with respect to each other:

$$D^{(1/2,0)}\left(B\right) = D^{(0,1/2)}\left(B\right)^{-1}\,. \tag{10.53}$$

In combination with the hermiticity (10.23) of these representation matrices it follows then that only the last two of the transformed combinations (10.51) match with their corresponding original combinations (10.51). In summary, we conclude that a Lorentz invariant action without space-time derivatives is only possible by combining the two Weyl spinors ξ and η.

In order to describe a particle, which moves in space-time, the action must also contain first partial derivatives of the Weyl spinors. To this end we consider at first spatial derivatives and form all possible combinations of two Weyl spinors

$$\xi^\dagger\sigma^k\partial_k\xi\,,\quad \eta^\dagger\sigma^k\partial_k\eta\,,\quad \eta^\dagger\sigma^k\partial_k\xi\,,\quad \xi^\dagger\sigma^k\partial_k\eta\,. \tag{10.54}$$

They are converted by a Lorentz transformation Λ into

$$\begin{aligned}\xi^\dagger D^{(1/2,0)}(\Lambda)^\dagger\,\sigma^k\,D^{(1/2,0)}(\Lambda)\,\partial'_k\xi\,,\quad \eta^\dagger D^{(0,1/2)}(\Lambda)^\dagger\,\sigma^k\,D^{(0,1/2)}(\Lambda)\,\partial'_k\eta\,,\\ \eta^\dagger D^{(0,1/2)}(\Lambda)^\dagger\,\sigma^k\,D^{1/2,0)}(\Lambda)\,\partial'_k\xi\,,\quad \xi^\dagger D^{1/2,0)}(\Lambda)^\dagger\,\sigma^k\,D^{(0,1/2)}(\Lambda)\,\partial'_k\eta\,.\end{aligned} \tag{10.55}$$

In case of a rotation $\Lambda = R$, the representation matrices $D^{(1/2,0)}(R)$ and $D^{(0,1/2)}(R)$ are identical, so that due to (10.55) only the expression

$$D(R)^\dagger\,\sigma^k\,D(R) \tag{10.56}$$

has to be examined in detail. Using (10.21) we arrive at first at

$$\begin{aligned}D(R)^\dagger\,\sigma^k\,D(R) &= \left[\cos\left(\frac{|\boldsymbol{\varphi}|}{2}\right)+i\frac{\boldsymbol{\sigma\varphi}}{|\boldsymbol{\varphi}|}\sin\left(\frac{|\boldsymbol{\varphi}|}{2}\right)\right]\sigma^k\left[\cos\left(\frac{|\boldsymbol{\varphi}|}{2}\right)-i\frac{\boldsymbol{\sigma\varphi}}{|\boldsymbol{\varphi}|}\sin\left(\frac{|\boldsymbol{\varphi}|}{2}\right)\right]\\ &= \cos^2\left(\frac{|\boldsymbol{\varphi}|}{2}\right)\sigma^k + i\sin\left(\frac{|\boldsymbol{\varphi}|}{2}\right)\cos\left(\frac{|\boldsymbol{\varphi}|}{2}\right)\frac{\varphi_l}{|\boldsymbol{\varphi}|}\left[\sigma^l,\sigma^k\right]_- + \sin^2\left(\frac{|\boldsymbol{\varphi}|}{2}\right)\frac{\varphi_l\varphi_m}{|\boldsymbol{\varphi}|^2}\,\sigma^l\sigma^k\sigma^m\,.\end{aligned} \tag{10.57}$$

In the last term the product of three Pauli matrices appears, which can be simplified by successively applying the calculation rule (10.7) and by taking into account the contraction rule of the three-dimensional Levi-Cività symbol (6.56):

$$\begin{aligned}\sigma^l\sigma^k\sigma^m &= (\delta_{lk}+i\epsilon_{lkn}\sigma^n)\,\sigma^m = \delta_{lk}\sigma^m + i\epsilon_{lkn}\sigma^n\sigma^m\\ &= \delta_{lk}\sigma^m + i\epsilon_{lkn}(\delta_{nm}+i\epsilon_{nmp}\sigma^p) = \delta_{lk}\sigma^m + i\epsilon_{lkm} - (\delta_{ml}\delta_{kp}-\delta_{lp}\delta_{km})\sigma^p\,.\end{aligned} \tag{10.58}$$

With this we end up with the result

$$\sigma^l\sigma^k\sigma^m = i\epsilon_{lkm} + \delta_{lk}\sigma^m + \delta_{km}\sigma^l - \delta_{lm}\sigma^k\,. \tag{10.59}$$

Inserting (10.5) and (10.59) in (10.57) and using trigonometric relations then yields

$$D(R)^\dagger\sigma^k D(R) = \sigma^k\cos|\boldsymbol{\varphi}| + \epsilon_{klm}\frac{\varphi_l}{|\boldsymbol{\varphi}|}\,\sigma^m\,\sin|\boldsymbol{\varphi}| + \frac{\varphi_k\boldsymbol{\sigma\varphi}}{|\boldsymbol{\varphi}|^2}\,(1-\cos|\boldsymbol{\varphi}|)\,. \tag{10.60}$$

This result can be concisely summarized as

$$D(R)^\dagger\sigma^k D(R) = R_{kl}\,\sigma^l\,, \tag{10.61}$$

where R_{kl} coincides with the representation matrix of rotations in three-dimensional space as already determined in (6.69). As the partial derivatives in (10.55) also transform like a vector

$$\partial_k \qquad\longrightarrow\qquad \partial'_k = R_{kl}\,\partial_l \tag{10.62}$$

and the representation matrix R is orthonormal due to (6.72), all combinations (10.55) turn out to be invariant under rotations:

$$D(R)^\dagger\,\sigma^k\,D(R)\,\partial'_k = R_{kl}\,\sigma^l\,R_{km}\,\partial_m = \delta_{lm}\,\sigma^l\,\partial_m = \sigma^k\,\partial_k\,. \tag{10.63}$$

Now the question arises, how the combinations of two Weyl spinors (10.54) can be extended to relativistic invariant combinations. To this end we remember that the Pauli matrices σ^k can be extended to four-vectors in two different ways, namely in the form of the four-vector of Pauli matrices σ^μ in (10.39) and in the form of the spatially inverted four-vector of Pauli matrices $\tilde{\sigma}^\mu$ in (10.43). Therefore we consider now the following eight combinations of two Weyl spinors:

$$\begin{aligned} &\xi^\dagger\sigma^\mu\partial_\mu\xi\,, \quad \eta^\dagger\sigma^\mu\partial_\mu\eta\,, \quad \eta^\dagger\sigma^\mu\partial_\mu\xi\,, \quad \xi^\dagger\sigma^\mu\partial_\mu\eta\,, \\ &\xi^\dagger\tilde{\sigma}^\mu\partial_\mu\xi\,, \quad \eta^\dagger\tilde{\sigma}^\mu\partial_\mu\eta\,, \quad \eta^\dagger\tilde{\sigma}^\mu\partial_\mu\xi\,, \quad \xi^\dagger\tilde{\sigma}^\mu\partial_\mu\eta\,. \end{aligned} \tag{10.64}$$

Here the additional term $\sigma^0\partial_0$ with the time derivative appears, which is trivially invariant under rotations

$$D(R)^\dagger\,\sigma^0\,D(R)\partial_0' = D(R)^\dagger D(R)\,\partial_0' = \partial_0 = \sigma^0\partial_0\,. \tag{10.65}$$

Thus it does not destroy the above discussed rotational invariance of the spatial derivative terms.

With this it remains to investigate, which of the eight combinations (10.64) are invariant under boost transformations. To this end expressions of the form

$$D(B)^\dagger\,\sigma^\mu\,D(B)\,, \qquad D(B)^\dagger\,\tilde{\sigma}^\mu\,D(B) \tag{10.66}$$

appear, where both representations (10.22) can occur in the left and the right factor, respectively. Let us first consider the case $\mu = 0$. In the case that the two representations in the left and right factor of (10.66) are different, then (10.66) is identical to σ^0 due to (10.23), (10.38), and (10.53). As this does not correspond to the transformation behavior, which is characteristic for boosts, we conclude that the 3rd, the 4th, the 7th, and the 8th combination in (10.64) is not invariant under boosts. In the case that both representations in the left and right factor of (10.66) are identical, then we obtain on the one hand for $\mu = 0$ together with (10.23), (10.26), and (10.38):

$$D(B)^\dagger\sigma^0 D(B) = D(B)^2 = \cosh|\boldsymbol{\xi}| \mp \frac{\boldsymbol{\sigma\xi}}{|\boldsymbol{\xi}|}\sinh|\boldsymbol{\xi}|\,. \tag{10.67}$$

On the other hand we get for $\mu = k$ due to (10.23) and (10.26)

$$\begin{aligned} D(B)^\dagger\sigma^k D(B) &= \left[\cosh\left(\frac{|\boldsymbol{\xi}|}{2}\right) \mp \frac{\boldsymbol{\sigma\xi}}{|\boldsymbol{\xi}|}\sinh\left(\frac{|\boldsymbol{\xi}|}{2}\right)\right]\sigma^k\left[\cosh\left(\frac{|\boldsymbol{\xi}|}{2}\right) \mp \frac{\boldsymbol{\sigma\xi}}{|\boldsymbol{\xi}|}\sinh\left(\frac{|\boldsymbol{\xi}|}{2}\right)\right] \\ = \cosh^2\left(\frac{|\boldsymbol{\xi}|}{2}\right)\sigma^k &\mp \sinh\left(\frac{|\boldsymbol{\xi}|}{2}\right)\cosh\left(\frac{|\boldsymbol{\xi}|}{2}\right)\frac{\xi_l}{|\boldsymbol{\xi}|}\left[\sigma^l,\sigma^k\right]_+ + \sinh^2\left(\frac{|\boldsymbol{\xi}|}{2}\right)\frac{\xi_l\xi_m}{|\boldsymbol{\xi}|^2}\sigma^l\sigma^k\sigma^m\,. \end{aligned} \tag{10.68}$$

Inserting (10.2) and (10.59) in (10.68) and using hyperbolic relations then yields

$$D(B)^\dagger\left(\mp\sigma^k\right)D(B) = \mp\sigma^k + \frac{\xi_k}{|\boldsymbol{\xi}|}\sinh|\boldsymbol{\xi}| + \frac{\xi_k}{|\boldsymbol{\xi}|}\frac{(\mp\boldsymbol{\sigma})\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\left(\cosh|\boldsymbol{\xi}| - 1\right)\,. \tag{10.69}$$

The two results (10.67) and (10.69) can be concisely summarized by

$$D^{(1/2,0)}(B)^\dagger\,\tilde{\sigma}^\mu\,D^{(1/2,0)}(B) = B^\mu{}_\nu\,\tilde{\sigma}^\nu\,, \tag{10.70}$$

$$D^{(0,1/2)}(B)^\dagger\,\sigma^\mu\,D^{(0,1/2)}(B) = B^\mu{}_\nu\,\sigma^\nu\,, \tag{10.71}$$

where $B^\mu{}_\nu$ coincides with the representation matrix of boost in the four-dimensional space-time as already determined in (6.79). As the partial derivatives in (10.64) also transform like a covariant four-vector

$$\partial_\mu \qquad \longrightarrow \qquad \partial'_\mu = B_\mu{}^\nu\, \partial_\nu\,, \tag{10.72}$$

and the representation matrix B fulfills the property (6.90), we can prove due to (10.70) and (10.71) that the 2nd and the 5th term in (10.64) is invariant:

$$\xi^\dagger \tilde{\sigma}^\mu \partial_\mu \xi \rightarrow \xi^\dagger D^{(1/2,0)}(B)^\dagger \tilde{\sigma}^\mu D^{(1/2,0)}(B) \partial'_\mu \xi = \xi^\dagger B^\mu{}_\nu \tilde{\sigma}^\nu B_\mu{}^\kappa \partial_\kappa \xi = \xi^\dagger \delta_\nu{}^\kappa \tilde{\sigma}^\nu \partial_\kappa \xi = \xi^\dagger \tilde{\sigma}^\nu \partial_\nu \xi, \tag{10.73}$$

$$\eta^\dagger \sigma^\mu \partial_\mu \eta \rightarrow \eta^\dagger D^{(0,1/2)}(B)^\dagger \sigma^\mu D^{(0,1/2)}(B) \partial'_\mu \eta = \eta^\dagger B^\mu{}_\nu \sigma^\nu B_\mu{}^\kappa \partial_\kappa \eta = \eta^\dagger \delta_\nu{}^\kappa \sigma^\nu \partial_\kappa \eta = \eta^\dagger \sigma^\nu \partial_\nu \eta. \tag{10.74}$$

For the two remaining combinations $\eta^\dagger \tilde{\sigma}^\mu \partial_\mu \eta$ and $\xi^\dagger \sigma^\mu \partial_\mu \xi$, i.e. the 1st and the 6th term in (10.64), a boost invariance can not be proved, because both $\tilde{\sigma}^\mu$ and σ^μ transform due to (10.70) and (10.71)) as a four-vector under the representations $D^{(1/2,0)}(B)$ and $D^{(0,1/2)}(B)$, respectively.

10.6 Dirac Action

From the considerations of the previous section follows the most general Lorentz-invariant action for describing a massive spin 1/2 particle

$$\mathcal{A} = \mathcal{A}[\xi(\bullet), \xi^\dagger(\bullet); \eta(\bullet), \eta^\dagger(\bullet)]\,, \tag{10.75}$$

which contains only quadratic terms in the Weyl spinors and their first partial derivatives:

$$\mathcal{A} = \frac{1}{c} \int d^4x\, \mathcal{L}\big(\xi(x), \partial_\mu \xi(x); \xi^\dagger(x), \partial_\mu \xi^\dagger(x); \eta(x), \partial_\mu \eta(x); \eta^\dagger(x)\,, \partial_\mu \eta^\dagger(x)\big)\,. \tag{10.76}$$

Here the Lagrange density

$$\mathcal{L} = A i \xi^\dagger \tilde{\sigma}^\mu \partial_\mu \xi + B i \eta^\dagger \sigma^\mu \partial_\mu \eta + C \xi^\dagger \eta + D \eta^\dagger \xi\,, \tag{10.77}$$

contains constants A, B, C, D, which are not yet defined. Below in Section 10.8 we show that the additional demand for an invariance of the Lagrange density under parity transformation leads to the fact, that both Weyl spinors ξ and η have to appear on equal footing. This reduces (10.77) to

$$\mathcal{L} = A \left(i \xi^\dagger \tilde{\sigma}^\mu \partial_\mu \xi + i \eta^\dagger \sigma^\mu \partial_\mu \eta - m \xi^\dagger \eta - m \eta^\dagger \xi \right)\,. \tag{10.78}$$

The still undetermined parameters A, m define the physical dimension of the action and are only fixed at a later stage by considering the non-relativistic limit. Due to the non-zero rest mass M of the particle, the action (10.78) necessarily contains both Weyl spinors ξ and η. Only in the case that the rest mass of the particle vanishes, a Lorentz-invariant action can be formed with just one of the two Weyl spinors, as is discussed below in Section 10.9.

Due to the action (10.78) the Weyl spinors ξ and η satisfy the equations of motion

$$\frac{\delta \mathcal{A}}{\delta \xi^\dagger(x)} = \frac{\partial \mathcal{L}}{\partial \xi^\dagger(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \xi^\dagger(x)\right)} = A\Big\{ i\tilde{\sigma}^\mu \partial_\mu \xi(x) - m\eta(x) \Big\} = 0\,, \qquad (10.79)$$

$$\frac{\delta \mathcal{A}}{\delta \eta^\dagger(x)} = \frac{\partial \mathcal{L}}{\partial \eta^\dagger(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \eta^\dagger(x)\right)} = A\Big\{ i\sigma^\mu \partial_\mu \eta(x) - m\xi(x) \Big\} = 0\,. \qquad (10.80)$$

In order to combine these two equations of motion one needs the calculation rules

$$\sigma^\mu \tilde{\sigma}^\nu + \sigma^\nu \tilde{\sigma}^\mu = 2g^{\mu\nu} I, \qquad (10.81)$$

$$\tilde{\sigma}^\mu \sigma^\nu + \tilde{\sigma}^\nu \sigma^\mu = 2g^{\mu\nu} I\,, \qquad (10.82)$$

which can be explicitly shown by specializing μ, ν to spatial and temporal indices. To this end one has to take into account the Clifford algebra property (10.2), the definitions (10.38), (10.39), and (10.43), as well as the components of the Minkowski metric in (6.3):

$$\sigma^0 \tilde{\sigma}^0 + \sigma^0 \tilde{\sigma}^0 = 2\sigma^0 = 2I = 2g^{00} I\,, \qquad (10.83)$$

$$\sigma^0 \tilde{\sigma}^k + \sigma^k \tilde{\sigma}^0 = -\sigma^0 \sigma^k + \sigma^k \sigma^0 = 0 = 2g^{0k} I\,, \qquad (10.84)$$

$$\sigma^k \tilde{\sigma}^l + \sigma^l \tilde{\sigma}^k = -\sigma^k \sigma^l - \sigma^l \sigma^k = -2\delta_{kl} I = 2g^{kl} I\,. \qquad (10.85)$$

Multiplying (10.79) with $i\sigma^\nu \partial_\nu$ and using (10.80) or, vice versa, multiplying (10.80) with $i\tilde{\sigma}^\nu \partial_\nu$ and using (10.79), we obtain due to (10.81) and (10.82)

$$-\sigma^\nu \tilde{\sigma}^\nu \partial_\nu \partial_\mu \xi(x) - mi\sigma^\nu \partial_\nu \eta(x) = -g^{\mu\nu} \partial_\mu \partial_\nu \xi(x) - m^2 \xi(x) = 0\,, \qquad (10.86)$$

$$-\tilde{\sigma}^\nu \sigma^\nu \partial_\nu \partial_\mu \eta(x) - mi\tilde{\sigma}^\nu \partial_\nu \xi(x) = -g^{\mu\nu} \partial_\mu \partial_\nu \eta(x) - m^2 \eta(x) = 0\,. \qquad (10.87)$$

Thus, both Weyl spinors ξ and η satisfy the Klein-Gordon equation of a particle (8.19), provided that the parameter m is identified according to

$$m = \frac{Mc}{\hbar}\,, \qquad (10.88)$$

i.e. being inversely proportional to the Compton wave length (8.21).

Since the description of a massive spin 1/2 particle necessarily involves both Weyl spinors ξ and η, it is suggestive to combine them to a Dirac spinor:

$$\psi(x) = \begin{pmatrix} \xi(x) \\ \eta(x) \end{pmatrix}. \qquad (10.89)$$

In view of that we rewrite the Lagrange density (10.78)

$$\mathcal{L} = A\left\{ (\xi^\dagger, \eta^\dagger) \begin{pmatrix} \tilde{\sigma}^\mu & O \\ O & \sigma^\mu \end{pmatrix} i\partial_\mu \begin{pmatrix} \xi \\ \eta \end{pmatrix} - (\xi^\dagger, \eta^\dagger) \begin{pmatrix} O & mI \\ mI & O \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\}, \qquad (10.90)$$

where we used the 2×2 unit matrix (10.3) and introduced in addition the 2×2 zero matrix

$$O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \qquad (10.91)$$

Furthermore, we define the Dirac adjoint of the Dirac spinor (10.89) according to

$$\overline{\psi}(x) = \left(\eta^\dagger(x), \xi^\dagger(x)\right) = \psi^\dagger(x) \begin{pmatrix} O & I \\ I & O \end{pmatrix} \leftrightarrow \psi^\dagger(x) = (\xi^\dagger(x), \eta^\dagger(x)) = \overline{\psi}(x) \begin{pmatrix} O & I \\ I & O \end{pmatrix} . \tag{10.92}$$

With this the Lagrange density (10.90) changes into

$$\mathcal{L} = A\left\{ \overline{\psi} \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} \tilde{\sigma}^\mu & O \\ O & \sigma^\mu \end{pmatrix} i\partial_\mu \psi - \overline{\psi} \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} O & mI \\ mI & O \end{pmatrix} \psi \right\} , \tag{10.93}$$

which finally reduces to

$$\mathcal{L} = A\overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi . \tag{10.94}$$

Here we have introduced the Dirac matrices

$$\gamma^\mu = \begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix} , \tag{10.95}$$

which turn out to obey the property of a Clifford algebra, see Eq. (10.4), due to the calculational rules (10.81) and (10.82):

$$\begin{aligned} [\gamma^\mu, \gamma^\nu]_+ &= \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = \begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix} \begin{pmatrix} O & \sigma^\nu \\ \tilde{\sigma}^\nu & O \end{pmatrix} + \begin{pmatrix} O & \sigma^\nu \\ \tilde{\sigma}^\nu & O \end{pmatrix} \begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix} \\ &= \begin{pmatrix} \sigma^\mu\tilde{\sigma}^\nu + \sigma^\nu\tilde{\sigma}^\mu & O \\ O & \tilde{\sigma}^\nu\sigma^\mu + \tilde{\sigma}^\nu\sigma^\mu \end{pmatrix} = 2g^{\mu\nu} \begin{pmatrix} I & O \\ O & I \end{pmatrix} . \end{aligned} \tag{10.96}$$

The action (10.75), (10.76) can, thus, be interpreted as a functional of the Dirac spinor $\psi(x)$ and the Dirac adjoint Dirac spinor $\overline{\psi}(x)$:

$$\mathcal{A}[\psi(\bullet); \overline{\psi}(\bullet)] = \frac{1}{c} \int d^4x \, \mathcal{L}\left(\psi(x), \partial_\mu\psi(x); \overline{\psi}(x); \partial_\mu\overline{\psi}(x)\right) . \tag{10.97}$$

The equation of motion of the Dirac spinor is thus given by

$$\frac{\delta \mathcal{A}}{\delta \overline{\psi}(x)} = \frac{\partial \mathcal{L}}{\partial \overline{\psi}(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \overline{\psi}(x))} = A\Big\{ i\gamma^\mu \partial_\mu \psi(x) - m\psi(x) \Big\} = 0 . \tag{10.98}$$

This reduces to

$$\big(i\partial\!\!\!/ - m\big)\,\psi(x) = 0 \tag{10.99}$$

with introducing the Feynman dagger as another widespread shortcut notation

$$\partial\!\!\!/ = \gamma^\mu \partial_\mu . \tag{10.100}$$

10.7 Spinor Representation of Lorentz Group

By construction the Dirac action (10.94), (10.97) is invariant under Lorentz transformations. Nevertheless we now aim for proving this again from a different point of view by studying the representation of the Lorentz group in the space of the Dirac spinors. To this end we deduce from the representations of the Lorentz group in the space of the Weyl spinors in (10.49) and (10.50)

$$\psi(x) = \begin{pmatrix} \xi(x) \\ \eta(x) \end{pmatrix} \qquad \longrightarrow \qquad \psi'(x') = \begin{pmatrix} \xi'(x') \\ \eta'(x') \end{pmatrix} = D(\Lambda)\psi(x)\,. \tag{10.101}$$

Here the representation matrices $D(\Lambda)$ of the Lorentz group for the Dirac spinor are composed of the respective representation matrices $D^{(1/2,0)}(\Lambda)$ and $D^{(0,1/2)}(\Lambda)$ for the Weyl spinors:

$$D(\Lambda) = \begin{pmatrix} D^{(1/2,0)}(\Lambda) & O \\ O & D^{(0,1/2)}(\Lambda) \end{pmatrix}. \tag{10.102}$$

Furthermore, we note that the relation (10.92) between the Dirac adjoint Dirac spinor $\overline{\psi}$ and the adjoint Dirac spinor $\psi^\dagger$ simplifies due to (10.38) and (10.95):

$$\overline{\psi}(x) = \psi^\dagger(x)\gamma^0 \qquad \Longleftrightarrow \qquad \psi^\dagger(x) = \overline{\psi}(x)\gamma^0\,. \tag{10.103}$$

Due to (10.101) and (10.103) the Lorentz transformation of the Dirac adjoint Dirac spinor reads

$$\overline{\psi}'(x') = \psi'^\dagger(x')\gamma^0 = \psi^\dagger(x)D(\Lambda)^\dagger\gamma^0 = \overline{\psi}(x)\overline{D}(\Lambda)\,. \tag{10.104}$$

Here we have introduced the Dirac adjoint representation matrices of the Lorentz group

$$\overline{D}(\Lambda) = \gamma^0 D(\Lambda)^\dagger \gamma^0\,, \tag{10.105}$$

for which we obtain due to (10.38), (10.95), and (10.102) the explicit result

$$\overline{D}(\Lambda) = \begin{pmatrix} D^{(0,1/2)}(\Lambda)^\dagger & O \\ O & D^{(1/2,0)}(\Lambda)^\dagger \end{pmatrix}. \tag{10.106}$$

Thus, taking into account (10.13), (10.14), (10.16), (10.102), and (10.106) we conclude

$$\overline{D}(\Lambda) = D(\Lambda)^{-1}\,. \tag{10.107}$$

Furthermore, we note that we showed in Section 10.5

$$D^{(1/2,0)}(\Lambda)^\dagger\,\tilde{\sigma}^\mu\,D^{(1/2,0)}(\Lambda) = \Lambda^\mu{}_\nu\tilde{\sigma}^\mu\,, \tag{10.108}$$

$$D^{(0,1/2)}(\Lambda)^\dagger\,\sigma^\mu\,D^{(0,1/2)}(\Lambda) = \Lambda^\mu{}_\nu\sigma^\mu \tag{10.109}$$

for $\Lambda = R$ and $\Lambda = B$ in (10.63), (10.65) and (10.70), (10.71), respectively. But since every Lorentz transformation can be understood as a successive execution of a boost and a rotation

$$\Lambda = BR\,, \tag{10.110}$$

the corresponding representation matrices factorize, i.e. we have

$$D^{(1/2,0)}(\Lambda) = D^{(1/2,0)}(B)D^{(1/2,0)}(R)\,, \qquad D^{(0,1/2)}(\Lambda) = D^{(0,1/2)}(B)D^{(0,1/2)}(R)\,. \quad (10.111)$$

With this we can show that (10.108) and (10.109) are even valid for any Lorentz transformation. At first we obtain for the representation $D^{(1/2,0)}$

$$\begin{aligned} D^{(1/2,0)}(\Lambda)^\dagger\, \tilde{\sigma}^\mu\, D^{(1/2,0)}(\Lambda) &= D^{(1/2,0)}(R)^\dagger\, D^{(1/2,0)}(B)^\dagger\, \tilde{\sigma}^\mu\, D^{(1/2,0)}(B)\, D^{(1/2,0)}(R) \\ &= B^\mu{}_\nu\, D^{(1/2,0)}(R)^\dagger\, \tilde{\sigma}^\nu\, D^{(1/2,0)}(R) = B^\mu{}_\nu\, R^\nu{}_\kappa\, \tilde{\sigma}^\kappa = \Lambda^\mu{}_\nu\, \tilde{\sigma}^\nu\,, \end{aligned} \quad (10.112)$$

and, correspondingly, we get for the representation $D^{(0,1/2)}$

$$\begin{aligned} D^{(0,1/2)}(\Lambda)^\dagger\, \sigma^\mu\, D^{(0,1/2)}(\Lambda) &= D^{(0,1/2)}(R)^\dagger\, D^{(0,1/2)}(B)^\dagger\, \sigma^\mu\, D^{(0,1/2)}(B)\, D^{(0,1/2)}(R) \\ &= B^\mu{}_\nu\, D^{(0,1/2)}(R)^\dagger\, \sigma^\nu\, D^{(0,1/2)}(R) = B^\mu{}_\nu\, R^\nu{}_\kappa\, \sigma^\kappa = \Lambda^\mu{}_\nu\, \sigma^\nu\,. \end{aligned} \quad (10.113)$$

Note that we have used (10.110) in the last step of both (10.112) and (10.113). The two transformation laws (10.108) and (10.109) can now be combined into one for the Dirac matrices (10.95). Taking into account (10.102) and (10.106) a direct multiplication of the involved 4×4 matrices yields

$$\begin{aligned} \overline{D}(\Lambda)\gamma^\mu D(\Lambda) &= \begin{pmatrix} D^{(0,1/2)}(\Lambda)^\dagger & O \\ O & D^{(1/2,0)}(\Lambda)^\dagger \end{pmatrix} \begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix} \begin{pmatrix} D^{(1/2,0)}(\Lambda) & O \\ O & D^{(0,1/2)}(\Lambda) \end{pmatrix} \\ &= \begin{pmatrix} O & D^{(0,1/2)}(\Lambda)^\dagger \sigma^\mu D^{(0,1/2)}(\Lambda) \\ D^{(1/2,0)}(\Lambda)^\dagger \tilde{\sigma}^\mu D^{(1/2,0)}(\Lambda) & O \end{pmatrix} = \Lambda^\mu{}_\nu \begin{pmatrix} O & \sigma^\nu \\ \tilde{\sigma}^\nu & O \end{pmatrix} = \Lambda^\mu{}_\nu \gamma^\nu \end{aligned} \quad (10.114)$$

After these preparations the invariance of the Dirac action can be shown as follows. At first we obtain for the Lorentz transformation of the action (10.94), (10.97)

$$\mathcal{A}' = \frac{A}{c} \int d^4x'\, \overline{\psi}'(x') \left(i\gamma^\mu \partial'_\mu - m\right) \psi'(x')\,, \quad (10.115)$$

which reads due to (10.101), (10.104), and the property $d^4x' = d^4x$ of special Lorentz transformations:

$$\mathcal{A}' = \frac{A}{c} \int d^4x\, \overline{\psi}(x) \Big[i\overline{D}(\Lambda)\gamma^\mu D(\Lambda)\partial'_\mu - m\overline{D}(\Lambda)D(\Lambda)\Big] \psi(x)\,. \quad (10.116)$$

Using (10.107) and (10.114) as well as taking into account that the partial derivatives in (10.116) transform like a covariant four vector

$$\partial_\mu \qquad \longrightarrow \qquad \partial'_\mu = \Lambda_\mu{}^\nu\, \partial_\nu\,, \quad (10.117)$$

we get

$$\mathcal{A}' = \frac{A}{c} \int d^4x\; \overline{\psi}(x) \left(i\Lambda^\mu{}_\nu \Lambda_\mu{}^\kappa \gamma^\nu \partial_\kappa - m\right) \psi(x)\,. \quad (10.118)$$

From (6.28) we then conclude that the Lorentz transformed action (10.118) coincides with the original action (10.94), (10.97).

Let us further investigate the representation (10.102) of the Lorentz group in the space of the Dirac spinors. To this end we use (10.13) as well as (10.14) and bring it to the following form:

$$D(\Lambda) = \exp\left[-i\begin{pmatrix}\sigma^k/2 & O \\ O & \sigma^k/2\end{pmatrix}\varphi_k - i\begin{pmatrix}-i\sigma^k/2 & O \\ O & i\sigma^k/2\end{pmatrix}\xi_k\right] . \tag{10.119}$$

Comparing this with a covariant formulation of the Lie theorem as in (6.61)–(6.64)

$$D(\Lambda) = \exp\left(-\frac{i}{2}\,\omega_{\mu\nu}S^{\mu\nu}\right) = \exp\left(-\frac{i}{2}\,\epsilon_{kij}S^{ij}\varphi_k - iS^{0k}\xi_k\right) , \tag{10.120}$$

the representation matrices for the generators of the boosts are given by

$$D(M_k) = S^{0k} = \begin{pmatrix}-i\sigma^k/2 & O \\ O & i\sigma^k/2\end{pmatrix} , \tag{10.121}$$

while the representation matrices for the generators of the rotations follow from

$$D(L_k) = S^k = \frac{1}{2}\epsilon_{kij}S^{ij} = \begin{pmatrix}\sigma^k/2 & O \\ O & \sigma^k/2\end{pmatrix} \tag{10.122}$$

and read

$$S^{ij} = \epsilon_{ijk}\begin{pmatrix}\sigma^k/2 & O \\ O & \sigma^k/2\end{pmatrix} . \tag{10.123}$$

According to (6.164) we read off that (10.122) just represents the spin vector for spin 1/2 particles. Furthermore, the two results (10.121) and (10.123) can be summarized in a covariant form with the help of the Dirac matrices (10.95) as follows:

$$S^{\mu\nu} = \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right]_- . \tag{10.124}$$

Indeed, whereas Eq. (10.121) follows directly from (10.124), the corresponding derivation of (10.123) needs to take into account the Lie algebra property of the Pauli matrices (10.5).

Now we aim for determining the commutator between two representation matrices $S^{\mu\nu}$ of the Lorentz algebra in the space of the Dirac spinors. To this end we apply the calculation rule (3.94), the definition (10.123) as well as the Clifford algebra property of the Dirac matrices in (10.96) and calculate at first the commutator

$$\left[S^{\mu\nu}, \gamma^\lambda\right]_- = i\left(g^{\nu\lambda}\gamma^\mu - g^{\mu\lambda}\gamma^\nu\right) . \tag{10.125}$$

Then we use (3.10) and (10.123)–(10.125) for obtaining

$$\left[S^{\mu\nu}, S^{\kappa\lambda}\right]_- = i\left(g^{\mu\lambda}S^{\nu\kappa} + g^{\nu\kappa}S^{\mu\lambda} - g^{\mu\kappa}S^{\nu\lambda} - g^{\nu\lambda}S^{\mu\kappa}\right) . \tag{10.126}$$

Thus, we read off from (10.126) that the representation matrices $S^{\mu\nu}$ satisfy, indeed, the usual commutation relations of the Lorentz algebra, see Eqs. (6.48) and (6.49). Furthermore, (10.125) and (10.126) show that γ^λ and $S^{\kappa\lambda}$ represent a tensor operator of rank $n = 1$ and $n = 2$ in the sense of (6.105).

10.8 Parity Transformation

Due to a parity transformation P the four-vector x is mapped to the spatially inverted four-vector $\tilde{x}$ introduced in (10.42):

$$x'_P = Px = \tilde{x}\,. \tag{10.127}$$

Performing a parity transformation P two times in a row, the original four-vector is reproduced. Thus, the parity transformation P is involutoric:

$$P^2 = 1 \qquad \Longleftrightarrow \qquad P^{-1} = P\,. \tag{10.128}$$

The representation matrix for such a parity transformation reads as follows:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{10.129}$$

Furthermore, it can be straight-forwardly shown that the representation matrix of the parity transformation (10.129) commutates with the matrix representations for the generators of rotations (6.53)

$$P^{-1}L_kP = L_k \qquad \Longleftrightarrow \qquad [P, L_k]_- = 0 \tag{10.130}$$

and anti-commutates with the matrix representations for the generators of boosts (6.54)

$$P^{-1}M_kP = -M_k \qquad \Longleftrightarrow \qquad [P, M_k]_+ = 0\,. \tag{10.131}$$

For instance, we have

$$P^{-1}L_1P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = L_1\,, \tag{10.132}$$

$$P^{-1}M_1P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = -M_1. \tag{10.133}$$

Performing a parity transformation upon a Dirac spinor yields

$$\psi(x) \longrightarrow \psi'_P(x) = D(P)\psi(\tilde{x})\,, \tag{10.134}$$

where $D(P)$ denotes the corresponding representation matrix of the parity transformation in the space of Dirac spinors. Thus, $D(P)$ must possess the same properties as P. For instance, due to (10.128), $D(P)$ must be involutoric:

$$D(P)^2 = 1\,. \tag{10.135}$$

Furthermore, $D(P)$ must satisfy both a commutator and an anti-commutator relation with the representation matrices $D(L_k)$ and $D(M_k)$ of the rotations and boosts in the space of Dirac spinors analogous to (10.130) and (10.131), respectively:

$$D(P)^{-1}D(L_k)D(P) = D(L_k)\,, \tag{10.136}$$

$$D(P)^{-1}D(M_k)D(P) = -D(M_k)\,. \tag{10.137}$$

We now determine the representation matrix $D(P)$ from the requirement that the Dirac equation is invariant under a parity transformation. To this end we rewrite at first the Dirac equation (10.98) by the applying the substitution $x \to \tilde{x}$:

$$\left(i\gamma^\mu\tilde{\partial}_\mu - m\right)\psi(\tilde{x}) = 0\,. \tag{10.138}$$

Then we replace $\psi(\tilde{x})$ in (10.138) with $\psi'_P(x)$ according to (10.134) and use the property of the scalar product that $\gamma^\mu\tilde{\partial}_\mu = \tilde{\gamma}^\mu\partial_\mu$ holds, yielding

$$\left[iD(P)\tilde{\gamma}^\mu D(P)^{-1}\partial_\mu - m\right]\psi'_P(x) = 0\,. \tag{10.139}$$

Thus, Eq. (10.139) reduces to the Dirac equation for the parity transformed mirrored Dirac spinor $\psi'_P(x)$, i.e.

$$\left(i\gamma^\mu\partial_\mu - m\right)\psi'_P(x) = 0\,, \tag{10.140}$$

provided that the representation matrix $D(P)$ satisfies the condition

$$D(P)\tilde{\gamma}^\mu D(P)^{-1} = \gamma^\mu\,. \tag{10.141}$$

Let us define the representation matrix $D(P)$ according to

$$D(P) = \gamma^0\,. \tag{10.142}$$

Then the involution property (10.135) is valid

$$D(P)^2 = (\gamma^0)^2 = \begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} O & I \\ I & O \end{pmatrix} = \begin{pmatrix} I & O \\ O & I \end{pmatrix} \tag{10.143}$$

and the condition (10.141) is fulfilled due to the Clifford algebra (10.96):

$$\gamma^0\tilde{\gamma}^0\gamma^0 = (\gamma^0)^3 = \gamma^0\,, \tag{10.144}$$

$$\gamma^0\tilde{\gamma}^k\gamma^0 = -\gamma^0\gamma^k\gamma^0 = \gamma^k\,. \tag{10.145}$$

Furthermore, taking into account (10.95), (10.121), (10.122) as well as (10.142) both the commutators (10.136) and the anti-commutators (10.137) can straight-forwardly be shown:

$$D(P)^{-1}D(L_k)D(P) = \begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} \sigma^k/2 & O \\ O & \sigma^k/2 \end{pmatrix}\begin{pmatrix} O & I \\ I & O \end{pmatrix} = D(L_k)\,, \tag{10.146}$$

$$D(P)^{-1}D(M_k)D(P) = \begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} -i\sigma^k/2 & O \\ O & i\sigma^k/2 \end{pmatrix}\begin{pmatrix} O & I \\ I & O \end{pmatrix} = -D(M_k)\,. \tag{10.147}$$

Additionally, we read off from the definition of γ^0 in (10.95) that a parity transformation (10.134) has the effect of interchanging the Weyl spinors ξ and η in the Dirac spinor (10.89):

$$\psi(x) = \begin{pmatrix} \xi(x) \\ \eta(x) \end{pmatrix} \longrightarrow \psi'_P(x) = \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} \xi(\tilde{x}) \\ \eta(\tilde{x}) \end{pmatrix} = \begin{pmatrix} \eta(\tilde{x}) \\ \xi(\tilde{x}) \end{pmatrix} . \tag{10.148}$$

Thus, in a theory, where both $\psi(x)$ and $\psi'_p(x)$ represent physically realized states, one needs both Weyl spinors ξ and η. And from the Lorentz invariance considerations in Section 10.5 follows then that the corresponding action must necessarily have a mass term. Furthermore, we conclude from (10.148) that in a parity transformation invariant theory both Weyl spinors ξ and η have to appear on equal footing. This result was already applied in Section 10.6 in order to simplify the Lagrange density (10.77) according to (10.78).

10.9 Neutrinos

A neutrino is an elementary particle with spin 1/2, which interacts only via the weak force and gravity. Historically, the neutrino was postulated first by Wolfgang Pauli in 1930 as an additional particle being involved in the beta decay of a neutron into a proton and an electron in order explain the conservation of energy, momentum, and angular momentum. The neutrino is so named because it is electrically neutral and its rest mass is so small that it was long thought to be zero, leading to the suffix -ino. Therefore, in accordance with previous experimental results, neutrinos were considered for decades to be massless spin 1/2-particles, which are described by a single Weyl spinor ξ or η. According to (10.77), their Lagrangian density is then given by either

$$\mathcal{L} = Ai\xi^\dagger \tilde{\sigma}^\mu \partial_\mu \xi \tag{10.149}$$

or by

$$\mathcal{L} = Ai\eta^\dagger \sigma^\mu \partial_\mu \eta \,. \tag{10.150}$$

Like in the Maxwell theory also the Lagrangians (10.149) and (10.150) of the Weyl theory do not contain a Planck constant but still represent a valid first-quantized theory due to the vanishing rest mass. In both cases, the Lagrangian density is invariant under Lorentz transformations according to Section 10.5 but not invariant under parity transformations due to Section 10.8. In order to describe neutrinos also with a Dirac spinor ψ, one must project out the upper or the lower Weyl spinor ξ or η. To this end one introduces the matrix

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{10.151}$$

for which we obtain due to the definition of the Dirac matrices in (10.95)

$$\gamma^5 = i \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} O & \sigma^1 \\ -\sigma^1 & I \end{pmatrix} \begin{pmatrix} O & \sigma^2 \\ -\sigma^2 & O \end{pmatrix} \begin{pmatrix} O & \sigma^3 \\ -\sigma^3 & O \end{pmatrix} = i \begin{pmatrix} \sigma^1\sigma^2\sigma^3 & O \\ O & -\sigma^1\sigma^2\sigma^3 \end{pmatrix} . \tag{10.152}$$

Here the product of the Pauli matrices (10.1) turns out to be

$$\sigma^1\sigma^2\sigma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \tag{10.153}$$

so that (10.152) reduces to

$$\gamma^5 = \begin{pmatrix} -I & O \\ O & I \end{pmatrix}. \tag{10.154}$$

Thus, we read off that also γ^5 is involutoric:

$$\left(\gamma^5\right)^2 = 1 \qquad \Longrightarrow \qquad \left(\gamma^5\right)^{-1} = \gamma^5 . \tag{10.155}$$

Furthermore, with the help of the γ^5 matrix we can construct projection matrices

$$P_u = \frac{1}{2}\left(1 - \gamma^5\right) \quad = \quad \begin{pmatrix} I & O \\ O & O \end{pmatrix}, \tag{10.156}$$

$$P_l = \frac{1}{2}\left(1 + \gamma^5\right) \quad = \quad \begin{pmatrix} O & O \\ O & I \end{pmatrix}, \tag{10.157}$$

which possess the desired effect:

$$P_u\psi \quad = \quad \begin{pmatrix} I & O \\ O & O \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi \\ 0 \end{pmatrix}, \tag{10.158}$$

$$P_l\psi \quad = \quad \begin{pmatrix} O & O \\ O & I \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ \eta \end{pmatrix}. \tag{10.159}$$

Thus, we read off that the Weyl spinors ξ and η represent in form of $(1 \mp \gamma^5)\psi/2$ eigenstates of the matrix γ^5 with the eigenvalues ∓ 1:

$$\gamma^5 \, \frac{1}{2}\left(1 \mp \gamma^5\right)\psi = \mp \frac{1}{2}\left(1 \mp \gamma^5\right)\psi . \tag{10.160}$$

As the neutrino states can be classified according to the eigenvalues of the matrix γ^5, it is of special importance. One calls γ^5 the chirality operator and speaks of left (-1) or right $(+1)$ chirality for the states $(1 \mp \gamma^5)\psi/2$.

We note that the chirality operator γ^5 from (10.151) can also be written as

$$\gamma^5 = \frac{i}{24}\,\epsilon_{\mu\nu\kappa\lambda}\,\gamma^\mu\gamma^\nu\gamma^\kappa\gamma^\lambda . \tag{10.161}$$

Indeed, due to the anti-symmetry (6.144) of the ϵ-tensor only $4! = 24$ non-vanishing terms contribute to (10.161), where each term consists of a product of 4 different Dirac matrices. Furthermore, all 24 terms agree due to the anti-symmetry $\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu$ for $\mu \neq \nu$ following from the Clifford algebra (10.96) and due to the anti-symmetry (6.144) of the ϵ-tensor, yielding

(10.151). Since the Dirac matrices γ^μ transform according to (10.114) like a contravariant four-vector under Lorentz transformations, Eq. (10.161) has due to (10.107) the consequence that the chirality operator γ^5 is Lorentz invariant:

$$\overline{D}(\Lambda)\gamma^5 D(\Lambda) = \frac{i}{24}\,\epsilon_{\mu\nu\kappa\lambda}\Big[\overline{D}(\Lambda)\gamma^\mu D(\Lambda)\Big]\Big[\overline{D}(\Lambda)\gamma^\nu D(\Lambda)\Big]\Big[\overline{D}(\Lambda)\gamma^\kappa D(\Lambda)\Big]\Big[\overline{D}(\Lambda)\gamma^\lambda D(\Lambda)\Big]$$
$$= \frac{i}{24}\,\epsilon_{\mu\nu\kappa\lambda}\Lambda^\mu{}_{\mu'}\Lambda^\nu{}_{\nu'}\Lambda^\kappa{}_{\kappa'}\Lambda^\lambda{}_{\lambda'}\gamma^{\mu'}\gamma^{\nu'}\gamma^{\kappa'}\gamma^{\lambda'} = \frac{i}{24}\,\epsilon_{\mu'\nu'\kappa'\lambda'}\gamma^{\mu'}\gamma^{\nu'}\gamma^{\kappa'}\gamma^{\lambda'} = \gamma^5\,. \tag{10.162}$$

Here we used the Weierstraß expansion of the determinant a 4×4-matrix $\Lambda = (\Lambda^\mu{}_\nu)$

$$(\mathrm{Det}\ \Lambda)\epsilon_{\mu'\nu'\kappa'\lambda'} = \epsilon_{\mu\nu\kappa\lambda}\Lambda^\mu{}_{\mu'}\Lambda^\nu{}_{\nu'}\Lambda^\kappa{}_{\kappa'}\Lambda^\lambda{}_{\lambda'}\,, \tag{10.163}$$

where the property Det $\Lambda = 1$ of the special Lorentz transformations implies that the four-dimensional Levi-Cività tensor has the same components in all inertial systems:

$$\epsilon'_{\mu\nu\kappa\lambda} = \epsilon_{\mu\nu\kappa\lambda}\,. \tag{10.164}$$

With the help of (10.156)–(10.159) the two neutrino Lagrangians (10.149) and (10.150) can be expressed by Dirac spinors:

$$\mathcal{L} = Ai\overline{\psi}(x)\gamma^\mu\partial_\mu\,\frac{1}{2}\,(1\mp\gamma^5)\psi(x)\,. \tag{10.165}$$

In fact, taking into account (10.89), (10.93), and (10.95) an explicit calculation yields for the upper Weyl spinor

$$\begin{aligned}\mathcal{L} &= Ai\overline{\psi}(x)\gamma^\mu\partial_\mu\,\frac{1}{2}\left(1-\gamma^5\right)\psi(x) = Ai(\xi^\dagger,\eta^\dagger)\begin{pmatrix} O & I\\ I & O\end{pmatrix}\begin{pmatrix} O & \sigma^\mu\\ \tilde\sigma^\mu & O\end{pmatrix}\partial_\mu\,\frac{1}{2}\left(1-\gamma^5\right)\begin{pmatrix}\xi\\ \eta\end{pmatrix}\\ &= Ai(\xi^\dagger,\eta^\dagger)\begin{pmatrix}\tilde\sigma^\mu & O\\ O & \sigma^\mu\end{pmatrix}\partial_\mu\begin{pmatrix}\xi\\ 0\end{pmatrix} = Ai\xi^\dagger\tilde\sigma^\mu\partial_\mu\xi\end{aligned} \tag{10.166}$$

and, correspondingly, for the lower Weyl spinor

$$\begin{aligned}\mathcal{L} &= Ai\overline{\psi}(x)\gamma^\mu\partial_\mu\,\frac{1}{2}(1+\gamma^5)\psi(x) = Ai(\xi^\dagger,\eta^\dagger)\begin{pmatrix} O & I\\ I & O\end{pmatrix}\begin{pmatrix} O & \sigma^\mu\\ \tilde\sigma^\mu & O\end{pmatrix}\partial_\mu\,\frac{1}{2}(1+\gamma^5)\begin{pmatrix}\xi\\ \eta\end{pmatrix}\\ &= Ai(\xi^\dagger,\eta^\dagger)\begin{pmatrix}\tilde\sigma^\mu & O\\ O & \sigma^\mu\end{pmatrix}\partial_\mu\begin{pmatrix}0\\ \eta\end{pmatrix} = Ai\eta^\dagger\sigma^\mu\partial_\mu\eta\,.\end{aligned} \tag{10.167}$$

The two neutrino Lagrangians (10.165) are manifestly Lorentz-invariant due to (10.101), (10.104), (10.114), and (10.162). Furthermore, we have due to (10.135), (10.141), (10.151), and (10.161)

$$D(P)^{-1}\gamma^5 D(P) = \frac{i}{24}\,\epsilon_{\mu\nu\kappa\lambda}\Big[D(P)^{-1}\gamma^\mu D(P)\Big]\Big[D(P)^{-1}\gamma^\nu D(P)\Big]\Big[D(P)^{-1}\gamma^\kappa D(P)\Big]$$
$$\times\Big[D(P)^{-1}\gamma^\lambda D(P)\Big] = \frac{i}{24}\,\epsilon_{\mu\nu\kappa\lambda}\tilde\gamma^\mu\tilde\gamma^\nu\tilde\gamma^\kappa\tilde\gamma^\lambda = \frac{-i}{24}\,\epsilon_{\mu\nu\kappa\lambda}\gamma^\mu\gamma^\nu\gamma^\kappa\gamma^\lambda = -\gamma^5\,, \tag{10.168}$$

so that a parity transformation maps the two neutrino Lagrangians (10.165) into each other.

We remark that the Lagrangians (10.165) were proposed for the first time by the mathematician Hermann Weyl in 1929 to describe massless spin 1/2-particles. But since the neutrino Lagrangians (10.165) are not invariant under parity transformations and at that time only interactions like the electromagnetic or the strong one were known, which are invariant under parity transformations, the Lagrangians (10.165) were not considered to be physical for a long time. Only in 1956 it was shown by Chien-Shiung Wu in a β-decay experiment of ${}^{60}_{27}$Co that the weak interaction is not invariant under parity transformations and, thus, violates parity. Since this discovery neutrinos were assumed to be described by the Lagrangians (10.165) for decades. But in 1987 one managed to resolve the flavour of sun neutrinos in the Kamiokande experiment and one showed that they oscillate between the electron, the myuon, and the tauon flavour. From this observation it was concluded that neutrinos must have finite masses although their precise values have not yet been determined. Therefore, the Lagrangians (10.165) have been abandoned for describing neutrinos. But, due to their charge neutrality, until today it has not yet been finally decided how to describe theoretically neutrinos as massive spin 1/2 particles. Currently there exist two alternative descriptions, which go back to proposals of Paul Dirac and Ettore Majorana, respectively. In the first case neutrinos and anti-matter neutrinos are considered to be different particles, whereas in the second case they are assumed to be one and the same particle masquerading as two. An experimental decision between both possible theoretical descriptions is still lacking.

Subsequently, we consider the Weyl equation, that is, i.e. the equation of motion for massless spin 1/2 particles, which follows from (10.165):

$$\frac{\delta \mathcal{A}}{\delta \overline{\psi}(x)} = \frac{\partial \mathcal{L}}{\partial \overline{\psi}(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \overline{\psi}(x))} = A i \gamma^\mu \partial_\mu \frac{1}{2}\left(1 \mp \gamma^5\right) \psi(x) = 0 \,. \tag{10.169}$$

In the case of a particle with a fixed four-momentum vector $p = (p^\mu)$

$$\psi(x) = \psi \, e^{-ipx} \tag{10.170}$$

the Weyl equation (10.169) changes into

$$\boldsymbol{\gamma}\mathbf{p} \, \frac{1}{2}\left(1 \mp \gamma^5\right) \psi = \gamma^0 p^0 \, \frac{1}{2}\left(1 \mp \gamma^5\right) \psi \,. \tag{10.171}$$

Multiplying (10.171) from the left by $\gamma^5\gamma^0$, we obtain due to (10.95) and (10.154)

$$\gamma^5\gamma^0\gamma^k = \begin{pmatrix} -I & O \\ O & I \end{pmatrix} \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} O & \sigma^k \\ -\sigma^k & O \end{pmatrix} = \begin{pmatrix} \sigma^k & O \\ O & \sigma^k \end{pmatrix}, \tag{10.172}$$

thus, taking into account the spin operator (10.122) the result is

$$\frac{\mathbf{S}\mathbf{p}}{|p^0|} \, \frac{1}{2}\left(1 \mp \gamma^5\right) = \frac{1}{2}\,\mathrm{sgn}(p^0)\,\gamma^5 \, \frac{1}{2}\left(1 \mp \gamma^5\right) \,. \tag{10.173}$$

Due to the energy-momentum dispersion relation $p^0 = \pm|\mathbf{p}|$ the eigenstates $(1 \mp \gamma^5)\psi/2$ of the chirality operator γ^5 with the eigenvalues ∓ 1, see Eq. (10.160), are also eigenstates of the helicity operator with the eigenvalues $\mp\mathrm{sgn}(p^0)/2$. Thus, we conclude that chirality and helicity are identical for massless spin 1/2 particles.

10.10 Charge conjugation

The Lagrange density (10.94) of the Dirac field is also invariant with respect to another discrete symmetry transformation, where the components of the Dirac spinor $\psi(x)$ are replaced by the components of the complex conjugate Dirac spinor $\psi^*(x)$. In order to perform such a symmetry transformation we make the ansatz

$$\psi'_C(x) = C\,\overline{\psi}^T(x) = C\gamma^0\psi^*(x)\,, \tag{10.174}$$

where the row spinor $\overline{\psi}(x)$ from (10.92) goes over into the corresponding column spinor $\overline{\psi}^T(x)$ by transposition and we have used that $(\gamma^0)^T = \gamma^0$ due to (10.95). Furthermore, C denotes a complex 4×4-matrix which mixes these components and is defined by the fact that the transformed Dirac spinor (10.174) obeys the same Dirac equation

$$(i\gamma^\mu\partial_\mu - m)\,\psi'_C(x) = 0 \tag{10.175}$$

as the original Dirac spinor $\psi(x)$ in (10.98). Inserting (10.174) into (10.175) and multiplying from the left by C^{-1}, then we obtain at first

$$iC^{-1}\gamma^\mu C\partial_\mu\overline{\psi}^T(x) - m\overline{\psi}^T(x) = 0\,, \tag{10.176}$$

which changes due to a subsequent transposition T into

$$i\partial_\mu\overline{\psi}(x)\left(C^{-1}\gamma^\mu C\right)^T - m\overline{\psi}(x) = 0\,. \tag{10.177}$$

This equation of motion is now compared with the Dirac equation for the Dirac adjoint Dirac spinor $\overline{\psi}(x)$. In order to derive it we start from the Dirac equation (10.98) and go over to the adjoint, yielding

$$-i\partial_\mu\psi^\dagger(x)(\gamma^\mu)^\dagger - m\psi^\dagger(x) = 0\,. \tag{10.178}$$

Taking into account the Clifford algebra (10.96) for $\mu=\nu=0$ and (10.103) changes (10.178) into

$$-i\partial_\mu\overline{\psi}(x)\gamma^0(\gamma^\mu)^\dagger\gamma^0 - m\overline{\psi}(x) = 0\,. \tag{10.179}$$

Here we note that the Dirac matrices (10.95) have due to (10.16) the property

$$\gamma^0(\gamma^\mu)^\dagger\gamma^0 = \begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} O & \tilde{\sigma}^\mu \\ \sigma^\mu & O \end{pmatrix}\begin{pmatrix} O & I \\ I & O \end{pmatrix} = \begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix} = \gamma^\mu\,, \tag{10.180}$$

so that the Dirac equation for the Dirac-adjoint spinor (10.179) reduces to

$$i\partial_\mu\overline{\psi}(x)\gamma^\mu + m\overline{\psi}(x) = 0\,. \tag{10.181}$$

We remark that this equation of motion for the Dirac adjoint Dirac spinor $\overline{\psi}(x)$ corresponds to the Euler-Lagrange equation of the Dirac Lagrange density (10.93):

$$\frac{\delta \mathcal{A}}{\delta \psi(x)} = \frac{\partial \mathcal{L}}{\partial \psi(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi(x))} = A\Big[i\partial_\mu \overline{\psi}(x)\gamma^\mu + m\overline{\psi}(x)\Big] = 0\,. \tag{10.182}$$

The comparison of (10.177) and (10.181) then leads to the following equation for determining the matrix C:

$$\left(C^{-1}\gamma^\mu C\right)^T = -\gamma^\mu \qquad \Longrightarrow \qquad C^{-1}\gamma^\mu C = -\left(\gamma^\mu\right)^T\,. \tag{10.183}$$

In order to solve (10.183) we make the following diagonal ansatz for the matrix C

$$C = \begin{pmatrix} c & O \\ O & -c \end{pmatrix}, \qquad C^{-1} = \begin{pmatrix} c^{-1} & O \\ O & -c^{-1} \end{pmatrix}. \tag{10.184}$$

With this we obtain from (10.95) for the left-hand side of (10.183)

$$\begin{pmatrix} c^{-1} & O \\ O & -c^{-1} \end{pmatrix}\begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix}\begin{pmatrix} c & O \\ O & -c \end{pmatrix} = \begin{pmatrix} O & -c^{-1}\sigma^\mu c \\ -c^{-1}\tilde{\sigma}^\mu c & O \end{pmatrix}, \tag{10.185}$$

so we conclude from (10.183)

$$c^{-1}\sigma^\mu c = \left(\tilde{\sigma}^\mu\right)^T, \qquad c^{-1}\tilde{\sigma}^\mu c = \left(\sigma^\mu\right)^T. \tag{10.186}$$

Splitting both equations (10.186) into $\mu = 0$ and $\mu = k$, they yield the conditions

$$c^{-1}\sigma^0 c = (\sigma^0)^T\,, \tag{10.187}$$

$$c^{-1}\sigma^k c = -(\sigma^k)^T\,. \tag{10.188}$$

Here the transposed Pauli matrices (10.1) and (10.38) are given by

$$\left(\sigma^0\right)^T = \sigma^0\,, \qquad \left(\sigma^1\right)^T = \sigma^1\,, \qquad \left(\sigma^2\right)^T = -\sigma^2\,, \qquad \left(\sigma^3\right)^T = \sigma^3\,. \tag{10.189}$$

Let us now define the matrix c according to

$$c = -i\sigma^2 = -i\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{10.190}$$

As it has the properties

$$c^\dagger = c^{-1} = c^T = -c = -c^*\,, \tag{10.191}$$

we read off that (10.187) and (10.188) are, indeed, fulfilled due to (10.2) and (10.189)–(10.191)

$$c^{-1}\sigma^0 c = i\sigma^2\sigma^0(-i\sigma^2) = \sigma^2\sigma^0\sigma^2 = (\sigma^2)^2 = \sigma^0 = (\sigma^0)^T\,, \tag{10.192}$$

$$c^{-1}\sigma^1 c = i\sigma^2\sigma^1(-i\sigma^2) = \sigma^2\sigma^1\sigma^2 = -(\sigma^2)^2\sigma^1 = -\sigma^1 = -(\sigma^1)^T\,, \tag{10.193}$$

$$c^{-1}\sigma^2 c = i\sigma^2\sigma^2(-i\sigma^2) = (\sigma^2)^2\sigma^2 = \sigma^2 = -(\sigma^2)^T\,, \tag{10.194}$$

$$c^{-1}\sigma^3 c = i\sigma^2\sigma^3(-i\sigma^2) = \sigma^2\sigma^3\sigma^2 = -(\sigma^2)^2\sigma^3 = -\sigma^3 = -(\sigma^3)^T\,. \tag{10.195}$$

Thus, in conclusion, taking into account (10.190) and (10.191) the matrix C defined in (10.184) has the properties

$$C^\dagger = C^{-1} = C^T = -C = -C^* \tag{10.196}$$

and can be represented as a product of Dirac matrices (10.95):

$$i\gamma^0\gamma^2 = i\begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} O & \sigma^2 \\ -\sigma^2 & O \end{pmatrix} = \begin{pmatrix} -i\sigma^2 & O \\ O & i\sigma^2 \end{pmatrix} = \begin{pmatrix} c & O \\ O & -c \end{pmatrix} = C\,. \tag{10.197}$$

Moreover, taking into account (10.95), (10.184), (10.191), and (10.196), it follows that also the discrete symmetry transformation (10.174) is involutoric:

$$\begin{aligned}\psi''_C(x) &= C\gamma^0\psi'^*_C(x) = C\gamma^0C^*\gamma^0\psi(x) = \begin{pmatrix} c & O \\ O & -c \end{pmatrix}\begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} c^* & O \\ O & -c^* \end{pmatrix}\begin{pmatrix} O & I \\ I & O \end{pmatrix}\psi(x) \\ &= \begin{pmatrix} O & c \\ -c & O \end{pmatrix}\begin{pmatrix} O & c^* \\ -c^* & O \end{pmatrix}\psi(x) = \begin{pmatrix} -cc^* & O \\ O & -cc^* \end{pmatrix}\psi(x) = \psi(x)\,.\end{aligned} \tag{10.198}$$

And, finally, we investigate how the discrete symmetry transformation (10.174) affects the four-vector current density of the Dirac field invariant. Multiplying the equations of motion (10.98) and (10.179) for $\psi(x)$ and $\overline{\psi}(x)$ with $\overline{\psi}(x)$ and $\psi(x)$, respectively, we yield

$$i\overline{\psi}(x)\gamma^\mu\partial_\mu\psi(x) - m\overline{\psi}(x)\psi(x) = 0\,, \tag{10.199}$$

$$i\partial_\mu\overline{\psi}(x)\gamma^\mu\psi(x) + m\overline{\psi}(x)\psi(x) = 0\,, \tag{10.200}$$

so we read off the continuity equation

$$i\partial_\mu\Big[\overline{\psi}(x)\gamma^\mu\psi(x)\Big] = 0 \qquad \Longrightarrow \qquad \partial_\mu j^\mu(x) = 0\,. \tag{10.201}$$

Here the four-vector current density $j^\mu(x)$ is fixed except for a constant K:

$$j^\mu(x) = K\,\overline{\psi}(x)\gamma^\mu\psi(x)\,. \tag{10.202}$$

Thus, the conserved charge reads due to (10.95) and (10.202)

$$Q = \int d^3x\, j^0(\mathbf{x},t) = K\int d^3x\,\psi^\dagger(\mathbf{x},t)\psi(\mathbf{x},t)\,. \tag{10.203}$$

In order to apply the discrete symmetry transformation (10.174) to the four-vector current density (10.202), we need to know how the Dirac adjoint Dirac spinor (10.103) is transformed. Thus, applying (10.95), (10.184), (10.191), and (10.196) we yield

$$\begin{aligned}\overline{\psi}'_C(x) &= \psi'^\dagger_C(x)\gamma^0 = \psi^T(x)(\gamma^0)^\dagger C^\dagger\gamma^0 = -\psi^T(x)\gamma^0C\gamma^0 = -\psi^T(x)\begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} c & O \\ O & -c \end{pmatrix} \\ &\times\begin{pmatrix} O & I \\ I & O \end{pmatrix} = -\psi^T(x)\begin{pmatrix} O & I \\ I & O \end{pmatrix}\begin{pmatrix} O & c \\ -c & O \end{pmatrix} = -\psi^T(x)\begin{pmatrix} -c & O \\ O & c \end{pmatrix} = \psi^T(x)C\,.\end{aligned} \tag{10.204}$$

Transforming the four-vector current density (10.202) with (10.174) and (10.204) we then conclude at first

$$j'^{\mu}_{C}(x) = K\overline{\psi}'_{C}\gamma^{\mu}\psi'_{C}(x) = K\psi^{T}(x)C\gamma^{\mu}C\gamma^{0}\psi^{*}(x)\,. \tag{10.205}$$

As each individual component of the transformed four-vector current density (10.205) coincides with its transposition, i.e. $j'^{\mu}_{C}(x) = (j'^{\mu}_{C}(x))^{T}$, it follows from (10.95), (10.103), (10.183), (10.196), and (10.202) that

$$j'^{\mu}_{C}(x) = K\psi^{\dagger}(x)(\gamma^{0})^{T}(C\gamma^{\mu}C)^{T}\psi(x) = K\psi^{\dagger}(x)\gamma^{0}\gamma^{\mu}\psi(x) = K\overline{\psi}(x)\gamma^{\mu}\psi(x) = j^{\mu}(x)\,. \tag{10.206}$$

Thus, we conclude that the discrete symmetry transformation (10.174) turns out not to change the four-vector current density. Note that the physical meaning of the discrete symmetry transformation (10.174) as a charge conjugation becomes clear only after having implemented the second quantization of the Dirac field, as then the four-vector density operator changes its sign in contrast to (10.206),

10.11 Time Inversion

Performing a time inversion T, the space-time four-vector x is mapped into the time-inverted space-time four-vector $-\tilde{x}$:

$$x'_{T} = Tx = -\tilde{x}\,. \tag{10.207}$$

Executing a time inversion T successively twice, one reproduces the original state, so the time inversion T is also involutoric:

$$T^{2} = 1 \qquad \Longleftrightarrow \qquad T^{-1} = T\,. \tag{10.208}$$

The representation matrix for such a time inversion reads as follows

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\,. \tag{10.209}$$

Thus, we conclude that the representation matrix of the time inversion (10.209) commutates with the matrix representations for the generators of rotations (6.53)

$$T^{-1}L_{k}T = L_{k} \tag{10.210}$$

and anti-commutates with the matrix representations for the generators of boosts (6.54)

$$T^{-1}M_{k}T = -M_{k}\,. \tag{10.211}$$

For instance, we have

$$T^{-1}L_1T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = L_1 \,, \tag{10.212}$$

$$T^{-1}M_1T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = -M_1 \,. \tag{10.213}$$

As the time inversion is more intriguing to interpret physicallys, we investigate at first its consequences for the Schrödinger equation

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2M} \Delta \right) \psi(\mathbf{x}, t) = 0 \,. \tag{10.214}$$

Obviously, the time inverted wave function

$$\psi'_T(\mathbf{x}, t) = \psi^*(\mathbf{x}, -t) \tag{10.215}$$

also obeys the Schrödinger equation:

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2M} \Delta \right) \psi'_T(\mathbf{x}, t) = 0 \,. \tag{10.216}$$

In analogy to (10.215) we now perform the time inversion for a Dirac spinor via

$$\psi(x) \longrightarrow \psi'_T(x) = D(T)\psi^*(-\tilde{x}) \,, \tag{10.217}$$

where $D(T)$ stands for the representation matrix of the time inversion in the space of Dirac spinors. Then $D(T)$ must also fulfill the involutoric property (10.208)

$$D(T)^2 = 1 \tag{10.218}$$

and we expect that also the commutator and anti-commutator relations (10.210) and (10.211) are satisfied by the representation matrices $D(L_k)$ and $D(M_k)$ of rotations and boosts in the space of Dirac spinors, respectively:

$$D(T)^{-1} D(L_k) D(T) = D(L_k) \,, \tag{10.219}$$

$$D(T)^{-1} D(M_k) D(T) = -D(M_k) \,. \tag{10.220}$$

In analogy with (10.216), we also require that the time inverted Dirac spinor (10.217) satisfies the Dirac equation (10.98):

$$(i\gamma^\mu \partial_\mu - m) \, \psi'_T(x) = 0 \,. \tag{10.221}$$

Inserting (10.217) into (10.221), we obtain

$$-i\Big[D(T)^{-1}\gamma^{\mu}D(T)\Big]^{*}\partial_{\mu}\psi(-\tilde{x})-m\psi(-\tilde{x})=0\,. \tag{10.222}$$

Comparing (10.222) with the time-inverted Dirac equation (10.98)

$$-i\tilde{\gamma}^{\mu}\partial_{\mu}\psi(-\tilde{x})-m\psi(-\tilde{x})=0\,, \tag{10.223}$$

where we used $\gamma^{\mu}\tilde{\partial}_{\mu}=\tilde{\gamma}^{\mu}\partial_{\mu}$, the representation matrix $D(T)$ of the time inversion is determined by the equation

$$D(T)^{-1}\gamma^{\mu}D(T)=(\tilde{\gamma}^{\mu})^{*}\,. \tag{10.224}$$

On the one hand we calculate the conjugate complex of the Dirac matrices (10.95) by taking into account the Pauli matrices (10.1), yielding

$$\begin{aligned}
\left(\gamma^{0}\right)^{*} &= \begin{pmatrix} O & \sigma^{0} \\ \sigma^{0} & O \end{pmatrix}, \qquad \left(\gamma^{1}\right)^{*} = \begin{pmatrix} O & \sigma^{1} \\ -\sigma^{1} & O \end{pmatrix}, \\
\left(\gamma^{2}\right)^{*} &= \begin{pmatrix} O & -\sigma^{2} \\ \sigma^{2} & O \end{pmatrix}, \qquad \left(\gamma^{3}\right)^{*} = \begin{pmatrix} O & \sigma^{3} \\ -\sigma^{3} & O \end{pmatrix}.
\end{aligned} \tag{10.225}$$

On the other hand we obtain for the quantities $(\tilde{\gamma}^{\mu})^{T}$:

$$\begin{aligned}
\left(\tilde{\gamma}^{0}\right)^{T} &= \left(\gamma^{0}\right)^{T} = \begin{pmatrix} O & (\sigma^{0})^{T} \\ (\sigma^{0})^{T} & O \end{pmatrix} = \begin{pmatrix} O & \sigma^{0} \\ \sigma^{0} & O \end{pmatrix}, \\
\left(\tilde{\gamma}^{1}\right)^{T} &= -\left(\gamma^{1}\right)^{T} = -\begin{pmatrix} O & -(\sigma^{1})^{T} \\ (\sigma^{1})^{T} & O \end{pmatrix} = \begin{pmatrix} O & \sigma^{1} \\ -\sigma^{1} & O \end{pmatrix}, \\
\left(\tilde{\gamma}^{2}\right)^{T} &= -\left(\gamma^{2}\right)^{T} = -\begin{pmatrix} O & -(\sigma^{2})^{T} \\ (\sigma^{2})^{T} & O \end{pmatrix} = \begin{pmatrix} O & -\sigma^{2} \\ \sigma^{2} & O \end{pmatrix}, \\
\left(\tilde{\gamma}^{3}\right)^{T} &= -\left(\gamma^{3}\right)^{T} = -\begin{pmatrix} O & -(\sigma^{3})^{T} \\ (\sigma^{3})^{T} & O \end{pmatrix} = \begin{pmatrix} O & \sigma^{3} \\ -\sigma^{3} & O \end{pmatrix}.
\end{aligned} \tag{10.226}$$

Thus, from (10.225) and (10.226) we read off the following identity

$$(\gamma^{\mu})^{*}=(\tilde{\gamma}^{\mu})^{T} \qquad \Longrightarrow \qquad (\gamma^{\mu})^{\dagger}=\tilde{\gamma}^{\mu}\,. \tag{10.227}$$

Inserting (10.227) into (10.224) then results in

$$D(T)^{-1}\gamma^{\mu}D(T)=(\gamma^{\mu})^{T}\,. \tag{10.228}$$

Now we take into account the property (10.183), which relates the Dirac matrices γ^{μ} with the representation matrix C of charge conjugation in the space of Dirac spinors. With this the equation (10.228) for determining $D(T)$ leads to

$$D(T)^{-1}\gamma^{\mu}D(T)=-C^{-1}\gamma^{\mu}C \qquad \Longrightarrow \qquad \Big[D(T)C^{-1}\Big]^{-1}\gamma^{\mu}\Big[D(T)C^{-1}\Big]=-\gamma^{\mu}\,. \tag{10.229}$$

A solution of (10.229) is given by

$$D(T)C^{-1} = -i\gamma^5 \tag{10.230}$$

together with its inverted matrix following from (10.155)

$$\left[D(T)C^{-1}\right]^{-1} = i\gamma^5 \,, \tag{10.231}$$

as is verified by an explicit calculation due to (10.95) and (10.154):

$$\gamma^5\gamma^\mu\gamma^5 = \begin{pmatrix} -I & O \\ O & I \end{pmatrix}\begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix}\begin{pmatrix} -I & O \\ O & I \end{pmatrix} = -\begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix} = -\gamma^\mu \,. \tag{10.232}$$

Note that (10.230) represents a quite subtle relation, which involves with the matrices γ^5, C, and $D(T)$ technical ingredients of all three discrete transformation, i.e. the parity, the charge conjugation, and the time inversion. Thus, taking into account (10.154) and (10.184), the representation matrix $D(T)$ follows from (10.230)

$$D(T) = -i\gamma^5 C = -i\begin{pmatrix} -I & O \\ O & I \end{pmatrix}\begin{pmatrix} c & O \\ O & -c \end{pmatrix} = i\begin{pmatrix} c & O \\ O & c \end{pmatrix} , \tag{10.233}$$

which has due to (10.191) the properties

$$D(T) = D(T)^{-1} = D(T)^\dagger = -D(T)^* = -D(T)^T \,. \tag{10.234}$$

According to (10.234) the representation matrix $D(T)$ satisfies the involutoric property (10.218), but the time inversion of the Dirac spinor is not involutoric due to (10.217) and (10.234):

$$\psi''_T(x) = D(T)\psi'^*_T(-\tilde{x}) = D(T)D(T)^*\psi(x) = -\psi(x) \,. \tag{10.235}$$

This behavior of Dirac spinors under time inversion corresponds to that under a rotation, where we read off from (10.21) and (10.102) that the original Dirac spinor is only recovered after a rotation with the angle 4π. Furthermore, we obtain for the commutators of $D(T)$ with the generators of rotation $D(L_k)$ due to (10.16), (10.122), (10.188), and (10.191)

$$\begin{aligned} D(T)^{-1}D(L_k)D(T) &= -\begin{pmatrix} c & O \\ O & c \end{pmatrix}\begin{pmatrix} \sigma^k/2 & O \\ O & \sigma^k/2 \end{pmatrix}\begin{pmatrix} c & O \\ O & c \end{pmatrix} = \begin{pmatrix} -c\sigma^k c/2 & O \\ O & -c\sigma^k c/2 \end{pmatrix} \\ &= \begin{pmatrix} (\sigma^k)^T/2 & O \\ O & (\sigma^k)^T/2 \end{pmatrix} = \begin{pmatrix} (\sigma^k)^*/2 & O \\ O & (\sigma^k)^*/2 \end{pmatrix} = D(L_k)^* \,, \end{aligned} \tag{10.236}$$

and, correspondingly, the commutators of $D(T)$ with $D(M_k)$ yield with (10.121) and (10.233)

$$\begin{aligned} D(T)^{-1}D(M_k)D(T) &= \frac{i}{2}\begin{pmatrix} c & O \\ O & c \end{pmatrix}\begin{pmatrix} \sigma^k & O \\ O & -\sigma^k/2 \end{pmatrix}\begin{pmatrix} c & O \\ O & c \end{pmatrix} = \frac{i}{2}\begin{pmatrix} c\sigma^k c & O \\ O & -c\sigma^k c \end{pmatrix} \\ = \frac{-i}{2}\begin{pmatrix} (\sigma^k)^T & O \\ O & -(\sigma^k)^T \end{pmatrix} &= \frac{-i}{2}\begin{pmatrix} (\sigma^k)^* & 0 \\ 0 & -(\sigma^k)^* \end{pmatrix} = -D(M_k)^* \,. \end{aligned} \tag{10.237}$$

The results (10.236) and (10.237) do not match the original expectations (10.219) and (10.220). Instead, they indicate that the time inversion represents an anti-linear operation as is further discussed in the exercises in the context of the second quantization of the Dirac field.

10.12 Dirac Representation

The representation (10.95) of the Dirac matrices used so far is called the chiral representation or the Weyl representation, as then the chirality operator γ^5 is diagonal according to (10.154). From a group-theoretical point of view this representation has the advantage that the representation matrices of the Lorentz transformation in the space of the Dirac spinors have a block diagonal shape according to (10.102), i.e. both Weyl spinors are treated on equal footing. Another common representation of the Dirac matrices is the so-called Dirac representation or the standard representation

$$\psi_D(x) = S_D \psi(x) \,, \tag{10.238}$$

where the transformation matrix S_D is given by

$$S_D = \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \tag{10.239}$$

with the inverse

$$S_D^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} = S_D^T \,. \tag{10.240}$$

Thus, the transformation matrix S_D is orthonormal or, more precisely, unitary. For the Dirac adjoint Dirac spinor $\overline{\psi}(x)$ one obtains in the Dirac representation from (10.92), (10.95), (10.239), and (10.240):

$$\begin{aligned} \overline{\psi}_D(x) &= \psi_D^\dagger(x)\gamma^0 = \psi^\dagger(x) S_D^\dagger \gamma^0 = \overline{\psi}(x)\gamma^0 S_D^\dagger \gamma^0 = \overline{\psi}(x) \frac{1}{\sqrt{2}} \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \begin{pmatrix} O & I \\ I & O \end{pmatrix} \\ &= \overline{\psi}(x) \frac{1}{\sqrt{2}} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} = \overline{\psi}(x) S_D^{-1} \,. \end{aligned} \tag{10.241}$$

In the same way one obtains for the Dirac matrices γ^μ in the Dirac representation

$$\gamma_D^0 = S_D \gamma^0 S_D^{-1} = \frac{1}{2} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} = \begin{pmatrix} I & O \\ O & -I \end{pmatrix} , \tag{10.242}$$

$$\gamma_D^k = S_D \gamma^k S_D^{-1} = \frac{1}{2} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \begin{pmatrix} O & \sigma^k \\ -\sigma^k & O \end{pmatrix} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} = \begin{pmatrix} O & \sigma^k \\ -\sigma^k & O \end{pmatrix} . \tag{10.243}$$

And, correspondingly, the chirality operator (10.154) in the Dirac representation turns out to be no longer diagonal:

$$\gamma_D^5 = S_D \gamma^5 S_D^{-1} = \frac{1}{2} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \begin{pmatrix} -I & O \\ O & I \end{pmatrix} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} = \begin{pmatrix} O & I \\ I & O \end{pmatrix} . \tag{10.244}$$

Conversely, the Dirac matrix γ^0 is not diagonal in the Weyl representation (10.95), while it is diagonal in the Dirac representation (10.242). Furthermore, the generators of the rotations in the spinor space (10.122) are invariant under the change of representation

$$D(L_k)_D = S_D D(L_k) S_D^{-1} = \frac{1}{4} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \begin{pmatrix} \sigma^k & O \\ O & \sigma^k \end{pmatrix} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sigma^k & O \\ O & \sigma^k \end{pmatrix} , \tag{10.245}$$

whereas the generators of the boosts in the spinor space (10.121) result in the Dirac representation to be given by

$$D(M_k)_D = S_D D(M_k) S_D^{-1} = \frac{i}{4} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \begin{pmatrix} -\sigma^k & O \\ O & \sigma^k \end{pmatrix} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} = \frac{i}{2} \begin{pmatrix} O & -\sigma^k \\ \sigma^k & O \end{pmatrix} . \tag{10.246}$$

10.13 Non-Relativistic Limit

The Dirac representation has the advantage that the non-relativistic limit is straight-forwardly carried out. To this end we transform the Dirac equation (10.98) according to (10.238) into the Dirac representation:

$$i\gamma_D^\mu \partial_\mu \psi_D(x) - m\psi_D(x) = 0 . \tag{10.247}$$

In this manifestly covariant formulation of the Dirac equation, we separate now explicitly the respective temporal and spatial contributions

$$i\gamma_D^0 \frac{1}{c}\frac{\partial}{\partial t} \psi_D(\mathbf{x},t) + i\gamma_D^k \partial_k \psi_D(\mathbf{x},t) - m\psi_D(\mathbf{x},t) = 0 . \tag{10.248}$$

The Dirac equation (10.248) can then be rewritten in the form of a Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \psi_D(\mathbf{x},t) = H_D(\mathbf{x})\psi_D(\mathbf{x},t) , \tag{10.249}$$

where the Dirac Hamiltonian is given by

$$H_D(\mathbf{x}) = -ic\hbar\, \boldsymbol{\alpha}\boldsymbol{\nabla} + c\hbar m\beta . \tag{10.250}$$

Here we have introduced the matrices

$$\beta = \gamma_D^0 = \begin{pmatrix} I & O \\ O & -I \end{pmatrix} , \tag{10.251}$$

$$\alpha^k = \gamma_D^0 \gamma_D^k = \begin{pmatrix} I & O \\ O & -I \end{pmatrix} \begin{pmatrix} O & \sigma^k \\ -\sigma^k & O \end{pmatrix} = \begin{pmatrix} O & \sigma^k \\ \sigma^k & O \end{pmatrix} , \tag{10.252}$$

where we used (10.242) and (10.243). With this we obtain the anti-commutator relations

$$[\beta,\beta]_+ = 2\begin{pmatrix} I & O \\ O & -I \end{pmatrix}\begin{pmatrix} I & O \\ O & -I \end{pmatrix} = 2\mathcal{I}\,, \tag{10.253}$$

$$\left[\alpha^k,\beta\right]_+ = \begin{pmatrix} O & \sigma^k \\ \sigma^k & O \end{pmatrix}\begin{pmatrix} I & O \\ O & -I \end{pmatrix} + \begin{pmatrix} I & O \\ O & -I \end{pmatrix}\begin{pmatrix} O & \sigma^k \\ \sigma^k & O \end{pmatrix} = \mathcal{O}\,, \tag{10.254}$$

$$\begin{aligned}\left[\alpha^k,\alpha^l\right]_+ &= \begin{pmatrix} O & \sigma^k \\ \sigma^k & O \end{pmatrix}\begin{pmatrix} O & \sigma^l \\ \sigma^l & O \end{pmatrix} + \begin{pmatrix} O & \sigma^l \\ \sigma^l & O \end{pmatrix}\begin{pmatrix} O & \sigma^k \\ \sigma^k & O \end{pmatrix} \\ &= \begin{pmatrix} \left[\sigma^k,\sigma^l\right]_+ & O \\ O & \left[\sigma^l,\sigma^k\right]_+ \end{pmatrix} = 2\delta_{kl}\,\mathcal{I}\,, \end{aligned} \tag{10.255}$$

where in the latter case we applied the Clifford algebra of the Pauli matrices (10.2). Furthermore, we introduced as new abbreviations both the 4×4 unit matrix

$$\mathcal{I} = \begin{pmatrix} I & O \\ O & I \end{pmatrix} \tag{10.256}$$

and the 4×4 zero matrix

$$\mathcal{O} = \begin{pmatrix} O & O \\ O & O \end{pmatrix} . \tag{10.257}$$

Thus, we read off from (10.253)–(10.255) that the 4×4 matrices β, α^k represent a Clifford algebra with $N = 4$ generators in the sense of (10.4).

In close analogy to the Weyl representation in (10.89), we now decompose also in the Dirac representation the four-component Dirac spinor into two two-component Weyl spinors

$$\psi_D(\mathbf{x},t) = \begin{pmatrix} \xi_D(\mathbf{x},t) \\ \eta_D(\mathbf{x},t) \end{pmatrix} . \tag{10.258}$$

Inserting (10.258) into (10.249) and (10.250) as well as taking into account (10.251) and (10.252) then leads to

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} \xi_D(\mathbf{x},t) \\ \eta_D(\mathbf{x},t) \end{pmatrix} = -ic\hbar\begin{pmatrix} O & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & O \end{pmatrix}\boldsymbol{\nabla}\begin{pmatrix} \xi_D(\mathbf{x},t) \\ \eta_D(\mathbf{x},t) \end{pmatrix} + c\hbar m\begin{pmatrix} I & O \\ O & -I \end{pmatrix}\begin{pmatrix} \xi_D(\mathbf{x},t) \\ \eta_D(\mathbf{x},t) \end{pmatrix} , \tag{10.259}$$

which reduces to two coupled equations of motion for these Weyl spinors in the Dirac representation:

$$i\hbar\frac{\partial}{\partial t}\xi_D(\mathbf{x},t) = -ic\hbar\boldsymbol{\sigma}\boldsymbol{\nabla}\eta_D(\mathbf{x},t) + c\hbar m\xi_D(\mathbf{x},t)\,, \tag{10.260}$$

$$i\hbar\frac{\partial}{\partial t}\eta_D(\mathbf{x},t) = -ic\hbar\boldsymbol{\sigma}\boldsymbol{\nabla}\xi_D(\mathbf{x},t) + c\hbar m\eta_D(\mathbf{x},t)\,. \tag{10.261}$$

As discussed already in Fig. 8.1 we now take into account that the relativistic and the non-relativistic energy scales are shifted against each other by the rest energy Mc^2, which leads to the ansatz

$$\psi_D(\mathbf{x},t) = \begin{pmatrix} \xi_D(\mathbf{x},t) \\ \eta_D(\mathbf{x},t) \end{pmatrix} = \begin{pmatrix} \tilde{\xi}_D(\mathbf{x},t)\,e^{-iMc^2t/\hbar} \\ \tilde{\eta}_D(\mathbf{x},t)\,e^{-iMc^2t/\hbar} \end{pmatrix} . \tag{10.262}$$

Thus the coupled equations of motion (10.260), (10.261) go over into

$$i\hbar\frac{\partial}{\partial t}\tilde{\xi}_D(\mathbf{x},t) = -ic\hbar\boldsymbol{\sigma}\boldsymbol{\nabla}\tilde{\eta}_D(\mathbf{x},t) + \left(c\hbar m - Mc^2\right)\tilde{\xi}_D(\mathbf{x},t)\,, \tag{10.263}$$

$$i\hbar\frac{\partial}{\partial t}\tilde{\eta}_D(\mathbf{x},t) = -ic\hbar\boldsymbol{\sigma}\boldsymbol{\nabla}\tilde{\xi}_D(\mathbf{x},t) + \left(-c\hbar m - Mc^2\right)\tilde{\eta}_D(\mathbf{x},t)\,. \tag{10.264}$$

As the parameter m was determined according to (10.88) to be inversely proportional to the Compton wave length (8.21), the rest energy Mc^2 turns out to appear only in the second equation of motion:

$$i\hbar\frac{\partial}{\partial t}\tilde{\xi}_D(\mathbf{x},t) = -ic\hbar\boldsymbol{\sigma}\boldsymbol{\nabla}\tilde{\eta}_D(\mathbf{x},t)\,, \tag{10.265}$$

$$i\hbar\frac{\partial}{\partial t}\tilde{\eta}_D(\mathbf{x},t) = -ic\hbar\boldsymbol{\sigma}\boldsymbol{\nabla}\tilde{\xi}_D(\mathbf{x},t) - 2Mc^2\tilde{\eta}_D(\mathbf{x},t)\,. \tag{10.266}$$

Performing now the non-relativistic limes $c \to \infty$ the kinetic energy of the Weyl spinor $\tilde{\eta}_D$ is negligible in comparison with its rest energy, i.e.

$$\left|i\hbar\frac{\partial}{\partial t}\tilde{\eta}_D(\mathbf{x},t)\right| \ll \left|Mc^2\tilde{\eta}_D(\mathbf{x},t)\right|\,, \tag{10.267}$$

so that the Weyl spinor $\tilde{\eta}_D$ can approximately be expressed by the Weyl spinor $\tilde{\xi}_D$:

$$\tilde{\eta}_D(\mathbf{x},t) = \frac{-i\hbar}{2Mc}\,\boldsymbol{\sigma}\boldsymbol{\nabla}\tilde{\xi}_D(\mathbf{x},t)\,. \tag{10.268}$$

Neglecting the temporal derivative in (10.266) thus leads to an adiabatic elimination of the Weyl spinor $\tilde{\eta}_D(\mathbf{x},t)$, i.e. it now longer has an independent dynamics but its temporal evolution follows quasi-instantaneously the corresponding one of the Weyl spinor $\tilde{\xi}_D(\mathbf{x},t)$. Note that similar applications of an adiabatic elimination of degrees of freedom are ubiquitous in theoretical physics:

- One prominent example is provided by the Born-Oppenheimer approximation in molecular physics. It is based on recognizing the large difference between the electron mass and the masses of atomic nuclei, and correspondingly the respective time scales of their motion. Given the same amount of kinetic energy, the nuclei move much more slowly than the electrons. Therefore, it is a valid assumption that the wave functions of atomic nuclei and electrons in a molecule can be treated separately. This enables a separation of the Hamiltonian operator into electronic and nuclear terms, where cross-terms between electrons and nuclei are neglected, so that the two smaller and decoupled systems can be solved more efficiently. As a result an effective electronic Hamilton operator for the electronic degrees of freedom is solved, where the positions of the nuclei are fixed quantities. In the second step of the Born-Oppenheimer approximation the Schrödinger equation for the nuclear motion is treated.

- Another important example is the semi-classical laser theory, where the electric field described by the Maxwell theory couples to the matter degrees of freedom, which are dealt with quantum mechanically. For the laser it turns out that the electric field evolves on a much larger time scale than the matter degrees of freedom. This allows to adiabatically eliminate the matter degrees of freedom from the dynamics and obtain an effective evolution equation for the electric field, which describes the spontaneous emergence of coherent laser light from an originally incoherent lamp light by increasing the pump power. This adiabatic elimination of fast (stable) degrees of freedom in favour of obtaining a resulting order parameter equation for slow (unstable) degrees of freedom was recognized by Hermann Haken in the realm of synergetics, which is a theory of self-organization. This fundamental discovery leads to many fascinating applications in natural and, partially, also in social sciences.

After this excursion we return to working out the non-relativistic limit of the Dirac equation. Substituting (10.268) into (10.265) leads to a Schrödinger equation for the Weyl spinor $\tilde{\xi}_D(\mathbf{x},t)$:

$$i\hbar\frac{\partial}{\partial t}\tilde{\xi}_D(\mathbf{x},t) = -\frac{\hbar^2}{2M}\sigma^k\partial_k\sigma^l\partial_l\tilde{\xi}_D(\mathbf{x},t) = -\frac{\hbar^2}{4M}\left[\sigma^k,\sigma^l\right]_+\partial_l\partial_k\tilde{\xi}_D(\mathbf{x},t) = -\frac{\hbar^2}{2M}\Delta\tilde{\xi}_D(\mathbf{x},t) \quad (10.269)$$

with applying the Clifford algebra of the Pauli matrices (10.2). In the exercises we work out the non-relativistic limit of the Dirac equation in the presence of a minimal coupling to the electromagnetic field in a more systematic way by performing the so-called Foldy-Wouthuysen transformation. This leads then not to the Schrödinger equation (10.269) but to the Pauli equation for the Weyl spinor $\tilde{\xi}_D(\mathbf{x},t)$ containing automatically the correct Landé factor $g_s = 2$ for a point-like massive spin 1/2 particle. Note that both the proton and the neutron are also massive spin 1/2 particles but measurements show that their respective Landé factors 2.79 and - 1.91 deviate significantly from 2.0 which indicates that they are not point-like but composite particles. Indeed, according to the standard model of elementary particle physics, each of these nucleons consists of three quarks, which are point-like massive spin 1/2 particles according to the present day knowledge.

Let us consider now the non-relativistic limit of the Dirac action (10.93), (10.97) in the Dirac representation

$$\mathcal{A} = \frac{A}{c}\int d^4x\,\overline{\psi}_D(x)\left(i\gamma_D^\mu\partial_\mu - m\right)\psi_D(x)\,. \quad (10.270)$$

As a first preparatory step we separate explicitly the respective temporal and spatial contributions:

$$\mathcal{A} = \frac{A}{c}\int d^4x\left[i\overline{\psi}_D(\mathbf{x},t)\gamma_D^0\frac{1}{c}\frac{\partial}{\partial t}\psi_D(\mathbf{x},t) + i\overline{\psi}_D(\mathbf{x},t)\boldsymbol{\gamma}_D\boldsymbol{\nabla}\psi_D(\mathbf{x},t) - m\overline{\psi}_D(\mathbf{x},t)\psi_D(\mathbf{x},t)\right]. \quad (10.271)$$

Then we take into account how the Dirac spinor decomposes into the Weyl spinors according to (10.262) and the corresponding expression for the Dirac adjoint Dirac spinor following from (10.103) and (10.242):

$$\overline{\psi}_D(\mathbf{x},t) = \psi_D^\dagger(\mathbf{x},t)\gamma_D^0 = \left(\tilde{\xi}_D^\dagger(\mathbf{x},t)e^{iMc^2t/\hbar}, -\tilde{\eta}_D^\dagger(\mathbf{x},t)e^{iMc^2t/\hbar}\right)\,. \quad (10.272)$$

Using in addition (10.88), (10.242), and (10.243) as well as (10.262) and (10.272), the Dirac action (10.270) reduces to

$$\mathcal{A} = A \int dt \int d^3x \left\{ \frac{i}{c} \left[\tilde{\xi}_D^\dagger(\mathbf{x},t) \, \frac{\partial \tilde{\xi}_D(\mathbf{x},t)}{\partial t} + \tilde{\eta}_D^\dagger(\mathbf{x},t) \, \frac{\partial \tilde{\eta}_D(\mathbf{x},t)}{\partial t} \right] \right.$$
$$\left. + i \left[\tilde{\xi}_D^\dagger(\mathbf{x},t) \boldsymbol{\sigma} \boldsymbol{\nabla} \tilde{\eta}_D(\mathbf{x},t) + \tilde{\eta}_D^\dagger(\mathbf{x},t) \boldsymbol{\sigma} \boldsymbol{\nabla} \tilde{\xi}_D(\mathbf{x},t) \right] + \frac{2Mc}{\hbar} \, \tilde{\eta}_D^\dagger(\mathbf{x},t) \tilde{\eta}_D(\mathbf{x},t) \right\} . \tag{10.273}$$

If one now expresses the Weyl spinor $\tilde{\eta}_D$ according to (10.268) by the Weyl spinor $\tilde{\xi}_D$ and takes into account the calculation rule (10.7), then (10.273) goes over in the non-relativistic limes $c \to \infty$ into

$$\mathcal{A} = A \int dt \int d^3x \left[\frac{i}{c} \, \tilde{\xi}_D^\dagger(\mathbf{x},t) \frac{\partial \tilde{\xi}(\mathbf{x},t)}{\partial t} + \frac{\hbar}{2Mc} \, \tilde{\xi}_D^\dagger(\mathbf{x},t) \Delta \tilde{\xi}_D(\mathbf{x},t) \right] . \tag{10.274}$$

Fixing the yet undetermined parameter A according to

$$\alpha = c\hbar \, , \tag{10.275}$$

then (10.274) reduces to the Schrödinger action for the Weyl spinor $\tilde{\xi}_D$:

$$\mathcal{A} = \int dt \int d^3x \left[i\hbar \, \tilde{\xi}_D^\dagger(\mathbf{x},t) \, \frac{\partial \tilde{\xi}_D(\mathbf{x},t)}{\partial t} + \frac{\hbar^2}{2M} \, \tilde{\xi}_D^\dagger(\mathbf{x},t) \Delta \tilde{\xi}_D(\mathbf{x},t) \right] . \tag{10.276}$$

Furthermore, according to (10.88) and (10.275), we then conclude that the Dirac Lagrange density in the Weyl representation (10.94) reads

$$\mathcal{L} = \overline{\psi}(x) \left(i\hbar c \gamma^\mu \partial_\mu - Mc^2 \right) \psi(x) \, . \tag{10.277}$$

And finally, inserting (10.262) and (10.268) into the conserved charge (10.203), we read off in the non-relativistic limit $c \to \infty$ that the yet undetermined parameter K has to be identified with

$$K = 1 \, , \tag{10.278}$$

so that we obtain in the Dirac representation the conserved quantity expected for a Schrödinger theory:

$$Q = \int d^3x \, \tilde{\xi}_D^\dagger(\mathbf{x},t) \tilde{\xi}_D(\mathbf{x},t) \, . \tag{10.279}$$

Thus, we conclude that the conserved charge (10.203) of the Dirac theory reads

$$Q = \int d^3x \, \psi^\dagger(\mathbf{x},t) \psi(\mathbf{x},t) \, . \tag{10.280}$$

10.14 Plane Waves

We now determine the fundamental solutions of the Dirac equation in the Weyl representation (10.98), which reads by taking into account (10.88):

$$\left(i\gamma^{\mu}\partial_{\mu} - \frac{Mc}{\hbar} \right) \psi(x) = 0 \,. \tag{10.281}$$

One solution method relies on performing a plane wave ansatz for the Dirac spinor $\psi(x)$, which converts the differential equation (10.281) into an algebraic equation for the corresponding spinor amplitudes. The latter would then have to be solved on the basis of the concrete form of the Dirac matrices in the Weyl representation. In this section, however, we work out a different solution method, which is group theoretically inspired. To this end we determine at first the trivial plane wave solutions in the rest frame of the massive spin 1/2 particle and then we boost them to a uniformly moving reference frame.

10.14.1 Rest Frame

In the rest frame of the massive spin 1/2 particle, the Dirac spinor can only depend on time t:

$$\psi_R(x) = \psi(t) \,. \tag{10.282}$$

Inserting (10.282) in (10.281) leads to

$$\left(i\gamma^{0}\frac{\partial}{\partial t} - \frac{Mc^2}{\hbar} \right) \psi(t) = 0 \,. \tag{10.283}$$

Multiplying (10.283) with the operator $(-i\gamma^0\partial/\partial t - Mc^2/\hbar)$ and taking into account $(\gamma^0)^2 = \mathcal{I}$ due to (10.95) then yields

$$\left(-i\gamma^{0}\frac{\partial}{\partial t} - \frac{Mc^2}{\hbar} \right)\left(i\gamma^{0}\frac{\partial}{\partial t} - \frac{Mc^2}{\hbar} \right) \psi(t) = \left[\frac{\partial^2}{\partial t^2} + \left(\frac{Mc^2}{\hbar} \right)^2 \right] \psi(t) = 0 \,. \tag{10.284}$$

Thus, we obtain the two solutions

$$\psi(t) = \psi \, e^{\mp i Mc^2 t/\hbar} \,, \tag{10.285}$$

where the spinor amplitude ψ satisfies due to (10.283) and (10.285) the algebraic equation

$$\left(\pm\gamma^0 - \mathcal{I} \right) \psi = 0 \,. \tag{10.286}$$

Taking into account the explicit form of the Dirac matrix γ^0 in the Weyl representation (10.95), then (10.286) reduces to

$$(\gamma^0 - \mathcal{I})\psi = \left[\begin{pmatrix} O & I \\ I & O \end{pmatrix} - \begin{pmatrix} I & O \\ O & I \end{pmatrix} \right] \psi = \begin{pmatrix} -I & I \\ I & -I \end{pmatrix} \psi = 0 \,, \tag{10.287}$$

$$(-\gamma^0 - \mathcal{I})\psi = \left[\begin{pmatrix} O & -I \\ -I & O \end{pmatrix} - \begin{pmatrix} I & O \\ O & I \end{pmatrix} \right] \psi = \begin{pmatrix} -I & -I \\ -I & -I \end{pmatrix} \psi = 0 \,. \tag{10.288}$$

Assuming that $\chi(+1/2)$ and $\chi(-1/2)$ are two orthonormal bi-spinors, i.e.

$$\chi^\dagger(\lambda)\chi(\lambda') = \delta_{\lambda\lambda'}\,, \tag{10.289}$$

the two solutions of (10.287) are given by

$$\psi^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \chi(1/2) \\ \chi(1/2) \end{pmatrix}, \qquad \psi^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \chi(-1/2) \\ \chi(-1/2) \end{pmatrix}. \tag{10.290}$$

Then we construct bi-spinors $\chi^c(\pm 1/2)$, which are charge conjugated with respect to $\chi(\pm 1/2)$, by defining analogous to (10.174) and (10.184)

$$\chi^c\left(\pm\frac{1}{2}\right) = c\,\chi^*\left(\pm\frac{1}{2}\right). \tag{10.291}$$

They turn out to be orthonormal as well due to (10.191), (10.289), and (10.291):

$$\chi^{c\dagger}(\lambda)\chi^c(\lambda') = \left(\chi^{c\dagger}(\lambda)\chi^c(\lambda')\right)^T = \left(\chi^T(\lambda)\,c^\dagger c\,\chi^*(\lambda')\right)^T = \chi^\dagger(\lambda')\chi(\lambda) = \delta_{\lambda\lambda'}\,. \tag{10.292}$$

With this we obtain also the two solutions of (10.288) according to

$$\psi^{(3)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \chi^c(1/2) \\ -\chi^c(1/2) \end{pmatrix}, \qquad \psi^{(4)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \chi^c(-1/2) \\ -\chi^c(-1/2) \end{pmatrix}. \tag{10.293}$$

We note that $\psi^{(3)}$ and $\psi^{(4)}$ just represent the charge conjugated Dirac spinors of $\psi^{(1)}$ and $\psi^{(2)}$. Namely the Dirac adjoint Dirac spinors

$$\overline{\psi}^{(1,2)} = \psi^{(1,2)\dagger}\gamma^0 \tag{10.294}$$

read explicitly with (10.290)

$$\overline{\psi}^{(1,2)} = \frac{1}{\sqrt{2}}\left(\chi^\dagger\left(\pm\frac{1}{2}\right), \chi^\dagger\left(\pm\frac{1}{2}\right)\right)\begin{pmatrix} O & I \\ I & O \end{pmatrix} = \frac{1}{\sqrt{2}}\left(\chi^\dagger\left(\pm\frac{1}{2}\right), \chi^\dagger\left(\pm\frac{1}{2}\right)\right), \tag{10.295}$$

so the charge conjugation yields due to (10.174), (10.184), (10.291), (10.293) and (10.295)

$$\psi_C'^{(1,2)} = C\overline{\psi}^{(1,2)T} = \begin{pmatrix} c & O \\ O & -c \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} \chi^*(\pm 1/2) \\ \chi^*(\pm 1/2) \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} c\chi^*(\pm 1/2) \\ -c\chi^*(\pm 1/2) \end{pmatrix} = \psi^{(3,4)}\,. \tag{10.296}$$

The finding (10.296) justifies a posteriori to define the charge conjugation of bi-spinors according to (10.291).

10.14.2 Boost to Uniformly Moving Reference Frame

Now we boost the fundamental solutions (10.285), (10.290), and (10.293) of the Dirac equation in the rest frame to a uniformly moving reference frame:

$$\psi^{(1,2)}\,e^{-iMc^2t/\hbar} \quad\longrightarrow\quad \psi_{\mathbf{p}}^{(1,2)}(x) = \psi_{\mathbf{p}}^{(1,2)}\,e^{-ipx/\hbar}\,, \tag{10.297}$$

$$\psi^{(3,4)}\,e^{+iMc^2t/\hbar} \quad\longrightarrow\quad \psi_{\mathbf{p}}^{(3,4)}(x) = \psi_{\mathbf{p}}^{(3,4)}e^{+ipx/\hbar}\,, \tag{10.298}$$

where the momentum four-vector is transferred from the rest frame (10.27) to the uniformly moving reference frame (10.28). Despite of such a boost transformation a scalar product remains invariant, so the time-like component of the boosted momentum four-vector (6.16) is fixed by its spatial components according to

$$p_R^\mu p_{R\mu} = p^\mu p_\mu \quad \Longrightarrow \quad M^2c^2 = (p^0)^2 - \mathbf{p}^2 \quad \Longrightarrow \quad E_\mathbf{p} = p^0 c = \sqrt{\mathbf{p}^2c^2 + M^2c^4}\,. \tag{10.299}$$

Note that this represents precisely the relativistic energy-momentum dispersion relation (6.13). Furthermore, the corresponding spinor amplitudes $\psi_\mathbf{p}^{(\nu)}$ for $\nu = 1,2,3,4$ in the uniformly moving reference frame emerge from boosting the spinor amplitudes $\psi^{(\nu)}$ in the rest frame:

$$\psi_\mathbf{p}^{(\nu)} = D(B)\psi^{(\nu)}\,. \tag{10.300}$$

Here the boost representation in the space of the Dirac spinors from (10.13), (10.14), (10.47), (10.48), and (10.102) reads in the Weyl representation:

$$D(B) = \begin{pmatrix} D^{(1/2,0)}(B) & O \\ O & D^{(0,1/2)}(B) \end{pmatrix} = \begin{pmatrix} e^{-\boldsymbol{\sigma\xi}/2} & O \\ O & e^{\boldsymbol{\sigma\eta}/2} \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}} & O \\ O & \sqrt{\frac{p\tilde{\sigma}}{Mc}} \end{pmatrix}. \tag{10.301}$$

Note that the spinor representations for boosts (10.47) and (10.48) represent here efficient shortcut notations for the more involved concrete expressions (10.40) and (10.44). Thus, applying (10.301) to both (10.290) and (10.293) yields

$$\psi_\mathbf{p}^{(1,2)} = D(B)\psi^{(1,2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi\left(\pm\frac{1}{2}\right) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi\left(\pm\frac{1}{2}\right) \end{pmatrix}, \tag{10.302}$$

$$\psi_\mathbf{p}^{(3,4)} = D(B)\psi^{(3,4)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi^c\left(\pm\frac{1}{2}\right) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c\left(\pm\frac{1}{2}\right) \end{pmatrix}. \tag{10.303}$$

With the side calculation following from (6.21) and (10.81)

$$(p\sigma)(p\tilde{\sigma}) = p_\mu\sigma^\mu p_\nu\tilde{\sigma}^\nu = \frac{1}{2}p_\mu p_\nu(\sigma^\mu\tilde{\sigma}^\nu + \sigma^\nu\tilde{\sigma}^\mu) = p_\mu p_\nu g^{\mu\nu} I = p^2 I = (Mc)^2 I \tag{10.304}$$

we see then explicitly that we have thus constructed solutions of the Dirac equation (10.281). At first we conclude from (10.297)

$$\left(i\gamma^\mu\partial_\mu - \frac{Mc}{\hbar}\right)\psi_\mathbf{p}^{(1,2)}(x) = 0 \qquad \Longrightarrow \qquad (\gamma^\mu p_\mu - Mc)\,\psi_\mathbf{p}^{(1,2)} = 0\,. \tag{10.305}$$

From (10.95), (10.302), and (10.304) follows then indeed:

$$\begin{aligned} &\begin{pmatrix} O & p\sigma \\ p\tilde{\sigma} & O \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi(\pm\frac{1}{2}) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi(\pm\frac{1}{2}) \end{pmatrix} = \frac{Mc}{\sqrt{2}}\begin{pmatrix} \frac{p\sigma}{Mc}\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi(\pm\frac{1}{2}) \\ \frac{p\tilde{\sigma}}{Mc}\sqrt{\frac{p\sigma}{Mc}}\,\chi(\pm\frac{1}{2}) \end{pmatrix} \\ &= \frac{Mc}{\sqrt{2}}\begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\sqrt{\frac{(p\sigma)(p\tilde{\sigma})}{(Mc)^2}}\,\chi(\pm\frac{1}{2}) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\sqrt{\frac{(p\tilde{\sigma})(p\sigma)}{(Mc)^2}}\,\chi(\pm\frac{1}{2}) \end{pmatrix} = \begin{pmatrix} Mc\,I & O \\ O & Mc\,I \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi(\pm\frac{1}{2}) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi(\pm\frac{1}{2}) \end{pmatrix}. \end{aligned} \tag{10.306}$$

In a similar way we read off from (10.298)

$$\left(i\gamma^\mu\partial_\mu - \frac{Mc}{\hbar}\right)\psi_{\mathbf{p}}^{(3,4)}(x) = 0 \qquad \Longrightarrow \qquad \left(\gamma^\mu p_\mu + Mc\right)\psi_{\mathbf{p}}^{(3,4)} = 0\,. \tag{10.307}$$

And from (10.95), (10.303), and (10.304) we get then indeed:

$$\begin{pmatrix} O & p\sigma \\ p\tilde{\sigma} & O \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi^c(\pm\frac{1}{2}) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c(\pm\frac{1}{2}) \end{pmatrix} = \frac{Mc}{\sqrt{2}} \begin{pmatrix} -\frac{p\sigma}{Mc}\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c(\pm\frac{1}{2}) \\ +\frac{p\tilde{\sigma}}{Mc}\sqrt{\frac{p\sigma}{Mc}}\,\chi^c(\pm\frac{1}{2}) \end{pmatrix} \tag{10.308}$$

$$= \frac{Mc}{\sqrt{2}} \begin{pmatrix} -\sqrt{\frac{p\sigma}{Mc}}\sqrt{\frac{(p\sigma)(p\tilde{\sigma})}{(Mc)^2}}\,\chi^c(\pm\frac{1}{2}) \\ +\sqrt{\frac{p\tilde{\sigma}}{Mc}}\sqrt{\frac{(p\tilde{\sigma})(p\sigma)}{(Mc)^2}}\,\chi^c(\pm\frac{1}{2}) \end{pmatrix} = -\begin{pmatrix} Mc\,I & O \\ O & Mc\,I \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi^c(\pm\frac{1}{2}) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c(\pm\frac{1}{2}) \end{pmatrix}.$$

We note that $\psi_{\mathbf{p}}^{(3)}$ and $\psi_{\mathbf{p}}^{(4)}$ just represent the charge-conjugate of the Dirac spinors $\psi_{\mathbf{p}}^{(1)}$ and $\psi_{\mathbf{p}}^{(2)}$. At first, we determine the Dirac adjoint Dirac spinors

$$\begin{aligned} \overline{\psi}_{\mathbf{p}}^{(1,2)} &= \psi_{\mathbf{p}}^{(1,2)\dagger}\gamma^0 = \frac{1}{\sqrt{2}}\left(\chi^\dagger\left(\pm\frac{1}{2}\right)\sqrt{\frac{p\sigma}{Mc}}, \chi^\dagger\left(\pm\frac{1}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}\right)\begin{pmatrix} O & I \\ I & O \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}\left(\chi^\dagger\left(\pm\frac{1}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}, \chi^\dagger\left(\pm\frac{1}{2}\right)\sqrt{\frac{p\sigma}{Mc}}\right). \end{aligned} \tag{10.309}$$

In addition, we summarize (10.187), (10.188) and conclude from (10.191)

$$c^{-1}\sigma^\mu c = (\tilde{\sigma}^\mu)^T \qquad \Longrightarrow \qquad c(\sigma^\mu)^T c^{-1} = \tilde{\sigma}^\mu\,, \qquad c(\tilde{\sigma}^\mu)^T c^{-1} = \sigma^\mu\,. \tag{10.310}$$

The latter two relations can be generalized to any function of Pauli matrices $f(\sigma^\mu)$ or $f(\tilde{\sigma}^\mu)$, which has a Taylor series:

$$cf(\sigma^\mu)^T c^{-1} = f(\tilde{\sigma}^\mu)\,, \qquad cf(\tilde{\sigma}^\mu)^T c^{-1} = f(\sigma^\mu)\,. \tag{10.311}$$

The charge conjugation of the Dirac spinors $\psi_{\mathbf{p}}^{(1)}$ and $\psi_{\mathbf{p}}^{(2)}$ then leads to due to (10.174), (10.184), (10.289), (10.309), and (10.311):

$$\begin{aligned} \psi_C^{(1,2)} &= C\overline{\psi}^{(1,2)T} = \begin{pmatrix} c & O \\ O & -c \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\tilde{\sigma}}{Mc}}^T \chi^*(\pm\frac{1}{2}) \\ \sqrt{\frac{p\sigma}{Mc}}^T \chi^*(\pm\frac{1}{2}) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} c\sqrt{\frac{p\tilde{\sigma}}{Mc}}^T c^{-1}\,c\chi^*(\pm\frac{1}{2}) \\ -c\sqrt{\frac{p\sigma}{Mc}}^T c^{-1}\,c\,\chi^*(\pm\frac{1}{2}) \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi^c(\pm\frac{1}{2}) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c(\pm\frac{1}{2}) \end{pmatrix} = \psi_{\mathbf{p}}^{(3,4)}\,. \end{aligned} \tag{10.312}$$

The spinor amplitudes (10.302) and (10.303) can now be written as

$$\psi_{\mathbf{p}}^{(\nu)} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{M}}\,\chi\left(\frac{(-1)^{\nu+1}}{2}\right) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi\left(\frac{(-1)^{\nu+1}}{2}\right) \end{pmatrix}; \qquad \nu = 1, 2\,, \tag{10.313}$$

$$\psi_{\mathbf{p}}^{(\nu)} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi^c\left(\frac{(-1)^{\nu+1}}{2}\right) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c\left(\frac{(-1)^{\nu+1}}{2}\right) \end{pmatrix}; \qquad \nu = 3, 4\,. \tag{10.314}$$

10.14.3 Orthonormality Relations

After having obtained the plane wave solutions, we embark now upon determining their respective orthonormality relations. To this end we start with mentioning the adjoint of the spinor amplitudes (10.313):

$$\psi_{\mathbf{p}}^{(\nu)\dagger} = \frac{1}{\sqrt{2}}\left(\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}, \chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}\right); \quad \nu = 1,2\,, \tag{10.315}$$

$$\psi_{\mathbf{p}}^{(\nu)\dagger} = \frac{1}{\sqrt{2}}\left(\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}, -\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}\right); \quad \nu = 3,4\,. \tag{10.316}$$

Furthermore, we remark that the spinor amplitudes $\psi_{\mathbf{p}}^{(\nu)}$ for $\nu = 1,2$ and $\psi_{-\mathbf{p}}^{(\nu)}$ for $\nu = 3,4$ satisfy the following orthonormality relations:

1. case: $\nu = 1,2$; $\nu' = 1,2$:

$$\psi_{\mathbf{p}}^{(\nu)\dagger}\psi_{\mathbf{p}}^{(\nu')} = \frac{1}{2}\left(\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}, \chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}\right)\begin{pmatrix}\sqrt{\frac{p\sigma}{Mc}}\,\chi(\frac{(-1)^{\nu'+1}}{2})\\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi(\frac{(-1)^{\nu'+1}}{2})\end{pmatrix} \tag{10.317}$$

$$= \chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\frac{p\sigma+p\tilde{\sigma}}{2Mc}\,\chi\left(\frac{(-1)^{\nu'+1}}{2}\right) = \frac{E_{\mathbf{p}}}{Mc^2}\,\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\chi\left(\frac{(-1)^{\nu'+1}}{2}\right) = \frac{E_{\mathbf{p}}}{Mc^2}\,\delta_{\nu\nu'}\,,$$

2. case: $\nu = 3,4$; $\nu' = 3,4$:

$$\psi_{-\mathbf{p}}^{(\nu)\dagger}\psi_{-\mathbf{p}}^{(\nu')} = \frac{1}{2}\left(\chi^c\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}, -\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}\right)\begin{pmatrix}\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c(\frac{(-1)^{\nu'+1}}{2})\\ -\sqrt{\frac{p\sigma}{Mc}}\,\chi^c(\frac{(-1)^{\nu'+1}}{2})\end{pmatrix} \tag{10.318}$$

$$= \chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\frac{p\tilde{\sigma}+p\sigma}{2Mc}\,\chi^c\left(\frac{(-1)^{\nu'+1}}{2}\right) = \frac{E_{\mathbf{p}}}{Mc^2}\,\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\chi^c\left(\frac{(-1)^{\nu'+1}}{2}\right) = \frac{E_{\mathbf{p}}}{Mc^2}\,\delta_{\nu\nu'}\,,$$

3. case: $\nu = 1,2$; $\nu' = 3,4$:

$$\begin{aligned}\psi_{\mathbf{p}}^{(\nu)\dagger}\psi_{-\mathbf{p}}^{(\nu')} &= \frac{1}{2}\left(\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}, \chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}\right)\begin{pmatrix}\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi^c(\frac{(-1)^{\nu'+1}}{2})\\ -\sqrt{\frac{p\sigma}{Mc}}\,\chi^c(\frac{(-1)^{\nu'+1}}{2})\end{pmatrix}\\ &= \frac{1}{2}\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\left(\sqrt{\frac{(p\sigma)(p\tilde{\sigma})}{(Mc)^2}} - \sqrt{\frac{(p\tilde{\sigma})(p\sigma)}{(Mc)^2}}\right)\chi^c\left(\frac{(-1)^{\nu'+1}}{2}\right) = 0\,,\end{aligned} \tag{10.319}$$

4. case $\nu = 3,4$; $\nu' = 1,2$:

$$\begin{aligned}\psi_{-\mathbf{p}}^{(\nu)\dagger}\psi_{\mathbf{p}}^{(\nu')} &= \frac{1}{2}\left(\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}, -\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}\right)\begin{pmatrix}\sqrt{\frac{p\sigma}{Mc}}\,\chi(\frac{(-1)^{\nu'+1}}{2})\\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi(\frac{(-1)^{\nu'+1}}{2})\end{pmatrix}\\ &= \frac{1}{2}\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\left(\sqrt{\frac{(p\tilde{\sigma})(p\sigma)}{(Mc)^2}} - \sqrt{\frac{(p\sigma)(p\tilde{\sigma})}{(Mc)^2}}\right)\chi\left(\frac{(-1)^{\nu'+1}}{2}\right) = 0\,.\end{aligned} \tag{10.320}$$

The orthonormality relations (10.317)–(10.320) can be summarized as follows

$$\psi^{(\nu)\dagger}_{\varepsilon_\nu \mathbf{p}}\,\psi^{(\nu')}_{\varepsilon_{\nu'}\mathbf{p}} = \frac{E_{\mathbf{p}}}{Mc^2}\,\delta_{\nu\nu'}\,, \tag{10.321}$$

where have introduced the abbreviation

$$\varepsilon_\nu = \begin{cases} +1\,; & \nu = 1,2 \\ -1\,; & \nu = 3,4 \end{cases} \,. \tag{10.322}$$

With this, we can check whether the fundamental solutions

$$\psi^{(\nu)}_{\mathbf{p}}(\mathbf{x},t) = \psi^{(\nu)}_{\mathbf{p}} e^{-\frac{i}{\hbar}\varepsilon_\nu(E_{\mathbf{p}}t - \mathbf{p}\mathbf{x})} \tag{10.323}$$

fulfill orthonormality relations. Taking into account (10.299) and (10.321) we obtain from (10.323)

$$\int d^3x\,\psi^{(\nu)\dagger}_{\mathbf{p}}(\mathbf{x},t)\psi^{(\nu')}_{\mathbf{p}'}(\mathbf{x},t) = \psi^{(\nu)\dagger}_{\mathbf{p}}\,\psi^{(\nu')}_{\mathbf{p}'}\,e^{\frac{i}{\hbar}(\varepsilon_{\nu'}E_{\mathbf{p}'} - \varepsilon_\nu E_{\mathbf{p}})t}(2\pi\hbar)^3\delta(\varepsilon_\nu\mathbf{p} - \varepsilon_{\nu'}\mathbf{p}') \tag{10.324}$$

$$= (2\pi\hbar)^3\psi^{(\nu)\dagger}_{\mathbf{p}}\,\psi^{(\nu')}_{\mathbf{p}\varepsilon_\nu\varepsilon_{\nu'}}\,e^{\frac{i}{\hbar}(\varepsilon_{\nu'}E_{\varepsilon_{\nu'}\varepsilon_\nu\mathbf{p}} - \varepsilon_\nu E_{\mathbf{p}})t}\delta(\mathbf{p}' - \varepsilon_\nu\varepsilon_{\nu'}\mathbf{p}) = \frac{(2\pi\hbar)^3E_{\mathbf{p}}}{Mc^2}\,\delta_{\nu\nu'}\delta(\mathbf{p}' - \mathbf{p})\,. \tag{10.325}$$

If we now replace (10.323) with

$$\psi^{(\nu)}_{\mathbf{p}}(\mathbf{x},t) = \sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}}}}\,\psi^{(\nu)}_{\mathbf{p}}\,e^{-\frac{i}{\hbar}\epsilon_\nu(E_{\mathbf{p}}t - \mathbf{p}\mathbf{x})}\,, \tag{10.326}$$

then the fundamental solutions of the Dirac equation satisfy the orthonormality relations

$$\int d^3x\,\psi^{(\nu)\dagger}_{\mathbf{p}}(\mathbf{x},t)\psi^{(\nu')}_{\mathbf{p}'}(\mathbf{x},t) = \delta_{\nu\nu'}\,\delta(\mathbf{p} - \mathbf{p}')\,. \tag{10.327}$$

10.14.4 Dirac Representation

For the sake of completeness, we finally determine the fundamental solutions (10.327) in the Dirac representation. To this end we have to calculate at first the spinor amplitudes (10.290) and (10.293) in the rest system in the Dirac representation:

$$\psi^{(1,2)}_D = S_D\psi^{(1,2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} I & I \\ -I & I\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}\chi(\pm\frac12) \\ \chi(\pm\frac12)\end{pmatrix} = \begin{pmatrix}\chi(\pm\frac12) \\ 0\end{pmatrix}, \tag{10.328}$$

$$\psi^{(3,4)}_D = S_D\psi^{(3,4)} = \frac{1}{\sqrt{2}}\begin{pmatrix} I & I \\ -I & I\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}\chi^c(\pm\frac12) \\ -\chi^c(\pm\frac12)\end{pmatrix} = \begin{pmatrix}0 \\ -\chi^c(\pm\frac12)\end{pmatrix}. \tag{10.329}$$

By boosting from the rest frame into the uniformly moving reference frame we then get

$$\psi^{(1,2)}_{\mathbf{p}D} = S_D\psi^{(1,2)}_{\mathbf{p}} = \frac{1}{\sqrt{2}}\begin{pmatrix} I & I \\ -I & I\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{\frac{p\sigma}{Mc}}\chi(\pm\frac12) \\ \sqrt{\frac{p\tilde\sigma}{Mc}}\chi(\pm\frac12)\end{pmatrix} = \frac12\begin{pmatrix}\left(\sqrt{\frac{p\sigma}{Mc}} + \sqrt{\frac{p\tilde\sigma}{Mc}}\right)\chi(\pm\frac12) \\ \left(-\sqrt{\frac{p\tilde\sigma}{Mc}} + \sqrt{\frac{p\sigma}{Mc}}\right)\chi(\pm\frac12)\end{pmatrix}$$

$$= \frac12\begin{pmatrix}\left(\frac{p\sigma + Mc}{\sqrt{2Mc(p^0+Mc)}} + \frac{p\tilde\sigma + Mc}{\sqrt{2Mc(p^0+Mc)}}\right)\chi(\pm\frac12) \\ \left(-\frac{p\sigma + Mc}{\sqrt{2mc(p^0+Mc)}} + \frac{p\tilde\sigma + Mc}{\sqrt{2Mc(p^0+Mc)}}\right)\chi(\pm\frac12)\end{pmatrix} = \begin{pmatrix}\sqrt{\frac{E_{\mathbf{p}}+Mc^2}{2Mc^2}}\,\chi(\pm\frac12) \\ \frac{\boldsymbol{\sigma}\mathbf{p}c}{\sqrt{2Mc^2(E_{\mathbf{p}}+Mc^2)}}\,\chi(\pm\frac12)\end{pmatrix} \tag{10.330}$$

and, correspondingly,

$$\psi_{\mathbf{p}D}^{(3,4)} = S_D \psi_{\mathbf{p}}^{(3,4)} = \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\chi^c(\pm\frac{1}{2}) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\chi^c(\pm\frac{1}{2}) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \left(\sqrt{\frac{p\sigma}{Mc}} - \sqrt{\frac{p\tilde{\sigma}}{Mc}}\right)\chi^c(\pm\frac{1}{2}) \\ \left(-\sqrt{\frac{p\sigma}{Mc}} - \sqrt{\frac{p\tilde{\sigma}}{Mc}}\right)\chi^c(\pm\frac{1}{2}) \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} \left(\frac{p\sigma+Mc}{\sqrt{2Mc(p^0+Mc)}} - \frac{p\tilde{\sigma}+Mc}{\sqrt{2Mc(p^0+Mc)}}\right)\chi^c(\pm\frac{1}{2}) \\ \left(-\frac{p\sigma+Mc}{\sqrt{2Mc(p^0+Mc)}} - \frac{p\tilde{\sigma}+Mc}{\sqrt{2Mc(p^0+Mc)}}\right)\chi^c(\pm\frac{1}{2}) \end{pmatrix} = \begin{pmatrix} \frac{-\boldsymbol{\sigma}\mathbf{p}c}{\sqrt{2Mc^2(E_{\mathbf{p}}+Mc^2)}}\chi^c(\pm\frac{1}{2}) \\ -\sqrt{\frac{E_{\mathbf{p}}+Mc^2}{2M^2}}\chi^c(\pm\frac{1}{2}) \end{pmatrix} . \tag{10.331}$$

Note that the results (10.330) and (10.331) are obtained in the exercises in a different way by invoking a Foldy-Wouthuysen transformation. Furthermore, we recognize in the non-relativistic limes $c \to \infty$ that the lower or upper components of the Dirac spinor are small at $\psi_{\mathbf{p}D}^{(1,2)}$ or $\psi_{\mathbf{p}D}^{(3,4)}$ in (10.330) or (10.331), respectively.

10.15 Helicity Spinors

In the considerations of the previous section, the two orthonormal bi-spinors $\chi(+1/2)$ and $\chi(-1/2)$ have not yet been specified. It is now time to catch up with this deficiency and to make a particular choice for those orthonormal bi-spinors. In the following we introduce even two possible choices, which depend on the quantization axis for the spin 1/2.

10.15.1 Rest Frame

At first, we consider spin 1/2 particles in the rest frame, where the spin is quantized with respect to the z-axis. In this case we define the orthonormal bi-spinors according to

$$\chi\left(+\frac{1}{2}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi\left(-\frac{1}{2}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{10.332}$$

as they represent the orthonormal eigenvectors of the generator $D(L_3) = \sigma^3/2$ for a rotation around the z-axis:

$$\frac{1}{2}\sigma^3 \chi\left(\pm\frac{1}{2}\right) = \pm\frac{1}{2}\chi\left(\pm\frac{1}{2}\right) . \tag{10.333}$$

From (10.190), (10.291), and (10.332) we then get the explicit form of the charge conjugated bi-spinors:

$$\chi^c\left(+\frac{1}{2}\right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{10.334}$$

$$\chi^c\left(-\frac{1}{2}\right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} . \tag{10.335}$$

Accordingly, the charge conjugated bi-spinors satisfy the eigenvalue problem

$$\frac{1}{2}\sigma^3\,\chi^c\left(\pm\frac{1}{2}\right) = \mp\frac{1}{2}\,\chi^c\left(\pm\frac{1}{2}\right)\,. \tag{10.336}$$

A comparison of (10.333) with (10.336) shows that the eigenvalues of $\chi(\pm 1/2)$ and $\chi^c(\pm 1/2)$ are just exchanged.

The Dirac spinors (10.290) and (10.293) formed with the bi-spinors $\chi(\pm 1/2)$ and $\chi^c(\pm 1/2)$ in the rest system of the particle turn out to represent eigenvectors of the generator $D(L_3)$ of the rotation about the z-axis:

$$D(L_3)\psi^{(\nu)} = \frac{(-1)^{\nu+1}}{2}\,\psi^{(\nu)}\,; \quad \nu = 1,2\,, \qquad D(L_3)\psi^{(\nu)} = \frac{(-1)^{\nu}}{2}\,\psi^{(\nu)}\,; \quad \nu = 3,4\,. \tag{10.337}$$

Namely, taking into account (10.122), the following holds:

$$\begin{pmatrix} \frac{1}{2}\sigma^3 & 0 \\ 0 & \frac{1}{2}\sigma^3 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \chi(\pm\frac{1}{2}) \\ \chi(\pm\frac{1}{2}) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{2}\sigma^3\,\chi(\pm\frac{1}{2}) \\ \frac{1}{2}\sigma^3\,\chi(\pm\frac{1}{2}) \end{pmatrix} = \pm\frac{1}{2}\,\frac{1}{\sqrt{2}} \begin{pmatrix} \chi(\pm\frac{1}{2}) \\ \chi(\pm\frac{1}{2}) \end{pmatrix}, \tag{10.338}$$

$$\begin{pmatrix} \frac{1}{2}\sigma^3 & 0 \\ 0 & \frac{1}{2}\sigma^3 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \chi^c(\pm\frac{1}{2}) \\ -\chi^c(\pm\frac{1}{2}) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{2}\sigma^3\,\chi^c(\pm\frac{1}{2}) \\ -\frac{1}{2}\sigma^3\,\chi^c(\pm\frac{1}{2}) \end{pmatrix} = \mp\frac{1}{2}\,\frac{1}{\sqrt{2}} \begin{pmatrix} \chi^c(\pm\frac{1}{2}) \\ -\chi^c(\pm\frac{1}{2}) \end{pmatrix}. \tag{10.339}$$

10.15.2 Helicity Operator

In the following we embark on considering spin 1/2 particles, whose spin is quantized with respect to the direction of the respective particle momentum $\mathbf{p}$. To this end we construct the corresponding helicity spinors analogous to Section 9.10, where the polarisation vectors of circularly polarised plane waves were determined in the realm of electrodynamics.

To this end we determine at first the helicity operator (6.188) in the space of bi-spinors, where the spin vector is given by $D(\mathbf{L}) = \boldsymbol{\sigma}/2$ due to (10.122):

$$h(\mathbf{p}) = \frac{D(\mathbf{L})\mathbf{p}}{p} = \frac{\boldsymbol{\sigma}\mathbf{p}}{2p}\,. \tag{10.340}$$

Taking into account the explicit form of the Pauli matrices (10.8) this yields

$$h(\mathbf{p}) = \frac{1}{2p}\left\{p_x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + p_y\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + p_z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right\} = \frac{1}{2p}\begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} \tag{10.341}$$

Now we define the helicity spinors $\chi_h(\mathbf{p}, \pm 1/2)$ as eigenvectors of the helicity operator (10.340) with the eigenvalues $\pm 1/2$:

$$h(\mathbf{p})\;\chi_h\left(\mathbf{p}, \pm\frac{1}{2}\right) = \pm\frac{1}{2}\,\chi_h\left(\mathbf{p}, \pm\frac{1}{2}\right)\,. \tag{10.342}$$

From (10.332) and (10.340) follows then that the bi-spinors $\chi(\pm 1/2)$ are eigenvectors of the helicity operator $h(p\,\mathbf{e}_z)$ to the eigenvalue $\pm 1/2$:

$$h(p\,\mathbf{e}_z)\chi\left(+\frac{1}{2}\right)=\frac{1}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\frac{1}{2}\chi\left(+\frac{1}{2}\right), \tag{10.343}$$

$$h(p\,\mathbf{e}_z)\chi\left(-\frac{1}{2}\right)=\frac{1}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=-\frac{1}{2}\chi\left(-\frac{1}{2}\right). \tag{10.344}$$

Thus, due to (10.342), we then conclude

$$\chi_h\left(p\,\mathbf{e}_z,\pm\frac{1}{2}\right)=\chi\left(\pm\frac{1}{2}\right). \tag{10.345}$$

10.15.3 Uniformly Moving Rest Frame

Now we consider a uniformly moving rest frame, where the spin is quantized with respect to the momentum vector, where $\mathbf{p}$ is described with the help of spherical coordinates p, $\phi\in[0,2\pi)$, $\theta\in[0,\pi)$:

$$\mathbf{p}=p\begin{pmatrix}\sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta\end{pmatrix}. \tag{10.346}$$

Then we know that the rotation matrix (9.129) determined in (9.132) yields (10.346) analogous to (9.133):

$$R(\theta,\phi)p\mathbf{e}_z=\mathbf{p}\,. \tag{10.347}$$

Therefore, we determine the rotation matrix $D(R(\theta,\phi)$ in the space of bi-spinors, where first the rotation $D(R_y(\theta))$ and then the rotation $D(R_z(\phi))$ is performed:

$$D(R(\theta,\phi))=D(R_z(\phi))\;D(R_y(\theta))\,. \tag{10.348}$$

The individual rotation matrices follow from (10.8), (10.10), (10.11), and (10.21):

$$\begin{aligned}D(R_z(\phi))&=e^{-iD(L_3)\phi}=e^{-\frac{i}{2}\sigma^3\phi}=\cos\left(\frac{\phi}{2}\right)I-i\sin\left(\frac{\phi}{2}\right)\sigma^3\\&=\cos\left(\frac{\phi}{2}\right)\begin{pmatrix}1&0\\0&1\end{pmatrix}-i\sin\left(\frac{\phi}{2}\right)\begin{pmatrix}1&0\\0&-1\end{pmatrix}=\begin{pmatrix}e^{-i\phi/2}&0\\0&e^{i\phi/2}\end{pmatrix},\end{aligned} \tag{10.349}$$

$$\begin{aligned}D(R_y(\theta))&=e^{-iD(L_2)\theta}=e^{-\frac{i}{2}\sigma^2\theta}=\cos\left(\frac{\theta}{2}\right)I-i\sin\left(\frac{\theta}{2}\right)\sigma^2\\&=\cos\left(\frac{\theta}{2}\right)\begin{pmatrix}1&0\\0&1\end{pmatrix}-i\sin\left(\frac{\theta}{2}\right)\begin{pmatrix}0&-i\\i&0\end{pmatrix}=\begin{pmatrix}\cos\left(\frac{\theta}{2}\right)&-\sin\left(\frac{\theta}{2}\right)\\\sin\left(\frac{\theta}{2}\right)&\cos\left(\frac{\theta}{2}\right)\end{pmatrix}.\end{aligned} \tag{10.350}$$

Thus, the resulting rotation matrix (10.348) is given by:

$$D(R(\theta,\phi)) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} & -\sin\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \\ \sin\left(\frac{\theta}{2}\right)\, e^{i\phi/2} & \cos\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \end{pmatrix} . \tag{10.351}$$

Now we map the bi-spinors $\chi(\pm 1/2)$, which describe a quantization of the spin 1/2 with respect to the z-axis, with the rotation matrix $D(R(\theta,\phi)$ and obtain the helicity bi-spinors, which describe a spin quantization with respect to the direction of the momentum $\mathbf{p}$:

$$\chi_h\left(\mathbf{p}, \pm\frac{1}{2}\right) = D(R(\theta,\phi))\,\chi\left(\pm\frac{1}{2}\right) . \tag{10.352}$$

With the explicit form of the dual spinors (10.332) and the rotation matrix (10.351), the helicity dual spinors are then:

$$\chi_h\left(\mathbf{p}, +\frac{1}{2}\right) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \\ \sin\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \end{pmatrix} , \qquad \chi_h\left(\mathbf{p}, -\frac{1}{2}\right) = \begin{pmatrix} -\sin\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \\ \cos\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \end{pmatrix} . \tag{10.353}$$

In special case $\mathbf{p} = p\,\mathbf{e}_z$, i.e. $\theta = \phi = 0$, the result (10.353) reduces to (10.332) according to (10.345). Furthermore, the charge conjugation of the helicity spinors $\chi_h(\mathbf{p}, \pm 1/2)$ leads to:

$$\chi_h^c\left(\mathbf{p}, +\frac{1}{2}\right) = c\,\chi_h^*\left(\mathbf{p}, +\frac{1}{2}\right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \\ \sin\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \end{pmatrix} = \begin{pmatrix} -\sin\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \\ \cos\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \end{pmatrix} , \tag{10.354}$$

$$\chi_h^c\left(\mathbf{p}, -\frac{1}{2}\right) = c\,\chi_h^*\left(\mathbf{p}, -\frac{1}{2}\right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -\sin\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \\ \cos\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \end{pmatrix} = \begin{pmatrix} -\cos\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \\ -\sin\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \end{pmatrix} . \tag{10.355}$$

In case $\mathbf{p} = p\,\mathbf{e}_z$, i.e. $\theta = \phi = 0$, (10.354) and (10.355) reduce to (10.332):

$$\chi_h^c\left(p\,\mathbf{e}_z, \pm\frac{1}{2}\right) = \chi^c\left(\pm\frac{1}{2}\right) . \tag{10.356}$$

Furthermore, we remark that the mapping of the charge conjugated bi-spinors (10.334) with the rotation matrix (10.351) leads to the charge conjugated helicity spinors (10.354) and (10.355):

$$\chi_h^c\left(\mathbf{p}, \pm\frac{1}{2}\right) = D(R(\theta,\phi))\,\chi^c\left(\pm\frac{1}{2}\right) . \tag{10.357}$$

As a crosscheck we also verify that the constructed helicity spinors $\chi_h(\mathbf{p}, \pm 1/2)$ are, indeed, eigenvectors of the helicity operator $h(\mathbf{p})$ from (10.341) with the eigenvalue $\pm\frac{1}{2}$:

$$\begin{aligned} h(\mathbf{p})\chi_h\left(\mathbf{p}, +\frac{1}{2}\right) &= \frac{1}{2}\begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix} \begin{pmatrix} \cos\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \\ \sin\left(\frac{\theta}{2}\right)\, e^{i\phi/2} \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} \cos\left(\frac{\theta}{2}\right) e^{-i\phi/2} \\ \sin\left(\frac{\theta}{2}\right) e^{i\phi/2} \end{pmatrix} = \frac{1}{2}\,\chi_h\left(\mathbf{p}, +\frac{1}{2}\right) , \end{aligned} \tag{10.358}$$

$$\begin{aligned} h(\mathbf{p})\chi_h\left(\mathbf{p}, -\frac{1}{2}\right) &= \frac{1}{2}\begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix} \begin{pmatrix} -\sin\left(\frac{\theta}{2}\right)\, e^{-i\phi/2} \\ \cos\left(\frac{\theta}{2}\right) e^{-i\phi/2} \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} \sin\left(\frac{\theta}{2}\right) e^{-i\phi/2} \\ -\cos\left(\frac{\theta}{2}\right) e^{i\phi/2} \end{pmatrix} = -\frac{1}{2}\,\chi_h\left(\mathbf{p}, -\frac{1}{2}\right) . \end{aligned} \tag{10.359}$$

Furthermore, we show that the constructed charge conjugated helicity spinors $\chi_h^c(\mathbf{p}, \pm 1/2)$ are eigenvectors of the helicity operator $h(\mathbf{p})$ with the eigenvalue $\mp 1/2$:

$$\begin{aligned} h(\mathbf{p})\chi_h^c\left(\mathbf{p}, +\frac{1}{2}\right) &= \frac{1}{2}\begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix}\begin{pmatrix} -\sin\left(\frac{\theta}{2}\right) e^{-i\phi/2} \\ \cos\left(\frac{\theta}{2}\right) e^{i\phi/2} \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} \sin\left(\frac{\theta}{2}\right) e^{-i\phi/2} \\ -\cos\left(\frac{\theta}{2}\right) e^{i\phi/2} \end{pmatrix} = -\frac{1}{2}\chi_h^c\left(\mathbf{p}, +\frac{1}{2}\right), \end{aligned} \tag{10.360}$$

$$\begin{aligned} h(\mathbf{p})\chi_h^c\left(\mathbf{p}, -\frac{1}{2}\right) &= \frac{1}{2}\begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}\begin{pmatrix} -\cos\left(\frac{\theta}{2}\right) e^{-i\phi/2} \\ -\sin\left(\frac{\theta}{2}\right) e^{i\phi/2} \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} -\cos\left(\frac{\theta}{2}\right) e^{-i\phi/2} \\ -\sin\left(\frac{\theta}{2}\right) e^{+i\phi/2} \end{pmatrix} = \frac{1}{2}\,\chi_h^c\left(\mathbf{p}, -\frac{1}{2}\right). \end{aligned} \tag{10.361}$$

Now we come back to the Dirac spinors (10.302) and (10.303) in the uniformly moving reference frame, where $\chi(\pm 1/2)$ and $\chi^c(\pm 1/2)$ denoted two sets of orthonormal bi-spinors, which are charge conjugated with respect to each other. Whereas we have discussed in the two previous subsections the case of choosing the z-axis as the quantization axis, we come now to another appropriate physical choice to identify $\chi(\pm 1/2)$ and $\chi^c(\pm 1/2)$. Namely we assume that the spin is quantized with respect to the direction of motion $\mathbf{p}/p$, which amounts to identifying $\chi(\pm 1/2)$ and $\chi^c(\pm 1/2)$ with the helicity spinors $\chi_h(\mathbf{p}, \pm 1/2)$ and $\chi_h^c(\mathbf{p}, \pm 1/2)$, respectively, yielding

$$\psi_{\mathbf{p}}^{(1,2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi_h\left(\mathbf{p}, \pm\frac{1}{2}\right) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi_h\left(\mathbf{p}, \pm\frac{1}{2}\right) \end{pmatrix}, \tag{10.362}$$

$$\psi_{\mathbf{p}}^{(3,4)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi_h^c\left(\mathbf{p}, \pm\frac{1}{2}\right) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi_h^c\left(\mathbf{p}, \pm\frac{1}{2}\right) \end{pmatrix}. \tag{10.363}$$

In order to justify this choice we define the helicity operator in the space of Dirac spinors due to (6.188) and (10.122):

$$H(\mathbf{p}) = \frac{D(\mathbf{L})\mathbf{p}}{p} = \frac{1}{2p}\begin{pmatrix} \boldsymbol{\sigma}\mathbf{p} & O \\ O & \boldsymbol{\sigma}\mathbf{p} \end{pmatrix} = \begin{pmatrix} h(\mathbf{p}) & O \\ O & h(\mathbf{p}) \end{pmatrix}. \tag{10.364}$$

According to (10.37), (10.46), and (10.340) as well as the Lie algebra of the Pauli matrices (10.5), the helicity operator $h(\mathbf{p})$ in the space of bi-spinors commutates with the boost representation in the space of bi-spinors:

$$\left[\sqrt{\frac{p\sigma}{Mc}}, h(\mathbf{p})\right]_- = \left[\sqrt{\frac{p\tilde{\sigma}}{Mc}}, h(\mathbf{p})\right]_- = 0\,. \tag{10.365}$$

Therefore, the Dirac spinors (10.362) and (10.363) are eigenstates of the helicity operator

$$H(\mathbf{p})\psi_{\mathbf{p}}^{(\nu)} = \eta_\nu\,\psi_{\mathbf{p}}^{(\nu)} \tag{10.366}$$

with the eigenvalues

$$\eta_\nu = \frac{(-1)^{\nu+1}}{2}, \quad \nu = 1, 2; \qquad \eta_\nu = \frac{(-1)^{\nu}}{2}, \quad \nu = 3, 4\,. \tag{10.367}$$

In detail, due to (10.285), (10.286), (10.341), (10.364), and (10.365) the following applies for $\nu = 1, 2$

$$H(\mathbf{p})\psi_{\mathbf{p}}^{(\nu)} = \begin{pmatrix} h(\mathbf{p}) & O \\ O & h(\mathbf{p}) \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi_h(\mathbf{p}, \pm\frac{1}{2}) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi_h(\mathbf{p}, \pm\frac{1}{2}) \end{pmatrix} = \frac{\pm 1}{2\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi_h(\mathbf{p}, \pm\frac{1}{2}) \\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi_h(\mathbf{p}, \pm\frac{1}{2}) \end{pmatrix} = \pm\frac{1}{2}\,\psi_{\mathbf{p}}^{(\nu)}\,,$$

and, correspondingly, we have for $\nu = 3, 4$

$$H(\mathbf{p})\psi_{\mathbf{p}}^{(\nu)} = \begin{pmatrix} h(\mathbf{p}) & O \\ O & h(\mathbf{p}) \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi_h^c(\mathbf{p}, \pm\frac{1}{2}) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi_h^c(\mathbf{p}, \pm\frac{1}{2}) \end{pmatrix} = \frac{\mp 1}{2\sqrt{2}} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}}\,\chi_h^c(\mathbf{p}, \pm\frac{1}{2}) \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}}\,\chi_h^c(\mathbf{p}, \pm\frac{1}{2}) \end{pmatrix} = \mp\frac{1}{2}\,\psi_{\mathbf{p}}^{(\nu)}\,.$$

Thus, in conclusion, we have determined in a group theoretically inspired approach the plane wave solutions of the Dirac equation (10.326), where the corresponding Dirac spinor amplitudes are given by (10.362) and (10.363). This result will turn out to be indispensable for the subsequent canonical field quantization of the Dirac theory.

10.16 Canonical Field Quantisation

In order to determine the Hamiltonian formulation of classical field theory from the Lagrangian formulation, one has to find at first the momentum fields, which are canonically conjugated to the independent field degrees of freedom. In case of the Dirac field, the canonically conjugated momentum fields are obtained for the Dirac spinor $\psi(\mathbf{x}, t)$ and the Dirac adjoint Dirac spinor $\overline{\psi}(\mathbf{x}, t)$, respectively:

$$\pi(\mathbf{x}, t) = \frac{\delta \mathcal{A}}{\delta\left(\frac{\partial\psi(\mathbf{x},t)}{\partial t}\right)} = \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial\psi(\mathbf{x},t)}{\partial t}\right)} = i\hbar\overline{\psi}(\mathbf{x}, t)\gamma^0 = i\hbar\psi^\dagger(\mathbf{x}, t)\,, \tag{10.368}$$

$$\overline{\pi}(\mathbf{x}, t) = \frac{\delta \mathcal{A}}{\delta\left(\frac{\partial\overline{\psi}(\mathbf{x},t)}{\partial t}\right)} = \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial\overline{\psi}(\mathbf{x},t)}{\partial t}\right)} = 0\,. \tag{10.369}$$

Note that the last equality in (10.368) follows from taking into account (10.103). Thus, in the Hamiltonian formulation of the Dirac theory, one can consider $\psi(\mathbf{x}, t)$ and $\pi(\mathbf{x}, t)$ or, equivalently, $\psi(\mathbf{x}, t)$ and $\psi^\dagger(\mathbf{x}, t)$ as the independent fields.

And, according to the Noether theorem applied to the Dirac field, any conserved physical quantity of the Dirac theory turns out to be bilinear in these independent fields. Namely, due to the sandwich principle, each conserved quantity follows from a spatial integral over the respective first-quantized operator, which is multiplied with $\psi^\dagger(\mathbf{x}, t)$ from the left and $\psi(\mathbf{x}, t)$ from the right. Indeed, the charge of the Dirac field is given by (10.280) and analogous expressions also hold for the energy, the momentum, and the helicity of the Dirac field:

$$E = \int d^3x\, \psi^\dagger(\mathbf{x},t) H_D(\mathbf{x}) \psi(\mathbf{x},t)\,, \tag{10.370}$$

$$\mathbf{P} = \int d^3x\, \psi^\dagger(\mathbf{x},t)\, \frac{\hbar}{i} \boldsymbol{\nabla}\, \psi(\mathbf{x},t)\,, \tag{10.371}$$

$$h = \int d^3x\, \psi^\dagger(\mathbf{x},t) \begin{pmatrix} \boldsymbol{\sigma}/2 & O \\ O & \boldsymbol{\sigma}/2 \end{pmatrix} \frac{\hbar\boldsymbol{\nabla}/i}{|\hbar\boldsymbol{\nabla}/i|}\, \psi(\mathbf{x},t)\,. \tag{10.372}$$

Note that the Dirac Hamiltonian $H_D(\mathbf{x})$ was already defined in (10.250) and reduces due to (10.88) to

$$H_D(\mathbf{x}) = -ic\hbar\, \boldsymbol{\alpha}\boldsymbol{\nabla} + Mc^2\beta\,. \tag{10.373}$$

Furthermore, we have used in (10.372) that the helicity (6.188) stems from the generators of the rotations (10.122) in the space of Dirac spinors.

In a canonical quantization of the Dirac field the independent fields $\psi(\mathbf{x},t)$ and $\pi(\mathbf{x},t)$ or $\psi(\mathbf{x},t)$ and $\psi^\dagger(\mathbf{x},t)$ of the Hamilton field theory become field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\pi}(\mathbf{x},t)$ or $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$. Since a bosonic quantisation of the Dirac field turns out to violate microcausality and, thus, leads inevitably to contradictions, one has to perform a fermionic quantisation. Therefore, the following equal-time anti-commutator algebra is required

$$\left[\hat{\psi}_\alpha(\mathbf{x},t), \hat{\psi}_\beta(\mathbf{x}',t)\right]_+ = \left[\hat{\pi}_\alpha(\mathbf{x},t), \hat{\pi}_\beta(\mathbf{x}',t)\right]_+ = 0\,,\; \left[\hat{\psi}_\alpha(\mathbf{x},t), \hat{\pi}_\beta(\mathbf{x}',t)\right]_+ = i\hbar\delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{x}') \tag{10.374}$$

where α, β denote the spinorial components. Due to the definition of the momentum field in (10.368) the anti-commutator algebra (10.374) reduces to

$$\left[\hat{\psi}_\alpha(\mathbf{x},t), \hat{\psi}_\beta(\mathbf{x}',t)\right]_+ = \left[\hat{\psi}^\dagger_\alpha(\mathbf{x},t), \hat{\psi}^\dagger_\beta(\mathbf{x}',t)\right]_+ = 0\,,\; \left[\hat{\psi}_\alpha(\mathbf{x},t), \hat{\psi}^\dagger_\beta(\mathbf{x}',t)\right]_+ = \delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{x}')\,. \tag{10.375}$$

Thus, the conserved quantities of the first quantized Dirac theory, i.e. the charge (10.280), the energy (10.370), the momentum (10.371), and the helicity (10.372), become second quantized operators due to the canonical field quantisation:

$$\hat{Q} = \int d^3x\, \hat{\psi}^\dagger(\mathbf{x},t)\hat{\psi}(\mathbf{x},t)\,, \tag{10.376}$$

$$\hat{H} = \int d^3x\, \hat{\psi}^\dagger(\mathbf{x},t) H_D(\mathbf{x}) \hat{\psi}(\mathbf{x},t)\,, \tag{10.377}$$

$$\hat{\mathbf{P}} = \int d^3x\, \hat{\psi}^\dagger(\mathbf{x},t)\, \frac{\hbar}{i} \nabla\, \hat{\psi}(\mathbf{x},t)\,, \tag{10.378}$$

$$\hat{h} = \int d^3x\, \hat{\psi}^\dagger(\mathbf{x},t) \begin{pmatrix} \boldsymbol{\sigma}/2 & O \\ O & \boldsymbol{\sigma}/2 \end{pmatrix} \frac{\hbar\boldsymbol{\nabla}/i}{|\hbar\boldsymbol{\nabla}/i|}\, \hat{\psi}(\mathbf{x},t)\,. \tag{10.379}$$

In order to determine the Heisenberg equations of motion (3.62), one needs to take into account both the first and the second quantized Hamilton operator (10.373) and (10.378) as well as to apply the calculation rule (3.94). With this the Heisenberg equations of motion of the field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$ result in

$$i\hbar\,\frac{\partial\hat{\psi}(\mathbf{x},t)}{\partial t} = \left[\hat{\psi}(\mathbf{x},t),\hat{H}\right]_- = H_D(\mathbf{x})\hat{\psi}(\mathbf{x},t) = \left(-ic\hbar\boldsymbol{\alpha}\nabla + Mc^2\beta\right)\hat{\psi}(\mathbf{x},t)\,, \tag{10.380}$$

$$i\hbar\,\frac{\partial\hat{\psi}^\dagger(\mathbf{x},t)}{\partial t} = \left[\hat{\psi}^\dagger(\mathbf{x},t),\hat{H}\right]_- = -\left\{H_D(\mathbf{x})\hat{\psi}(\mathbf{x},t)\right\}^\dagger = \left(-ic\hbar\boldsymbol{\alpha}\nabla - Mc^2\beta\right)\hat{\psi}^\dagger(\mathbf{x},t)\,. \tag{10.381}$$

Thus, the field operators $\hat{\psi}(\mathbf{x},t)$ and $\hat{\psi}^\dagger(\mathbf{x},t)$ satisfy the Dirac equation (10.249) and the adjoint Dirac equation, respectively.

10.17 Decomposition Into Plane Waves

The field operator $\hat{\psi}(\mathbf{x},t)$ is now decomposed with respect to the fundamental plane wave solutions $\psi_{\mathbf{p}}^{(\nu)}(\mathbf{x},t)$ of the Dirac equation defined in (10.326). The expansion coefficients in this decomposition are then operators of second quantisation:

$$\hat{\psi}(\mathbf{x},t) = \sum_{\nu=1}^{4}\int d^3p\,\psi_{\mathbf{p}}^{(\nu)}(\mathbf{x},t)\hat{a}_{\mathbf{p}}^{(\nu)}\,. \tag{10.382}$$

Correspondingly, one obtains for the adjoint field operator:

$$\hat{\psi}^\dagger(\mathbf{x},t) = \sum_{\nu=1}^{4}\int d^3p\,\psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\,. \tag{10.383}$$

With the help of the orthonormality relation (10.327) of the fundamental plane wave solutions, the expansions (10.382) and (10.383) can be inverted, yielding:

$$\int d^3x\,\psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\hat{\psi}(\mathbf{x},t) = \hat{a}_{\mathbf{p}}^{(\nu)}\,, \tag{10.384}$$

$$\int d^3x\,\hat{\psi}^\dagger(\mathbf{x},t)\psi_{\mathbf{p}}^{(\nu)}(\mathbf{x},t) = \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\,. \tag{10.385}$$

From the equal-time anti-commutator algebra (10.375) of the field operator $\hat{\psi}(\mathbf{x},t)$ and its adjoint $\hat{\psi}^\dagger(\mathbf{x},t)$, a corresponding anti-commutator algebra can then be determined for the expansion coefficients $\hat{a}_{\mathbf{p}}^{(\nu)}$ and $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$:

$$\left[\hat{a}_{\mathbf{p}}^{(\nu)},\hat{a}_{\mathbf{p}'}^{(\nu')}\right]_+ = \int d^3x\int d^3x'\sum_{\alpha,\alpha'=1}^{4}\psi_{\mathbf{p},\alpha}^{(\nu)}(\mathbf{x},t)\psi_{\mathbf{p}',\alpha'}^{(\nu')}(\mathbf{x}',t)\left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\psi}_{\alpha'}(\mathbf{x}',t)\right]_+ = 0\,, \tag{10.386}$$

$$\left[\hat{a}_{\mathbf{p}}^{(\nu)\dagger},\hat{a}_{\mathbf{p}'}^{(\nu')\dagger}\right]_+ = \int d^3x\int d^3x'\sum_{\alpha,\alpha'=1}^{4}\psi_{\mathbf{p},\alpha}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}',\alpha'}^{(\nu')\dagger}(\mathbf{x}',t)\left[\hat{\psi}_\alpha^\dagger(\mathbf{x},t),\hat{\psi}_{\alpha'}^\dagger(\mathbf{x}',t)\right]_+ = 0\,, \tag{10.387}$$

$$\left[\hat{a}_{\mathbf{p}}^{(\nu)},\hat{a}_{\mathbf{p}'}^{(\nu')\dagger}\right]_+ = \int d^3x\int d^3x'\sum_{\alpha,\alpha'=1}^{4}\psi_{\mathbf{p},\alpha}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}',\alpha'}^{(\nu')}(\mathbf{x}',t)\left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\psi}_{\alpha'}^\dagger(\mathbf{x}',t)\right]_+$$

$$= \int d^3x\sum_{\alpha=1}^{4}\psi_{\mathbf{p},\alpha}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}',\alpha}^{(\nu')}(\mathbf{x},t) = \int d^3x\,\psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) = \delta_{\nu\nu'}\delta(\mathbf{p}-\mathbf{p}')\,. \tag{10.388}$$

Note that in (10.388) the orthonormality relation (10.327) is applied. As the operators $\hat{a}_{\mathbf{p}}^{(\nu)}$, $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$ fulfill according to (10.386)–(10.388) the canonical anti-commutator algebra, they are interpreted for the time being as annihilation and creation operators of fermionic particles.

10.18 Second Quantizied Operators

Inserting (10.382) and (10.383) into (10.376) and taking into account the orthonormality relation (10.327) the charge operator $\hat{Q}$ in second quantisation can be expressed in terms of the creation and annihilation operators $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$ and $\hat{a}_{\mathbf{p}}^{(\nu)}$:

$$\begin{aligned}\hat{Q} &= \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p'\, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')} \int d^3x\, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) \\ &= \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p'\, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')}\delta_{\nu\nu'}\delta(\mathbf{p}-\mathbf{p}') = \sum_{\nu=1}^{4}\int d^3p\, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)} . \end{aligned} \tag{10.389}$$

Since the particle number operator $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)}$ is positive definite, also the charge operator $\hat{Q}$ is positive definite due to (10.389). Thus, it looks like as if the fermionic particles seem to have only a positive charge.

Accordingly, inserting (10.382) and (10.383) into (10.377) one obtains for the Hamilton operator $\hat{H}$ of second quantisation at first

$$\hat{H} = \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p'\, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')} \int d^3x\, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)H_D(\mathbf{x})\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) . \tag{10.390}$$

Here we can take into account that the plane waves $\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)$ from (10.326) are eigenfunctions of the Dirac Hamiltonian operator of the first quantisation (10.373) as they were determined in Section 10.14 to solve the Dirac equation (10.281):

$$H_D(\mathbf{x})\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) = i\hbar\frac{\partial}{\partial t}\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) = \varepsilon_{\nu'}E_{\mathbf{p}'}\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) . \tag{10.391}$$

With the help of the orthonormality relation (10.327) the Hamilton operator of second quantisation (10.390) then results in

$$\begin{aligned}\hat{H} &= \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p'\, \varepsilon_{\nu'}E_{\mathbf{p}'}\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')} \int d^3x\, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) \\ &= \sum_{\nu=1}^{4}\int d^3p\, \varepsilon_\nu E_{\mathbf{p}}\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)} = \int d^3p \left(\sum_{\nu=1}^{2}E_{\mathbf{p}}\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)} - \sum_{\nu=3}^{4}E_{\mathbf{p}}\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)}\right) , \end{aligned} \tag{10.392}$$

where we have used the abbreviation (10.322) in the last step. Thus, the fermionic particles with $\nu = 1, 2$ appear to have positive energies $E_{\mathbf{p}}$, while those with $\nu = 3, 4$ seem to have correspondingly negative energies $-E_{\mathbf{p}}$.

Subsequently, we insert (10.382) and (10.383) into (10.378), so the momentum operator $\hat{\mathbf{P}}$ of second quantisation results at first in

$$\hat{\mathbf{P}} = \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p' \, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')} \int d^3x \, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\, \frac{\hbar}{i}\boldsymbol{\nabla}\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)\,. \tag{10.393}$$

Here we use the fact that the plane waves $\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)$ from (10.326) are eigenfunctions of the momentum operator of first quantisation:

$$\frac{\hbar}{i}\boldsymbol{\nabla}\,\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) = \varepsilon_{\nu'}\mathbf{p}'\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)\,. \tag{10.394}$$

Thus, with the orthonormality relation (10.327) the momentum operator of second quantisation (10.379) reduces to

$$\begin{aligned}\hat{\mathbf{P}} &= \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p' \, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')}\varepsilon_{\nu'}\mathbf{p}' \int d^3x \, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)\\ &= \sum_{\nu=1}^{4}\int d^3p \, \varepsilon_\nu \mathbf{p}\, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)} = \int d^3p \left(\sum_{\nu=1}^{2}\mathbf{p}\,\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\,\hat{a}_{\mathbf{p}}^{(\nu)} - \sum_{\nu=3}^{4}\mathbf{p}\,\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)}\right). \end{aligned}\tag{10.395}$$

We conclude that the fermionic particles with $\nu = 1,2$ seem to have the momentum $\mathbf{p}$ and, correspondingly, those with $\nu = 3,4$ the momentum $-\mathbf{p}$.

In a similar way we also proceed for the helicity operator (10.379), where we insert (10.382) and (10.383), yielding

$$\hat{h} = \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p' \, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')} \int d^3x \, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t) \begin{pmatrix} \boldsymbol{\sigma}/2 & O \\ O & \boldsymbol{\sigma}/2 \end{pmatrix} \frac{\hbar\boldsymbol{\nabla}/i}{|\hbar\boldsymbol{\nabla}/i|}\,\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)\,. \tag{10.396}$$

Applying the eigenvalue problem (10.394) and the first quantized helicity operator (10.364) this reduces to

$$\hat{h} = \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p' \, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')}\varepsilon_{\nu'} \int d^3x \, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\, H(\mathbf{p}')\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)\,. \tag{10.397}$$

Here we use the fact that the plane waves $\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)$ from (10.326) are eigenfunctions of the helicity operator of first quantisation according to (10.366):

$$H(\mathbf{p}')\,\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) = \eta_{\nu'}\,\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t)\,. \tag{10.398}$$

Thus, with this and the orthonormality relation (10.327) the helicity operator of second quantisation (10.397) reads

$$\begin{aligned}\hat{h} &= \sum_{\nu=1}^{4}\sum_{\nu'=1}^{4}\int d^3p \int d^3p' \, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}'}^{(\nu')}\varepsilon_{\nu'}\eta_{\nu'} \int d^3x \, \psi_{\mathbf{p}}^{(\nu)\dagger}(\mathbf{x},t)\psi_{\mathbf{p}'}^{(\nu')}(\mathbf{x},t) \qquad (10.399)\\ &= \sum_{\nu=1}^{4}\int d^3p \, \varepsilon_\nu\eta_\nu\, \hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)} = \int d^3p \left(\sum_{\nu=1}^{2}\frac{(-1)^{\nu+1}}{2}\,\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\,\hat{a}_{\mathbf{p}}^{(\nu)} + \sum_{\nu=3}^{4}\frac{(-1)^{\nu+1}}{2}\,\hat{a}_{\mathbf{p}}^{(\nu)\dagger}\hat{a}_{\mathbf{p}}^{(\nu)}\right). \end{aligned}$$

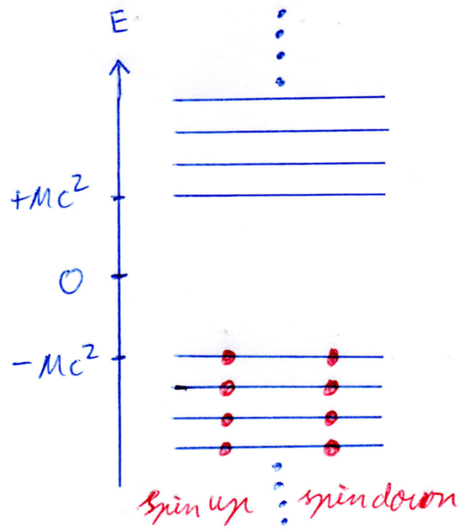


Figure 10.2: Schematic sketch of the Dirac sea, which models the physical vacuum as an infinite sea of particles with negative energy.

Note that we have used in the last step the abbreviations (10.322) and (10.367). The result (10.399) means that the fermionic particles with $\nu = 1, 3$ ($\nu = 2, 4$) have supposedly the helicity $+1/2$ ($-1/2$).

Finally, we conclude this section by summarizing that, indeed, the second quantized operators for the charge (10.389), the energy (10.392), the momentum (10.395), and the helicity (10.399) have turned out to not explicitly depend on time. This reflects that these second quantized operators correspond to conserved quantities.

10.19 Dirac Sea

Within the framework of the canonical field quantisation, the vacuum state $|0\rangle_V$ is usually defined by the fact that it does not contain any particle. This is guaranteed provided that all annihilation operators $\hat{a}_{\mathbf{p}}^{(\nu)}$ annul the vacuum state $|0\rangle_V$:

$$\hat{a}_{\mathbf{p}}^{(\nu)} \, |0\rangle_V = 0 \quad \text{for all } \nu, \mathbf{p} \, . \tag{10.400}$$

On the other, in the second quantized Dirac theory we are confronted with the fact that particles with both positive and negative energies appear, see Eq. (10.392). In order to provide a physical interpretation for the latter observation, Paul Dirac assumed in 1930 that instead of the vacuum state $|0\rangle_V$ a physical vacuum state $|0\rangle_P$ is realised in nature. It is defined by the condition that all states with negative energies, i.e. those with $\nu = 3, 4$, are occupied, forming the so-called Fermi sea, see Fig. 10.2:

$$|0\rangle_P = \prod_{\nu=3,4} \prod_{\mathbf{p}} \hat{a}_{\mathbf{p}}^{(\nu)\dagger} \, |0\rangle_V \, . \tag{10.401}$$

Here, a continuous product is formed with respect to all momenta $\mathbf{p}$. Dirac justifies this transition from the vacuum state $|0\rangle_V$ to the physical vacuum state $|0\rangle_P$ by the argument that the Dirac sea is always present and can, therefore, not be measured in any experiment, Thus, the infinitely large energy or charge of the Dirac sea can be renormalised.

An immediate consequence of the definition of the physical vacuum state $|0\rangle_P$ in (10.401) is that it is annulled by the annihilation operators $\hat{a}_{\mathbf{p}}^{(\nu)}$ for $\nu = 1, 2$ because of (10.400) and by the creation operators $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$ for $\nu = 3, 4$ due to the anti-commutator algebra (10.387):

$$\hat{a}_{\mathbf{p}}^{(\nu)} |0\rangle_P = 0 \quad \text{for all } \nu = 1, 2 \text{ and } \mathbf{p}\,; \qquad \hat{a}_{\mathbf{p}}^{(\nu)\dagger} |0\rangle_P = 0 \quad \text{for all } \nu = 3, 4 \text{ and } \mathbf{p}\,. \tag{10.402}$$

If one takes into account the anti-commutator algebra (10.386)–(10.388) and the property (10.402) of the physical vacuum $|0\rangle_P$, a reinterpretation of the annihilation and creation operators becomes possible. While $\hat{a}_{\mathbf{p}}^{(\nu)}$ and $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$ for $\nu = 1, 2$ continue to be considered as annihilation and creation operators of particles, $\hat{a}_{\mathbf{p}}^{(\nu)}$ and $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$ for $\nu = 3, 4$ can now be interpreted inversely as the creation and annihilation operators of particles. For instance, applying $\hat{a}_{\mathbf{p}}^{(\nu)}$ for $\nu = 3, 4$ to the physical vacuum state (10.401) annihilates a particle in the Dirac sea of Fig. 10.2, which corresponds to the creation of a hole.

Consequently, by convention we consider in the Dirac hole theory that the indices $\nu = 1, 2$ ($\nu = 3, 4$) describe particles (antiparticles), for instance electrons (positrons) with spin up and down. The double role of the expansion operators $\hat{a}_{\mathbf{p}}^{(\nu)}$ and $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$ as creation and annihilation operators, respectively, makes the theory at a first glance confusing. Therefore, it is suggestive to introduce different symbols in order to discriminate already from the notation between the operators of particles and antiparticles. For the particles we use from now on the following definition for the creation operators

$$\hat{a}_{\mathbf{p}}^{(1)\dagger} = \hat{b}_{\mathbf{p}}^{(1)\dagger}\,, \qquad \hat{a}_{\mathbf{p}}^{(2)\dagger} = \hat{b}_{\mathbf{p}}^{(2)\dagger} \tag{10.403}$$

and for the annihilation operators

$$\hat{a}_{\mathbf{p}}^{(1)} = \hat{b}_{\mathbf{p}}^{(1)}\,, \qquad \hat{a}_{\mathbf{p}}^{(2)} = \hat{b}_{\mathbf{p}}^{(2)}\,. \tag{10.404}$$

Correspondingly, we introduce for the antiparticles the creation operators

$$\hat{a}_{\mathbf{p}}^{(3)} = \hat{d}_{\mathbf{p}}^{(1)\dagger}\,, \qquad \hat{a}_{\mathbf{p}}^{(4)} = \hat{d}_{\mathbf{p}}^{(2)\dagger} \tag{10.405}$$

and the annihilation operators

$$\hat{a}_{\mathbf{p}}^{(3)\dagger} = \hat{d}_{\mathbf{p}}^{(1)}\,, \qquad \hat{a}_{\mathbf{p}}^{(4)\dagger} = \hat{d}_{\mathbf{p}}^{(2)}\,. \tag{10.406}$$

For $\nu = 1, 2$ this redefinition just corresponds to a simple renaming. But for $\nu = 3, 4$ the creation and annihilation operators exchange their roles. Note that the anti-commutator algebra (10.386)–(10.388) remains invariant due to this redefinition, since creation and annihilation operators appear there on equal footing:

$$\left[\hat{b}_{\mathbf{p}}^{(\nu)}, \hat{b}_{\mathbf{p}'}^{(\nu')}\right]_{+} = \left[\hat{b}_{\mathbf{p}}^{(\nu)}, \hat{d}_{\mathbf{p}'}^{(\nu')}\right]_{+} = \left[\hat{d}_{\mathbf{p}}^{(\nu)}, \hat{d}_{\mathbf{p}'}^{(\nu')}\right]_{+} = \left[\hat{b}_{\mathbf{p}}^{(\nu)}, \hat{d}_{\mathbf{p}'}^{(\nu')\dagger}\right]_{+} = 0\,, \tag{10.407}$$

$$\left[\hat{b}_{\mathbf{p}}^{(\nu)\dagger}, \hat{b}_{\mathbf{p}'}^{(\nu')\dagger}\right]_{+} = \left[\hat{b}_{\mathbf{p}}^{(\nu)\dagger}, \hat{d}_{\mathbf{p}'}^{(\nu')\dagger}\right]_{+} = \left[\hat{d}_{\mathbf{p}}^{(\nu)\dagger}, \hat{d}_{\mathbf{p}'}^{(\nu')\dagger}\right]_{+} = \left[\hat{b}_{\mathbf{p}}^{(\nu)\dagger}, \hat{d}_{\mathbf{p}'}^{(\nu')}\right]_{+} = 0\,, \tag{10.408}$$

$$\left[\hat{b}_{\mathbf{p}}^{(\nu)}, \hat{b}_{\mathbf{p}'}^{(\nu')\dagger}\right]_{+} = \left[\hat{d}_{\mathbf{p}}^{(\nu)}, \hat{d}_{\mathbf{p}'}^{(\nu')\dagger}\right]_{+} = \delta_{\nu\nu'}\delta(\mathbf{p}-\mathbf{p}')\,. \tag{10.409}$$

The physical vacuum state $|0\rangle_P$ is now determined by the fact that it is annulled by the annihilation operators $\hat{b}_{\mathbf{p}}^{(\nu)}$, $\hat{d}_{\mathbf{p}}^{(\nu)}$ of both the particles and the antiparticles:

$$\hat{b}_{\mathbf{p}}^{(\nu)}\,|0\rangle_P = 0\,, \tag{10.410}$$

$$\hat{d}_{\mathbf{p}}^{(\nu)}\,|0\rangle_P = 0\,. \tag{10.411}$$

The Hamilton operator (10.392) of the second quantisation has both positive and negative energy values. Due to the redefinition of second quantized operators (10.403)–(10.406) it changes into

$$\hat{H} = \sum_{\nu=1}^{2}\int d^3p\, E_{\mathbf{p}}\left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger}\hat{b}_{\mathbf{p}}^{(\nu)} - \hat{d}_{\mathbf{p}}^{(\nu)}\hat{d}_{\mathbf{p}}^{(\nu)\dagger}\right)\,. \tag{10.412}$$

But, taking into account the anti-commutator algebra (10.409), the expression (10.412) for the Hamilton operator is transformed into:

$$\hat{H} = \sum_{\nu=1}^{2}\int d^3p\, E_{\mathbf{p}}\left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger}\hat{b}_{\mathbf{p}}^{(\nu)} + \hat{d}_{\mathbf{p}}^{(\nu)\dagger}\hat{d}_{\mathbf{p}}^{(\nu)}\right) - \sum_{\nu=1}^{2}\int d^3p\, E_{\mathbf{p}}\,\delta(\mathbf{0})\,. \tag{10.413}$$

The expectation value of this Hamilton operator with respect to the physical vacuum state $|0\rangle_P$ reads due to (10.410) and (10.411)

$${}_P\langle 0|\,\hat{H}\,|0\rangle_P = -\sum_{\nu=1}^{2}\int d^3p\, E_{\mathbf{p}}\,\delta(\mathbf{0})\,. \tag{10.414}$$

First of all we note that the vacuum energy for the fermions of the Dirac theory turns out to be negative in contrast to the bosonic cases of the Klein-Gordon theory in (8.118) and the Maxwell theory in (9.159). This is an immediate consequence of having an underlying anti-commutator algebra instead of a commutator algebra. But also in the fermionic case the vacuum energy (10.414) is divergent due to two reasons. On the one hand the respective momentum integral over the relativistic energy-momentum dispersion (10.299) is divergent and on the other hand the factor $\delta(\mathbf{0})$ is divergent as well. The renormalisation of the Hamilton operator (10.413) is performed by simply subtracting this infinitely large expectation value (10.414), yielding the normal-ordered Hamilton operator

$$:\hat{H}: = \hat{H} - {}_P\langle 0|\,\hat{H}\,|0\rangle_P = \sum_{\nu=1}^{2}\int d^3p\, E_{\mathbf{p}}\left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger}\hat{b}_{\mathbf{p}}^{(\nu)} + \hat{d}_{\mathbf{p}}^{(\nu)\dagger}\hat{d}_{\mathbf{p}}^{(\nu)}\right)\,. \tag{10.415}$$

This physical Hamilton operator is positive definite as both particles and antiparticles have the same energy $E_{\mathbf{p}} > 0$.

Quite correspondingly, the charge operator $\hat{Q}$, the momentum operator $\hat{P}$, and the helicity operator from (10.389), (10.395), and (10.399) change due to the redefinition of second quantized operators (10.403)–(10.406) to

$$\hat{Q} = \sum_{\nu=1}^{2} \int d^3p \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} + \hat{d}_{\mathbf{p}}^{(\nu)} \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \right) , \tag{10.416}$$

$$\hat{\mathbf{P}} = \sum_{\nu=1}^{2} \int d^3p \, \mathbf{p} \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} - \hat{d}_{\mathbf{p}}^{(\nu)} \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \right) , \tag{10.417}$$

$$\hat{h} = \sum_{\nu=1}^{2} \int d^3p \, \frac{(-1)^{\nu+1}}{2} \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} + \hat{d}_{\mathbf{p}}^{(\nu)} \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \right) . \tag{10.418}$$

Applying the anti-commutator algebra (10.409) yields

$$\hat{Q} = \sum_{\nu=1}^{2} \int d^3p \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} - \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \hat{d}_{\mathbf{p}}^{(\nu)} \right) + \sum_{\nu=1}^{2} \int d^3p \, \delta(\mathbf{0}) , \tag{10.419}$$

$$\hat{\mathbf{P}} = \sum_{\nu=1}^{2} \int d^3p \, \mathbf{p} \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} + \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \hat{d}_{\mathbf{p}}^{(\nu)} \right) - \sum_{\nu=1}^{2} \int d^3p \, \mathbf{p} \, \delta(\mathbf{0}) , \tag{10.420}$$

$$\hat{h} = \sum_{\nu=1}^{2} \int d^3p \, \frac{(-1)^{\nu+1}}{2} \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} - \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \hat{d}_{\mathbf{p}}^{(\nu)} \right) + \sum_{\nu=1}^{2} \frac{(-1)^{\nu+1}}{2} \int d^3p \, \delta(\mathbf{0}) . \tag{10.421}$$

The charge operator $\hat{Q}$ can be renormalised by subtracting its divergency, which amount to going over to the normal ordered charge operator

$$: \hat{Q} : = \hat{Q} - {}_P \langle 0 | \, \hat{Q} \, | 0 \rangle_P = \sum_{\nu=1}^{2} \int d^3p \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} - \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \hat{d}_{\mathbf{p}}^{(\nu)} \right) . \tag{10.422}$$

In contrast to that a renormalisation of the momentum operator $\hat{\mathbf{P}}$ is not necessary, since the expectation value of (10.420) with respect to the physical vacuum state $|0\rangle_P$ vanishes due to symmetry reasons in the momentum integral. Thus, the momentum operator (10.420) is already normal ordered:

$$: \hat{\mathbf{P}} : = \hat{\mathbf{P}} = \sum_{\nu=1}^{2} \int d^3p \, \mathbf{p} \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} + \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \hat{d}_{\mathbf{p}}^{(\nu)} \right) . \tag{10.423}$$

We conclude that particles carry the charge $+1$ and possess the momentum $\mathbf{p}$, while antiparticles have the negative charge -1 and also possess the momentum $\mathbf{p}$. And, finally, we recognize that also a renormalization of the helicity operator $\hat{h}$ is superfluous as the expectation value of (10.421) with respect to the physical vacuum state $|0\rangle_P$ vanishes due to symmetry reasons in the discrete sum. Thus, the helicity operator (10.421) is already normal ordered:

$$: \hat{h} : = \hat{h} = \sum_{\nu=1}^{2} \int d^3p \, \frac{(-1)^{\nu+1}}{2} \left(\hat{b}_{\mathbf{p}}^{(\nu)\dagger} \hat{b}_{\mathbf{p}}^{(\nu)} - \hat{d}_{\mathbf{p}}^{(\nu)\dagger} \hat{d}_{\mathbf{p}}^{(\nu)} \right) . \tag{10.424}$$

This means that particles with $\nu = 1$ ($\nu = 2$) and antiparticles with $\nu = 2$ ($\nu = 1$) have a positive (negative) helicity.

10.20 Propagator as Green Function

Analogous to the Klein-Gordon or the Maxwell propagator, also the Dirac propagator is defined as the expectation value of the time-ordered product of the field operators $\hat{\psi}_\alpha(\mathbf{x},t)$ and $\hat{\overline{\psi}}_\beta(\mathbf{x}',t')$ with respect to the physical vacuum $|0\rangle_P$:

$$S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = {}_P\left\langle 0 \left| \hat{T}\left(\hat{\psi}_\alpha(\mathbf{x},t)\hat{\overline{\psi}}_\beta(\mathbf{x}',t')\right)\right| 0\right\rangle_P . \tag{10.425}$$

We emphasize that the definition (8.123) of two time-dependent operators $\hat{A}(t)$ and $\hat{B}(t')$ in the context of bosonic operators is not valid for fermionic operators, but is given instead by

$$\hat{T}\left(\hat{A}(t)\,\hat{B}(t')\right) = \Theta(t-t')\,\hat{A}(t)\,\hat{B}(t') - \Theta(t'-t)\,\hat{B}(t')\,\hat{A}(t) \tag{10.426}$$

with the Heaviside function (8.124). Note the appearance of the minus in (10.426), which reflects the anti-commutativity of fermionic operators. Due to (10.426) the Dirac propagator (10.425) reads explicitly

$$S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \Theta(t-t')_P\left\langle 0\left|\hat{\psi}_\alpha(\mathbf{x},t)\hat{\overline{\psi}}_\beta(\mathbf{x}',t')\right|0\right\rangle_P - \Theta(t'-t)_P\left\langle 0\left|\hat{\overline{\psi}}_\beta(\mathbf{x}',t')\hat{\psi}_\alpha(\mathbf{x},t)\right|0\right\rangle_P . \tag{10.427}$$

At first we derive the equation of motion for the Dirac propagator by performing the time derivative of (10.427) and by taking into account (8.127):

$$i\hbar\frac{\partial}{\partial t}S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = i\hbar\delta(t-t')_P\left\langle 0\left|\left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\overline{\psi}}_\beta(\mathbf{x}',t')\right]_+\right|0\right\rangle_P \tag{10.428}$$
$$+\Theta(t-t')_P\left\langle 0\left|i\hbar\frac{\partial\hat{\psi}_\alpha(\mathbf{x},t)}{\partial t}\hat{\overline{\psi}}_\beta(\mathbf{x}',t')\right|0\right\rangle_P - \Theta(t'-t)_P\left\langle 0\left|\hat{\overline{\psi}}_\beta(\mathbf{x}',t')i\hbar\frac{\partial\hat{\psi}_\alpha(\mathbf{x},t)}{\partial t}\right|0\right\rangle_P .$$

With the definition of the Dirac adjoint Dirac spinor (10.92), the equal-time anti-commutator algebra (10.375), the Heisenberg equation of the Dirac spinor (10.380), and (10.425) we then yield

$$i\hbar\frac{\partial}{\partial t}S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \sum_{\gamma=1}^{4}\left(-i\hbar c\,\boldsymbol{\alpha}_{\alpha\gamma}\boldsymbol{\nabla} + Mc^2\beta_{\alpha\gamma}\right)S_{\gamma\beta}(\mathbf{x},t;\mathbf{x}',t') + i\hbar\gamma^0_{\alpha\beta}\delta(t-t')\delta(\mathbf{x}-\mathbf{x}') . \tag{10.429}$$

Thus, the Dirac propagator is just the Green function of the Dirac equation, which follows from (10.88), (10.249), and (10.250). Multiplying (10.429) from the left by γ^0/c and taking into account (10.251), (10.252) the equation of motion of the Dirac propagator can also be rewritten in a manifestly covariant form:

$$(i\hbar\gamma^\mu\partial_\mu - Mc)\,S(x;x') = i\hbar\delta(x-x') . \tag{10.430}$$

10.21 Propagator Calculation

In order to derive a Fourier representation for the Dirac propagator, we must first transfer the Dirac reinterpretation for the creation and annihilation operators $\hat{a}_{\mathbf{p}}^{(\nu)\dagger}$ and $\hat{a}_{\mathbf{p}}^{(\nu)}$ to the plane wave expansions (10.382) and (10.383) of the field operator $\hat{\psi}(\mathbf{x},t)$ and its adjoint $\hat{\psi}^{\dagger}(\mathbf{x},t)$. To this end we introduce the following notation for the plane waves of the particles

$$u_{\mathbf{p}}^{(1)}(\mathbf{x},t)=\psi_{\mathbf{p}}^{(1)}(\mathbf{x},t)\,,\qquad u_{\mathbf{p}}^{(2)}(\mathbf{x},t)=\psi_{\mathbf{p}}^{(2)}(\mathbf{x},t) \tag{10.431}$$

and, correspondingly, for the antiparticles

$$v_{\mathbf{p}}^{(1)}(\mathbf{x},t)=\psi_{\mathbf{p}}^{(3)}(\mathbf{x},t)\,,\qquad v_{\mathbf{p}}^{(2)}(\mathbf{x},t)=\psi_{\mathbf{p}}^{(4)}(\mathbf{x},t)\,. \tag{10.432}$$

Taking into account (10.403)–(10.406) the expansions (10.382), (10.383) then merge into

$$\hat{\psi}(\mathbf{x},t)=\sum_{\nu=1}^{2}\int d^3p\left[u_{\mathbf{p}}^{(\nu)}(\mathbf{x},t)\hat{b}_{\mathbf{p}}^{(\nu)}+v_{\mathbf{p}}^{(\nu)}(\mathbf{x},t)\hat{d}_{\mathbf{p}}^{(\nu)\dagger}\right]\,, \tag{10.433}$$

$$\hat{\overline{\psi}}(\mathbf{x},t)=\sum_{\nu=1}^{2}\int d^3p\left[\overline{u}_{\mathbf{p}}^{(\nu)}(\mathbf{x},t)\hat{b}_{\mathbf{p}}^{(\nu)\dagger}+\overline{v}_{\mathbf{p}}^{(\nu)}(\mathbf{x},t)\hat{d}_{\mathbf{p}}^{(\nu)}\right]\,. \tag{10.434}$$

Now we can insert (10.433) and (10.434) into (10.427). As the annihilation operators $\hat{b}_{\mathbf{p}}^{(\nu)}$, $\hat{d}_{\mathbf{p}}^{(\nu)}$ annul the ket vacuum state $|0\rangle_P$ according to (10.410), (10.411) and, correspondingly, the creation operators $\hat{b}_{\mathbf{p}}^{(\nu)\dagger}$, $\hat{d}_{\mathbf{p}}^{(\nu)\dagger}$ annul the bra vacuum state ${}_P\langle 0|$, we get

$$\begin{aligned} S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') \;=\;& \Theta(t-t')\sum_{\nu=1}^{2}\sum_{\nu'=1}^{2}\int d^3p\int d^3p'\,u_{\mathbf{p}\alpha}^{(\nu)}(\mathbf{x},t)\overline{u}_{\mathbf{p}'\beta}^{(\nu')}(\mathbf{x}',t')\,{}_P\langle 0|\,\hat{b}_{\mathbf{p}}^{(\nu)}\hat{b}_{\mathbf{p}'}^{(\nu')\dagger}\,|0\rangle_P \\ &-\Theta(t'-t)\sum_{\nu=1}^{2}\sum_{\nu'=1}^{2}\int d^3p\int d^3p'\,\overline{v}_{\mathbf{p}'\beta}^{(\nu')}(\mathbf{x}',t')v_{\mathbf{p}\alpha}^{(\nu)}(\mathbf{x},t)\,{}_P\langle 0|\,\hat{d}_{\mathbf{p}'}^{(\nu')}\hat{d}_{\mathbf{p}}^{(\nu)\dagger}\,|0\rangle_P\,. \end{aligned} \tag{10.435}$$

Due to the anti-commutator algebra (10.409) this reduces to

$$S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t')=\sum_{\nu=1}^{2}\int d^3p\left[\Theta(t-t')u_{\mathbf{p}\alpha}^{(\nu)}(\mathbf{x},t)\overline{u}_{\mathbf{p}\beta}^{(\nu)}(\mathbf{x}',t')-\Theta(t'-t)v_{\mathbf{p}\alpha}^{(\nu)}(\mathbf{x},t)\overline{v}_{\mathbf{p}\beta}^{(\nu)}(\mathbf{x}',t')\right]. \tag{10.436}$$

Inserting the plane waves (10.326) into (10.436) and considering (10.431) as well as (10.432) one obtains for the Fourier representation of the Dirac propagator

$$\begin{aligned} S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') =& \int d^3p\,\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}}} \\ &\times\left\{\Theta(t-t')e^{-i[E_{\mathbf{p}}(t-t')-\mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]}P_{\alpha\beta}^{u}(p)-\Theta(t'-t)e^{+i[E_{\mathbf{p}}(t-t')-\mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]}P_{\alpha\beta}^{v}(p)\right\}\,, \end{aligned} \tag{10.437}$$

where the following polarisation sums for both particles and antiparticles are introduced:

$$P_{\alpha\beta}^{u}(p) \;=\; \sum_{\nu=1}^{2}u_{\mathbf{p}\alpha}^{(\nu)}\,\overline{u}_{\mathbf{p}\beta}^{(\nu)}=\sum_{\nu=1}^{2}\psi_{\mathbf{p}\alpha}^{(\nu)}\,\overline{\psi}_{\mathbf{p}\beta}^{(\nu)}\,, \tag{10.438}$$

$$P_{\alpha\beta}^{v}(p) \;=\; \sum_{\nu=1}^{2}v_{\mathbf{p}\alpha}^{(\nu)}\,\overline{v}_{\mathbf{p}\beta}^{(\nu)}=\sum_{\nu=3}^{4}\psi_{\mathbf{p}\alpha}^{(\nu)}\,\overline{\psi}_{\mathbf{p}\beta}^{(\nu)}\,. \tag{10.439}$$

In order to evaluate these polarisation sums we have to perform several auxiliary calculations. To this end we start with the Dirac adjoint spinor amplitudes resulting from (10.316) with the help of (10.95) and (10.103):

$$\overline{\psi}_{\mathbf{p}}^{(\nu)} = \psi_{\mathbf{p}}^{(\nu)\dagger}\gamma^0 = \frac{1}{\sqrt{2}}\left(\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}, \chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}\right) \quad \text{for } \nu = 1,2\,, \tag{10.440}$$

$$\overline{\psi}_{\mathbf{p}}^{(\nu)} = \psi_{\mathbf{p}}^{(\nu)\dagger}\gamma^0 = \frac{1}{\sqrt{2}}\left(-\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\tilde{\sigma}}{Mc}}, \chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right)\sqrt{\frac{p\sigma}{Mc}}\right) \quad \text{for } \nu = 3,4\,. \tag{10.441}$$

We also note that the bi-spinors $\chi(\pm 1/2)$ are complete:

$$\sum_{\nu=1}^{2}\chi\left(\frac{(-1)^{\nu+1}}{2}\right)\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right) = \chi\left(\frac{1}{2}\right)\chi^\dagger\left(\frac{1}{2}\right) + \chi\left(-\frac{1}{2}\right)\chi^\dagger\left(-\frac{1}{2}\right) = I\,. \tag{10.442}$$

In fact, for the quantisation of the spin 1/2 with respect to the direction of the momentum $\mathbf{p}$ we obtain according to (10.353)

$$\begin{aligned}&\begin{pmatrix}\cos(\frac{\theta}{2})e^{-i\phi/2}\\ \sin(\frac{\theta}{2})e^{+i\phi/2}\end{pmatrix}\left(\cos\left(\frac{\theta}{2}\right)e^{+i\phi/2}, \sin\left(\frac{\theta}{2}\right)e^{-i\phi/2}\right)\\ &+\begin{pmatrix}-\sin(\frac{\theta}{2})e^{-i\phi/2}\\ \cos(\frac{\theta}{2})e^{+i\phi/2}\end{pmatrix}\left(-\sin\left(\frac{\theta}{2}\right)e^{+i\phi/2}, \cos\left(\frac{\theta}{2}\right)e^{-i\phi/2}\right) = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = I\,.\end{aligned} \tag{10.443}$$

Furthermore, from the completeness of the bi-spinors $\chi(\pm 1/2)$ in (10.442) we then conclude the completeness of the charge-conjugated bi-spinors $\chi^c(\pm 1/2)$:

$$\begin{aligned}&\sum_{\nu=1}^{2}\chi^c\left(\frac{(-1)^{\nu+1}}{2}\right)\chi^{c\dagger}\left(\frac{(-1)^{\nu+1}}{2}\right) = \sum_{\nu=1}^{2}c\chi^*\left(\frac{(-1)^{\nu+1}}{2}\right)\chi^T\left(\frac{(-1)^{\nu+1}}{2}\right)c^\dagger\\ &= c\left[\sum_{\nu=1}^{2}\chi\left(\frac{(-1)^{\nu+1}}{2}\right)\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\right]^T c^\dagger = cIc^\dagger = cc^\dagger = I\,.\end{aligned} \tag{10.444}$$

After these preparations, the polarisation sum of the particles is calculated as follows. At first, we insert (10.315) and (10.440) in (10.438):

$$P^u(p) = \frac{1}{2}\begin{pmatrix}\sqrt{\frac{p\sigma}{Mc}}\\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\end{pmatrix}\sum_{\nu=1}^{2}\chi\left(\frac{(-1)^{\nu+1}}{2}\right)\chi^\dagger\left(\frac{(-1)^{\nu+1}}{2}\right)\left(\sqrt{\frac{p\tilde{\sigma}}{Mc}}, \sqrt{\frac{p\sigma}{Mc}}\right)\,. \tag{10.445}$$

Due to the completeness relation (10.442) this reduces to

$$P^u(p) = \frac{1}{2}\begin{pmatrix}\sqrt{\frac{p\sigma}{Mc}}\\ \sqrt{\frac{p\tilde{\sigma}}{Mc}}\end{pmatrix}\left(\sqrt{\frac{p\tilde{\sigma}}{Mc}}, \sqrt{\frac{p\sigma}{Mc}}\right) = \frac{1}{2}\begin{pmatrix}\sqrt{\frac{p\sigma\, p\tilde{\sigma}}{(Mc)^2}} & \frac{p\sigma}{Mc}\\ \frac{p\tilde{\sigma}}{Mc} & \sqrt{\frac{p\tilde{\sigma}\, p\sigma}{(Mc)^2}}\end{pmatrix}\,. \tag{10.446}$$

And, finally, using the side calculation (10.304) we yield with the Dirac matrices (10.95) and the shortcut notation with the Feynman dagger (10.100)

$$P^u(p) = \frac{1}{2}\left[\frac{p_\mu}{Mc}\begin{pmatrix}O & \sigma^\mu\\ \tilde{\sigma}^\mu & O\end{pmatrix} + \begin{pmatrix}I & O\\ O & I\end{pmatrix}\right] = \frac{p_\mu\gamma^\mu + Mc}{2Mc} = \frac{\not{p} + Mc}{2Mc}\,. \tag{10.447}$$

The polarisation sum of the antiparticles is calculated along similar lines. Inserting (10.316) and (10.441) in (10.439) we get

$$P^v(p) = \frac{1}{2} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}} \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}} \end{pmatrix} \sum_{\nu=1}^{2} \chi^c \left(\frac{(-1)^{\nu+1}}{2} \right) \chi^{c\dagger} \left(\frac{(-1)^{\nu+1}}{2} \right) \left(-\sqrt{\frac{p\tilde{\sigma}}{Mc}}, \sqrt{\frac{p\sigma}{Mc}} \right) , \quad (10.448)$$

which reduces according to the completeness relation (10.444)

$$P^v(p) = \frac{1}{2} \begin{pmatrix} \sqrt{\frac{p\sigma}{Mc}} \\ -\sqrt{\frac{p\tilde{\sigma}}{Mc}} \end{pmatrix} \left(-\sqrt{\frac{p\tilde{\sigma}}{Mc}}, \sqrt{\frac{p\sigma}{Mc}} \right) = \frac{1}{2} \begin{pmatrix} -\sqrt{\frac{p\sigma\, p\tilde{\sigma}}{(Mc)^2}} & \frac{p\sigma}{Mc} \\ \frac{p\tilde{\sigma}}{Mc} & -\sqrt{\frac{p\tilde{\sigma}\, p\sigma}{(Mc)^2}} \end{pmatrix} . \quad (10.449)$$

With the side calculation (10.304), the Dirac matrices (10.95), and the shortcut notation with the Feynman dagger (10.100) we finally obtain

$$P^v(p) = \frac{1}{2} \left[\frac{p_\mu}{Mc} \begin{pmatrix} O & \sigma^\mu \\ \tilde{\sigma}^\mu & O \end{pmatrix} - \begin{pmatrix} I & O \\ O & I \end{pmatrix} \right] = \frac{p_\mu \gamma^\mu - Mc}{2Mc} = \frac{\not{p} - Mc}{2Mc} . \quad (10.450)$$

A comparison of (10.447) and (10.450) reveals that there is a simple relationship between the polarisation sums of the particles and the antiparticles:

$$P^v(p) = -P^u(-p) . \quad (10.451)$$

Using (10.451) in (10.437), the minus sign between the polarisation sums of the particles and antiparticles compensates the minus sign, which originally stems from the definition of the time-ordered product of fermionic operators in (10.426), yielding

$$S_{\alpha\beta}(\mathbf{x}, t; \mathbf{x}', t') = \int d^3p \, \frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}}} \times \left\{ \Theta(t-t') e^{-i[E_{\mathbf{p}}(t-t') - \mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]} P^u_{\alpha\beta}(p) + \Theta(t'-t) e^{+i[E_{\mathbf{p}}(t-t') - \mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]} P^u_{\alpha\beta}(-p) \right\} . \quad (10.452)$$

It turns out that this form of the Dirac propagator is universally valid for massive particles with arbitrary spin. The respective spin dependencies are hidden in the polarisation sum of the particles. For example, the result (10.452) agrees with the Klein-Gordon propagator (8.138) with the plane waves (8.112) provided that the polarisation sum is identified according to $P^u_{\alpha\beta}(p) = 1$.

10.22 Four-Dimensional Fourier Representation

Substituting the explicit form of the polarisation sum of the particles (10.447) into (10.452), one obtains

$$S_{\alpha\beta}(\mathbf{x}, t; \mathbf{x}', t') = \int d^3p \, \frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}}} \left\{ \Theta(t-t') e^{-i[E_{\mathbf{p}}(t-t') - \mathbf{p}(\mathbf{x}-\mathbf{x}')]/\hbar} \, \frac{p_\mu \gamma^\mu_{\alpha\beta} + Mc\,\delta_{\alpha\beta}}{2Mc} + \Theta(t'-t) e^{+i[E_{\mathbf{p}}(t-t') - \mathbf{p}(\mathbf{x}-\mathbf{x}')]/\hbar} \, \frac{-p_\mu \gamma^\mu_{\alpha\beta} + Mc\,\delta_{\alpha\beta}}{2Mc} \right\} . \quad (10.453)$$

The four-momentum vector in the polarisation sum of the particles can now be understood as the effect of applying the four-momentum operator on the plane waves, see (6.99):

$$\begin{aligned} S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') &= \int d^3p\,\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}}}\left\{\Theta(t-t')\,\frac{i\hbar\partial_\mu\gamma^\mu_{\alpha\beta}+Mc\,\delta_{\alpha\beta}}{2Mc}\,e^{-i[E_{\mathbf{p}}(t-t')-\mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]}\right. \\ &\left. +\Theta(t'-t)\,\frac{i\hbar\partial_\mu\gamma^\mu_{\alpha\beta}+Mc\,\delta_{\alpha\beta}}{2Mc}\,e^{+i[E_{\mathbf{p}}(t-t')-\mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]}\right\}. \end{aligned} \tag{10.454}$$

As now both terms involve the same differential operator it is suggestive to bring it in front of the momentum integral. Note that this manipulation leads to an additional term due to applying the time derivative upon the Heavside functions. But one can convince oneself that this additional term vanishes due to the odd symmetry of the respective momentum integral. With this we yield

$$\begin{aligned} S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') &= \frac{i\hbar\partial_\mu\gamma^\mu_{\alpha\beta}+Mc\,\delta_{\alpha\beta}}{2Mc}\int d^3p\,\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}}} \\ &\times\left\{\Theta(t-t')e^{-i[E_{\mathbf{p}}(t-t')-\mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]}+\Theta(t'-t)e^{+i[E_{\mathbf{p}}(t-t')-\mathbf{p}(\mathbf{x}-\mathbf{x}')/\hbar]}\right\}. \end{aligned} \tag{10.455}$$

The remaining momentum integral just represents the Klein-Gordon propagator as discussed below Eq. (10.452). Thus, the Dirac propagator can be obtained directly from the Klein-Gordon propagator by applying the following differential rule:

$$S_{\alpha\beta}(\mathbf{x},t;\mathbf{x}',t') = \frac{i\hbar\partial_\mu\gamma^\mu_{\alpha\beta}+Mc\,\delta_{\alpha\beta}}{2Mc}\,G(\mathbf{x},t;\mathbf{x}',t')\,. \tag{10.456}$$

Since we have already found a covariant formulation for the Klein-Gordon propagator in Section 8.12, also the Dirac propagator can be formulated covariantly according to (10.456):

$$S(x;x') = \frac{i\hbar\partial_\mu\gamma^\mu+Mc}{2Mc}\,G(x;x')\,. \tag{10.457}$$

Note that also (10.457) can be generalized to any massive particles with arbitrary spin according to the remarks below (10.452):

$$S(x;x') = P^u(i\hbar\partial)\,G(x;x')\,. \tag{10.458}$$

Indeed, inserting the explicit form of the polarisation sum of the particles (10.447) for the Dirac theory in (10.458) yields back (10.457). Substituting the four-dimensional Fourier representation of the Klein-Gordon propagator (8.168) into (10.457), we obtain a corresponding four-dimensional Fourier representation of the Dirac propagator:

$$\begin{aligned} S(x;x') &= \frac{i\hbar\partial_\mu\gamma^\mu+Mc}{2Mc}\,i\hbar 2Mc\lim_{\eta\downarrow 0}\int\frac{d^4p}{(2\pi\hbar)^4}\,\frac{1}{p^2-M^2c^2+i\eta}\,e^{-ip(x-x')/\hbar} \\ &= i\hbar\lim_{\eta\downarrow 0}\int\frac{d^4p}{(2\pi\hbar)^4}\,\frac{p_\mu\gamma^\mu+Mc}{p^2-M^2c^2+i\eta}\,e^{-ip(x-x')/\hbar}\,. \end{aligned} \tag{10.459}$$

With the help of the Clifford algebra (10.96) of the Dirac matrices, the denominator of (10.459) can be transformed as follows:

$$\begin{aligned} p^2-M^2c^2 &= p_\mu p_\nu g^{\mu\nu}-M^2c^2 = \frac{1}{2}\,p_\mu p_\nu\left(\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu\right)-M^2c^2 \\ &= \left(p_\mu\gamma^\mu\right)\left(p_\nu\gamma^\nu\right)-(Mc)^2 = \left(p_\mu\gamma^\mu-Mc\right)\left(p_\nu\gamma^\nu+Mc\right)\,. \end{aligned} \tag{10.460}$$

In the $\eta \downarrow 0$ limit, the numerator in (10.459) can be cancelled by a factor of the denominator in (10.460). With this the Dirac propagator has the following compact four-dimensional Fourier representation:

$$S(x;x') = \lim_{\eta \downarrow 0} \int \frac{d^4p}{(2\pi\hbar)^4} \frac{i\hbar}{p_\mu \gamma^\mu - Mc + i\eta} \, e^{-ip(x-x')/\hbar} \, . \tag{10.461}$$

In this form, the Dirac propagator obviously satisfies the equation of motion (10.430):

$$(i\hbar\gamma^\mu \partial_\mu - Mc) \, S(x;x') \;\;=\;\; i\hbar \int \frac{d^4p}{(2\pi\hbar)^4} \, e^{-ip(x-x')/\hbar} = i\hbar \, \delta(x-x') \, . \tag{10.462}$$

Part III:

Interacting Relativistic Fields and Their Quantization

Chapter 11

Relativistic Light-Matter Interaction

Quantum electrodynamics is the relativistic quantum field theory of electrodynamics. It describes how light and matter interact and represents historically the first successful many-body theory, which unites quantum mechanics and special relativity. It involves all phenomena, where electrically charged particles interact by means of an exchange of photons. This chapter focuses on working out the relativistic light-matter interaction consecutively at first on a classical, then on a first quantized, and, finally, on a second quantized description level. At all three stages the common guiding principle to introduce an interaction between the free theories of light and matter consists of a minimal coupling scheme, which is based on a local gauge theory. As the main result we derive the second quantized Hamilton operator underlying quantum electrodynamics. Apart from the free Maxwell and the free Dirac theory, which have already been discussed in the two previous chapters, we also obtain an interaction term, whose physical consequences have to be studied perturbatively. To this end we concisely review the Dirac interaction picture, which allows to treat the relativistic light-matter interaction systematically order by order. As a special case we outline how to analyse a generic scattering problem with the help of a corresponding perturbative expansion of the scattering operator, whose matrix elements determine the cross section.

11.1 Relativistic Mechanics

After having summarized concisely the basic principles of relativistic mechanics, we discuss first a free particle and then we introduce the description of a charged particle.

11.1.1 Basic Principles

The trajectory of a classical relativistic particle is described by specifying both the time coordinate t and the space coordinates $\mathbf{x}$ as a function of some parameter σ:

$$(x^\lambda(\sigma)) = (ct(\sigma), \mathbf{x}(\sigma)) \,. \tag{11.1}$$

Thus, the velocity with respect to this trajectory parameter σ reads

$$\left(\dot{x}^\lambda(\sigma)\right) = \left(\frac{dx^\lambda(\sigma)}{d\sigma}\right) = \left(c\,\frac{dt(\sigma)}{d\sigma}, \frac{d\mathbf{x}(\sigma)}{d\sigma}\right) . \tag{11.2}$$

The action represents a functional of the trajectory in the four-dimensional space-time

$$\mathcal{A} = \mathcal{A}\left[x^\lambda(\bullet)\right] \tag{11.3}$$

and is defined as the integral of the Lagrange function with respect to the chosen trajectory parameter σ:

$$\mathcal{A} = \int_{\sigma_i}^{\sigma_f} d\sigma L\left(x^\lambda(\sigma); \dot{x}^\lambda(\sigma)\right) . \tag{11.4}$$

Then the Hamilton principle leads to the underlying equations of motion in form of the Euler-Lagrange equations:

$$\frac{\delta \mathcal{A}}{\delta x^\mu(\sigma)} = \frac{\partial L}{\partial x^\mu(\sigma)} - \frac{d}{d\sigma}\frac{\partial L}{\partial \dot{x}^\mu(\sigma)} = 0 . \tag{11.5}$$

Note that the Hamilton principle exhibits a mechanical gauge invariance. Namely, regauging the Lagrange function according to

$$L'\left(x^\lambda; \dot{x}^\lambda\right) = L\left(x^\lambda; \dot{x}^\lambda\right) + \frac{d}{d\sigma}\chi\left(x^\lambda\right) = L\left(x^\lambda; \dot{x}^\lambda\right) + \partial_\nu \chi\left(x^\lambda\right)\dot{x}^\nu \tag{11.6}$$

only leads to additional surface terms of the action (11.3)

$$\mathcal{A}' = \mathcal{A} + \chi\left(x^\lambda(\sigma_f)\right) - \chi\left(x^\lambda(\sigma_i)\right) \tag{11.7}$$

and, therefore, does not change the equations of motion. In fact, for the transformed Lagrange function (11.6) one obtains the same Euler-Lagrange equations

$$\frac{\partial L'}{\partial x^\mu(\sigma)} - \frac{d}{d\sigma}\frac{\partial L'}{\partial \dot{x}^\mu(\sigma)} = \frac{\partial L}{\partial x^\mu(\sigma)} + \partial_\mu\partial_\nu\chi(x^\lambda(\sigma))\dot{x}^\nu(\sigma) - \frac{d}{d\sigma}\frac{\partial L}{\partial \dot{x}^\mu(\sigma)} - \partial_\nu\partial_\mu\chi(x^\lambda(\sigma))\dot{x}^\nu(\sigma), \tag{11.8}$$

since the gauge function $\chi\left(x^\lambda\right)$ is supposed to be twice continuously differentiable and therefore satisfies the Schwarz theorem:

$$\left(\partial_\mu\partial_\nu - \partial_\nu\partial_\mu\right)\chi\left(x^\lambda\right) = 0 . \tag{11.9}$$

In addition to this mechanical gauge invariance, relativistic mechanics has even further symmetries that take into account the principles of special relativity. Since the laws of physics are supposed to have the same form in all inertial frames, the action must be invariant under Lorentz transformations. In addition, however, the description of the trajectory should also be independent of the choice of the parameter σ, so that the action must also be form invariant under any transformation of the trajectory parameter:

$$\sigma = \sigma\left(\sigma'\right) . \tag{11.10}$$

This reparametrisation invariance is guaranteed by the fact that the Lagrange function represents a homogeneous function of the velocities of first order:

$$L\left(x^\lambda;\alpha\dot{x}^\lambda\right)=\alpha L\left(x^\lambda;\dot{x}^\lambda\right)\,. \tag{11.11}$$

Then applying (11.10) and (11.11) to the action (11.4) yields

$$\begin{aligned}\mathcal{A} &= \int_{\sigma_i}^{\sigma_f} d\sigma L\left(x^\lambda(\sigma);\dot{x}^\lambda(\sigma)\right)=\int_{\sigma_i'}^{\sigma_f'} d\sigma'\,\frac{d\sigma}{d\sigma'}\,L\left(x^\lambda(\sigma');\dot{x}^\lambda(\sigma')\frac{d\sigma'}{d\sigma}\right)\\ &= \int_{\sigma_i'}^{\sigma_f'} d\sigma' L(x^\lambda\left(\sigma'\right);\dot{x}^\lambda(\sigma'))\,. \end{aligned}\tag{11.12}$$

Furthermore, differentiating the condition (11.11) with respect to α and evaluating it then at the point $\alpha=1$ yields the corresponding Euler theorem. It states that the Hamilton function of relativistic mechanics vanishes:

$$H=\frac{\partial L}{\partial\dot{x}^\mu}\,\dot{x}^\mu-L=0\,. \tag{11.13}$$

This result is at first glance puzzling in view of the question how a relativistic mechanical systems is supposed to be quantized. The generic operator approach to determine from the underlying Hamilton function a Hamilton operator seems not to be possible due to (11.13). This can be considered as a motivation of Richard Feynman to work out an alternative formulation of quantum mechanics, which does not rely on Hamilton mechanics but is based instead on Lagrange mechanics.

11.1.2 Free Particle

Let us consider at first a free relativistic particle of mass M, whose action is motivated as follows. With the help of the Minkowski metric $g_{\mu\nu}$ one can determine the distance between two inifinitesimally adjacent space-time points x^μ and $x^\mu+dx^\mu$ according to

$$ds=\sqrt{g_{\mu\nu}\,dx^\mu dx^\nu}\,. \tag{11.14}$$

Decomposing this Lorentz invariant length element ds into the temporal and spatial contributions

$$ds=\sqrt{c^2dt^2-d\mathbf{x}^2}\,, \tag{11.15}$$

its physical meaning becomes apparent. Considering two infinitesimally adjacent space-time points in the rest frame of the particle, where we have $d\mathbf{x}_R=\mathbf{0}$, then ds becomes the proper length and, correspondingly, $\tau=ds/c$ denotes the proper time. The length of a trajectory between two different space-time points follows then from integrating the proper length (11.14) with respect to the chosen trajectory parameter σ:

$$\int_{s_i}^{s_f} ds=\int_{\sigma_i}^{\sigma_f} d\sigma\,\frac{ds}{d\sigma}=\int_{\sigma_i}^{\sigma_f} d\sigma\,\sqrt{g_{\mu\nu}\,\dot{x}^\mu(\sigma)\dot{x}^\nu(\sigma)}\,. \tag{11.16}$$

We remark that the proper length of the trajectory (11.16) is a Lorentz invariant quantity, which is also reparametrization invariant as the integrand is homogeneous in the velocities of first order in the sense of (11.11). This suggests that (11.16) is a viable candidate for an action in relativistic mechanics. Therefore, we argue now that

$$\mathcal{A}^{(0)} = -Mc \int_{s_i}^{s_f} ds = -Mc \int_{\sigma_i}^{\sigma_f} d\sigma \sqrt{g_{\mu\nu}\, \dot{x}^\mu(\sigma)\dot{x}^\nu(\sigma)} \tag{11.17}$$

represents the action of a free relativistic particle of mass M. We justify our choice by proving that it has the correct non-relativistic limit. Using the time t as the trajectory parameter σ, (11.17) namely leads to

$$\mathcal{A}^{(0)} = -Mc^2 \int_{t_i}^{t_f} dt \sqrt{1 - \frac{\dot{\mathbf{x}}(t)^2}{c^2}}\,, \tag{11.18}$$

so that the limes $c \to \infty$ yields the leading contribution

$$\mathcal{A}^{(0)} = \int_{t_i}^{t_f} dt \left[\frac{M}{2}\, \dot{x}(t)^2 - Mc^2\right] + \ldots\,. \tag{11.19}$$

This is the action of a free non-relativistic particle of mass M for which the energy scale is just shifted by the rest energy Mc^2.

11.1.3 Charged Particle

If a non-relativistic particle has a charge q, its interaction with a scalar potential $\varphi(\mathbf{x}, t)$ reads

$$\mathcal{A}^{(\mathrm{int})} = -q \int_{t_i}^{t_f} dt\, \varphi(\mathbf{x}(t), t)\,. \tag{11.20}$$

Taking into account (9.34) and (11.1), this can also be written as

$$\mathcal{A}^{(\mathrm{int})} = -q \int_{\sigma_i}^{\sigma_f} d\sigma\, \dot{x}^0(\sigma) A_0(x^\lambda(\sigma))\,. \tag{11.21}$$

Generalising (11.21) in a relativistic covariant way yields the interaction of a relativistic particle with the entire electromagnetic field, which is described by the four-vector potential $A_\mu(x^\lambda)$:

$$\mathcal{A}^{(\mathrm{int})} = -q \int_{\sigma_i}^{\sigma_f} d\sigma\, \dot{x}^\mu(\sigma) A_\mu(x^\lambda(\sigma))\,. \tag{11.22}$$

Thus, one can consider the charge q as a formal coupling constant, which measures the strength of interaction between the particle four-velocity and the four-vector potential. Note that also the interaction (11.22) is reparametrisation invariant as its integrand is homogeneous in the velocities of first order in the sense of (11.11) like the free action (11.17). Adding the free action (11.17) and the interaction (11.22) leads to a resulting action (11.3) with the Lagrange function

$$L\,(x^\mu; \dot{x}^\mu) = -Mc\sqrt{g_{\mu\nu}\, \dot{x}^\mu \dot{x}^\nu} - q\, A_\mu(x^\lambda)\dot{x}^\mu\,. \tag{11.23}$$

An electromagnetic gauge transformation

$$A'_\mu(x^\lambda) = A_\mu(x^\lambda) + \partial_\mu \Lambda(x^\lambda) \tag{11.24}$$

see (9.42), leads according to (11.4) and (11.23) to a mechanical gauge transformation (11.6), where the mechanical gauge function χ and the electromagnetic gauge function Λ turn out to be proportional to each other:

$$\chi\left(x^\lambda\right) = -q\Lambda(x^\lambda)\,. \tag{11.25}$$

Let us form the partial derivatives of the Lagrange function (11.23)

$$\frac{\partial L}{\partial x^\mu} = -q\,\partial_\mu A_\nu(x^\lambda)\dot{x}^\nu\,, \tag{11.26}$$

$$\frac{\partial L}{\partial \dot{x}^\mu} = -Mc\,\frac{g_{\mu\nu}\,\dot{x}^\nu}{\sqrt{g_{\kappa\lambda}\,\dot{x}^\kappa\dot{x}^\lambda}} - qA_\mu(x^\lambda)\,. \tag{11.27}$$

Furthermore, we introduce for the derivative of the proper length s with respect to the trajectory parameter σ according to (11.14) the shortcut notation

$$\dot{s}(\sigma) = \sqrt{g_{\mu\nu}\,\dot{x}^\mu(\sigma)\dot{x}^\nu(\sigma)}\,. \tag{11.28}$$

Then the Euler-Lagrange equations (11.5) following from (11.26) and (11.27) read as follows

$$M\ddot{x}^\mu = M\,\frac{\ddot{s}}{\dot{s}}\,\dot{x}^\mu + \frac{q\dot{s}}{c}\,g^{\mu\kappa}\left[\partial_\kappa A_\nu(x^\lambda) - \partial_\nu A_\kappa(x^\lambda)\right]\dot{x}^\nu\,. \tag{11.29}$$

Due to the reparametrization invariance of relativistic mechanics we are free to make a physically reasonable choice for the trajectory parameter. To this end we choose the trajectory parameter σ to be the proper time $\tau = s/c$. On the one hand this corresponds to the time which passes in the rest frame of the moving particle. On the other hand this simplifies the equations of motion (11.29) due to $\dot{s} = c$ and $\ddot{s} = 0$:

$$M\ddot{x}^\mu = q\,F^\mu{}_\nu\left(x^\lambda\right)\dot{x}^\nu\,. \tag{11.30}$$

Here the electrodynamic field strength sensor

$$F^\mu{}_\nu\left(x^\lambda\right) = g^{\mu\kappa}\,F_{\kappa\nu}\left(x^\lambda\right) \tag{11.31}$$

was introduced as an abbreviation. Note with (9.20) and (11.31) we recognize at the right-hand side of (11.30) the relativistic generalization of the Lorentz force.

11.1.4 Minimal Coupling

In order to investigate in more detail the description of a charged particle in relativistic mechanics we choose the trajectory parameter σ to be the time t in the laboratory frame. Then the action (11.3), (11.4) reduces to

$$\mathcal{A} = \mathcal{A}[\mathbf{x}(\bullet)] = \int_{t_i}^{t_f} dt\,L\left(\mathbf{x}(t);\dot{\mathbf{x}}(t);t\right) \tag{11.32}$$

and the Lagrange function (11.23) specializes due to (9.34) and (11.1) to

$$L = -Mc^2\sqrt{1-\frac{\dot{\mathbf{x}}^2}{c^2}} - q\varphi(\mathbf{x},t) + q\,\dot{\mathbf{x}}\,\mathbf{A}(\mathbf{x},t)\,. \tag{11.33}$$

Thus, the canonical momentum reads

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = \frac{M\dot{\mathbf{x}}}{\sqrt{1-\dot{\mathbf{x}}^2/c^2}} + q\mathbf{A}(\mathbf{x},t)\,. \tag{11.34}$$

Here the first term represents the kinetic momentum and the second term the corresponding contribution of the vector potential, so (11.34) corresponds to

$$\mathbf{p} = \mathbf{p}_{\text{kin}} + q\mathbf{A}(\mathbf{x},t)\,. \tag{11.35}$$

In view of performing a Legendre transformation from the Lagrange function to the Hamilton function, we have to invert the relation (11.34) between the momentum $\mathbf{p}$ and the velocity $\dot{\mathbf{x}}$. A straight-forward algebraic calculation yields

$$\dot{\mathbf{x}} = \frac{c\,[\mathbf{p}-q\mathbf{A}(\mathbf{x},t)]}{\sqrt{[\mathbf{p}-q\mathbf{A}(\mathbf{x},t)]^2 + M^2c^2}}\,. \tag{11.36}$$

Thus, evaluating the Legendre transformation

$$H = \dot{\mathbf{x}}\,\frac{\partial L}{\partial \dot{\mathbf{x}}} - L \tag{11.37}$$

by using (11.33) and (11.36) we obtain for the Hamilton function of a relativistic charged particle in the electromagnetic field

$$H = c\sqrt{[\mathbf{p}-q\mathbf{A}(\mathbf{x},t)]^2 + M^2c^2} + q\varphi(\mathbf{x},t)\,. \tag{11.38}$$

We remark that this result reduces in the limit $c \to \infty$ apart from the rest energy Mc^2 to the familiar non-relativistic expression

$$H = Mc^2 + \frac{[\mathbf{p}-q\mathbf{A}(\mathbf{x},t)]^2}{2M} + q\varphi(\mathbf{x},t) + \ldots\,. \tag{11.39}$$

Furthermore, analogous to (11.35), we interpret (11.38) such that the first term describes the kinetic energy and the second term the potential energy:

$$H = H_{\text{kin}} + q\varphi(\mathbf{x},t)\,. \tag{11.40}$$

Conversely, we read off from (11.35) and (11.40) that a free theory with $q = 0$ can formally be transferred into the corresponding interacting one with $q \neq 0$ by substituting momentum (energy) via

$$\mathbf{p} \quad\to\quad \mathbf{p} - q\mathbf{A}(\mathbf{x},t)\,, \qquad H \quad\to\quad H - q\varphi(\mathbf{x},t)\,. \tag{11.41}$$

This so-called minimal coupling of charged particle to the electromagnetic field can now be covariantly formulated in terms of a covariant momentum four-vector (6.16) by taking into account (9.34) according to

$$p_\mu \quad\to\quad p_\mu - qA_\mu(x^\lambda)\,. \tag{11.42}$$

In the following the minimal coupling rule (11.42) is applied to the realm of relativistic quantum field theory by combining it with the Jordan rule.

11.2 QED Actions

Quantum electrodynamics describes the interaction between charged massive particles and the electromagnetic field. One distinguishes, in principle, between scalar quantum electrodynamics for charged spin 0-particles as, for instance, pions $\pi^{\pm}$, and spinor quantum electrodynamics for charged spin 1/2-particles as, for instance, electrons e^- or positrons e^+. As a consequence, the underlying equations of motion for massive particles and for the electromagnetic field are coupled by additional interaction terms. In the Lagrangian density this leads to an additional interaction term in addition to the free Lagrangian density, whose strength depends on the coupling constant of electrodynamics, i.e. the charge q. In the following, we examine at first scalar quantum electrodynamics.

11.2.1 Scalar QED

We start with the relativistic covariant action of the free Klein-Gordon field

$$\mathcal{A}\left[\Psi(\bullet);\Psi^*(\bullet)\right] = \frac{1}{c}\int d^4x\,\mathcal{L}\left(\Psi(x^\lambda),\partial_\mu\Psi(x^\lambda);\Psi^*(x^\lambda),\partial_\mu\Psi^*(x^\lambda)\right)\,, \tag{11.43}$$

where the Lagrange density reads according to Section 8.1

$$\mathcal{L} = \frac{\hbar^2}{2M}\,g^{\mu\nu}\partial_\mu\Psi^*(x^\lambda)\partial_\nu\Psi(x^\lambda) \;-\; \frac{Mc^2}{2}\,\Psi^*(x^\lambda)\Psi(x^\lambda)\,. \tag{11.44}$$

A minimal coupling of the Klein-Gordon field to the electromagnetic field is now implemented by combining the substitution rule (11.42) with the Jordan rule, see (6.99):

$$p_\mu \quad\rightarrow\quad i\hbar\partial_\mu\,. \tag{11.45}$$

This leads to the catchy substitution rule

$$\partial_\mu\Psi(x^\lambda) \quad\rightarrow\quad D_\mu\Psi(x^\lambda)\,, \qquad \partial_\mu\Psi^*(x^\lambda) \quad\rightarrow\quad D_\mu^*\Psi^*(x^\lambda)\,, \tag{11.46}$$

where D_μ denotes the so-called gauge covariant derivative:

$$D_\mu = \partial_\mu + \frac{iq}{\hbar}\,A_\mu(x^\lambda)\,. \tag{11.47}$$

Applying (11.46) and (11.47) to the Lagrangian density (11.44) we get

$$\mathcal{L} = \frac{\hbar}{2M}\,g^{\mu\nu}\left[\partial_\mu - \frac{iq}{\hbar}\,A_\mu(x^\lambda)\right]\Psi^*(x^\lambda)\left[\partial_\nu + \frac{iq}{\hbar}\,A_\nu(x^\lambda)\right]\Psi(x^\lambda) - \frac{Mc^2}{2}\Psi^*(x^\lambda)\Psi(x^\lambda)\,, \tag{11.48}$$

which can be rewritten in a form which resembles that of a free Lagrangian density of the Klein-Gordon field:

$$\mathcal{L} = \frac{\hbar}{2M}\,g^{\mu\nu}D_\mu^*\Psi^*(x^\lambda)D_\nu\Psi(x^\lambda) \;-\; \frac{Mc^2}{2}\Psi^*(x^\lambda)\Psi(x^\lambda)\,. \tag{11.49}$$

We now examine the consequences of an electrodynamic gauge transformation (11.24). As the fields in quantum mechanics are only uniquely determined up to a phase factor, an electrodynamic gauge transformation can only change the phase. Supplementing an electrodynamic gauge transformation (11.24) accordingly with a quantum mechanical gauge transformation

$$\Psi'(x^\lambda) = \exp\left[-\frac{iq}{\hbar}\,\Lambda(x^\lambda)\right]\Psi(x^\lambda)\,, \tag{11.50}$$

the gauge covariant derivative (11.47) turns out to transform like the Klein-Gordon field $\Psi(x^\lambda)$:

$$\begin{aligned} D'_\mu\Psi'(x^\lambda) &= \left[\partial_\mu + \frac{iq}{\hbar}\,A_\mu(x^\lambda) + \frac{iq}{\hbar}\,\partial_\mu\,\Lambda(x^\lambda)\right]\exp\left[-\frac{iq}{\hbar}\,\Lambda(x^\lambda)\right]\Psi(x^\lambda) \\ &= \exp\left[-\frac{iq}{\hbar}\,\Lambda(x^\lambda)\right]D_\mu\Psi(x^\lambda)\,. \end{aligned} \tag{11.51}$$

Analogously, one obtains for the adjoint field

$$\Psi'^*(x^\lambda) = \exp\left[\frac{iq}{\hbar}\,\Lambda(x^\lambda)\right]\Psi^*(x^\lambda)\,, \tag{11.52}$$

$$D'^*_\mu\Psi'^*(x^\lambda) = \exp\left[\frac{iq}{\hbar}\,\Lambda(x^\lambda)\right]D^*_\mu\Psi^*(x^\lambda)\,. \tag{11.53}$$

Then it follows straight-forwardly from (11.50)–(11.52) that the Lagrangian density (11.49) is invariant under an electrodynamic gauge transformation:

$$\begin{aligned} \mathcal{L}' &= \frac{\hbar}{2M}\,g^{\mu\nu}D'^*_\mu\Psi'^*(x^\lambda)D'_\nu\Psi'(x^\lambda) - \frac{Mc^2}{2}\,\Psi'^*(x^\lambda)\Psi'(x^\lambda) \\ &= \frac{\hbar}{2M}\,g^{\mu\nu}D^*_\mu\Psi(x^\lambda)D_\nu\Psi(x^\lambda) - \frac{Mc^2}{2}\Psi^*(x^\lambda)\,\Psi(x^\lambda) = \mathcal{L}\,. \end{aligned} \tag{11.54}$$

If we consider the four-vector potential $A_\mu(x^\lambda)$ not as a given quantity but as a dynamic field, we must add to the Lagrangian density (11.48) the Lagrangian density of the free Maxwell field from Subsections 9.5 and 9.6, which is also invariant under the local gauge transformation (11.24). Accordingly, scalar quantum electrodynamics has the gauge invariant action

$$\mathcal{A}\left[\Psi(\bullet);\Psi^*(\bullet);A_\nu(\bullet)\right] = \frac{1}{c}\int d^4x\,\mathcal{L}\,, \tag{11.55}$$

where the Lagrange density is of the form

$$\mathcal{L} = \mathcal{L}\left(\Psi(x^\lambda),\partial_\mu\Psi(x^\lambda);\Psi^*(x^\lambda),\partial_\mu\Psi^*(x^\lambda);A_\nu(x^\lambda),\partial_\mu A_\nu(x^\lambda)\right) \tag{11.56}$$

and reads explicitly

$$\mathcal{L} = \frac{\hbar}{2M}\,g^{\mu\nu}\left(\partial_\mu - \frac{iq}{\hbar}\,A_\mu\right)\Psi^*\left(\partial_\nu + \frac{iq}{\hbar}\,A_\nu\right)\Psi - \frac{Mc^2}{2}\Psi^*\Psi - \frac{1}{4\mu_0}\,F_{\mu\nu}F^{\mu\nu}\,. \tag{11.57}$$

The Lagrange density thus decomposes according to

$$\mathcal{L} = \mathcal{L}^{(0)} + \mathcal{L}^{(\mathrm{int})}\,. \tag{11.58}$$

Here $\mathcal{L}^{(0)}$ includes the free Lagrange densities of the Klein-Gordon field and the Maxwell field

$$\mathcal{L}^{(0)} = \frac{\hbar}{2M}\, g^{\mu\nu}\partial_\mu\Psi^*(x^\lambda)\partial_\nu\Psi(x^\lambda) - \frac{Mc^2}{2}\Psi^*(x^\lambda)\Psi(x^\lambda) - \frac{1}{4\mu_0}\, F_{\mu\nu}(x^\lambda)F^{\mu\nu}(x^\lambda) \tag{11.59}$$

and the interaction term turns out to have the structure

$$\mathcal{L}^{(\mathrm{int})} = -j^\mu(x^\lambda)A_\mu(x^\lambda)\,. \tag{11.60}$$

The four-vector potential thus couples to the four-vector current density, which follows from applying to the free four-vector current density

$$j^\mu(x^\lambda) = \frac{i\hbar q}{2M}\, g^{\mu\nu}\Big[\Psi^*(x^\lambda)\partial_\nu\Psi(x^\lambda) - \Psi(x^\lambda)\partial_\nu\Psi^*(x^\lambda)\Big]\,, \tag{11.61}$$

see Section 8.2, the catchy substitution rule (11.46), (11.47), yielding

$$j^\mu(x^\lambda) = \frac{i\hbar q}{2M}\, g^{\mu\nu}\Big[\Psi^*(x^\lambda)\partial_\nu\Psi(x^\lambda) - \Psi(x^\lambda)\partial_\nu\Psi^*(x^\lambda)\Big] - \frac{q^2}{M}\, g^{\mu\nu}\, A_\nu(x^\lambda)\Psi^*(x^\lambda)\Psi(x^\lambda)\,. \tag{11.62}$$

The respective partial derivatives of the Lagrange density (11.57) read as follows

$$\frac{\partial\mathcal{L}}{\partial A_\nu} = \frac{i\hbar q}{2M}\, g^{\nu\kappa}\,\Psi\left(\partial_\kappa - \frac{iq}{\hbar}\,A_\kappa\right)\Psi^* - \frac{i\hbar q}{2M}\, g^{\nu\kappa}\,\Psi^*\left(\partial_\kappa + \frac{iq}{\hbar}\,A_\kappa\right)\Psi\,, \tag{11.63}$$

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{\mu_0}\,F^{\mu\nu}\,, \tag{11.64}$$

$$\frac{\partial\mathcal{L}}{\partial\Psi^*} = -\frac{i\hbar q}{2M} g^{\mu\nu}\,A_\mu\left(\partial_\nu + \frac{iq}{\hbar}\,A_\nu\right)\Psi - \frac{Mc^2}{2}\Psi\,, \tag{11.65}$$

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi^*)} = \frac{\hbar^2}{2M}\, g^{\mu\nu}\left(\partial_\nu + \frac{iq}{\hbar}\,A_\nu\right)\Psi\,, \tag{11.66}$$

$$\frac{\partial\mathcal{L}}{\partial\Psi} = \frac{i\hbar q}{2M}\, g^{\mu\nu}\,A_\mu\left(\partial_\nu - \frac{iq}{\hbar}\,A_\nu\right)\Psi - \frac{Mc^2}{2}\Psi^*\,, \tag{11.67}$$

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi)} = \frac{\hbar^2}{2M}\, g^{\mu\nu}\left(\partial_\nu - \frac{iq}{\hbar}\,A_\nu\right)\Psi^*\,. \tag{11.68}$$

With this we obtain the Euler-Lagrange equations of scalar quantum electrodynamics. For the Maxwell field the Euler-Lagrange equations

$$\frac{\partial\mathcal{L}}{\partial A_\nu} - \partial_\mu\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} = 0 \tag{11.69}$$

are specified as follows:

$$\partial_\mu F^{\mu\nu} = \mu_0\, g^{\nu\kappa}\,\frac{i\hbar q}{2M}\left[\Psi^*\left(\partial_\kappa + \frac{iq}{\hbar}\,A_\kappa\right)\Psi - \Psi\left(\partial_\kappa - \frac{iq}{\hbar}\,A_\kappa\right)\Psi^*\right]\,, \tag{11.70}$$

and for the Klein-Gordon field we get

$$\frac{\partial\mathcal{L}}{\partial\Psi^*} - \partial_\mu\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi^*)} = 0 \quad\Rightarrow\quad g^{\mu\nu}\left(\partial_\mu + \frac{iq}{\hbar}\,A_\mu\right)\left(\partial_\nu + \frac{iq}{\hbar}\,A_\nu\right)\Psi + \frac{M^2c^2}{\hbar^2}\,\Psi = 0\,, \tag{11.71}$$

$$\frac{\partial\mathcal{L}}{\partial\Psi} - \partial_\mu\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi)} = 0 \quad\Rightarrow\quad g^{\mu\nu}\left(\partial_\mu - \frac{iq}{\hbar}\,A_\mu\right)\left(\partial_\nu - \frac{iq}{\hbar}\,A_\nu\right)\Psi^* + \frac{M^2c^2}{\hbar^2}\,\Psi^* = 0\,. \tag{11.72}$$

The equations of motion (11.70) represent inhomogeneous Maxwell equations (9.29) with the current density (11.62). Furthermore, the equations of motion (11.71) and (11.72) arise from the free Klein-Gordon equations by applying the catchy substitution rule (11.46), (11.47).

11.2.2 Spinor QED

Now we construct the corresponding Lagrange density of spinor quantum electrodynamics by applying the principle of local gauge invariance. The starting point is the Lagrange density of the free Dirac field, see Sections 10.6 and 10.13:

$$\mathcal{L} = \overline{\psi}(x) \left(i\hbar c \gamma^\mu \partial_\mu - Mc^2 \right) \psi(x) \,. \tag{11.73}$$

Obviously, this Lagrange density is invariant with respect to a global phase transformation of the form

$$\psi'(x) = e^{-iq\Lambda/\hbar}\, \psi(x) \,, \qquad \overline{\psi}'(x) = e^{iq\Lambda/\hbar}\, \overline{\psi}(x) \,, \tag{11.74}$$

where q denotes the charge of the massive spin 1/2-particle. This global $U(1)$ invariance implies via the Noether theorem the derivation of the continuity equation of charge conservation for the free Dirac theory. We now try to achieve that one can choose any phase at any space-time point, so that the above global phase Λ becomes a space- and time-dependent quantity $\Lambda(x)$. Accordingly, we consider the local $U(1)$ phase transformation

$$\psi'(x) = e^{-iq\Lambda(x)/\hbar}\, \psi(x) \,, \qquad \overline{\psi}'(x) = e^{iq\Lambda(x)/\hbar}\, \overline{\psi}(x) \,. \tag{11.75}$$

The Lagrange density of the free Dirac field (11.73) is then no longer invariant under such a local phase transformation, since an additional term appears due to the partial derivative of the Dirac spinor:

$$\partial_\mu \psi'(x) \;=\; e^{-iq\Lambda(x)/\hbar} \left[\partial_\mu \psi(x) - \frac{iq}{\hbar}\, \partial_\mu \Lambda(x) \psi(x) \right] \tag{11.76}$$

and we get

$$\mathcal{L}' = \overline{\psi}'(x) \left(i\hbar c \gamma^\mu \partial_\mu - Mc^2 \right) \psi'(x) = \mathcal{L} + qc\overline{\psi}(x)\gamma^\mu \partial_\mu \Lambda(x) \psi(x) \,. \tag{11.77}$$

In order to establish a local gauge invariance, additional fields must be introduced and the Lagrange density (11.73) must be extended correspondingly. Since the additional term in (11.77) depends on the gradient of the phase $\partial_\mu \Lambda(x)$ and, therefore, represents a Lorentz vector, we introduce a gauge field $A_\mu(x)$, which couples to the spinor with the coupling constant q. To this end we replace the partial derivative of the spinor by

$$\partial_\mu \psi(x) \quad \to \quad \mathcal{D}_\mu \psi(x) \,, \tag{11.78}$$

where the gauge covariant derivative of the spinor is defined by

$$\mathcal{D}_\mu = \partial_\mu + \frac{iq}{\hbar} A_\mu(x) \,. \tag{11.79}$$

Then we determine the transformation behaviour of the gauge field by requiring that the gauge covariant derivative of the spinor transforms like the spinor itself:

$$\mathcal{D}'_\mu \psi'(x) = e^{-iq\Lambda(x)/\hbar}\, \mathcal{D}_\mu \psi(x) \,. \tag{11.80}$$

Substituting (11.79) into (11.80) then leads to the condition

$$\partial_\mu \psi'(x) + \frac{iq}{\hbar}\, A'_\mu(x)\psi'(x) = e^{iq\Lambda(x)/\hbar} \left[\partial_\mu \psi(x) + \frac{iq}{\hbar}\, A_\mu(x)\psi(x)\right] . \tag{11.81}$$

With the help of (11.75) and (11.76) this reduces, finally, to the gauge transformation

$$A'_\mu(x) = A_\mu(x) + \partial_\mu \Lambda(x) . \tag{11.82}$$

Since the gauge field $A_\mu(x)$ transforms just like the four-vector potential of electrodynamics in (11.24), it is identified with the latter in the following. The substitution rule (11.80), (11.81) then corresponds to the minimal coupling of the Dirac field to the Maxwell field. For the sake of completeness we note that the substitution rule for the Dirac adjoint spinor is given analogous to (11.78) by

$$\partial_\mu \overline{\psi}(x) \quad \to \quad \mathcal{D}^*_\mu\, \overline{\psi}(x) . \tag{11.83}$$

The gauge-covariant derivative of the Dirac-adjoint spinor transforms then via

$$\mathcal{D}'^*_\mu\, \overline{\psi}'(x) \quad = \quad \left[\partial_\mu - \frac{iq}{\hbar}\, A_\mu(x) - \frac{iq}{\hbar}\, \partial_\mu \Lambda(x)\right] e^{iq\Lambda(x)/\hbar}\, \overline{\psi}(x) = e^{iq\Lambda(x)/\hbar}\, \mathcal{D}^*_\mu \overline{\psi}(x) . \tag{11.84}$$

Performing the substitution (11.78) in the Lagrange density of the free Dirac field (11.73), we obtain

$$\mathcal{L} = \overline{\psi}(x) \left(i\hbar c \gamma^\mu \mathcal{D}_\mu - Mc^2\right) \Psi(x) . \tag{11.85}$$

Decomposing the gauge covariant derivative D_μ according to (11.81), then in addition to the original free Lagrangian density of the Dirac field (11.73) also an interaction term arises:

$$\mathcal{L} = \overline{\psi}(x) \left\{ i\hbar c \gamma^\mu \left[\partial_\mu + \frac{iq}{\hbar}\, A_\mu(x)\right] - Mc^2 \right\} \Psi(x) . \tag{11.86}$$

If we also consider the vector potential $A_\mu(x)$ as a dynamic field, we must add to the Lagrangian density (11.86) the Lagrangian density of the free Maxwell field. The resulting Lagrangian density turns out to be then manifestly local gauge invariant due to (11.75), (11.80), and (11.82). It represents the Lagrangian density of spinor quantum electrodynamics:

$$\mathcal{L} = \overline{\psi}(x) \left\{ i\hbar c \gamma^\mu \left[\partial_\mu + \frac{iq}{\hbar} A_\mu(x)\right] - Mc^2 \right\} \psi(x) - \frac{1}{4\mu_0} F_{\mu\nu}(x) F^{\mu\nu}(x) . \tag{11.87}$$

This Lagrange density decomposes according to

$$\mathcal{L} = \mathcal{L}^{(0)} + \mathcal{L}^{(\mathrm{int})} , \tag{11.88}$$

where $\mathcal{L}^{(0)}$ representing the free Lagrangian density including both the Dirac field and the Maxwell field

$$\mathcal{L}^{(0)} = \overline{\psi}(x) \left(i\hbar c \gamma^\mu \partial_\mu - Mc^2\right) \psi(x) - \frac{1}{4\mu_0} F_{\mu\nu}(x) F^{\mu\nu}(x) \tag{11.89}$$

and the interaction term turns out to have the structure

$$\mathcal{L}^{(\mathrm{int})} = -j^{\mu}(x) A_{\mu}(x) \,. \tag{11.90}$$

The four-vector potential thus couples to the four-current density of the free Dirac field, see Section 10.10 and 10.13:

$$j^{\mu}(x) = qc\,\overline{\psi}(x)\gamma^{\mu}\psi(x) \,. \tag{11.91}$$

The respective partial derivatives of the Lagrange density (11.87) lead to

$$\frac{\partial \mathcal{L}}{\partial A_{\nu}} = -qc\,\overline{\psi}\gamma^{\nu}\psi \,, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_{\mu}A_{\nu})} = -\frac{1}{\mu_0}F^{\mu\nu} \,, \tag{11.92}$$

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}} = \left(i\hbar c\gamma^{\mu}\partial_{\mu} - Mc^2\right)\psi - qc\gamma^{\mu}\Psi A_{\mu} \,, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\overline{\psi})} = 0 \,, \tag{11.93}$$

$$\frac{\partial \mathcal{L}}{\partial \psi} = -Mc^2\,\overline{\psi} - qc\,\overline{\psi}\,\gamma^{\mu}A_{\mu} \,, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\psi)} = i\hbar c\overline{\psi}(x)\gamma^{\mu} \,. \tag{11.94}$$

The Euler-Lagrange equations of spinor quantum electrodynamics thus result in

$$\frac{\partial \mathcal{L}}{\partial A_{\nu}} - \partial_{\mu}\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}A_{\nu})} = 0 \qquad \Rightarrow \qquad \partial_{\mu}F^{\mu\nu} = \mu_0 qc\,\overline{\psi}\,\gamma^{\mu}\psi \,, \tag{11.95}$$

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}} - \partial_{\mu}\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\overline{\psi})} = 0 \qquad \Rightarrow \qquad i\gamma^{\mu}\left(\partial_{\mu} + \frac{iq}{\hbar}A_{\mu}\right)\psi - \frac{Mc}{\hbar}\,\psi = 0 \,, \tag{11.96}$$

$$\frac{\partial \mathcal{L}}{\partial \psi} - \partial_{\mu}\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\psi)} = 0 \qquad \Rightarrow \qquad i\left(\partial_{\mu} - \frac{iq}{\hbar}\,A_{\mu}\right)\overline{\psi}\,\gamma^{\mu} + \frac{Mc}{\hbar}\overline{\psi} = 0 \,. \tag{11.97}$$

The equations of motion (11.95) agree with the inhomogeneous Maxwell equations (9.29) with the current density (11.91). Furthermore, the equations of motion (11.96) and (11.97) emerge from the free Dirac equations by applying the minimal couplings (11.78) and (11.83), which involve the gauge-covariant derivative (11.79) via (11.78) and (11.83).

11.3 QED Hamilton Function

Starting from the Lagrange density of spinor quantum electrodynamics in (11.87), we now calculate the corresponding Hamilton density. At first, we express the contribution of the free Maxwell field in terms of the electric field strength **E** and the magnetic field strength **B**, see Section 9.6:

$$\mathcal{L} = \overline{\psi}\left(i\hbar c\gamma^{\mu}\partial_{\mu} - Mc^2\right)\psi + \frac{\epsilon_0}{2}\,\mathbf{E}^2 - \frac{1}{2\mu_0}\,\mathbf{B}^2 - qc\,\overline{\psi}\,\gamma^{\mu}\psi A_{\mu} \,. \tag{11.98}$$

Then we express the electric field strength **E** and the magnetic field strength **B** by the scalar potential φ and the vector potential **A** due to (9.7) and (9.8), yielding

$$\begin{aligned}\mathcal{L} = {} & \overline{\psi}(\mathbf{x},t)\left(i\hbar c\gamma^{\mu}\partial_{\mu} - Mc^2\right)\psi(\mathbf{x},t) + \frac{\epsilon_0}{2}\left[\frac{\partial \mathbf{A}(\mathbf{x},t)}{\partial t}\right]^2 + \epsilon_0\,\frac{\partial \mathbf{A}(\mathbf{x},t)}{\partial t}\,\boldsymbol{\nabla}\varphi(\mathbf{x},t) \\ & + \frac{\epsilon_0}{2}\,\left[\boldsymbol{\nabla}\varphi(\mathbf{x},t)\right]^2 - \frac{1}{2\mu_0}\,\left[\boldsymbol{\nabla}\times\mathbf{A}(\mathbf{x},t)\right]^2 - qc\,\overline{\psi}(\mathbf{x},t)\,\gamma^{\mu}\psi(\mathbf{x},t)A_{\mu}(\mathbf{x},t) \,.\end{aligned} \tag{11.99}$$

Note that the Coulomb gauge (9.13) yields a scalar potential, which no longer represents a dynamic field but is determined by the charge density following from (9.16) and (10.280):

$$\varphi(\mathbf{x},t) = \int d^3x' \, \frac{q\,\psi^\dagger(\mathbf{x}',t)\psi(\mathbf{x}',t)}{4\pi\epsilon_0|\mathbf{x}-\mathbf{x}'|} \,. \tag{11.100}$$

Thus, with the charge density also the scalar potential does not vanish, so the radiation gauge (9.58) is no longer valid here. The canonical conjugated momentum fields then follow from (11.99) to be

$$\pi(\mathbf{x},t) = \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \psi(\mathbf{x},t)}{\partial t}\right)} = i\hbar\overline{\psi}(\mathbf{x},t)\gamma^0 = i\hbar\psi^\dagger(\mathbf{x},t)\,, \tag{11.101}$$

$$\overline{\pi}(\mathbf{x},t) = \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \overline{\psi}(\mathbf{x},t)}{\partial t}\right)} = 0\,, \tag{11.102}$$

$$\boldsymbol{\pi}(\mathbf{x},t) = \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \mathbf{A}(\mathbf{x},t)}{\partial t}\right)} = \epsilon_0\left[\frac{\partial \mathbf{A}(\mathbf{x},t)}{\partial t} + \boldsymbol{\nabla}\varphi(\mathbf{x},t)\right] . \tag{11.103}$$

Note that the last term in (11.103) did not appear in Section 9.6, as there we considered the free Maxwell field in vacuum. A subsequent Legendre transformation leads then to the corresponding Hamilton density:

$$\mathcal{H} = \pi(\mathbf{x},t)\,\frac{\partial\psi(\mathbf{x},t)}{\partial t} + \frac{\partial\overline{\psi}(\mathbf{x},t)}{\partial t}\,\overline{\pi}(\mathbf{x},t) + \boldsymbol{\pi}(\mathbf{x},t)\,\frac{\partial \mathbf{A}(\mathbf{x},t)}{\partial t} - \mathcal{L}\,. \tag{11.104}$$

Thus, using (11.99) and (11.101)–(11.104) as well as (10.251) and (10.252) we obtain:

$$\begin{aligned}\mathcal{H} \;=\;& \psi^\dagger(\mathbf{x},t)\left(-i\hbar c\,\boldsymbol{\alpha}\boldsymbol{\nabla} + Mc^2\beta\right)\psi(\mathbf{x},t) + \frac{\epsilon_0}{2}\left[\frac{\partial\mathbf{A}(\mathbf{x},t)}{\partial t}\right]^2 + \frac{1}{2\mu_0}\left[\boldsymbol{\nabla}\times\mathbf{A}(\mathbf{x},t)\right]^2 \\ & -\frac{\epsilon_0}{2}\left[\boldsymbol{\nabla}\varphi(\mathbf{x},t)\right]^2 + q\psi^\dagger(\mathbf{x},t)\psi(\mathbf{x},t)\varphi(\mathbf{x},t) \; - qc\,\psi^\dagger(\mathbf{x},t)\,\boldsymbol{\alpha}\psi(\mathbf{x},t)\mathbf{A}(\mathbf{x},t)\,.\end{aligned} \tag{11.105}$$

Going over to the Hamiltonian function, we yield by partial integration and by taking into account the Coulomb gauge (9.13), see Section 9.6:

$$\begin{aligned} H = \int d^3x \Big\{ & \psi^\dagger(\mathbf{x},t)\left(-i\hbar c\boldsymbol{\alpha}\boldsymbol{\nabla} + Mc^2\beta\right)\psi(\mathbf{x},t) + \frac{\epsilon_0}{2}\left[\frac{\partial\mathbf{A}(\mathbf{x},t)}{\partial t}\right]^2 + \frac{1}{2\mu_0}\partial_k A_l(\mathbf{x},t)\partial_k A_l(\mathbf{x},t) \\ & + \frac{\epsilon_0}{2}\,\varphi(\mathbf{x},t)\Delta\varphi(\mathbf{x},t) + q\psi^\dagger(\mathbf{x},t)\psi(\mathbf{x},t)\varphi(\mathbf{x},t) - qc\psi^\dagger(\mathbf{x},t)\boldsymbol{\alpha}\psi(\mathbf{x},t)\mathbf{A}(\mathbf{x},t)\Big\}\,. \end{aligned} \tag{11.106}$$

At this stage we use the Poisson equation for a point charge for determining that the Green function of the Poisson equation is given by the Coulomb potential:

$$\Delta\,\frac{q}{4\pi\epsilon_0\,|\mathbf{x}-\mathbf{x}'|} = -\frac{q}{\epsilon_0}\,\delta(\mathbf{x}-\mathbf{x}') \qquad \Rightarrow \qquad \Delta\,\frac{1}{|\mathbf{x}-\mathbf{x}'|} = -4\pi\delta(\mathbf{x}-\mathbf{x}')\,. \tag{11.107}$$

Thus, taking into account (11.100) and (11.107) we yield the auxiliary calculation

$$\begin{aligned} &\frac{\epsilon_0}{2}\int d^3x\,\varphi(\mathbf{x},t)\Delta\varphi(\mathbf{x},t) = \frac{\epsilon_0}{2}\int d^3x \int d^3x'\,\varphi(\mathbf{x},t)\,\frac{q\psi^\dagger(\mathbf{x}',t)\psi(\mathbf{x}',t)}{4\pi\epsilon_0}\,\Delta\,\frac{1}{|\mathbf{x}-\mathbf{x}'|} \\ &-\frac{\epsilon_0}{2}\int d^3x\,\varphi(\mathbf{x},t)\,\frac{q\psi^\dagger(\mathbf{x}',t)\psi(\mathbf{x}',t)}{4\pi\epsilon_0}\,4\pi\delta(\mathbf{x}-\mathbf{x}') = -\frac{q}{2}\int d^3x\,\varphi(\mathbf{x},t)\psi^\dagger(\mathbf{x},t)\psi(\mathbf{x},t)\,. \end{aligned} \tag{11.108}$$

Substituting (11.100) and (11.108) into (11.106), the Hamilton function of spinor quantum electrodynamics decomposes according to

$$H = H^{(0)} + H^{(\mathrm{int})} . \tag{11.109}$$

where $H^{(0)}$ represents the free contributions of both the Dirac field and the Maxwell field:

$$\begin{aligned} H^{(0)} &= \int d^3x \Bigg[\psi^\dagger(\mathbf{x},t) \left(-i\hbar c \boldsymbol{\alpha}\boldsymbol{\nabla} + Mc^2\beta\right) \psi(\mathbf{x},t) + \frac{\epsilon_0}{2} \frac{\partial A_k(\mathbf{x},t)}{\partial t} \frac{\partial A_k(\mathbf{x},t)}{\partial t} \\ &\quad + \frac{1}{2\mu_0}\, \partial_k A_l(\mathbf{x},t) \partial_k A_l(\mathbf{x},t) \Bigg] . \end{aligned} \tag{11.110}$$

The term $H^{(\mathrm{int})}$ represents the interaction between the Dirac and the Maxwell field:

$$\begin{aligned} H^{(\mathrm{int})} &= -qc\, \overline{\psi}(\mathbf{x},t)\, \boldsymbol{\gamma}\, \psi(\mathbf{x},t) \mathbf{A}(\mathbf{x},t) \\ &\quad + \frac{q^2}{8\pi\epsilon_0} \int d^3x \int d^3x'\, \frac{\overline{\psi}(\mathbf{x},t)\gamma^0\psi(\mathbf{x},t)\overline{\psi}(\mathbf{x}',t)\gamma^0\psi(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|} . \end{aligned} \tag{11.111}$$

The first term in (11.111) arises from the free Hamilton function of the Dirac field in (11.110) by performing a minimal coupling to the vector potential in accordance with (11.79):

$$\boldsymbol{\nabla} \quad \rightarrow \quad \boldsymbol{\nabla} - \frac{iq}{\hbar}\, \mathbf{A}(\mathbf{x},t) . \tag{11.112}$$

The second term in (11.111) represents an instantaneous Coulomb self-interaction of the Dirac field. It is non-trivial to prove that such an instantaneous self-interaction does not contradict the principles of special relativity. Later we show by a concrete example of a scattering process that the instantaneous Coulomb self-interaction in (11.111) turns out to compensate an unwanted contribution (9.207) of the Maxwell propagator (9.205), which comes from the Coulomb gauge, thus yielding at the end manifestly covariant physical results.

11.4 Dirac Picture

In quantum field theory the quantisation of free fields is basically trivial, since the Hamilton function and, thus, the second quantized Hamilton operator is quadratic in the fields and the field operators, respectively. This has the consequence that the Fourier operators occurring in plane wave expansions of the field operators represent physically the creation and the annihilation of individual particles with well-defined properties. But the quantisation of interacting fields is non-trivial as it leads to interesting physical processes due to the involved nonlinearities. The Hamilton operator contains higher powers of the same field in the case of a self-interaction or products of different fields as in quantum electrodynamics. The resulting dynamics of the field operators is, thus, complicated because, at each instant, the Fourier operators correspond to the creation and annihilation of particles with different properties. For instance, preparing an annihilation operator at in initial time t_0, it may happen that at a later time instant $t > t_0$ it evolves into a certain superposition of creation and annihilation operators.

Basically, it is not possible to solve exactly an interacting quantum field theory. However, provided that the interaction is sufficiently weak, reliable approximations can be obtained with the help of perturbation theory. Then neither the Schrödinger picture, in which the state vectors are time-dependent and the operators time-independent, nor the Heisenberg picture, in which conversely the state vectors are time-independent and the operators time-dependent, is suitable, see Section 3.4. Instead, in perturbation theory the Dirac or interaction picture turns out to be more favorable, since the time dependencies are then distributed appropriately between both the state vectors and the operators.

11.4.1 Derivation

The starting point of perturbation theory is the assumption that the Hamilton operator of the system under consideration can be split into two parts in the Schrödinger picture:

$$\hat{H}_S = \hat{H}_S^{(0)} + \hat{H}_S^{(\text{int})} . \tag{11.113}$$

Here $\hat{H}_S^{(0)}$ represents the Hamilton operator of a system of free fields and $\hat{H}_S^{(\text{int})}$ denotes the interacting part of the Hamilton operator. In the Schrödinger picture the time-dependent state vector $|\psi_S(t)\rangle$ fulfills the Schrödinger equation (3.53), which has the formal solution (3.55). Thus, the time dependence of $|\psi_S(t)\rangle$ is determined by the mutual influence of both the unperturbed and the perturbed Hamilton operator $\hat{H}_S^{(0)}$ and $\hat{H}_S^{(\text{int})}$. The idea for introducing the Dirac picture is now to redo the temporal evolution with the free Hamilton operator $\hat{H}_S^{(0)}$ according to

$$|\psi_D(t)\rangle = e^{i\hat{H}_S^{(0)}t/\hbar}\,|\psi_S(t)\rangle \qquad \Longleftrightarrow \qquad |\psi_S(t)\rangle = e^{-i\hat{H}_S^{(0)}t/\hbar}\,|\psi_D(t)\rangle . \tag{11.114}$$

In order to determine the operator $\hat{O}_D(t)$ in the Dirac picture, we require that the expectation values do not change during the transition from the Schrödinger picture to the Dirac picture:

$$\langle\psi_D(t)|\hat{O}_D(t)|\psi_D(t)\rangle = \langle\psi_S(t)|\hat{O}_S|\psi_S(t)\rangle . \tag{11.115}$$

Inserting (11.114) into (11.115) then actually leads to determine the operator $\hat{O}_D(t)$ in the Dirac picture

$$\begin{aligned} \langle\psi_D(t)|\, e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\, e^{-i\hat{H}_S^{(0)}t/\hbar}\,|\psi_D(t)\rangle &= \langle\psi_D(t)|\hat{O}_D(t)|\psi_D(t)\rangle \\ \Rightarrow \qquad \hat{O}_D(t) &= e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\, e^{-i\hat{H}_S^{(0)}t/\hbar} . \end{aligned} \tag{11.116}$$

For example, for the free Hamilton operator $\hat{O}_S = \hat{H}_S^{(0)}$ follows that it does not change its shape during the transition from the Schrödinger picture to the Dirac picture:

$$\hat{H}_D^{(0)}(t) = e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{H}_S^{(0)}\, e^{-i\hat{H}_S^{(0)}t/\hbar} = \hat{H}_S^{(0)} . \tag{11.117}$$

With (11.114) and (11.116) we have, thus, defined the Dirac picture both for the state vectors and the operators. It remains to investigate their respective equations of motion. Based on the equation of motion of a state vector in the Schrödinger picture (3.53) together with (11.113)

$$i\hbar\frac{\partial}{\partial t}|\psi_S(t)\rangle = \hat{H}_S|\psi_S(t)\rangle = \left(\hat{H}_S^{(0)} + \hat{H}_S^{(\mathrm{int})}\right)|\psi_S(t)\rangle \tag{11.118}$$

and taking into account (11.114) we then obtain the equation of motion of the corresponding state vector in the Dirac picture, which is called the Tomonaga-Schwinger equation:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi_D(t)\rangle &= e^{i\hat{H}_S^{(0)}t/\hbar}\left[i\hbar\frac{\partial}{\partial t}|\psi_S(t)\rangle - \hat{H}_S^{(0)}|\psi_S(t)\rangle\right] = e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{H}_S^{(\mathrm{int})}|\psi_S(t)\rangle \\ &\Longrightarrow\; i\hbar\frac{\partial}{\partial t}|\psi_D(t)\rangle = \hat{H}_D^{(\mathrm{int})}(t)\,|\psi_D(t)\rangle\,. \end{aligned} \tag{11.119}$$

Here the interacting part of the Hamilton operator is transferred from the Schrödinger picture to the Dirac picture according to (11.116):

$$\hat{H}_D^{(\mathrm{int})}(t) = e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{H}_S^{(\mathrm{int})}\,e^{-i\hat{H}_S^{(0)}t/\hbar}\,. \tag{11.120}$$

Furthermore, starting from the equation of motion of an operator in the Schrödinger picture

$$i\hbar\frac{\partial}{\partial t}\hat{O}_S = 0\,, \tag{11.121}$$

we use (11.116) in order to derive the equation of motion of the corresponding operator in the Dirac picture as follows:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{O}_D(t) &= e^{i\hat{H}_S^{(0)}t/\hbar}\left[\hat{O}_S\hat{H}_S^{(0)} - \hat{H}_S^{(0)}\hat{O}_S\right]e^{-i\hat{H}_S^{(0)}t/\hbar} \\ &= e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\,e^{-i\hat{H}_S^{(0)}t/\hbar}\hat{H}_S^{(0)} - \hat{H}_S^{(0)}e^{i\hat{H}_S^{(0)}t/\hbar}\hat{O}_S\,e^{-i\hat{H}_S^{(0)}t/\hbar} = [\hat{O}_D(t), \hat{H}_S^{(0)}]_-\,. \end{aligned} \tag{11.122}$$

While in the Dirac picture the dynamics of the state vectors is determined by the interacting part of the Hamilton operator according to (11.119), only the free Hamilton operator enters the dynamics of the operators according to (11.122). The latter result has the consequence that the field operators in the Dirac picture still retain their respective properties of a free theory to create and annihilate particles.

11.4.2 Example

In order to illustrate the latter point we consider the quantum field-theoretic description of non-relativistic bosons, see Chapter 4. In the Schrödinger picture, the field operators $\hat{\psi}_S(\mathbf{x})$ and $\hat{\psi}_S^\dagger(\mathbf{x})$ satisfy the canonical commutator relations

$$\left[\hat{\psi}_S(\mathbf{x}), \hat{\psi}_S(\mathbf{x}')\right]_- = \left[\hat{\psi}_S^\dagger(\mathbf{x}), \hat{\psi}_S^\dagger(\mathbf{x}')\right]_- = 0\,, \quad \left[\hat{\psi}_S(\mathbf{x}), \hat{\psi}_S^\dagger(\mathbf{x}')\right]_- = \delta(\mathbf{x}-\mathbf{x}')\,. \tag{11.123}$$

Thus, $\hat{\psi}_S(\mathbf{x})$ and $\hat{\psi}_S^\dagger(\mathbf{x})$ describe the annihilation and creation of a bosonic particle at space point $\mathbf{x}$, respectively. With the help of basis functions $u_\mathbf{p}(\mathbf{x})$, which fulfill the orthonormality relation

$$\int d^3x\, u_\mathbf{p}^*(\mathbf{x})\, u_{\mathbf{p}'}(\mathbf{x}) = \delta(\mathbf{p}-\mathbf{p}') \tag{11.124}$$

and the completeness relation

$$\int d^3p\, u_\mathbf{p}^*(\mathbf{x})\, u_\mathbf{p}(\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}')\,, \tag{11.125}$$

the field operators $\hat{\psi}_S(\mathbf{x})$ and $\hat{\psi}_S^\dagger(\mathbf{x})$ can be expanded as follows:

$$\hat{\psi}_S(\mathbf{x}) = \int d^3p\, u_\mathbf{p}(\mathbf{x})\hat{a}_S(\mathbf{p})\,, \tag{11.126}$$

$$\hat{\psi}_S^\dagger(\mathbf{x}) = \int d^3p\, u_\mathbf{p}^*(\mathbf{x})\hat{a}_S^\dagger(\mathbf{p})\,. \tag{11.127}$$

Using (11.124), the expansions (11.126) and (11.127) are then inverted according to

$$\int d^3x\, u_\mathbf{p}^*(\mathbf{x})\hat{\psi}_S(\mathbf{x}) = \hat{a}_S(\mathbf{p})\,, \tag{11.128}$$

$$\int d^3x\, u_\mathbf{p}(\mathbf{x})\hat{\psi}_S^\dagger(\mathbf{x}) = \hat{a}_S^\dagger(\mathbf{p})\,. \tag{11.129}$$

With this the commutator relations (11.123) of the expansion operators $\hat{a}_S(\mathbf{p})$ and $\hat{a}_S^\dagger(\mathbf{p})$ result in

$$\left[\hat{a}_S(\mathbf{p}), \hat{a}_S(\mathbf{p}')\right]_- = \left[\hat{a}_S^\dagger(\mathbf{p}), \hat{a}_S^\dagger(\mathbf{p}')\right]_- = 0\,, \quad \left[\hat{a}_S(\mathbf{p}), \hat{a}_S^\dagger(\mathbf{p}')\right]_- = \delta(\mathbf{p}-\mathbf{p}')\,. \tag{11.130}$$

Accordingly, the operator $\hat{a}_S(\mathbf{p})$ ($\hat{a}_S^\dagger(\mathbf{p})$) describes the annihilation (creation) of a particle of momentum $\mathbf{p}$. Let us assume for the sake of simplicity that the free Hamiltonian operator has already a diagonal form with an energy-momentum dispersion $E_\mathbf{p}$ in the Schrödinger picture:

$$\hat{H}_S^{(0)} = \int d^3p\, E_\mathbf{p}\, \hat{a}_S^\dagger(\mathbf{p})\hat{a}_S(\mathbf{p})\,. \tag{11.131}$$

The Heisenberg equation for the evolution of the annihilation operator in the Dirac picture (11.122) results then in

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\,\hat{a}_D(\mathbf{p},t) &= \left[\hat{a}_D(\mathbf{p},t), \hat{H}_S^{(0)}\right]_- = e^{i\hat{H}_S^{(0)}t/\hbar}\left[\hat{a}_S(\mathbf{p}), \hat{H}_S^{(0)}\right]_- e^{-i\hat{H}_S^{(0)}t/\hbar} \\ &= \int d^3p'\, E_{\mathbf{p}'}\, e^{i\hat{H}_S^{(0)}t/\hbar}\left[\hat{a}_S(\mathbf{p}), \hat{a}_S^\dagger(\mathbf{p}')\hat{a}_S(\mathbf{p}')\right]_- e^{-i\hat{H}_S^{(0)}t/\hbar} = E_\mathbf{p}\,\hat{a}_D(\mathbf{p},t)\,. \end{aligned} \tag{11.132}$$

The solution of this operator-valued first-order differential equation with the initial condition

$$\hat{a}_D(\mathbf{p},0) = \hat{a}_S(\mathbf{p}) \tag{11.133}$$

is given by

$$\hat{a}_D(\mathbf{p},t) = e^{-iE_\mathbf{p}t/\hbar}\,\hat{a}_S(\mathbf{p})\,. \tag{11.134}$$

Correspondingly, the time evolution of the creation operator yields

$$\hat{a}_D^\dagger(\mathbf{p},t) = e^{iE_\mathbf{p}t/\hbar}\,\hat{a}_S^\dagger(\mathbf{p})\,. \tag{11.135}$$

Due to (11.116), (11.131), (11.134), and (11.135) we then prove (11.117) as expected:

$$\hat{H}_D^{(0)}(t) = e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{H}_S^{(0)}\,e^{-i\hat{H}_S^{(0)}t/\hbar} = \int d^3p\,E_\mathbf{p}\,e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{a}_S^\dagger(\mathbf{p})\,e^{-i\hat{H}_S^{(0)}t/\hbar}\,e^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{a}_S(\mathbf{p})\,e^{-i\hat{H}_S^{(0)}t/\hbar}$$
$$= \int d^3p\,E_\mathbf{p}\,\hat{a}_D^\dagger(\mathbf{p},t)\hat{a}_D(\mathbf{p},t) = \int d^3p\,E_\mathbf{p}\,\hat{a}_S^\dagger(\mathbf{p})\hat{a}_S(\mathbf{p}) = \hat{H}_S^{(0)}\,. \tag{11.136}$$

From (11.134) and (11.135) we read off that the creation and annihilation operators in the Dirac picture differ only by one additional phase factor from their counterparts in the Schrödinger picture. This means that the creation and annihilation operators in the Dirac picture do not change their character as single-particle operators during the time evolution. In particular, it follows directly from (11.130), (11.134), and (11.135) that the equal-time commutator relations in the Dirac picture coincide with those in the Schrödinger picture:

$$\left[\hat{a}_D(\mathbf{p},t),\hat{a}_D(\mathbf{p}',t)\right]_- = \left[\hat{a}_D^\dagger(\mathbf{p},t),\hat{a}_D^\dagger(\mathbf{p}',t)\right]_- = 0\,,\;\; \left[\hat{a}_D(\mathbf{p},t),\hat{a}_D^\dagger(\mathbf{p}',t)\right]_- = \delta(\mathbf{p}-\mathbf{p}')\,. \tag{11.137}$$

This means that $\hat{a}_D(\mathbf{p},t)$ and $\hat{a}_D^\dagger(\mathbf{p},t)$ annihilate and create a particle with momentum $\mathbf{p}$ at time t. Furthermore, the field operators (11.126) and (11.127) in the Schrödinger picture change in the Dirac picture into

$$\hat{\psi}_D(\mathbf{x},t) = e^{i\hat{H}_S^{(0)}t/\hbar}\hat{\psi}_S(\mathbf{x})e^{-i\hat{H}_S^{(0)}t/\hbar} = \int d^3p\,u_\mathbf{p}(\mathbf{x})\,\hat{a}_D(\mathbf{p},t)\,, \tag{11.138}$$

$$\hat{\psi}_D^\dagger(\mathbf{x},t) = e^{i\hat{H}_S^{(0)}t/\hbar}\hat{\psi}_S^\dagger(\mathbf{p})e^{-i\hat{H}_S^{(0)}t/\hbar} = \int d^3p\,u_\mathbf{p}^*(\mathbf{x})\,\hat{a}_D^\dagger(\mathbf{p},t)\,. \tag{11.139}$$

Thus, according to (11.138) and (11.139), the field operators in the Dirac picture can be expanded with respect to creation and annihilation operators in exactly the same way as in the Heisenberg picture, see Section 3.4. Moreover, we obtain for the equal-time commutator relations of the field operators in the Dirac picture:

$$\left[\hat{\psi}_D(\mathbf{x},t),\hat{\psi}_D(\mathbf{x}',t)\right]_- = \left[\hat{\psi}_D^\dagger(\mathbf{x},t),\hat{\psi}_D^\dagger(\mathbf{x}',t)\right]_- = 0\,,\;\; \left[\hat{\psi}_D(\mathbf{x},t),\hat{\psi}_D^\dagger(\mathbf{x}',t)\right]_- = \delta(\mathbf{x}-\mathbf{x}')\,. \tag{11.140}$$

Thus, we have in the Dirac picture the same equal-time commutator relations for the field operators as in the Heisenberg picture for free particles. This means that $\hat{\psi}_D(\mathbf{x},t)$ and $\hat{\psi}_D^\dagger(\mathbf{x},t)$ annihilate and create a particle at space point $\mathbf{x}$ at time t.

11.5 Canonical Field Quantisation

We now perform the canonical field quantisation of spinor quantum electrodynamics in the Dirac picture. According to the previous section, this means that we demand the same equal-time commutator or anti-commutator relations for the interacting theory in the Dirac picture as for the free theory in the Heisenberg picture. As we work from now on only in the Dirac picture we simplify our notation by omitting the index D, which indicates the Dirac picture. Concerning the Dirac field, equal-time anti-commutator relations are required for the independent field operators $\hat{\psi}_\alpha(\mathbf{x},t)$ and $\hat{\pi}_\beta(\mathbf{x},t)$:

$$\left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\psi}_\beta(\mathbf{x}',t)\right]_+ = \left[\hat{\pi}_\alpha(\mathbf{x},t),\hat{\pi}_\beta(\mathbf{x}',t)\right]_+ = 0, \left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\pi}_\beta(\mathbf{x}',t)\right]_+ = i\hbar\delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{x}'). \tag{11.141}$$

Concerning the Maxwell field, equal-time commutator relations are used for the independent field operators $\hat{A}_k(\mathbf{x},t)$ and $\hat{\pi}_l(\mathbf{x},t)$:

$$\left[\hat{A}_k(\mathbf{x},t),\hat{A}_l(\mathbf{x}',t)\right]_+ = \left[\hat{\pi}_k(\mathbf{x},t),\hat{\pi}_l(\mathbf{x}',t)\right]_+ = 0\,, \left[\hat{A}_k(\mathbf{x},t),\hat{\pi}_l(\mathbf{x}',t)\right]_+ = i\hbar\delta^T_{kl}(\mathbf{x}-\mathbf{x}')\,, \tag{11.142}$$

where the transveral delta function (9.86) ensures analogous to Section 9.7 that the Coulomb gauge also holds for the field operators $\hat{A}_k(\mathbf{x},t)$ and $\hat{\pi}_l(\mathbf{x},t)$. And, due to their independence, equal-time commutator relations are required between the field operators of the Dirac and the Maxwell fields:

$$\begin{aligned}\left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{A}_k(\mathbf{x}',t)\right]_- &= \left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\pi}_k(\mathbf{x}',t)\right]_- \\ = \left[\hat{\pi}_\alpha(\mathbf{x},t),\hat{A}_k(\mathbf{x}',t)\right]_- &= \left[\hat{\pi}_\alpha(\mathbf{x},t),\hat{\pi}_k(\mathbf{x}',t)\right]_- = 0\,.\end{aligned} \tag{11.143}$$

Applying the field quantization to the momentum fields (11.101) and (11.103) yields for the corresponding momentum operators:

$$\hat{\pi}(\mathbf{x},t) = i\hbar\hat{\overline{\psi}}(\mathbf{x},t)\gamma^0 = i\hbar\hat{\psi}^\dagger(\mathbf{x},t)\,, \tag{11.144}$$

$$\hat{\boldsymbol{\pi}}(\mathbf{x},t) = \epsilon_0\left[\frac{\partial\hat{\mathbf{A}}(\mathbf{x},t)}{\partial t}+\boldsymbol{\nabla}\hat{\varphi}(\mathbf{x},t)\right]\,, \tag{11.145}$$

where the scalar field operator follows from (11.100):

$$\hat{\varphi}(\mathbf{x},t) = \int d^3x'\,\frac{q\,\hat{\psi}^\dagger(\mathbf{x}',t)\hat{\psi}(\mathbf{x}',t)}{4\pi\epsilon_0|\mathbf{x}-\mathbf{x}'|}\,. \tag{11.146}$$

Thus, we can also use instead of the momentum field operators $\hat{\pi}_\alpha(\mathbf{x},t)$ and $\hat{\pi}_l(\mathbf{x},t)$ the field operators $\hat{\psi}^\dagger_\alpha(\mathbf{x},t)$ and $\partial\hat{A}_k(\mathbf{x},t)/\partial t$ in order to define the underlying equal-time (anti-)commutator relations of spinor QED. For instance, (11.141) can be directly rewritten as

$$\left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\psi}_\beta(\mathbf{x}',t)\right]_+ = \left[\hat{\psi}^\dagger_\alpha(\mathbf{x},t),\hat{\psi}^\dagger_\beta(\mathbf{x}',t)\right]_+ = 0, \left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{\psi}^\dagger_\beta(\mathbf{x}',t)\right]_+ = \delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{x}'). \tag{11.147}$$

Accordingly, we obtain from (11.143) straight-forwardly

$$\left[\hat{\psi}_\alpha(\mathbf{x},t),\hat{A}_k(\mathbf{x}',t)\right]_- = \left[\hat{\psi}^\dagger_\alpha(\mathbf{x},t),\hat{A}_k(\mathbf{x}',t)\right]_- = 0\,. \tag{11.148}$$

Furthermore, taking into account (9.86), (11.142), (11.145), and (11.146) we get at first

$$\left[\hat{\psi}_\alpha(\mathbf{x},t), \frac{\partial \hat{A}_k(\mathbf{x}',t)}{\partial t}\right]_- = \left[\hat{\psi}_\alpha(\mathbf{x},t), \frac{1}{\epsilon_0}\,\hat{\pi}_k(\mathbf{x}',t)\right]_- - \left[\hat{\psi}_\alpha(\mathbf{x},t), \partial'_k \hat{\varphi}(\mathbf{x}',t)\right]_- \tag{11.149}$$
$$= -\frac{q}{4\pi\epsilon_0}\partial'_k \int d^3x''\,\frac{1}{|\mathbf{x}'-\mathbf{x}''|}\,\left[\hat{\psi}_\alpha(\mathbf{x},t), \hat{\psi}^\dagger_\beta(\mathbf{x}'',t)\hat{\psi}_\beta(\mathbf{x}'',t)\right]_- .$$

Applying the operator identity (3.94) and (11.141) this reduces to

$$\left[\hat{\psi}_\alpha(\mathbf{x},t), \frac{\partial \hat{A}_k(\mathbf{x}',t)}{\partial t}\right]_- = \frac{q}{4\pi\epsilon_0}\left[\partial_k\,\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right]\hat{\psi}_\alpha(\mathbf{x},t) = -\frac{q}{4\pi\epsilon_0}\,\frac{(\mathbf{x}-\mathbf{x}')_k}{|\mathbf{x}-\mathbf{x}'|^3}\,\hat{\psi}_\alpha(\mathbf{x},t)\,. \tag{11.150}$$

Similarly we also yield

$$\left[\hat{\psi}^\dagger_\alpha(\mathbf{x},t), \frac{\partial \hat{A}_k(\mathbf{x}',t)}{\partial t}\right]_- = -\frac{q}{4\pi\epsilon_0}\,\frac{(\mathbf{x}-\mathbf{x}')_k}{|\mathbf{x}-\mathbf{x}'|^3}\,\hat{\psi}^\dagger_\alpha(\mathbf{x},t)\,. \tag{11.151}$$

Note that the non-locality of the commutator relations (11.150) and (11.151) is typical for the Coulomb gauge used here. Finally, we also convert (11.142) correspondingly. At first we get

$$\left[\hat{A}_k(\mathbf{x},t), \hat{A}_l(\mathbf{x}',t)\right]_- = 0 \tag{11.152}$$

and then we take into account (11.142), (11.145), and (11.146) in order to yield

$$\left[\hat{A}_k(\mathbf{x},t), \frac{\partial \hat{A}_l(\mathbf{x}',t)}{\partial t}\right]_- = \left[\hat{A}_k(\mathbf{x},t), \frac{1}{\epsilon_0}\,\hat{\pi}_l(\mathbf{x}',t)\right]_- - \left[\hat{A}_k(\mathbf{x},t), \partial'_l \hat{\varphi}_l(\mathbf{x}',t)\right]_-$$
$$= \frac{i\hbar}{\epsilon_0}\,\delta^T_{kl}(\mathbf{x}-\mathbf{x}') - \partial'_l \int d^3x''\,\frac{q}{4\pi\epsilon_0|\mathbf{x}'-\mathbf{x}''|}\,\left[\hat{A}_k(\mathbf{x},t), \hat{\psi}^\dagger_\alpha(\mathbf{x}'',t)\hat{\psi}_\alpha(\mathbf{x}'',t)\right]_- . \tag{11.153}$$

Applying (3.43) and (11.148) this reduces finally to

$$\left[\hat{A}_k(\mathbf{x},t), \frac{\partial \hat{A}_l(\mathbf{x}',t)}{\partial t}\right]_- = \frac{i\hbar}{\epsilon_0}\,\delta^T_{kl}(\mathbf{x}-\mathbf{x}')\,. \tag{11.154}$$

In the same way we also obtain

$$\left[\frac{\partial \hat{A}_k(\mathbf{x},t)}{\partial t}, \frac{\partial \hat{A}_l(\mathbf{x}',t)}{\partial t}\right]_- = \frac{1}{\epsilon_0^2}\left[\hat{\pi}_k(\mathbf{x},t), \hat{\pi}_l(\mathbf{x}',t)\right]_- + \partial_k \int d^3x''\,\frac{q}{4\pi\epsilon_0^2|\mathbf{x}-\mathbf{x}''|} \tag{11.155}$$
$$\times\left[\hat{\pi}_l(\mathbf{x}',t), \hat{\psi}^\dagger_\alpha(\mathbf{x}'',t)\hat{\psi}_\alpha(\mathbf{x}'',t)\right]_- - \partial'_l \int d^3x'''\frac{q}{4\pi\epsilon_0^2|\mathbf{x}'-\mathbf{x}'''|}\left[\hat{\pi}_k(\mathbf{x},t), \hat{\psi}^\dagger_\beta(\mathbf{x}''',t)\hat{\psi}_\beta(\mathbf{x}''',t)\right]_-$$
$$+\partial_k\partial'_l \int d^3x'' \int d^3x'''\frac{q}{4\pi\epsilon_0^2|\mathbf{x}-\mathbf{x}''|}\,\frac{q}{4\pi\epsilon_0^2|\mathbf{x}'-\mathbf{x}'''|}\left[\hat{\psi}^\dagger_\alpha(\mathbf{x}'',t)\hat{\psi}_\alpha(\mathbf{x}'',t), \hat{\psi}^\dagger_\beta(\mathbf{x}''',t)\hat{\psi}_\beta(\mathbf{x}''',t)\right]_- .$$

Thus, finally, after applying (3.43), (11.142), (11.147), and (11.148) we end up with

$$\left[\frac{\partial \hat{A}_k(\mathbf{x},t)}{\partial t}, \frac{\partial \hat{A}_l(\mathbf{x}',t)}{\partial t}\right]_- = 0\,. \tag{11.156}$$

In the canonical field quantisation in the Dirac picture, the dynamics of the state vectors is determined according to (11.119) by the interacting part of the Hamilton operator. In spinor quantum electrodynamics it consists of two parts due to (11.111):

$$\begin{aligned}\hat{H}_D^{(\text{int})}(t) &= -qc\int d^3x : \hat{\bar{\psi}}(\mathbf{x},t)\boldsymbol{\gamma}\,\hat{\psi}(\mathbf{x},t) : \hat{\mathbf{A}}(\mathbf{x},t) \\ &\quad + \frac{q^2}{8\pi\epsilon_0}\int d^3x \int d^3x' \,\frac{: \hat{\bar{\psi}}(\mathbf{x},t)\gamma^0\hat{\psi}(\mathbf{x},t)\hat{\bar{\psi}}(\mathbf{x}',t)\gamma^0\hat{\psi}(\mathbf{x}',t) :}{|\mathbf{x}-\mathbf{x}'|}\,. \end{aligned} \tag{11.157}$$

Here, the normal ordering of the field operators was additionally used.

11.6 Time Evolution Operator

In the Dirac picture the interaction affects the dynamics of the state vectors according to (11.119). In order to investigate this in more detail we introduce the time evolution operator $\hat{U}(t_2, t_1)$, which connects the state vectors $|\psi_D(t_1)\rangle$ and $|\psi_D(t_2)\rangle$ at two consecutive times t_1 and t_2, respectively:

$$|\psi_D(t_2)\rangle = \hat{U}(t_2, t_1)|\psi_D(t_1)\rangle\,. \tag{11.158}$$

With the help of (11.114) and the formal solution of the Schrödinger equation (11.118)

$$|\psi_S(t_2)\rangle = e^{-i\hat{H}_S(t_2-t_1)/\hbar}|\psi_S(t_1)\rangle \tag{11.159}$$

we conclude

$$\begin{aligned}|\psi_D(t_2)\rangle &= e^{i\hat{H}_S^{(0)}t_2/\hbar}\,|\psi_S(t_2)\rangle = e^{i\hat{H}_S^{(0)}t_2/\hbar}\,e^{-i\hat{H}_S(t_2-t_1)/\hbar}\,|\psi_S(t_1)\rangle \\ &= e^{i\hat{H}_S^{(0)}t_2/\hbar}e^{-i\hat{H}_S(t_2-t_1)/\hbar}e^{-i\hat{H}_S^{(0)}t_1/\hbar}|\psi_D(t_1)\rangle\,. \end{aligned} \tag{11.160}$$

Thus, a comparison with (11.158) leads to a formal expression for the time evolution operator $\hat{U}(t_2, t_1)$:

$$\hat{U}(t_2, t_1) = e^{i\hat{H}_S^{(0)}t_2/\hbar}\,e^{-i\hat{H}_S(t_2-t_1)/\hbar}\,e^{-i\hat{H}_S^{(0)}t_1/\hbar}\,. \tag{11.161}$$

Since the Hamilton operators $\hat{H}_S^{(0)}$ and $\hat{H}_S$ generally do not commute with each other, it is important to take into account the particular operator ordering in (11.161). With the help of the formal expression (11.161), various properties of the time evolution operator can be proved. It has the initial condition

$$\hat{U}(t_1, t_1) = 1 \tag{11.162}$$

and fulfills the group property

$$\hat{U}(t_3, t_2)\hat{U}(t_2, t_1) = \hat{U}(t_3, t_1)\,. \tag{11.163}$$

Indeed, we obtain from applying (11.161)

$$\begin{aligned}\hat{U}(t_3,t_2)\hat{U}(t_2,t_1) &= e^{i\hat{H}_S^{(0)}t_3/\hbar}\,e^{-i\hat{H}_S(t_3-t_2)/\hbar}\,e^{-i\hat{H}_S^{(0)}t_2/\hbar}\,e^{i\hat{H}_S^{(0)}t_2/\hbar}\,e^{-i\hat{H}_S(t_2-t_1)/\hbar}\,e^{-i\hat{H}_S^{(0)}t_1/\hbar}\\ &= e^{i\hat{H}_S^{(0)}t_3/\hbar}\,e^{-i\hat{H}_S(t_3-t_1)/\hbar}\,e^{-i\hat{H}_S^{(0)}t_1/\hbar} = \hat{U}(t_3,t_1)\,. \end{aligned} \tag{11.164}$$

Furthermore, we read off from evaluating (11.163) for $t_3 = t_1$ together with (11.162) the inverse time evolution operator

$$\hat{U}^{-1}(t_2,t_1) = \hat{U}(t_1,t_2)\,. \tag{11.165}$$

And we deduce from (11.161) and (11.165) that the time evolution operator is unitary:

$$\hat{U}^{\dagger}(t_2,t_1) = e^{i\hat{H}_S^{(0)}t_1/\hbar}\,e^{-i\hat{H}_S(t_1-t_2)/\hbar}\,e^{-i\hat{H}_S^{(0)}t_2/\hbar} = \hat{U}(t_1,t_2) = \hat{U}^{-1}(t_2,t_1)\,. \tag{11.166}$$

Finally, we determine which differential equation the time evolution operator $\hat{U}(t_2,t_1)$ solves. Differentiating (11.161) with respect to t_2 and taking into account (11.113) yields

$$i\hbar\frac{\partial}{\partial t_2}\,\hat{U}^{\dagger}(t_2,t_1) = e^{i\hat{H}_S^{(0)}t_2/\hbar}\,\hat{H}_S^{(\mathrm{int})}\,e^{-i\hat{H}_S^{(0)}t_2/\hbar}\,e^{i\hat{H}_S^{(0)}t_2/\hbar}\,e^{-i\hat{H}_S(t_2-t_1)/\hbar}\,e^{-i\hat{H}_S^{(0)}t_1/\hbar}\,. \tag{11.167}$$

Thus, we conclude from (11.120), (11.161), and (11.167) that $\hat{U}(t_2,t_1)$ fulfills the differential equation

$$i\hbar\frac{\partial}{\partial t_2}\,\hat{U}(t_2,t_1) = \hat{H}_D^{(\mathrm{int})}(t_2)\hat{U}(t_2,t_1)\,. \tag{11.168}$$

The initial value problem (11.162) and (11.168) can be formally rewritten in form of an integral equation:

$$\hat{U}(t_2,t_1) = 1 - \frac{i}{\hbar}\int_{t_1}^{t_2} dt_1'\,\hat{H}_D^{(\mathrm{int})}(t_1')\,\hat{U}(t_1',t_1)\,. \tag{11.169}$$

Successively reinserting the left-hand side of (11.169) into the right-hand side, one obtains the von Neumann series

$$\begin{aligned}\hat{U}(t_2,t_1) = 1 &- \frac{i}{\hbar}\int_{t_1}^{t_2} dt_1'\,\hat{H}_D^{(\mathrm{int})}(t_1') + \left(\frac{-i}{\hbar}\right)^2\int_{t_1}^{t_2} dt_1'\int_{t_1}^{t_1'} dt_2'\,\hat{H}_D^{(\mathrm{int})}(t_1')\,\hat{H}_D^{(\mathrm{int})}(t_2') + \ldots\\ &+ \left(\frac{-i}{\hbar}\right)^n\int_{t_1}^{t_2} dt_1'\int_{t_1}^{t_1'} dt_2'\ldots\int_{t_1}^{t_{n-1}'} dt_n'\,\hat{H}_D^{(\mathrm{int})}(t_1')\,\hat{H}_D^{(\mathrm{int})}(t_2')\cdots\hat{H}_D^{(\mathrm{int})}(t_n') + \ldots\,. \end{aligned} \tag{11.170}$$

It is noticeable in the nth summand of the von Neumann series that the time arguments of the multiple integrals are ordered in decreasing order: $t_1' > t_2' > \ldots > t_n'$. According to an idea of Freeman Dyson, all n integrals can be rewritten such that they are all performed over the same interval $[t_1,t_2]$ by using the time-ordered product of operators. Although the time ordering of operators has already been introduced previously for calculating the propagators of the Klein-Gordon field, the Maxwell field, and the Dirac field in the Chapters 8–10, its original

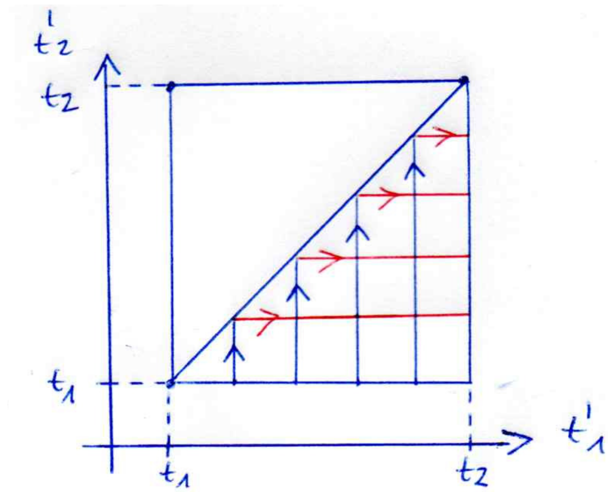


Figure 11.1: The hatched triangle can be integrated in two ways, which allows to rearrange the integral (11.171).

motivation becomes apparent only now. To this end we consider exemplarily the second term in the von Neumann series (11.170) and reorganize it as follows:

$$\int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_1'} dt_2' \, \hat{H}_D^{(\mathrm{int})}(t_1') \, \hat{H}_D^{(\mathrm{int})}(t_2') = \int_{t_1}^{t_2} dt_2' \int_{t_2'}^{t_2} dt_1' \, \hat{H}_D^{(\mathrm{int})}(t_1') \, \hat{H}_D^{(\mathrm{int})}(t_2') \,. \tag{11.171}$$

Here we use the fact that the hatched triangle in Fig. 11.1 can be integrated in two ways. Either we first integrate over t_2' and then over t_1' or, conversely, first over t_1' and then over t_2'. Exchanging both integration variables at the right-hand side of (11.171) we conclude

$$\begin{aligned} 2\int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_1'} dt_2' \, \hat{H}_D^{(\mathrm{int})}(t_1') \, \hat{H}_D^{(\mathrm{int})}(t_2') &= \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_1'} dt_2' \, \hat{H}_D^{(\mathrm{int})}(t_1') \, \hat{H}_D^{(\mathrm{int})}(t_2') \\ + \int_{t_1}^{t_2} dt_1' \int_{t_1'}^{t_2} dt_2' \, \hat{H}_D^{(\mathrm{int})}(t_2') \, \hat{H}_D^{(\mathrm{int})}(t_1') &= \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_2} dt_1' \, \Theta(t_1' - t_2') \, \hat{H}_D^{(\mathrm{int})}(t_1') \, \hat{H}_D^{(\mathrm{int})}(t_2') \\ + \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_2} dt_1' \, \Theta(t_2' - t_1') \, \hat{H}_D^{(\mathrm{int})}(t_2') \, \hat{H}_D^{(\mathrm{int})}(t_1') &= \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_2} dt_1' \, \hat{T}\left(\hat{H}_D^{(\mathrm{int})}(t_1') \, \hat{H}_D^{(\mathrm{int})}(t_2') \right) . \end{aligned} \tag{11.172}$$

In the last step we assumed that the interacting Hamilton operator in the Dirac picture $\hat{H}_D^{(\mathrm{int})}(t)$ is bosonic, so the time ordering was used for two bosonic operators whose time order is not yet fixed:

$$\hat{T}\left(\hat{H}_D^{(\mathrm{int})}(t_1') \hat{H}_D^{(\mathrm{int})}(t_2') \right) = \Theta(t_1' - t_2') \hat{H}_D^{(\mathrm{int})}(t_1') \hat{H}_D^{(\mathrm{int})}(t_2') + \Theta(t_2' - t_1') \hat{H}_D^{(\mathrm{int})}(t_2') \hat{H}_D^{(\mathrm{int})}(t_1') \,. \tag{11.173}$$

Analogous to (11.172), also all other terms in the von Neumann series (11.170) can be rewritten as multiple integrals over the entire interval $[t_1, t_2]$ with the help of the time-ordered product of operators. In the case of the nth-order term, one has to take into account in total $n!$

permutations of the time arguments. Therefore the generalisation of (11.172) reads

$$
\begin{aligned}
& n! \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_1'} dt_2' \cdots \int_{t_1}^{t_{n-1}'} dt_n' \, \hat{H}_D^{(\mathrm{int})}(t_1')\, \hat{H}_D^{(\mathrm{int})}(t_2') \cdots \hat{H}_D^{(\mathrm{int})}(t_n') \\
& = \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_2} dt_2' \cdots \int_{t_1}^{t_2} dt_n' \, \hat{T} \left(\hat{H}_D^{(\mathrm{int})}(t_1')\, \hat{H}_D^{(\mathrm{int})}(t_2') \cdots \hat{H}_D^{(\mathrm{int})}(t_n') \right) . \qquad (11.174)
\end{aligned}
$$

This result can be proven by complete induction. With the help of (11.174) the von Neumann series (11.170) for the time evolution operator is finally given by

$$
\hat{U}(t_2, t_1) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-i}{\hbar} \right)^n \int_{t_1}^{t_2} dt_1' \cdots \int_{t_1}^{t_2} dt_n' \, \hat{T} \left(\hat{H}_D^{(\mathrm{int})}(t_1') \cdots \hat{H}_D^{(\mathrm{int})}(t_n') \right) . \qquad (11.175)
$$

We can explicitly verify that the von Neumann series (11.175) solves the differential equation (11.168). Differentiating (11.175) with respect to t_2 we obtain due to the symmetry of the integrand with respect to the integration variables $t_1', t_2', ..., t_n'$:

$$
\begin{aligned}
i\hbar \frac{\partial}{\partial t_2} \hat{U}(t_2, t_1) \quad = \quad & \sum_{n=1}^{\infty} \frac{i\hbar}{n!} \left(\frac{-i}{\hbar} \right)^n n \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_2} dt_2' \cdots \int_{t_1}^{t_2} dt_{n-1}' \\
& \times \hat{T} \left(\hat{H}_D^{(\mathrm{int})}(t_1')\, \hat{H}_D^{(\mathrm{int})}(t_2') \cdots \hat{H}_D^{(\mathrm{int})}(t_{n-1}') \hat{H}_D^{(\mathrm{int})}(t_2) \right) . \qquad (11.176)
\end{aligned}
$$

Due to the fact that the time t_2 is larger than all remaining integration variables t_1', t_2', ..., t_{n-1}' and using the definition (11.173) of the time-ordered product of operators, one can pull the operator $\hat{H}_D^{(\mathrm{int})}(t_2)$ out of the time ordering and obtain together with (11.175)

$$
\begin{aligned}
& i\hbar \frac{\partial}{\partial t_2} \hat{U}(t_2, t_1) = \hat{H}_D^{(\mathrm{int})}(t_2') \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \left(\frac{-i}{\hbar} \right)^{n-1} \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_2} dt_2' \cdots \int_{t_1}^{t_2} dt_{n-1}' \\
& \times \hat{T} \left(\hat{H}_D^{(\mathrm{int})}(t_1') \hat{H}_D^{(\mathrm{int})}(t_2') \cdots \hat{H}_D^{(\mathrm{int})}(t_{n-1}') \right) = \hat{H}_D^{(\mathrm{int})}(t_2) \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-i}{\hbar} \right)^n \int_{t_1}^{t_2} dt_1' \int_{t_1}^{t_2} dt_2' \cdots \int_{t_1}^{t_2} dt_n' \\
& \times \hat{T} \left(\hat{H}_D^{(\mathrm{int})}(t_1') \hat{H}_D^{(\mathrm{int})}(t_2') \cdots \hat{H}_D^{(\mathrm{int})}(t_n') \right) \; = \; \hat{H}_D^{(\mathrm{int})}(t_2) \hat{U}(t_2, t_1) . \qquad (11.177)
\end{aligned}
$$

Formally, the von Neumann series (11.175) can be summed up to a time-ordered exponential function:

$$
\hat{U}(t_2, t_1) \; = \; \hat{T} \, \exp \left\{ \frac{-i}{\hbar} \int_{t_1}^{t_2} dt \, \hat{H}_D^{(\mathrm{int})}(t) \right\} . \qquad (11.178)
$$

By taking into account that the time evolution operator (11.178) is defined by the von Neumann series (11.175) one can calculate perturbatively the cross sections of scattering processes.

11.7 Scattering Operator

We now consider a generic scenario for a scattering problem in the realm of relativistic quantum field theory. To this end we denote with $|\psi(t)\rangle$ a time-dependent state vector, which evolves starting from an initial state $|\psi_i\rangle$ in the limit $t \to -\infty$:

$$
|\psi(-\infty)\rangle = |\psi_i\rangle . \qquad (11.179)
$$

The time evolution of the state vector $|\psi(t)\rangle$ under the influence of the interaction is determined in the Dirac picture by the time evolution operator $\hat{U}(t,-\infty)$ according to (11.158):

$$|\psi(t)\rangle = \hat{U}(t,-\infty)|\psi_i\rangle \,. \tag{11.180}$$

The scattering matrix S_{fi} denotes then the projection of the state vector $|\psi(t)\rangle$ in the limit $t \to +\infty$ onto the final state $|\psi_f\rangle$:

$$S_{fi} = \lim_{t\to+\infty} \langle\psi_f|\psi(t)\rangle \,. \tag{11.181}$$

From the knowledge of the scattering matrix (11.181) all observable quantities such as the scattering cross sections and decay rates can be calculated from the square of its absolute values and some kinetic considerations. According to (11.180) and (11.181), the probability amplitude S_{fi} for the transition from $|\psi_i\rangle$ to $|\psi_f\rangle$ can also be calculated as the matrix element

$$S_{fi} = \langle\psi_f|\hat{S}|\psi_i\rangle \tag{11.182}$$

of the scattering operator

$$\hat{S} = \hat{U}(+\infty,-\infty)\,. \tag{11.183}$$

According to (11.178) the scattering operator is explicitly given by

$$\hat{S} = \hat{T}\,\exp\left\{\frac{-i}{\hbar}\int_{-\infty}^{+\infty} dt\,\hat{H}_D^{(\mathrm{int})}(t)\right\}\,. \tag{11.184}$$

In spinor quantum electrodynamics, the scattering operator reads according to (11.157) and (11.184):

$$\begin{aligned}\hat{S} &= \hat{T}\exp\left\{\frac{iq}{\hbar}\int d^4x : \hat{\overline{\psi}}(x)\boldsymbol{\gamma}\hat{\psi}(x) : \hat{\mathbf{A}}(x)\right.\\ &\left.-\frac{iq^2}{8\pi\hbar\epsilon_0}\int dt\int d^3x\int d^3x'\,\frac{:\hat{\overline{\psi}}(\mathbf{x},t)\gamma^0\hat{\psi}(\mathbf{x},t)\hat{\overline{\psi}}(\mathbf{x}',t)\gamma^0\hat{\psi}(\mathbf{x}',t):}{|\mathbf{x}-\mathbf{x}'|}\right\}\,.\end{aligned} \tag{11.185}$$

Expanding the scattering operator up to the second order in the charge q, we obtain:

$$\begin{aligned}\hat{S} &= 1 + \frac{iq}{\hbar}\int d^4x : \hat{\overline{\psi}}(x)\boldsymbol{\gamma}\hat{\psi}(x) : \hat{\mathbf{A}}(x)\\ &-\frac{iq^2}{8\pi\hbar\epsilon_0}\int dt\int d^3x\int d^3x'\,\frac{:\hat{\overline{\psi}}(\mathbf{x},t)\gamma^0\hat{\psi}(\mathbf{x},t)\hat{\overline{\psi}}(\mathbf{x}',t)\gamma^0\hat{\psi}(\mathbf{x}',t):}{|\mathbf{x}-\mathbf{x}'|}\\ &+\frac{1}{2}\left(\frac{iq}{\hbar}\right)^2\int d^4x\int d^4x'\,\hat{T}\left\{\left[:\hat{\overline{\psi}}(x)\boldsymbol{\gamma}\overline{\psi}(x) : \hat{\mathbf{A}}(x)\right]\left[:\hat{\overline{\psi}}(x')\boldsymbol{\gamma}\hat{\psi}(x') \;: \hat{\mathbf{A}}(x')\right]\right\} + \ldots\,.\end{aligned} \tag{11.186}$$

We summarize that (11.182) and (11.186) represent the starting point for determining the cross sections of scattering processes in the realm of spinor quantum electrodynamics.

Chapter 12

Møller Scattering

In the last chapter we apply our previous findings in order to calculate the cross section for the concrete example of an elastic scattering of two electrons:

$$e^- e^- \to e^- e^- . \tag{12.1}$$

This represents a paradigmatic scattering process in quantum field theory, which is named after the Danish physicist Christian Møller. The interaction between two electrons, that is idealized in the Møller scattering, forms the theoretical basis of many familiar physical phenomena such as, for instance, the repulsion between the two electrons of the helium atom. Furthermore, Møller scattering is a fundamental, purely pointlike process in quantum electrodynamics, which provides an important means to test the standard model of elementary particle physics. In addition, it is the dominant physical process in low-energy (< 100 MeV) electron scattering experiments. Thus, it is an important constraint in the design of electron scattering experiments that search for new physics beyond the standard model.

First we apply the perturbative technique worked out in Chapter 11 and determine the scattering matrix in the leading non-vanishing order, which turns out to be the quadratic one. Due to an intriguing cancellation of non-covariant terms the result is finally manifestly covariant and consists of two expressions. Taking into account the Feyman rules these two analytic expressions can be graphically represented in terms of Feynman diagrams. Secondly, we assume that the polarization is unknown for both the initial and the final electrons. This allows to average the square of the scattering matrix with respect to the polarizations of the involved electrons. The corresponding evaluation is quite technical and relies basically on the Clifford algebra of the Dirac matrices. Thirdly we analyze in detail the kinematics of such a two-particle scattering process by introducing the Lorentz-invariant Mandelstam variables. In particular, we specialize the relativistic scattering problem for two particles to the center of mass reference frame. This allows to express the Mandelstam variables just in terms of the scattering energy and the scattering angle. And, finally, we determine the scattering cross section for the Møller scattering and discuss both the ultra-relativistic and the non-relativistic limit. In the latter case we find that the Rutherford scattering formula is recovered for the forward peak.

12.1 Scattering Matrix

In the case of Møller scattering, one investigates a scattering process, where two electrons in the initial state

$$|\psi_i\rangle = |\mathbf{p}_{i_1}, s_{i_1}; \mathbf{p}_{i_2}, s_{i_2}\rangle \tag{12.2}$$

change into two electrons in the final state

$$|\psi_f\rangle = |\mathbf{p}_{f_1}, s_{f_1}; \mathbf{p}_{f_2}, s_{f_2}\rangle \,. \tag{12.3}$$

In the following we determine the matrix element of the scattering operator (11.186) up to the second order in the charge $q = -e$ with respect to the initial state (12.2) and the final state (12.3) according to (11.182). We observe that the zeroth order vanishes, since both states are orthogonal to each other for different momenta $\mathbf{p}_{i_1}, \mathbf{p}_{i_2} \neq \mathbf{p}_{f_1}, \mathbf{p}_{f_2}$:

$$\langle \mathbf{p}_{f_1}, s_{f_1}; \mathbf{p}_{f_2}, s_{f_2}|\mathbf{p}_{i_1}, s_{i_1}; \mathbf{p}_{i_2}, s_{i_2}\rangle = 0 \,. \tag{12.4}$$

Furthermore, also the first order disappears, since both the initial and the final state (12.2) and (12.3) do not contain any photon and the first-order term in the scattering operator (11.186) involves the operator of the vector potential, whose plane wave decomposition (9.155) contains the annihilation and the creation of a photon. Therefore, the lowest non-vanishing perturbative order is the quadratic one, which turns out to consist of two contributions:

$$S_{fi}^{(2)} = S_{fi}^{(2,\text{inst})} + S_{fi}^{(2,\text{rad})} \,. \tag{12.5}$$

The first contribution stems from the instantaneous Coulomb self-interaction of the Dirac field

$$S_{fi}^{(2,\text{inst})} = \frac{-ie^2}{8\pi\hbar\epsilon_0 c^2} \int dt \int d^3x \int d^3x' \, \frac{\langle\psi_f| : \hat{j}^0(\mathbf{x},t)\hat{j}^0(\mathbf{x}',t) : |\psi_i\rangle}{|\mathbf{x}-\mathbf{x}'|} \,, \tag{12.6}$$

while the second contribution represents an interaction between the Dirac and the Maxwell field:

$$S_{fi}^{(2,\text{rad})} = -\frac{e^2}{2\hbar^2 c^2} \int d^4x \int d^4x' \, \langle\psi_f|\hat{T}\{ : \hat{j}^k(x)\hat{A}_k(x) : : \hat{j}^l(x')\hat{A}_l(x') : \} |\psi_i\rangle \,. \tag{12.7}$$

Note that in (12.6) the time-like and in (12.7) the space-like components of the four-vector current density operator (11.91) occur, respectively:

$$\hat{j}^\mu(x) = c\hat{\overline{\psi}}(x)\gamma^\mu\hat{\psi}(x) \,. \tag{12.8}$$

Here we take into account the plane wave decompositions of the spinor field operators (10.433) and (10.434), which we rewrite according to

$$\hat{\overline{\psi}}(x) = \int d^3p_2 \sum_{s_2} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_2}}} \left\{ e^{ip_2x/\hbar}\, \overline{u}(\mathbf{p}_2, s_2)\hat{b}^\dagger_{\mathbf{s}_2,s_2} + e^{-ip_2x/\hbar}\, \overline{v}(\mathbf{p}_2, s_2)\hat{d}_{\mathbf{p}_2,s_2} \right\} , \tag{12.9}$$

$$\hat{\psi}(x) = \int d^3p_1 \sum_{s_1} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_1}}} \left\{ e^{-ip_1x/\hbar}\, u(\mathbf{p}_1, s_1)\hat{b}_{\mathbf{p}_1,s_1} + e^{ip_1x/\hbar}\, v(\mathbf{p}_1, s_1)\hat{d}^\dagger_{\mathbf{p}_1,s_1} \right\} , \tag{12.10}$$

where $s = \pm 1/2$ denotes the helicity. With this one obtains for the four-vector current density operator (12.8) the decomposition

$$\hat{j}^\mu(x) = c \int d^3p_1 \int d^3p_2 \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_1}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_2}}} \left\{ e^{i(p_2-p_1)x/\hbar}\, \overline{u}(\mathbf{p}_2, s_2)\gamma^\mu u(\mathbf{p}_1, s_1) \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{b}_{\mathbf{p}_1,s_1} \right.$$
$$+ e^{i(p_2+p_1)x/\hbar}\, \overline{u}(\mathbf{p}_2, s_2)\gamma^\mu v(\mathbf{p}_1, s_1) \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{d}^\dagger_{\mathbf{p}_1,s_1} + e^{-i(p_1+p_2)x/\hbar}\, \overline{v}(\mathbf{p}_2, s_2)\gamma^\mu u(\mathbf{p}_1, s_1) \hat{d}_{\mathbf{p}_2,s_2} \hat{b}_{\mathbf{p}_1,s_1}$$
$$\left. + e^{-i(p_2-p_1)x/\hbar}\, \overline{v}(\mathbf{p}_2, s_2)\gamma^\mu v(\mathbf{p}_1, s_1) \hat{d}_{\mathbf{p}_2,s_2} \hat{d}^\dagger_{\mathbf{p}_1,s_1} \right\} . \tag{12.11}$$

Evaluating the matrix element of the normal ordered operator $:\hat{j}^0(\mathbf{x},t)\hat{j}^0(\mathbf{x}',t):$ with the states

$$\langle \mathbf{p}_{f_1}, s_{f_1}; \mathbf{p}_{f_2}, s_{f_2}| = \langle 0|\, \hat{b}_{\mathbf{p}_{f_2},s_{f_2}} \hat{b}_{\mathbf{p}_{f_1},s_{f_1}} , \tag{12.12}$$
$$|\mathbf{p}_{i_1}, s_{i_1}; \mathbf{p}_{i_2}, s_{i_2}\rangle = \hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}} \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}} |0\rangle , \tag{12.13}$$

then only the first summand in (12.11) leads to a non-vanishing contribution. For the instantaneous self-interaction of the Dirac field (12.6) this results in

$$S_{fi}^{(2,\text{inst})} = \frac{-ie^2}{8\pi\hbar\epsilon_0 c^2} \int dt \int d^3x \int d^3x' \frac{1}{|\mathbf{x}-\mathbf{x}'|} \int d^3p_1 \int d^3p_2 \int d^3p_3 \int d^3p_4 \sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4}$$
$$\times \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_1}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_2}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_3}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_4}}}\, e^{i\left(E_{\mathbf{p}_2}-E_{\mathbf{p}_1}\right)t/\hbar} e^{-i(\mathbf{p}_2-\mathbf{p}_1)\mathbf{x}/\hbar}$$
$$\times \overline{u}(\mathbf{p}_2, s_2)\gamma^0 u(\mathbf{p}_1, s_1)\, e^{i\left(E_{\mathbf{p}_4}-E_{\mathbf{p}_3}\right)t/\hbar} e^{-i(\mathbf{p}_4-\mathbf{p}_3)\mathbf{x}'/\hbar}\, \overline{u}(\mathbf{p}_4, s_4)\gamma^0 u(\mathbf{p}_3, s_3)$$
$$\times C(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4; s_1, s_2, s_3, s_4) . \tag{12.14}$$

Here we have introduced a vacuum expectation value of creation and annihilation operators as an abbreviation:

$$C(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4; s_1, s_2, s_3, s_4) = \langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}} \hat{b}_{\mathbf{p}_{f_1},s_{f_1}} : \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{b}_{\mathbf{p}_1,s_1} \hat{b}^\dagger_{\mathbf{p}_4,s_4} \hat{b}_{\mathbf{p}_3,s_3} : \hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}} \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}} |0\rangle$$
$$= -\langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}} \hat{b}_{\mathbf{p}_{f_1},s_{f_1}} \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{b}^\dagger_{\mathbf{p}_4,s_4} \hat{b}_{\mathbf{p}_1,s_1} \hat{b}_{\mathbf{p}_3,s_3} \hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}} \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}} |0\rangle , \tag{12.15}$$

where the evaluation of the normal ordering led to a minus sign due to the anti-commutator algebra of the fermionic operators (10.407). Afterwards, we evaluate the interaction (12.7) between the Dirac and the Maxwell fields. Here we use the bosonic definition of the time-ordering operator (8.123) and note that the operators $\hat{j}^k(x)$ and $\hat{A}_k(x)$ interchange with each other. Furthermore, taking into account the initial and the final state defined according to (12.2), (12.3), (12.12), and (12.13) yields

$$S_{fi}^{(2,\text{rad})} = -\frac{e^2}{2\hbar^2 c^2} \int d^4x \int d^4x' \left\{ \Theta(x^0 - x'^0)\, \langle \psi_f| : \hat{j}^k(x)\hat{A}_k(x) : : \hat{j}^l(x')\hat{A}_l(x')|\psi_i\rangle \right.$$
$$\left. + \Theta(x'^0 - x^0)\, \langle \psi_f| : \hat{j}^l(x')\hat{A}_l(x') : : \hat{j}^k(x)\hat{A}_k(x) : |\psi_i\rangle \right\} = -\frac{e^2}{2\hbar^2 c^2} \int d^4x \int d^4x'$$
$$\times \left\{ \Theta(x^0 - x'^0)\, \langle 0|\hat{A}_\mu(x)\hat{A}_\nu(x')|0\rangle\, \langle \mathbf{p}_{f_1}, s_{f_1}; \mathbf{p}_{f_2}, s_{f_2}| : \hat{j}^\mu(x) : : \hat{j}^\nu(x') : |\mathbf{p}_{i_1}, s_{i_1}; \mathbf{p}_{i_2}, s_{i_2}\rangle \right.$$
$$\left. + \Theta(x'^0 - x^0)\, \langle 0|\hat{A}_\nu(x')\hat{A}_\mu(x)|0\rangle\, \langle \mathbf{p}_{f_1}, s_{f_1}; \mathbf{p}_{f_2}, s_{f_2}| : \hat{j}^\nu(x') : : \hat{j}^\mu(x) : |\mathbf{p}_{i_1}, s_{i_1}; \mathbf{p}_{i_2}, s_{i_2}\rangle \right\} . \tag{12.16}$$

In the last step, we replaced the summations over the spatial indices k, l by summations over the spatio-temporal indices μ, ν, since we have $\hat{A}_0(x) = 0$ in the radiation gauge. The normal ordering of the four-current density operator (12.11) leads to

$$
\begin{aligned}
&: \hat{j}^\mu(x) := c \int d^3p_1 \int d^3p_2 \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_1}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_2}}} \qquad (12.17)\\
&\times \Big\{ e^{i(p_2-p_1)x/\hbar}\, \overline{u}(\mathbf{p}_2, s_2)\gamma^\mu u(\mathbf{p}_1, s_1) \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{b}_{\mathbf{p}_1,s_1} + e^{i(p_2+p_1)x/\hbar}\, \overline{u}(\mathbf{p}_2, s_2)\gamma^\mu v(\mathbf{p}_1, s_1) \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{d}^\dagger_{\mathbf{p}_1,s_1} \\
&+ e^{-i(p_1+p_2)x/\hbar}\, \overline{v}(\mathbf{p}_2, s_2)\gamma^\mu u(\mathbf{p}_1, s_1) \hat{d}_{\mathbf{p}_2,s_2} \hat{b}_{\mathbf{p}_1,s_1} - e^{-i(p_2-p_1)x/\hbar}\, \overline{v}(\mathbf{p}_2, s_2)\gamma^\mu v(\mathbf{p}_1, s_1) \hat{d}^\dagger_{\mathbf{p}_1,s_1} \hat{d}_{\mathbf{p}_2,s_2} \Big\} .
\end{aligned}
$$

Note that the normal ordering affected only the last term by changing its sign. Evaluating the matrix element for the product of two normally ordered four-vector current density operators $: \hat{j}^\mu(x) : : \hat{j}^\nu(x) :$ with the states (12.12) and (12.13), then only the first summand in (12.17) leads in both cases to a non-vanishing contribution:

$$
\begin{aligned}
&\langle \mathbf{p}_{f_1}, s_{f_1}; \mathbf{p}_{f_2}, s_{f_2} | : \hat{j}^\mu(x) : : \hat{j}^\nu(x') : | \mathbf{p}_{i_1}, s_{i_1}; \mathbf{p}_{i_2}, s_{i_2} \rangle = c^2 \int d^3p_1 \int d^3p_2 \int d^3p_3 \int d^3p_4 \\
&\times \sum_{s_1} \sum_{s_2} \sum_{s_3} \sum_{s_4} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_1}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_2}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_3}}} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_4}}}\, e^{i(p_2-p_1)x/\hbar} \\
&\times \overline{u}(\mathbf{p}_2, s_2)\gamma^\mu u(\mathbf{p}_1, s_1)\, e^{i(p_4-p_3)x'/\hbar}\, \overline{u}(\mathbf{p}_4, s_4)\gamma^\mu u(\mathbf{p}_3, s_3)\, \tilde{C}(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4; s_1, s_2, s_3, s_4) . \qquad (12.18)
\end{aligned}
$$

The vacuum expectation value introduced here reads

$$
\begin{aligned}
&\tilde{C}(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4; s_1, s_2, s_3, s_4) = \langle 0 | \hat{b}_{\mathbf{p}_{f_2},s_{f_2}} \hat{b}_{\mathbf{p}_{f_1},s_{f_1}} \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{b}_{\mathbf{p}_1,s_1} \hat{b}^\dagger_{\mathbf{p}_4,s_4} \hat{b}_{\mathbf{p}_3,s_3} \hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}} \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}} | 0 \rangle \\
&= - \langle 0 | \hat{b}_{\mathbf{p}_{f_2},s_{f_2}} \hat{b}_{\mathbf{p}_{f_1},s_{f_1}} \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{b}^\dagger_{\mathbf{p}_4,s_4} \hat{b}_{\mathbf{p}_1,s_1} \hat{b}_{\mathbf{p}_3,s_3} \hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}} \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}} | 0 \rangle \\
&+ \delta(\mathbf{p}_1 - \mathbf{p}_4) \delta_{s_1,s_4} \langle 0 | \hat{b}_{\mathbf{p}_{f_2},s_{f_2}} \hat{b}_{\mathbf{p}_{f_1},s_{f_1}} \hat{b}^\dagger_{\mathbf{p}_2,s_2} \hat{b}_{\mathbf{p}_3,s_3} \hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}} \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}} | 0 \rangle \ , \qquad (12.19)
\end{aligned}
$$

where we have applied the anti-commutator algebra of the fermionic operators (10.407). In (12.19) the second term disappears due to the different momenta of the initial and the final state (12.12), and (12.13). Indeed, as (12.19) contains two creation (annihilation) operators for the initial (final) states but only one annihilation (creation) operator for an intermediate state, there always remains one creation (annihilation) operator, which finally annihilates the bra (ket) vacuum. Thus, a comparison with (12.15) yields:

$$
\tilde{C}(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4; s_1, s_2, s_3, s_4) = C(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4; s_1, s_2, s_3, s_4) . \qquad (12.20)
$$

We conclude from (12.14), (12.16), (12.18), and (12.20) that both contributions of the scattering matrix (12.5) depend on the same vacuum expectation value (12.15). We now evaluate the latter by iteratively applying the underlying anti-commutator relations (10.407):

$$\begin{aligned}
&C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4) = \langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}_{\mathbf{p}_{f_1},s_{f_1}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}^\dagger_{\mathbf{p}_4,s_4}\left(\hat{b}_{\mathbf{p}_1,s_1}\hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}\right)\left(\hat{b}_{\mathbf{p}_3,s_3}\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\right)|0\rangle \\
&-\delta(\mathbf{p}_3-\mathbf{p}_{i_1})\delta_{s_3,s_{i_1}}\,\langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}_{\mathbf{p}_{f_1},s_{f_1}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}^\dagger_{\mathbf{p}_4,s_4}\left(\hat{b}_{\mathbf{p}_1,s_1}\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\right)|0\rangle = \langle 0|\,\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}_{\mathbf{p}_{f_1},s_{f_1}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}^\dagger_{\mathbf{p}_4,s_4} \\
&\times\left[-\cancel{\hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}\hat{b}_{\mathbf{p}_1,s_1}}+\delta(\mathbf{p}_1-\mathbf{p}_{i_1})\delta_{s_1,s_{i_1}}\right]\left[-\cancel{\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\hat{b}_{\mathbf{p}_3,s_3}}+\delta(\mathbf{p}_1-\mathbf{p}_{i_1})\delta_{s_1,s_{i_1}}\right]|0\rangle-\delta(\mathbf{p}_3-\mathbf{p}_{i_1})\delta_{s_3,si_1} \\
&\times\langle 0|\,\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}_{\mathbf{p}_{f_1},s_{f_1}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}^\dagger_{\mathbf{p}_4,s_4}\left[\cancel{\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\hat{b}_{\mathbf{p}_1,s_1}}+\delta(\mathbf{p}_1-\mathbf{p}_{i_2})\delta_{s_1,s_{i_2}}\right]|0\rangle = \Big\{\delta(\mathbf{p}_1-\mathbf{p}_{i_1})\delta_{s_1,s_{i_1}} \\
&\times\delta(\mathbf{p}_3-\mathbf{p}_{i_2})\delta_{s_3,s_{i_2}}-\delta(\mathbf{p}_1-\mathbf{p}_{i_2})\delta_{s_1,s_{i_2}}\delta(\mathbf{p}_3-\mathbf{p}_{i_1})\delta_{s_3,s_{i_1}}\Big\}\,\langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}_{\mathbf{p}_{f_1},s_{f_1}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}^\dagger_{\mathbf{p}_4,s_4}|0\rangle\;. \qquad (12.21)
\end{aligned}$$

Here the crossed out terms do not contribute as the creation operator of an initial state annihilates the bra vacuum due to $\mathbf{p}_{i_1},\mathbf{p}_{i_2}\neq\mathbf{p}_{f_1},\mathbf{p}_{f_2}$. The remaining vacuum expectation value (12.21) results in

$$\begin{aligned}
&\langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}_{\mathbf{p}_{f_1},s_{f_1}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}^\dagger_{\mathbf{p}_4,s_4}|0\rangle = -\,\langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}_{\mathbf{p}_{f_1},s_{f_1}}\hat{b}^\dagger_{\mathbf{p}_4,s_4}|0\rangle+\delta(\mathbf{p}_2-\mathbf{p}_{f_1})\delta_{s_2,s_{f_1}} \qquad (12.22)\\
&\times\langle 0|\hat{b}_{\mathbf{p}_{f_2},s_{f_2}}\hat{b}^\dagger_{\mathbf{p}_4,s_4}|0\rangle = \delta(\mathbf{p}_2-\mathbf{p}_{f_1})\delta_{s_2,s_{f_1}}\delta(\mathbf{p}_4-\mathbf{p}_{f_2})\delta_{s_4,s_{f_2}}-\delta(\mathbf{p}_2-\mathbf{p}_{f_2})\delta_{s_2,s_{f_2}}\delta(\mathbf{p}_4-\mathbf{p}_{f_1})\delta_{s_4,s_{f_1}}\;.
\end{aligned}$$

Inserting (12.22) into (12.21) yields in total four terms:

$$\begin{aligned}
&C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4) = \delta(\mathbf{p}_{f_1}-\mathbf{p}_2)\delta_{s_{f_1},s_2}\delta(\mathbf{p}_{f_2}-\mathbf{p}_4)\delta_{s_{f_2},s_4}\delta(\mathbf{p}_{i_1}-\mathbf{p}_1)\delta_{s_{i_1},s_1} \\
&\times\delta(\mathbf{p}_{i_2}-\mathbf{p}_3)\delta_{s_{i_2},s_3}+\delta(\mathbf{p}_{f_1}-\mathbf{p}_4)\delta_{s_{f_1},s_4}\delta(\mathbf{p}_{f_2}-\mathbf{p}_2)\delta_{s_{f_2},s_2}\delta(\mathbf{p}_{i_1}-\mathbf{p}_3)\delta_{s_{i_1},s_3}\delta(\mathbf{p}_{i_2}-\mathbf{p}_1)\delta_{s_{i_2},s_1} \\
&-\delta(\mathbf{p}_{f_1}-\mathbf{p}_2)\delta_{s_{f_1},s_2}\delta(\mathbf{p}_{f_2}-\mathbf{p}_4)\delta_{s_{f_2},s_4}\delta(\mathbf{p}_{i_1}-\mathbf{p}_3)\delta_{s_{i_1},s_3}\delta(\mathbf{p}_{i_2}-\mathbf{p}_1)\delta_{s_{i_2},s_1} \\
&-\delta(\mathbf{p}_{f_1}-\mathbf{p}_4)\delta_{s_{f_1},s_4}\delta(\mathbf{p}_{f_2}-\mathbf{p}_4)\delta_{s_{f_2},s_4}\delta(\mathbf{p}_{i_1}-\mathbf{p}_1)\delta_{s_{i_1},s_1}\delta(\mathbf{p}_{i_2}-\mathbf{p}_3)\delta_{s_{i_2},s_3}\;. \qquad (12.23)
\end{aligned}$$

We recognize that the vacuum expectation value (12.23) turns out to have the symmetry

$$C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4) = C(\mathbf{p}_3,\mathbf{p}_4,\mathbf{p}_1,\mathbf{p}_2;s_3,s_4,s_1,s_2)\,, \qquad (12.24)$$

where both the initial and the final momenta as well as the helicities are exchanged with respect to each other. Therefore, the substitutions $\mathbf{p}_1,s_1\leftrightarrow\mathbf{p}_3,s_3$ and $\mathbf{p}_2,s_2\leftrightarrow\mathbf{p}_4,s_4$ in (12.18) lead with (12.20) to a corresponding symmetry of the matrix element

$$\begin{aligned}
&\langle\mathbf{p}_{f_1},s_{f_1};\mathbf{p}_{f_2},s_{f_2}|:\hat{j}^\mu(x)::\hat{j}^\nu(x'):|\mathbf{p}_{i_1},s_{i_1};\mathbf{p}_{i_2},s_{i_2}\rangle \\
&=\langle\mathbf{p}_{f_1},s_{f_1};\mathbf{p}_{f_2},s_{f_2}|:\hat{j}^\nu(x')::\hat{j}^\mu(x):|\mathbf{p}_{i_1},s_{i_1};\mathbf{p}_{i_2},s_{i_2}\rangle\;. \qquad (12.25)
\end{aligned}$$

Using (12.25) in (12.16), the latter reduces to

$$\begin{aligned}
S_{fi}^{(2,\mathrm{rad})} &= -\frac{e^2}{2\hbar^2c^2}\int d^4x\int d^4x'\langle\mathbf{p}_{f_1},s_{f_1};\mathbf{p}_{f_2},s_{f_2}|:\hat{j}^\mu(x): \\
&\quad\times:\hat{j}^\nu(x'):|\mathbf{p}_{i_1},s_{i_1};\mathbf{p}_{i_2},s_{i_2}\rangle\,D_{\mu\nu}(x,x')\,, \qquad (12.26)
\end{aligned}$$

where we have introduced as an abbreviation the Maxwell propagator

$$D_{\mu\nu}(x,x') = \Theta(x^0-x'^0)\,\langle 0|\hat{A}_\mu(x)\hat{A}_\nu(x')|0\rangle+\Theta(x'^0-x^0)\,\langle 0|\hat{A}_\nu(x')\hat{A}_\mu(x)|0\rangle\;. \qquad (12.27)$$

Substituting (12.18) and (12.20) into (12.26), we obtain for the interaction between the Dirac and the Maxwell field:

$$S_{fi}^{(2,\mathrm{rad})} = -\frac{e^2}{2\hbar^2}\int d^4x \int d^4x' \int d^3p_1 \int d^3p_2 \int d^3p_3 \int d^3p_4 \sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4}$$
$$\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_1}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_2}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_3}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_4}}}\,D_{\mu\nu}(x,x')\,e^{i(p_2-p_1)x/\hbar}$$
$$\times\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\,e^{i(p_4-p_3)x'/\hbar}\,\overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)\,C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4)\,. \tag{12.28}$$

Based on the previous results we now establish an intriguing connection between both contributions (12.14) and (12.28) of the scattering matrix (12.5). To this end we first use the Fourier expansion of the Coulomb potential in (12.14)

$$\frac{1}{|\mathbf{x}-\mathbf{x}'|} = \int \frac{d^3k}{(2\pi)^3}\,\frac{4\pi}{\mathbf{k}^2}\,e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')}\,, \tag{12.29}$$

so that the scattering matrix contribution from the instantaneous Coulomb self-interaction of the Dirac field (12.14) reduces to

$$S_{fi}^{(2,\mathrm{inst})} = \frac{-ie^2}{8\pi\hbar\epsilon_0}\int d^3p_1 \int d^3p_2 \int d^3p_3 \int d^3p_4 \sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_1}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_2}}}$$
$$\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_3}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_4}}}\,\overline{u}(\mathbf{p}_2,s_2)\gamma^0u(\mathbf{p}_1,s_1)\,\overline{u}(\mathbf{p}_4,s_4)\gamma^0u(\mathbf{p}_3,s_3)\,C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4)$$
$$\times\int d^3k\,\frac{1}{2\pi^2\mathbf{k}^2}\int dt\,e^{i(E_{\mathbf{p}_2}+E_{\mathbf{p}_4}-E_{\mathbf{p}_1}-E_{\mathbf{p}_3})t/\hbar}\int d^3x\,e^{i(\hbar\mathbf{k}-\mathbf{p}_2+\mathbf{p}_1)\mathbf{x}/\hbar}\int d^3x'\,e^{i(-\hbar\mathbf{k}-\mathbf{p}_4+\mathbf{p}_3)\mathbf{x}'/\hbar}\,, \tag{12.30}$$

where the evaluation of the respective spatial and temporal integrals yields

$$\int dt\,e^{i(E_{\mathbf{p}_2}+E_{\mathbf{p}_4}-E_{\mathbf{p}_1}-E_{\mathbf{p}_3})t/\hbar} = 2\pi\hbar c\,\delta(p_2^0+p_4^0-p_1^0-p_3^0)\,, \tag{12.31}$$

$$\int d^3x\,e^{i(\hbar\mathbf{k}-\mathbf{p}_2+\mathbf{p}_1)\mathbf{x}/\hbar} = (2\pi\hbar)^3\,\delta(\hbar\mathbf{k}-\mathbf{p}_2+\mathbf{p}_1)\,, \tag{12.32}$$

$$\int d^3x'\,e^{i(-\hbar\mathbf{k}-\mathbf{p}_4+\mathbf{p}_3)\mathbf{x}'/\hbar} = (2\pi\hbar)^3\,\delta(\hbar\mathbf{k}-\mathbf{p}_4+\mathbf{p}_3)\,. \tag{12.33}$$

Substituting (12.31)–(12.33) into (12.30) and evaluating the **k**-integral finally leads to

$$S_{fi}^{(2,\mathrm{inst})} = \frac{-i\hbar e^2}{2\epsilon_0 c}\int d^3p_1 \int d^3p_2 \int d^3p_3 \int d^3p_4 \sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_1}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_2}}}$$
$$\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_3}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_4}}}\,(2\pi\hbar)^4\delta(p_2+p_4-p_1-p_3)\,\frac{1}{(\mathbf{p}_2-\mathbf{p}_1)^2}\,\overline{u}(\mathbf{p}_2,s_2)\gamma^0u(\mathbf{p}_1,s_1)$$
$$\times\overline{u}(\mathbf{p}_4,s_4)\gamma^0u(\mathbf{p}_3,S_3)\,C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4)\,. \tag{12.34}$$

On the other hand, with the help of the four-dimensional Fourier representation of the Maxwell propagator (9.199)

$$D_{\mu\nu}(x,x') = \frac{i\hbar}{c\epsilon_0}\int \frac{d^4k}{(2\pi)^4}\,\frac{1}{k^2}\,e^{ik(x-x')}\,P_{\mu\nu}(k) \tag{12.35}$$

the scattering matrix contribution (12.28) stemming from the interaction between the Dirac and the Maxwell field yields

$$S_{fi}^{(2,\mathrm{rad})} = -\frac{ie^2}{2\pi\hbar c\epsilon_0}\int d^3p_1\int d^3p_2\int d^3p_3\int d^3p_4\sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_1}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_2}}}$$
$$\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_3}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_4}}}\,\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)\int\frac{d^4k}{(2\pi)^4}\frac{1}{k^2}P_{\mu\nu}(k)$$
$$\times C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4)\int d^4x\,e^{i(\hbar k+p_2-p_1)x/\hbar}\int d^4x' e^{i(-\hbar k+p_4-p_3)x'/\hbar}\,. \tag{12.36}$$

The evaluation of the two spatio-temporal integrals results in

$$\begin{aligned}\int d^4x\,e^{i(\hbar k+p_2-p_1)x/\hbar} &= (2\pi\hbar)^4\,\delta(\hbar k+p_2-p_1)\\ \int d^4x' e^{i(-\hbar k+p_4-p_3)x'/\hbar} &= (2\pi\hbar)^4\,\delta(-\hbar k+p_4-p_3)\,,\end{aligned} \tag{12.37}$$

so that the k-integral in (12.36) can be evaluated as follows:

$$S_{fi}^{(2,\mathrm{rad})} = -\frac{ie^2\hbar}{2\epsilon_0 c}\int d^3p_1\int d^3p_2\int d^3p_3\int d^3p_4\sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_1}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_2}}}$$
$$\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_3}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_4}}}\,(2\pi\hbar)^4\delta(p_2+p_4-p_1-p_3)\,\frac{P_{\mu\nu}(p_2-p_1)}{(p_2-p_1)^2}$$
$$\times\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\,\overline{u}(_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)\,C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4)\,. \tag{12.38}$$

Inserting the polarization sum from (9.204)

$$P_{\mu\nu}(k) = -g_{\mu\nu} - k^2\frac{\xi_\mu\xi_\nu}{(k\xi)^2-k^2} - \frac{k_\mu k_\nu + (k\xi)(k_\mu\xi_\nu + k_\nu\xi_\mu)}{(k\xi)^2-k^2} \tag{12.39}$$

into (12.38), it turns out that its last term does not contribute. Namely, due to the algebraic equations (10.305) and (10.307) determining the Dirac spinor $u(\mathbf{p},s)$ and the Dirac adjoint Dirac spinor $\overline{u}(\mathbf{p},s)$, we conclude

$$\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)(p_{2\mu}-p_{1\mu}) = \Big\{\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu p_{2\mu}\Big\}u(\mathbf{p}_1,s_1) - \overline{u}(\mathbf{p}_2,s_2)\Big\{\gamma^\mu p_{1\mu}u(\mathbf{p}_1,s_1)\Big\}$$
$$= -Mc\,\overline{u}(\mathbf{p}_2,s_2)u(\mathbf{p}_1,s_1) + Mc\,\overline{u}(\mathbf{p}_2,s_2)u(\mathbf{p}_1,s_1) = 0 \tag{12.40}$$

and, analogously, we also obtain

$$\begin{aligned}&\delta(p_2+p_4-p_1-p_3)\,\overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)(p_{2\nu}-p_{1\nu})\\ &= \delta(p_2+p_4-p_1-p_3)\,\overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)(p_{4\nu}-p_{3\nu}) = 0\,.\end{aligned} \tag{12.41}$$

Note that the identities (12.40) and (12.41) are a consequence of the charge conservation at a vertex and can be studied in more detail in the framework of the so-called Ward-Takahashi identities. From (12.38)–(12.41) we then conclude

$$S_{fi}^{(2,\mathrm{rad})} = -\frac{i\hbar e^2}{2\epsilon_0 c}\int d^3p_1 \int d^3p_2 \int d^3p_3 \int d^3p_4 \sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_1}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_2}}}$$
$$\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_3}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_4}}}\frac{(2\pi\hbar)^4\delta(p_2+p_4-p_1-p_3)}{(p_2-p_1)^2}\left\{-g_{\mu\nu}-\frac{(p_2-p_1)^2\xi_\mu\xi_\nu}{[(p_2-p_1)\xi]^2-(p_2-p_1)^2}\right\}$$
$$\times\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\,\overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)\,C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4)\,. \tag{12.42}$$

Adding now both contributions (12.34) and (12.42) to the scattering matrix element (12.5) and taking into account the explicit form of the time-like vector ξ according to (9.201) yields a manifestly covariant result:

$$S_{fi}^{(2)} = \frac{i\hbar e^2}{2\epsilon_0 c}\int d^3p_1 \int d^3p_2 \int d^3p_3 \int d^3p_4 \sum_{s_1}\sum_{s_2}\sum_{s_3}\sum_{s_4} \sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_1}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_2}}}$$
$$\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_3}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_4}}}\,(2\pi\hbar)^4\delta(p_2+p_4-p_1-p_3)\,\frac{g_{\mu\nu}}{(p_2-p_1)^2}$$
$$\times\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\,\overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)\,C(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4;s_1,s_2,s_3,s_4)\,. \tag{12.43}$$

Substituting the vacuum expectation value (12.23) into (12.43), the first two and the last two terms yield the same contribution, respectively, due to the obvious identity

$$\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\,\overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3) = \overline{u}(\mathbf{p}_4,s_4)\gamma^\nu u(\mathbf{p}_3,s_3)\,\overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1) \tag{12.44}$$

and the symmetry of the integrand with respect to the substitutions $\mathbf{p}_1, s_1 \leftrightarrow \mathbf{p}_3, s_3$ and $\mathbf{p}_2, s_2 \leftrightarrow \mathbf{p}_4, s_4$. This results in a factor of 2, which just compensates for the factor 1/2 stemming from the second order in the Taylor expansion of the exponential function:

$$S_{fi}^{(2)} = \frac{i\hbar e^2}{\epsilon_0 c}(2\pi\hbar)^4\delta(p_{f_1}+p_{f_2}-p_{i_1}-p_{i_2})\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_{i_1}}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_{i_2}}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_{f_1}}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_{f_2}}}}$$
$$\times\left\{\frac{g_{\mu\nu}}{(p_{f_1}-p_{i_1})^2}\,\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\mu u(\mathbf{p}_{i_1},s_{i_1})\,\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu u(\mathbf{p}_{i_2},s_{i_2})\right.$$
$$\left.-\frac{g_{\mu\nu}}{(p_{f_1}-p_{i_2})^2}\,\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\mu u(\mathbf{p}_{i_2},s_{i_2})\,\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu u(\mathbf{p}_{i_1},s_{i_1})\right\}. \tag{12.45}$$

This perturbative result for the scattering matrix element of the Møller scattering can be represented in the form of two Feynman diagrams, which are depicted in Fig. 12.1. Note that no momentum integrals occur in (12.45), which would correspond to internal loops in the Feynman diagrams. Therefore, one calls the graphs in Fig. 12.1 to be tree-level graphs. The corresponding manifestly covariant Feynman rules for converting the scattering matrix element (12.45) into the Feynman diagrams of Fig. 12.1 and vice versa read in momentum space as follows:

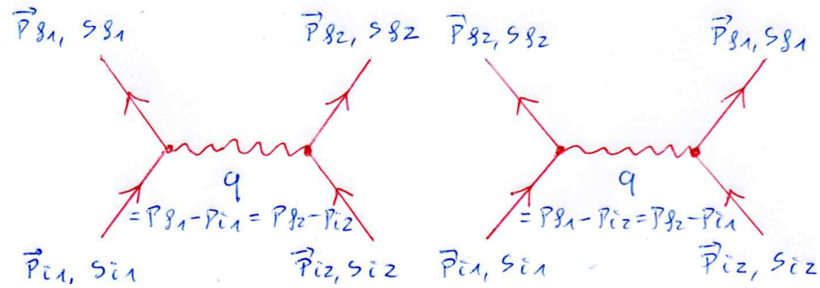


Figure 12.1: Direct (left) and exchange (right) Feynman diagram for the Møller scattering of two electrons.

(F1) The prefactor $(2\pi\hbar)^4\,\delta(p_{f_1} + p_{f_2} - p_{i_1} - p_{i_2})$ guarantees the conservation of energy and momentum in the scattering process.

(F2) An incoming electron corresponds to the factor $\sqrt{\dfrac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_i}}}\,u(\mathbf{p}_i, s_i)$.

(F3) An outgoing electron leads to the factor $\sqrt{\dfrac{Mc^2}{(2\pi\hbar)^3 E_{\mathbf{p}_f}}}\,\overline{u}(\mathbf{p}_f, s_f)$.

(F4) A vertex yields the factor $e\,\gamma^\mu$.

(F5) The Maxwell propagator corresponds to the covariant factor $\hbar g_{\mu\nu}/(\epsilon_0 c q^2)$, where q denotes the momentum transfer, see Fig. 12.1.

(F6) The phase of the scattering matrix element is calculated according to the following rule: $(-i)^{\text{number of vertices}}\,(-i)^{\text{number of inner lines}}$. Here, the minus sign for the number of vertices comes from the negative charge of the electron, while the minus sign for the inner line stems from the Maxwell propagator.

The phase rule (F6) leads directly to the correct phase of the direct graph: $(-i)^2(-i)^1 = +i$. Due to the indistinguishability of the two incoming and outgoing electrons, apart from the direct graph also the exchange graph contributes, where in the latter the two outgoing electrons are swapped. Due to the Fermi-Dirac statistics of the electrons, the exchange graph has an additional minus sign. Consequently, the entire scattering matrix is anti-symmetric with respect to the exchange of the two incoming or outgoing electrons. If we had calculated the scattering of identical bosons, the exchange graph would have the same sign as the direct graph and the total scattering amplitude would be symmetrical with respect to the exchange of the two incoming and outgoing bosons. Note that the Feynman diagrams in quantum electrodynamics always have the multiplicities ± 1 in contrast to other field theories such as the ϕ^4-theory of critical phenomena, where the multiplicities are highly non-trivial as they follow from involved combinatorial reasons.

12.2 Polarization Averaging

The second perturbative order of the Møller scattering matrix element in (12.45) factorises according to

$$\begin{aligned} S_{fi}^{(2)} &= \frac{i\hbar e^2}{\epsilon_0 c}\,(2\pi\hbar)^4\delta(p_{f_1}+p_{f_2}-p_{i_1}-p_{i_2})\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{i_1}}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{i_2}}}} \\ &\times\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{f_1}}}}\sqrt{\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{f_2}}}}\,M_{fi}^{(2)}\,, \end{aligned} \tag{12.46}$$

where we introduced the matrix element

$$\begin{aligned} M_{fi}^{(2)} &= \frac{g_{\mu\nu}}{(p_{f_1}-p_{i_1})^2}\,\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\mu u(\mathbf{p}_{i_1},s_{i_1})\,\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu u(\mathbf{p}_{i_2},s_{i_2}) \\ &-\frac{g_{\mu\nu}}{(p_{f_1}-p_{i_2})^2}\,\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\mu u(\mathbf{p}_{i_2},s_{i_2})\,\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu u(\mathbf{p}_{i_1},s_{i_1})\,. \end{aligned} \tag{12.47}$$

Provided that the polarizations of both the incoming and the outgoing electrons are not detected during the scattering process, we have to calculate the scattering cross section from averaging the squared matrix element over all these polarisations:

$$\overline{|M_{fi}^{(2)}|^2} = \frac{1}{4}\sum_{s_{i_1}}\sum_{s_{i_2}}\sum_{s_{f_1}}\sum_{s_{f_2}}|M_{fi}^{(2)}|^2\,. \tag{12.48}$$

Substituting (12.47) into (12.48) leads in total to four terms:

$$\begin{aligned} \overline{|M_{fi}^{(2)}|^2} &= \frac{1}{4}\sum_{s_{i_1}}\sum_{s_{i_2}}\sum_{s_{f_1}}\sum_{s_{f_2}}\Bigg\{\frac{1}{(p_{f_1}-p_{i_1})^4}\left[\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\mu u(\mathbf{p}_{i_1},s_{i_1})\right]^*\left[\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma_\mu u(_{i_2},s_{i_2})\right]^* \\ &\times\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\nu u(\mathbf{p}_{i_1},s_{i_1})\,\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma_\nu u(\mathbf{p}_{i_2},s_{i_2}) \\ &-\frac{1}{(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\left[\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\mu u(\mathbf{p}_{i_1},s_{i_1})\right]^*\left[\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma_\mu u(\mathbf{p}_{i_2},s_{i_2})\right]^* \\ &\times\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu u(\mathbf{p}_{i_1},s_{i_1})\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma_\nu u(\mathbf{p}_{i_2},s_{i_2})+(p_{f_1}\leftrightarrow p_{f_2})\Bigg\}\,. \end{aligned} \tag{12.49}$$

Calculating the expression $[\overline{u}(\mathbf{p}_1,s_1)\gamma^\mu u(\mathbf{p}_2,s_2)]^*$, we note that $\overline{u}(\mathbf{p}_1,s_1)\gamma^\mu u(\mathbf{p}_2,s_2)$ coincides with its transpose as it is a scalar:

$$\begin{aligned} &[\overline{u}(\mathbf{p}_1,s_1)\gamma^\mu\,u(\mathbf{p}_2,s_2)]^* = [\overline{u}(\mathbf{p}_1,s_1)\gamma^\mu u(\mathbf{p}_2,s_2)]^\dagger \\ &= u^\dagger(\mathbf{p}_2,s_2)(\gamma^\mu)^\dagger\,\overline{u}^\dagger(\mathbf{p}_1,s_1) = \overline{u}(\mathbf{p}_2,s_2)\gamma^0(\gamma^\mu)^\dagger\gamma^0 u(\mathbf{p}_1,s_1)\,. \end{aligned} \tag{12.50}$$

From the chiral representation of the Dirac matrices (10.95) follows due to the hermiticity of the four Pauli matrices σ^μ:

$$(\gamma^\mu)^\dagger = \begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix}^\dagger = \begin{pmatrix} 0 & \tilde{\sigma}^\mu \\ \sigma^\mu & 0 \end{pmatrix} \qquad \Longrightarrow \qquad \begin{cases} (\gamma^0)^\dagger = \gamma^0 \\ (\gamma^i)^\dagger = -\gamma^i \end{cases}\,. \tag{12.51}$$

With this we conclude by taking into account the Clifford algebra (10.96):

$$\begin{cases}\gamma^0(\gamma^0)^\dagger\gamma^0 = \gamma^0\gamma^0\gamma^0 = \gamma^0 \\ \gamma^0(\gamma^i)^\dagger\gamma^0 = -\gamma^0\gamma^i\gamma^0 = \gamma^i\gamma^0\gamma^0 = \gamma^i\end{cases} \quad \Longrightarrow \quad \gamma^0(\gamma^\mu)^\dagger\gamma^0 = \gamma^\mu\,. \tag{12.52}$$

Substituting (12.52) into (12.50) then leads to the result

$$\left[\overline{u}(\mathbf{p}_1,s_1)\gamma^\mu u(\mathbf{p}_2,s_2)\right]^* = \overline{u}(\mathbf{p}_2,s_2)\gamma^\mu\, u(\mathbf{p}_1,s_1)\,. \tag{12.53}$$

Using (12.53) in (12.49) yields

$$\begin{aligned}\overline{\left|M_{fi}^{(2)}\right|^2} \;=\;& \frac{1}{4}\sum_{s_{i_1}}\sum_{s_{i_2}}\sum_{s_{f_1}}\sum_{s_{f_2}}\left\{\frac{1}{(p_{f_1}-p_{i_1})^4}\,\overline{u}(\mathbf{p}_{i_1},s_{i_1})\gamma^\mu u(\mathbf{p}_{f_1},s_{f_1})\,\overline{u}(\mathbf{p}_{i_2},s_{i_2})\gamma_\mu u(\mathbf{p}_{f_2},s_{f_2})\right.\\ &\times\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\nu u(\mathbf{p}_{i_1},s_{i_1})\,\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma_\nu u(\mathbf{p}_{i_2},s_{i_2})\\ &-\frac{1}{(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\,\overline{u}(\mathbf{p}_{i_1},s_{i_1})\gamma^\mu u(\mathbf{p}_{f_1},s_{f_1})\,\overline{u}(\mathbf{p}_{i_2},s_{i_2})\gamma_\mu u(\mathbf{p}_{f_2},s_{f_2})\\ &\left.\times\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu u(\mathbf{p}_{i_1},s_{i_1})\,\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma_\nu u(\mathbf{p}_{i_2},s_{i_2}) + (p_{f_1}\leftrightarrow p_{f_2})\right\}.\end{aligned} \tag{12.54}$$

As the factors $\overline{u}(\mathbf{p}_1,s_1)\gamma^\mu u(\mathbf{p}_2,s_2)$ are scalars, their order can be changed:

$$\begin{aligned}\overline{\left|M_{fi}^{(2)}\right|^2} \;=\;& \frac{1}{4}\sum_{s_{f_1}}\sum_{s_{f_2}}\left\{\frac{1}{(p_{f_1}-p_{i_1})^4}\,\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma^\nu\left[\sum_{s_{i_1}}u(\mathbf{p}_{i_1},s_{i_1})\overline{u}(\mathbf{p}_{i_1},s_{i_1})\right]\gamma^\mu u(\mathbf{p}_{f_1},s_{f_1})\right.\\ &\times\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma_\nu\left[\sum_{s_{i_2}}u(\mathbf{p}_{i_2},s_{i_2})\overline{u}(\mathbf{p}_{i_2},s_{i_2})\right]\gamma_\mu u(\mathbf{p}_{f_2},s_{f_2})\\ &-\frac{1}{(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\,\overline{u}(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu\left[\sum_{s_{i_1}}u(\mathbf{p}_{i_1},s_{i_1})\overline{u}(\mathbf{p}_{i_1},s_{i_1})\right]\gamma^\mu u(\mathbf{p}_{f_1},s_{f_1})\\ &\left.\times\overline{u}(\mathbf{p}_{f_1},s_{f_1})\gamma_\nu\left[\sum_{s_{i_2}}u(\mathbf{p}_{i_2},s_{i_2})\overline{u}(\mathbf{p}_{i_2},s_{i_2})\right]\gamma_\mu u(\mathbf{p}_{f_2},s_{f_2}) + (p_{f_1}\leftrightarrow p_{f_2})\right\}.\end{aligned} \tag{12.55}$$

The polarisation sums occurring here with respect to s_{i_1}, s_{i_2} were already calculated according to (10.438) and (10.447). We implement now this result by introducing for the sake of clarity spinorial indices and by using for notational brevity the Einstein summation convention that implies summation over identical spinorial indices:

$$\begin{aligned}\overline{\left|M_{fi}^{(2)}\right|^2} \;=\;& \frac{1}{4}\sum_{s_{f_1}}\sum_{s_{f_2}}\left\{\frac{1}{(p_{f_1}-p_{i_1})^4}\,\overline{u}_\alpha(\mathbf{p}_{f_1},s_{f_1})\gamma^\nu_{\alpha\beta}\left(\frac{\not{p}_{i_1}+Mc}{2Mc}\right)_{\beta\gamma}\gamma^\mu_{\gamma\delta}\,u_\delta(\mathbf{p}_{f_1},s_{f_1})\right.\\ &\times\overline{u}_{\alpha'}(\mathbf{p}_{f_2},s_{f_2})\gamma_{\nu\alpha'\beta'}\left(\frac{\not{p}_{i_2}+Mc}{2Mc}\right)_{\beta'\gamma'}\gamma_{\mu\gamma'\delta'}\,u_{\delta'}(\mathbf{p}_{f_2},s_{f_2})\\ &-\frac{1}{(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\,\overline{u}_\alpha(\mathbf{p}_{f_2},s_{f_2})\gamma^\nu_{\alpha\beta}\left(\frac{\not{p}_{i_1}+Mc}{2Mc}\right)_{\beta\gamma}\gamma^\mu_{\gamma\delta}\,u_\delta(\mathbf{p}_{f_1},s_{f_1})\\ &\left.\times\overline{u}_{\alpha'}(\mathbf{p}_{f_1},s_{f_1})\gamma_{\nu\alpha'\beta'}\left(\frac{\not{p}_{i_2}+Mc}{2Mc}\right)_{\beta'\gamma'}\gamma_{\mu\gamma'\delta'}\,u_{\delta'}(\mathbf{p}_{f_2},s_{f_2}) + (p_{f_1}\leftrightarrow p_{f_2})\right\}.\end{aligned} \tag{12.56}$$

Paying attention to the respective spinorial indices, the individual terms can be rearranged as follows

$$\overline{|M_{fi}^{(2)}|^2} = \frac{1}{4}\Bigg\{ \frac{1}{(p_{f_1}-p_{i_1})^4}\,\gamma^\nu_{\alpha\beta}\left(\frac{\not p_{i_1}+Mc}{2Mc}\right)_{\beta\gamma}\gamma^\mu_{\gamma\delta}\left[\sum_{s_{f_1}} u_\delta(\mathbf{p}_{f_1},s_{f_1})\overline{u}_\alpha(\mathbf{p}_{f_1},s_{f_1})\right]$$
$$\times\gamma_{\nu\alpha'\beta'}\left(\frac{\not p_{i_2}+Mc}{2Mc}\right)_{\beta'\gamma'}\gamma_{\mu\gamma'\delta'}\left[\sum_{s_{f_2}} u_{\delta'}(\mathbf{p}_{f_2},s_{f_2})\overline{u}_{\alpha'}(\mathbf{p}_{f_2},s_{f_2})\right]$$
$$-\frac{1}{(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\,\gamma^\nu_{\alpha\beta}\left(\frac{\not p_{i_1}+Mc}{2Mc}\right)_{\beta\gamma}\gamma^\mu_{\gamma\delta}\left[\sum_{s_{f_1}} u_\delta(\mathbf{p}_{f_1},s_{f_1})u_{\delta'}(\mathbf{p}_{f_1},s_{f_1})\right]$$
$$\times\gamma_{\nu\alpha'\beta'}\left(\frac{\not p_{i_2}+Mc}{2Mc}\right)_{\beta'\gamma'}\gamma_{\mu\gamma'\delta'}\left[\sum_{s_{f_2}} u_{\delta'}(\mathbf{p}_{f_2},s_{f_2})\overline{u}_{\alpha}(\mathbf{p}_{f_2},s_{f_2})\right] + (p_{f_1}\leftrightarrow p_{f_2})\Bigg\}. \quad (12.57)$$

Here we take into account that also the polarisation sums with respect to s_{f_1}, s_{f_2} were already calculated according to (10.438) and (10.447), yielding:

$$\overline{|M_{fi}^{(2)}|^2} = \frac{1}{4}\Bigg\{ \frac{1}{(p_{f_1}-p_{i_1})^4}\,\gamma^\nu_{\alpha\beta}\left(\frac{\not p_{i_1}+Mc}{2Mc}\right)_{\beta\gamma}\gamma^\mu_{\gamma\delta}\left(\frac{\not p_{f_1}+Mc}{2Mc}\right)_{\delta\alpha}\gamma_{\nu\alpha'\beta'}\left(\frac{\not p_{i_2}+Mc}{2Mc}\right)_{\beta'\gamma'}$$
$$\times\gamma_{\mu\gamma'\delta'}\left(\frac{\not p_{f_2}+Mc}{2Mc}\right)_{\delta'\alpha'} - \frac{1}{(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\,\gamma^\nu_{\alpha\beta}\left(\frac{\not p_{i_1}+Mc}{2Mc}\right)_{\beta\gamma}\gamma^\mu_{\gamma\delta}\left(\frac{\not p_{f_1}+Mc}{2Mc}\right)_{\delta\alpha'}$$
$$\times\gamma_{\nu\alpha'\beta'}\left(\frac{\not p_{i_2}+Mc}{2Mc}\right)_{\beta'\gamma'}\gamma_{\mu\gamma'\delta'}\left(\frac{\not p_{f_2}+Mc}{2Mc}\right)_{\delta'\alpha} + (p_{f_1}\leftrightarrow p_{f_2})\Bigg\}. \quad (12.58)$$

The sums with respect to the spinorial indices can be interpreted as traces:

$$\overline{|M_{fi}^{(2)}|^2} = \frac{1}{4}\Bigg\{ \frac{1}{(p_{f_1}-p_{i_1})^4}\,\mathrm{Tr}\left[\gamma^\nu\frac{\not p_{i_1}+Mc}{2Mc}\gamma^\mu\frac{\not p_{f_1}+Mc}{2Mc}\right]\,\mathrm{Tr}\left[\gamma_\nu\frac{\not p_{i_2}+Mc}{2Mc}\gamma_\mu\frac{\not p_{f_2}+Mc}{2Mc}\right] \quad (12.59)$$
$$-\frac{1}{(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\,\mathrm{Tr}\left[\gamma^\nu\frac{\not p_{i_1}+Mc}{2Mc}\gamma^\mu\frac{\not p_{f_1}+Mc}{2Mc}\gamma_\nu\frac{\not p_{i_2}+Mc}{2Mc}\gamma_\mu\frac{\not p_{f_2}+Mc}{2Mc}\right] + (p_{f_1}\leftrightarrow p_{f_2})\Bigg\}$$

The first contribution in (12.59) is called the direct term

$$\overline{|M_{fi}^{(2)}|^2_{\mathrm{d}}} = \frac{\mathrm{Tr}\left[\gamma^\nu(\not p_{i_1}+Mc)\gamma^\mu(\not p_{f_1}+Mc)\right]\,\mathrm{Tr}\left[\gamma_\nu(\not p_{i_2}+Mc)\gamma_\mu(\not p_{f_2}+Mc)\right]}{64M^4c^4(p_{f_1}-p_{i_1})^4}\,. \quad (12.60)$$

It consists of the product of two traces of the same design type, which reads due to the shortcut notation with the Feynman dagger (10.100) as follows:

$$\mathrm{Tr}\left[\gamma^\mu(\not p_{i_1}+Mc)\gamma^\nu(\not p_{f_1}+Mc)\right] = \mathrm{Tr}\left[\gamma^\mu\not p_{i_1}\gamma^\nu\not p_{f_1} + Mc\gamma^\mu\not p_{i_1}\gamma^\nu + Mc\gamma^\mu\gamma^\nu\not p_{f_1} + M^2c^2\gamma^\mu\gamma^\nu\right]$$
$$= p_{i_1\kappa}p_{f_1\lambda}\mathrm{Tr}\left[\gamma^\mu\gamma^\kappa\gamma^\nu\gamma^\lambda\right] + Mc\,p_{i_1\kappa}\mathrm{Tr}\left[\gamma^\mu\gamma^\kappa\gamma^\nu\right] + Mc\,p_{i_1\lambda}\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\gamma^\lambda\right] + M^2c^2\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right]\,. \quad (12.61)$$

12.3 Traces of Product of Dirac Matrices

Thus, according to (12.61), we have now to calculate traces over different products of γ-matrices. Due to the explicit form of the Dirac matrices (10.95), the trace over each individual γ-matrix disappears:

$$\mathrm{Tr}\left[\gamma^{\mu_1}\right] = 0\,. \tag{12.62}$$

The trace over the product of two γ-matrices can be calculated by using their property of representing a Clifford algebra (10.96):

$$\mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2}\right] = \frac{1}{2}\,\mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2} + \gamma^{\mu_2}\gamma^{\mu_1}\right] = g^{\mu_1\mu_2}\mathrm{Tr}[1] = 4g^{\nu_1\mu_2}\,. \tag{12.63}$$

We show now that the trace vanishes over a product of any odd number of γ-matrices. To this end we consider the γ^5-matrix defined in (10.151) that has the explicit form (10.154) and, thus, the property to be involutoric according to (10.155) as well as anti-commuting with any Dirac matrix according to (10.232). With this follows then for the trace of a product of γ-matrices:

$$\begin{aligned}
&\mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2}\,\cdots\,\gamma^{\mu_n}\right] = \mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2}\,\cdots\,\gamma^{\mu_n}\gamma^5\gamma^5\right] = \mathrm{Tr}\left[\gamma^5\gamma^{\mu_1}\gamma^{\mu_2}\,\cdots\,\gamma^{\mu_n}\gamma^5\right] \\
&= (-1)^n\,\mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2}\,\cdots\,\gamma^{\mu_n}\gamma^5\gamma^5\right] = (-1)^n\,\mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2}\,\cdots\,\gamma^{\mu_n}\right]\,,
\end{aligned} \tag{12.64}$$

so we obtain for n being odd:

$$\mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2}\,\cdots\,\gamma^{\mu_{2n+1}}\right] = 0\,. \tag{12.65}$$

Thus, only the traces over a product of an even number of γ-matrices can be non-vanishing. Let us consider now the trace over a product of four γ-matrices. Successively applying the Clifford algebra property (10.96) together with (12.63) yields $4!! = 3$ terms:

$$\begin{aligned}
&\mathrm{Tr}\left[\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3}\gamma^{\mu_4}\right] = -\mathrm{Tr}\left[\gamma^{\mu_2}\gamma^{\mu_1}\gamma^{\mu_3}\gamma^{\mu_4}\right] + 2g^{\mu_1\mu_2}\mathrm{Tr}\left[\gamma^{\mu_3}\gamma^{\mu_4}\right] = \mathrm{Tr}\left[\gamma^{\mu_2}\gamma^{\mu_3}\gamma^{\mu_1}\gamma^{\mu_4}\right] + 8g^{\mu_1\mu_2}g^{\mu_3\mu_4} \\
&-2g^{\mu_1\mu_3}\mathrm{Tr}\left[\gamma^{\mu_2}\gamma^{\mu_4}\right] = -\mathrm{Tr}[\gamma^{\mu_2}\gamma^{\mu_3}\gamma^{\mu_4}\gamma^{\mu_1}] + 2g^{\mu_1\mu_4}\mathrm{Tr}[\gamma^{\mu_2}\gamma^{\mu_3}] + 8g^{\mu_1\mu_2}g^{\mu_3\mu_4} - 8g^{\mu_1\mu_3}g^{\mu_2\mu_4} \\
&\Longrightarrow \qquad \mathrm{Tr}[\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3}\gamma^{\mu_4}] = 4\left(g^{\mu_1\mu_2}g^{\mu_3\mu_4} - g^{\mu_1\mu_3}g^{\mu_2\mu_4} + g^{\mu_1\mu_4}g^{\mu_2\mu_3}\right)\,.
\end{aligned} \tag{12.66}$$

With the help of the auxiliary calculations (12.62)–(12.66) we obtain for (12.61) the result

$$\begin{aligned}
\mathrm{Tr}[\gamma^\mu(\not{p}_{i_1} + Mc)\gamma^\nu(\not{p}_{f1} + Mc)] &= 4p_{i_1\kappa}p_{f_1\lambda}(g^{\mu\kappa}g^{\nu\lambda} - g^{\mu\nu}g^{\kappa\lambda} + g^{\mu\lambda}g^{\kappa\nu}) + 4M^2c^2g^{\mu\nu} \\
&= 4(p^\mu_{i_1}p^\nu_{f_1} - p_{i_1}p_{f_1}g^{\mu\nu} + p^\nu_{i_1}p^\mu_{f_1} + M^2c^2g^{\mu\nu})\,.
\end{aligned} \tag{12.67}$$

Using (12.67) the direct term (12.60) yields

$$\overline{|M^{(2)}_{fi}|^2_{\mathrm{d}}} = \frac{[p^\mu_{i_1}p^\nu_{f_1} + p^\nu_{i_1}p^\mu_{f_1} + (M^2c^2 - p_{i_1}p_{f_1})g^{\mu\nu}][p_{i_2\mu}p_{f_2\nu} + p_{i_2\nu}p_{f_2\mu} + (M^2c^2 - p_{i_2}p_{f_2})g_{\mu\nu}]}{4M^4c^4(p_{f_1} - p_{i_1})^4}\,,$$

which finally reduces to

$$\overline{|M^{(2)}_{fi}|^2_{\mathrm{d}}} = \frac{(p_{i_1}p_{i_2})(p_{f_1}p_{f_2}) + (p_{f_1}p_{i_2})(p_{i_1}p_{f_2}) - M^2c^2p_{i_1}p_{f_1} - M^2c^2p_{i_2}p_{f_2} + 2M^4c^4}{2M^4c^4(p_{f_1} - p_{i_1})^4}\,. \tag{12.68}$$

The exchange term in (12.59) is formally obtained from the direct term (12.60) by interchanging the final momenta p_{f_1} and p_{f_2}:

$$\overline{\left|M_{fi}^{(2)}\right|^2_{\rm ex}} = \frac{\mathrm{Tr}\left[\gamma^\nu(\not{p}_{i_1}+Mc)\gamma^\mu(\not{p}_{f_2}+Mc)\right]\;\mathrm{Tr}\left[\gamma_\nu(\not{p}_{i_1}+Mc)\gamma_\mu(\not{p}_{f_1}+Mc)\right]}{64M^4c^4(p_{f_2}-p_{i_1})^4}\,. \qquad (12.69)$$

Therefore, we obtain the result for evaluating the traces in (12.69) from (12.68) by interchanging the final momenta p_{f_1} and p_{f_2}:

$$\overline{\left|M_{fi}^{(2)}\right|^2_{\rm ex}} = \frac{(p_{i_1}p_{i_2})(p_{f_1}p_{f_2})+(p_{f_2}p_{i_2})(p_{i_1}p_{f_1})-M^2c^2p_{i_1}p_{f_2}-M^2c^2p_{i_2}p_{f_1}+2M^4c^4}{2M^4c^4(p_{f_2}-p_{i_1})^4}\,. \qquad (12.70)$$

Thus, it only remains to consider the interference term between the direct and the exchange scattering in (12.59):

$$\overline{\left|M_{fi}^{(2)}\right|^2_{\rm i}} = \frac{-\{\mathrm{Tr}[\gamma^\mu(\not{p}_{i_2}+Mc)\gamma^\nu(\not{p}_{f_1}+Mc)\gamma_\mu(\not{p}_{i_2}+Mc)\gamma_\nu(\not{p}_{f_2}+Mc)]+(p_{f_1}\leftrightarrow p_{f_2})\}}{64M^4c^4(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\,. \qquad (12.71)$$

Let us restrict us for the time being to the evaluation of the first term in (12.71). The corresponding trace can be simplified due to (12.65) such that only the trace over products of an even number of γ-matrices occurs:

$$\begin{aligned}
&\mathrm{Tr}[...] = \mathrm{Tr}[(\gamma^\mu\not{p}_{i_1}\gamma^\nu\not{p}_{f_1}+Mc\gamma^\mu\gamma^\nu\not{p}_{f_1}+Mc\gamma^\mu\not{p}_{i_1}\gamma^\nu+M^2c^2\gamma^\mu\gamma^\nu) \qquad (12.72)\\
&\times(\gamma_\mu\not{p}_{i_1}\gamma_\nu\not{p}_{f_2}+Mc\gamma_\mu\gamma_\nu\not{p}_{f_2}+Mc\gamma_\mu\not{p}_{i_2}\gamma_\nu+M^2c^2\gamma_\mu\gamma_\nu)] = \mathrm{Tr}[\gamma^\mu\not{p}_{i_1}\gamma^\nu\not{p}_{f_1}\gamma_\mu\not{p}_{i_2}\gamma_\nu\not{p}_{f_2}\\
&+M^2c^2\gamma^\mu\not{p}_{i_1}\gamma^\nu\not{p}_{f_1}\gamma_\mu\gamma_\nu+m^2c^2\gamma^\mu\gamma^\nu\not{p}_{f_1}\gamma^\nu\gamma_\mu\gamma_\nu\not{p}_{f_2}+M^2c^2\gamma^\mu\gamma^\nu\not{p}_{f_1}\gamma^\mu\not{p}_{i_2}\gamma_\nu\\
&+M^2c^2\gamma^\mu\not{p}_{i_1}\gamma^\nu\gamma_\mu\gamma_\nu\not{p}_{f_2}+M^2c^2\gamma^\mu\not{p}_{i_1}\gamma^\nu\gamma_\mu\not{p}_{i_2}\gamma_\nu+M^2c^2\gamma^\mu\gamma^\nu\gamma_\mu\not{p}_{i_2}\gamma_\nu\not{p}_{f_2}+M^4c^4\gamma^\mu\gamma^\nu\gamma_\mu\gamma_\nu]\,.
\end{aligned}$$

These traces over products of an even number of γ-matrices should actually be calculated analogously to (12.63) and (12.66). However, the trace over a product of six (eight) γ-matrices, which appear here for the first time, leads in total to $6!!=15$ $(8!!=105)$ terms. Thus evaluating (12.72) with the previous calculational technique would be too involved. Instead we use the observation, that the contractions of γ-matrices occur in (12.72) within the trace, to our advantage. Namely it turns out that this circumstance drastically simplifies the trace calculation. With the help of the Clifford algebra (10.96) the contracted product of two γ-matrices can be calculated as follows:

$$\gamma^\mu\gamma_\mu = g_{\mu\nu}\gamma^\mu\gamma^\nu = \frac{1}{2}\,g_{\mu\nu}(\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu) = g_{\mu\nu}g^{\mu\nu} = \delta^\mu{}_\mu = 4\,. \qquad (12.73)$$

In case of one γ-matrix between the two contracted γ-matrices we get by applying the Clifford algebra (10.96)

$$\gamma^\mu\gamma^\nu\gamma_\mu = (-\gamma^\nu\gamma^\mu\gamma_\mu+2g^{\mu\nu})\gamma_\mu = -\gamma^\nu\gamma^\mu\gamma_\mu+2g^{\mu\nu}\gamma_\mu = -2\gamma^\nu\,. \qquad (12.74)$$

This result can be used to deal with two γ-matrices lying in between

$$\gamma^\mu\gamma^\nu\gamma^\kappa\gamma_\mu \;=\; (-\gamma^\nu\gamma^\mu+2g^{\mu\nu})\gamma^\kappa\gamma_\mu = -\gamma^\nu(\gamma^\mu\gamma^\kappa\gamma_\mu)+2g^{\mu\nu}\gamma^\kappa\gamma_\mu = 2[\gamma^\nu,\gamma^\kappa]_+ = 4g^{\nu\kappa}\,. \qquad (12.75)$$

And, correspondingly, we yield for three γ-matrices:

$$\gamma^\mu\gamma^\nu\gamma^\kappa\gamma^\lambda\gamma_\mu = (-\gamma^\nu\gamma^\mu + 2g^{\mu\nu})\gamma^\kappa\gamma\lambda\gamma_\mu = -\gamma^\nu(\gamma^\mu\gamma^\kappa\gamma^\lambda\gamma_\mu) + 2g^{\mu\nu}\gamma^\kappa\gamma^\lambda\gamma_\mu$$
$$= -4\gamma^\nu g^{\kappa\lambda} + 2(\gamma^\kappa\gamma^\lambda)\gamma^\nu = -4\gamma^\nu g^{\kappa\lambda} + 2(-\gamma^\lambda\gamma^\kappa + 2g^{\kappa\lambda})\gamma^\nu = -2\gamma^\lambda\gamma^\kappa\gamma^\nu \,. \tag{12.76}$$

These contraction rules for γ-matrices can now be iteratively applied to the respective terms in the trace (12.72) of the interference term (12.71):

$$1)(\gamma^\mu\gamma^\nu\gamma_\mu)\gamma_\nu = -2\gamma^\nu\gamma_\nu = -8\,, \tag{12.77}$$
$$2)\gamma^\mu \not{p}_{i_1}\gamma^\nu \not{p}_{f_1}\gamma_\mu\gamma_\nu = p_{i_1\kappa}p_{f_1\lambda}(\gamma^\mu\gamma^\kappa\gamma^\nu\gamma^\lambda\gamma_\mu)\gamma_\nu = p_{i_1\kappa}p_{f_1\lambda}(-2)\gamma^\lambda(\gamma^\nu\gamma^\kappa\gamma_\nu) = 4p_{i_1\kappa}p_{f_1\lambda}\gamma^\lambda\gamma^\kappa, \tag{12.78}$$
$$3)\gamma^\mu \not{p}_{i_1}\gamma^\nu\gamma_\mu \not{p}_{i_2}\gamma_\nu = p_{i_1\kappa}p_{i_2\lambda}(\gamma^\mu\gamma^\kappa\gamma^\nu\gamma_\mu)\gamma^\lambda\gamma_\nu = 4p_{i_1\kappa}p_{i_2\lambda}g^{\kappa\nu}\gamma^\lambda\gamma_\nu = 4p_{i_1\kappa}p_{i_2\lambda}\gamma^\lambda\gamma^\kappa\,, \tag{12.79}$$
$$4)\gamma^\mu\gamma^\nu \not{p}_{f_1}\gamma_\mu \not{p}_{i_2}\gamma_\nu = p_{i_1\kappa}p_{i_2\lambda}(\gamma^\mu\gamma^\kappa\gamma^\nu\gamma_\mu)\gamma^\lambda\gamma_\nu = 4p_{i_1\kappa}p_{i_2\lambda}\gamma^\lambda\gamma^\kappa\,, \tag{12.80}$$
$$5)\gamma^\mu \not{p}_{i_1}\gamma^\nu\gamma_\mu\gamma_\nu \not{p}_{f_2} = p_{i_1\kappa}p_{f_2\lambda}(\gamma^\mu\gamma^\kappa\gamma^\nu\gamma_\mu)\gamma_\nu\gamma^\lambda = 4p_{i_1\kappa}p_{f_2\lambda}g^{\kappa\nu}\gamma_\nu\gamma^\lambda = 4p_{i_1\kappa}p_{f_2\lambda}\gamma^\kappa\gamma^\lambda\,, \tag{12.81}$$
$$6)\gamma^\mu\gamma^\nu\gamma_\mu \not{p}_{i_2}\gamma_\nu \not{p}_{f_2} = p_{i_2\kappa}p_{f_2\lambda}(\gamma^\mu\gamma^\nu\gamma_\mu)\gamma^\kappa\gamma_\nu\gamma^\lambda = -2p_{i_2\kappa}p_{f_2\lambda}(\gamma^\nu\gamma^\kappa\gamma_\nu)\gamma^\lambda = 4p_{i_2\kappa}p_{f_2\lambda}\gamma^\kappa\gamma^\lambda\,, \tag{12.82}$$
$$7)\gamma^\mu\gamma^\nu \not{p}_{f_1}\gamma_\mu\gamma_\nu \not{p}_{f_2} = p_{f_1\kappa}p_{f_2\lambda}(\gamma^\mu\gamma^\nu\gamma^\kappa\gamma_\mu)\gamma_\nu\gamma^\lambda = 4p_{f_1\kappa}p_{f_2\lambda}\gamma^\kappa\gamma^\lambda\,, \tag{12.83}$$
$$8)\gamma^\mu \not{p}_{i_1}\gamma^\nu \not{p}_{f_1}\gamma_\mu \not{p}_{i_2}\gamma_\nu \not{p}_{f_2} \;= p_{i_1\kappa}p_{f_1\lambda}p_{i_2\sigma}p_{f_2\tau}(\gamma^\mu\gamma^\kappa\gamma^\nu\gamma^\lambda\gamma_\mu)\gamma^\tau\gamma_\nu\gamma^\tau \tag{12.84}$$
$$= -2p_{i_1\kappa}p_{f_1\lambda}p_{i_2\sigma}p_{f_2\tau}\gamma^\lambda(\gamma^\nu\gamma^\kappa\gamma^\sigma\gamma_\nu)\gamma^\tau = -8p_{i_1\kappa}p_{f_1\lambda}p_{i_2\sigma}p_{f_2\tau}g^{\kappa\sigma}\gamma^\lambda\gamma^\tau = -8(p_{i_1}p_{i_2})p_{f_1\lambda}p_{f_2\tau}\gamma^\lambda\gamma^\tau\,.$$

Using the auxiliary calculations (12.77)–(12.84) and taking into account (12.63) we obtain for (12.72) the following result

$$\mathrm{Tr}[...] = -8(p_{i_1}p_{i_2})p_{f_1\lambda}p_{f_2\kappa}\mathrm{Tr}[\gamma^\lambda\gamma^\kappa] + 4M^2c^2\left\{p_{i_1\kappa}p_{f_1\lambda}\mathrm{Tr}[\gamma^\lambda\gamma^\kappa] + p_{i_1\kappa}p_{i_2\lambda}\mathrm{Tr}[\gamma^\lambda\gamma^\kappa]\right. \tag{12.85}$$
$$\left.+p_{f_1\kappa}p_{i_2\lambda}\mathrm{Tr}[\gamma^\lambda\gamma^\kappa]\right\} + \left\{p_{i_1\kappa}p_{f_2\lambda}\mathrm{Tr}[\gamma^\kappa\gamma^\lambda] + p_{f_1\kappa}p_{f_2\lambda}\mathrm{Tr}[\gamma^\kappa\gamma^\lambda] + p_{i_2\kappa}p_{f_1\lambda}\mathrm{Tr}[\gamma^\kappa\gamma^\lambda]\right\} - 8M^4c^4\mathrm{Tr}[1]$$
$$= -32(p_{i_1}p_{i_2})(p_{f_1}p_{f_2}) - 32M^4c^4 + 16M^2c^2(p_{i_1}p_{f_1} + p_{i_1}p_{i_2} + p_{f_1}p_{i_2} + p_{i_1}p_{f_2} + p_{f_1}p_{f_2} + p_{i_2}p_{f_1})\,.$$

Substituting (12.85) into (12.71) leads to the final expression for the interference term between the direct and the exchange scattering:

$$\overline{\left|M_{fi}^{(2)}\right|_{\mathrm{i}}^2} \;=\; \frac{1}{4M^4c^4(p_{f_1}-p_{i_1})^2(p_{f_2}-p_{i_1})^2}\left[2(p_{i_1}p_{i_2})(p_{f_1}p_{f_2}) + M^4c^4\right. \tag{12.86}$$
$$\left.-M^2c^2(p_{i_1}p_{f_1} + p_{i_1}p_{i_2} + p_{f_1}p_{i_2} + p_{i_1}p_{f_2} + p_{f_1}p_{f_2} + p_{i_2}p_{f_1}) + (p_{f_1} \leftrightarrow p_{f_2})\right]\,.$$

We conclude that the direct term (12.68), the exchange term (12.70), and the interference term (12.86) have the common property of having a manifestly covariant form as they only depend on the scalar product of momenta. Thus, it only remains to relate these scalar product of momenta to observable properties of the scattering process. This is achieved by introducing the Lorentz invariant Mandelstam variables.

12.4 Mandelstam Variables

Let us investigate now the kinematics of a general two-body scattering process

$$A + B \qquad \Longrightarrow \qquad C + D\,, \tag{12.87}$$

which is described by the four-vector momenta p_a, p_b, p_c, and p_d with a total of 16 components. The equivalence principle of special relativity requires that observable quantities, such as the scattering cross section, can be expressed by Lorentz invariants.

12.4.1 General Case

With the four-vector momenta p_i with $i = a, b, c, d$, one can form ten different scalar products $p_i p_j$ with $i \leq j$, four of which are fixed by the relativistic energy-momentum dispersion relations

$$p_i^2 = M_i^2 c^2 \,. \tag{12.88}$$

The remaining six degrees of freedom are still interdependent, as each scattering process must satisfy the energy-momentum conservation law:

$$p_a + p_b = p_c + p_d \,. \tag{12.89}$$

These four additional conditions lead to the fact that, ultimately, two kinematic variables are sufficient to describe the two-body scattering process (12.87), provided that one can perform an average over the polarisations of both the initial and the final particles. For historical reasons, one describes the two-body scattering process (12.87) by the following three Lorentz-invariant Mandelstam variables

$$s = (p_a + p_b)^2 = (p_c + p_d)^2 \,, \tag{12.90}$$

$$t = (p_c - p_a)^2 = (p_d - p_b)^2 \,, \tag{12.91}$$

$$u = (p_c - p_b)^2 = (p_d - p_a)^2 \,. \tag{12.92}$$

Due to (12.88) and (12.90)–(12.91) each of the six scalar products $p_i p_j$ with $i < j$ can be expressed by the three Mandelstam variables:

$$p_a p_b = \frac{1}{2}\left(s - M_a^2 c^2 - M_b^2 c^2\right) , \qquad p_c p_d = \frac{1}{2}\left(s - M_c^2 c^2 - M_d^2 c^2\right) \tag{12.93}$$

$$p_a p_c = -\frac{1}{2}\left(t - M_a^2 c^2 - M_c^2 c^2\right) , \qquad p_b p_d = -\frac{1}{2}\left(t - M_b^2 c^2 - M_d^2 c^2\right) , \tag{12.94}$$

$$p_b p_c = -\frac{1}{2}\left(u - M_b^2 c^2 - M_c^2 c^2\right) , \qquad p_a p_d = -\frac{1}{2}\left(u - M_a^2 c^2 - M_d^2 c^2\right) . \tag{12.95}$$

Furthermore, it is possible to derive a restriction for the three Mandelstam variables. At first we obtain from (12.90)–(12.92)

$$s + t + u = (p_a + p_b)^2 + (p_a - p_c)^2 + (p_a - p_d)^2 = 3p_a^2 + p_b^2 + p_c^2 + p_d^2 + 2p_a(p_b - p_c - p_d) , \tag{12.96}$$

which reduces then with (12.88) and (12.89) to

$$s + t + u = p_a^2 + p_b^2 + p_c^2 + p_d^2 = (M_a^2 + M_b^2 + M_c^2 + M_d^2) c^2 \,. \tag{12.97}$$

Obviously, one of the three Mandelstam variables s, t, u can be eliminated with the help of (12.97). Nevertheless, all the three Mandelstam variables are often used, as the results for scattering cross sections turn out to acquire then a symmetrical form.

12.4.2 Equal Masses

Various simplifications occur for two-body scattering processes (12.87), where the involved particles have an equal mass:

$$M_a = M_b = M_c = M_d = M \, . \tag{12.98}$$

With the help of the identifications

$$p_a = p_{i_1} \, , \quad p_b = p_{i_2} \, , \quad p_c = p_{f_1} \, , \quad p_d = p_{f_2} \tag{12.99}$$

the relativistic energy-momentum dispersion relations (12.88) go over to

$$p_{i_1}^2 = p_{i_2}^2 = p_{f_1}^2 = p_{f_2}^2 = M^2c^2 \, . \tag{12.100}$$

Additionally the corresponding scalar products (12.93)–(12.95) read then as follows:

$$p_{i_1}p_{i_2} = p_{f_1}p_{f_2} = \frac{1}{2}\left(s - 2M^2c^2\right) \, , \tag{12.101}$$

$$p_{i_1}p_{f_1} = p_{i_2}p_{f_2} = -\frac{1}{2}\left(t - 2M^2c^2\right) \, , \tag{12.102}$$

$$p_{i_2}p_{f_1} = p_{i_1}p_{f_2} = -\frac{1}{2}\left(u - 2M^2c^2\right) \, . \tag{12.103}$$

And the definitions of the Mandelstam variables (12.90)–(12.92) take now the form

$$s = (p_{i_1} + p_{i_2})^2 = (p_{f_1} + p_{f_2})^2 \, , \tag{12.104}$$

$$t = (p_{f_1} - p_{i_1})^2 = (p_{f_2} - p_{i_2})^2 \, , \tag{12.105}$$

$$u = (p_{f_1} - p_{i_2})^2 = (p_{f_2} - p_{i_1})^2 \, , \tag{12.106}$$

whereby the restriction (12.97) coverts into

$$s + t + u = 4M^2c^2 \, . \tag{12.107}$$

12.4.3 Matrix Element

Now we return to the polarisation averaged matrix element of the Møller scattering and express the individual contributions with the help of (12.101)–(12.107) by the three Mandelstam variables s, t, u. For the direct term (12.68) we obtain

$$\overline{\left|M_{fi}^{(2)}\right|_{\mathrm{d}}^2} = \frac{(s - 2M^2c^2)^2 + (u - 2M^2c^2)^2 + 4M^2c^2t}{8M^4c^4t^2} \, . \tag{12.108}$$

The exchange term (12.70) follows from the direct term (12.68) by exchanging the final momenta p_{f_1} and p_{f_2}. At the level of the Mandelstam variables (12.101)–(12.107) this corresponds to an exchange of t and u, so we get

$$\overline{\left|M_{fi}^{(2)}\right|_{\mathrm{ex}}^2} = \frac{(s - 2M^2c^2)^2 + (t - 2M^2c^2)^2 + 4M^2c^2u}{8M^4c^4u^2} \, . \tag{12.109}$$

The Feynman diagrams in Fig. 12.1, whose absolute square and a subsequent polarization average leads to the terms (12.108) and (12.109), are also called after the Mandelstam variable in the denominator to graphically represent the t- and the u-channel, respectively. Correspondingly, the interference term (12.85) yields

$$\overline{\left|M_{fi}^{(2)}\right|_{\mathrm{i}}^{2}} = \frac{1}{2M^4c^4tu}\left\{\frac{1}{2}(s-2M^2c^2)^2 - M^2c^2\left[(s-2M^2c^2)-(t-2M^2c^2)-(u-2M^2c^2)\right]+2M^4c^4+(u\leftrightarrow t)\right\}. \tag{12.110}$$

Both contributions in (12.110) are apparently identical, we obtain

$$\overline{\left|M_{fi}^{(2)}\right|_{\mathrm{i}}^{2}} = \frac{1}{2M^4c^4tu}\left[\frac{1}{2}(s-2M^2c^2)^2 - M^2c^2(s-t-u)\right]. \tag{12.111}$$

Taking into account the restriction (12.107) this reduces to

$$\overline{\left|M_{fi}^{(2)}\right|_{\mathrm{i}}^{2}} = \frac{(s-2M^2c^2)(s-6M^2c^2)}{4M^4c^4tu}. \tag{12.112}$$

12.5 Center-of-Mass System

Now we specialize the kinematic analysis to the center of mass reference frame for two particles of equal mass.

12.5.1 Kinematics

Here the four-momentum vectors

$$p_{i_1} = \begin{pmatrix} E_{i_1}/c \\ \mathbf{p}_{i_1} \end{pmatrix}, \quad p_{i_2} = \begin{pmatrix} E_{i_2}/c \\ \mathbf{p}_{i_2} \end{pmatrix}, \quad p_{f_1} = \begin{pmatrix} E_{f_1}/c \\ \mathbf{p}_{f_1} \end{pmatrix}, \quad p_{f_2} = \begin{pmatrix} E_{f_2}/c \\ \mathbf{p}_{f_2} \end{pmatrix} \tag{12.113}$$

simplify even further. Namely, the center of mass system is distinguished from other inertial systems by the fact that the total momentum of the two incoming particles disappears:

$$\mathbf{p}_{i_1} + \mathbf{p}_{i_2} = \mathbf{0} \qquad \Longrightarrow \qquad \mathbf{p}_{i_1} = \mathbf{p}\,, \quad \mathbf{p}_{i_2} = -\mathbf{p}\,. \tag{12.114}$$

From their respective energy-momentum dispersion relations (12.100)

$$E_{i_1} = \sqrt{\mathbf{p}_{i_1}^2c^2 + M^2c^4}\,, \qquad E_{i_2} = \sqrt{\mathbf{p}_{i_2}^2c^2 + M^2c^4} \tag{12.115}$$

then follows that the energies of the two incoming particles coincide:

$$E_{i_1} = E_{i_2} = E\,. \tag{12.116}$$

From the momentum conservation (12.89) in the center of mass reference frame follows with (12.99) and (12.114) for the momenta of the two outgoing particles

$$\mathbf{p}_{f_1} + \mathbf{p}_{f_2} = \mathbf{0} \qquad \Longrightarrow \qquad \mathbf{p}_{f_1} = \mathbf{p}' \,, \quad \mathbf{p}_{f_2} = -\mathbf{p}' \,. \tag{12.117}$$

Thus, the corresponding energy-momentum dispersion relations (12.100)

$$E_{f_1} = \sqrt{\mathbf{p}_{f_1}^2 c^2 + M^2 c^4} \,, \qquad E_{f_2} = \sqrt{\mathbf{p}_{f_2}^2 c^2 + M^2 c^4} \tag{12.118}$$

imply that also the energies of the two outgoing particles are equal:

$$E_{f_1} = E_{f_2} = E' \,. \tag{12.119}$$

And from the energy conservation (12.89) in the center of mass reference frame

$$E_a + E_b = E_c + E_d \tag{12.120}$$

then follows with (12.99), (12.116), and (12.119) that the energy of the incoming and the outgoing particles E and E' coincide:

$$E_{i_1} + E_{i_2} = E_{f_1} + E_{f_2} \qquad \Longrightarrow \qquad E = E' \,. \tag{12.121}$$

We conclude from (12.114), (12.116), (12.117), (12.119), and (12.121) that the four-momentum vectors (12.113) are given in the center of mass reference frame as follows:

$$p_{i_1} = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix} , \quad p_{i_2} = \begin{pmatrix} E/c \\ -\mathbf{p} \end{pmatrix} , \quad p_{f_1} = \begin{pmatrix} E/c \\ \mathbf{p}' \end{pmatrix} , \quad p_{f_2} = \begin{pmatrix} E/c \\ -\mathbf{p}' \end{pmatrix} . \tag{12.122}$$

For the Mandelstam variables (12.104)–(12.106) this has due to (12.115), (12.118), (12.121), and (12.122) the consequence

$$s = (p_{i_1} + p_{i_2})^2 = \begin{pmatrix} 2E/c \\ \mathbf{0} \end{pmatrix}^2 = \frac{4E^2}{c^2} \,, \tag{12.123}$$

$$t = (p_{f_1} - p_{i_1})^2 = \begin{pmatrix} 0 \\ \mathbf{p}' - \mathbf{p} \end{pmatrix}^2 = -(\mathbf{p}' - \mathbf{p})^2 = -2\mathbf{p}^2(1 - \cos\theta) \,, \tag{12.124}$$

$$u = (p_{f_1} - p_{i_2})^2 = \begin{pmatrix} 0 \\ \mathbf{p}' + \mathbf{p} \end{pmatrix}^2 = -(\mathbf{p}' + \mathbf{p})^2 = -2\mathbf{p}^2(1 + \cos\theta) \,. \tag{12.125}$$

Here θ denotes the angle between the incoming and the outgoing electrons, which coincides with the angle between the momenta $\mathbf{p}$ and $\mathbf{p}'$ as illustrated in Fig. 12.2. Obviously, the Mandelstam variables s, t, u in the center of mass reference frame (12.123)–(12.125) satisfy the restriction (12.107) due to the relativistic energy-momentum dispersion relation (12.114)–(12.116):

$$s + t + u = \frac{4E^2}{c^2} - 2\mathbf{p}^2(1 - \cos\theta) - 2\mathbf{p}^2(1 + \cos\theta) = \frac{4}{c^2}\left(E^2 - \mathbf{p}^2 c^2\right) = 4M^2 c^2 \,. \tag{12.126}$$

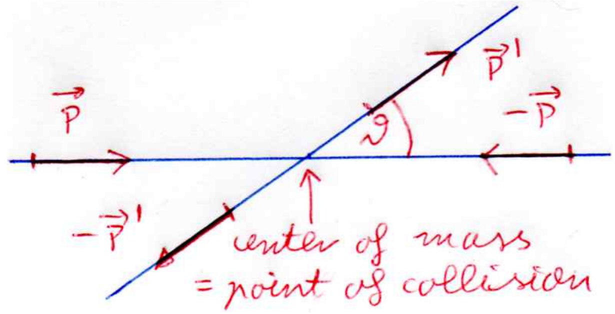


Figure 12.2: Geometry of the elastic Møller scattering in the center of mass reference frame with two incoming (outgoing) electrons of momenta $\pm\mathbf{p}$ ($\pm\mathbf{p}'$).

Furthermore, we read off from (12.114)–(12.116) that the two Mandelstam variables (12.124) and (12.125) can be rewritten as

$$t = -2\,\frac{E^2 - M^2c^4}{c^2}\,(1-\cos\theta)\,, \tag{12.127}$$

$$u = -2\,\frac{E^2 - M^2c^4}{c^2}\,(1+\cos\theta)\,. \tag{12.128}$$

Thus, for a scattering process of two particles with equal masses the Mandelstam variables (12.123), (12.127), and (12.128) in the center of mass reference frame depend on both the scattering energy E and the scattering angle θ.

12.5.2 Matrix Element

With the help of (12.123), (12.127), and (12.128) the individual contributions to the polarisation-averaged squared matrix element for the Møller scattering can be expressed as follows. The direct term (12.108) goes over into

$$\overline{|M_{fi}^{(2)}|^2_{\mathrm{d}}} = \frac{1}{8M^4c^4(E^2-M^2c^4)^2(1-\cos\theta)^2} \times\Big\{(2E^2-M^2c^4)^2 + \big[(E^2-M^2c^4)(1+\cos\theta)+2M^2c^4\big]^2 - 2M^2c^4(E^2-M^2c^4)(1-\cos\theta)\Big\}\,, \tag{12.129}$$

the exchange term (12.109) reads

$$\overline{|M_{fi}^{(2)}|^2_{\mathrm{ex}}} = \frac{1}{8M^4c^4(E^2-M^2c^4)^2(1+\cos\theta)^2} \times\Big\{(2E^2-M^2c^4)^2 + \big[(E^2-M^2c^4)(1-\cos\theta)+M^2c^4\big]^2 - 2M^2c^4(E^2-M^2c^4)(1+\cos\theta)\Big\}\,, \tag{12.130}$$

and the interference term (12.112) results in

$$\overline{|M_{fi}^{(2)}|^2_{\mathrm{i}}} = \frac{(2E^2-M^2c^4)(2E^2-3M^2c^4)}{4M^4c^4(E^2-M^2c^4)^2(1-\cos\theta)(1+\cos\theta)}\,. \tag{12.131}$$

These three contributions are now to added:

$$\overline{|M_{fi}^{(2)}|^2} = \overline{|M_{fi}^{(2)}|^2_{\mathrm{d}}} + \overline{|M_{fi}^{(2)}|^2_{\mathrm{ex}}} + \overline{|M_{fi}^{(2)}|^2_{\mathrm{i}}} = \frac{f(\theta)}{8M^4c^4(E^2-M^2c^4)^2(1-\cos\theta)^2(1+\cos\theta)^2}\,. \tag{12.132}$$

Due to straight-forward but lengthy manipulations the angle-dependent numerator results in

$$\begin{aligned} f(\theta) &= (1+2\cos\theta+\cos^2\theta)\left[(2E^2-M^2c^4)^2+E^4+2E^2(E^2-M^2c^4)\cos\theta\right. \\ &\quad \left.+(E^2-M^2c^4)\cos^2\theta-2M^2c^4(E^2-M^2c^4)(1-\cos\theta)\right]+(1-2\cos\theta+\cos^2\theta) \\ &\quad \times\left[(2E^2-M^2c^4)^2+E^4-2E^2(E^2-M^2c^4)\cos\theta+(E^2-M^2c^4)\cos^2\theta\right. \\ &\quad \left.-2M^2c^4(E^2-M^2c^4)(1+\cos\theta)\right]+2(1-\cos^2\theta)(2E^2-M^2c^4)(2E^2-3M^2c^4)\,. \end{aligned} \tag{12.133}$$

It turns out to be useful to take into account the trigonometric Pythagoras

$$\sin^2\theta+\cos^2\theta=1 \tag{12.134}$$

in order to further simplify the expression (12.133), yielding after some further algebraic manipulations the concise result:

$$f(\theta)=2\left[4(2E^2-M^2c^4)^2-(8E^4-4M^2c^4E^2-M^4c^8)\sin^2\theta+(E^2-M^2c^4)^2\sin^4\theta\right]\,. \tag{12.135}$$

Inserting (12.135) into (12.132) leads together with (12.134) the following angular dependence of the polarisation-averaged squared matrix element of the Møller scattering in the center of mass reference frame:

$$\overline{\left|M_{fi}^{(2)}\right|^2}=\frac{4(2E^2-M^2c^4)^2-(8E^4-4M^2c^4E^2-M^4c^8)\sin^2\theta+(E^2-M^2c^4)^2\sin^4\theta}{4M^4c^4(E^2-M^2c^4)^2\sin^4\theta}\,. \tag{12.136}$$

12.6 Transition Rate Per Volume

Now we return to the perturbative result for the scattering matrix of the Møller scattering (12.46) and evaluate its absolute square:

$$\begin{aligned} \left|S_{fi}^{(2)}\right|^2 &= \frac{\hbar^2e^4}{\epsilon_0^2c^2}\,(2\pi\hbar)^8\delta(0)\delta(p_{f_1}+p_{f_2}-p_{i_1}-p_{i_2}) \\ &\quad \times\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{i_1}}}\,\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{i_2}}}\,\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{f_1}}}\,\frac{Mc^2}{(2\pi\hbar)^3E_{\mathbf{p}_{f_2}}}\left|M_{fi}^{(2)}\right|^2. \end{aligned} \tag{12.137}$$

The transition probability (12.137) is formally infinite due to the appearance of the singular factor $\delta(0)$. In order to deal with this singularity we reconsider the decomposition of the field operator $\hat{\psi}(x)$ into plane waves according to (12.10). However, instead we now assume, as is usual in solid-state physics, that an electron is located in a finite box with volume V. Then we have instead of (12.10) the following plane wave decomposition:

$$\hat{\psi}(x)=\sum_{\mathbf{p}}\sum_s\sqrt{\frac{Mc^2}{VE_{\mathbf{p}}}}\left\{e^{-ipx/\hbar}\,u(\mathbf{p},s)\,\hat{b}_{\mathbf{p},s}+e^{ipx/\hbar}\,v(\mathbf{p},s)\,\hat{d}^\dagger_{\mathbf{p},s}\right\}\,. \tag{12.138}$$

While the orthonormality relation of the plane waves in the continuum reads

$$\int d^4x\,e^{i(p-p')x/\hbar}=(2\pi\hbar)^4\,\delta(p-p')\,, \tag{12.139}$$

it reads in a finite box V within a finite observation time T

$$\int_V d^3x \int_{-Tc/2}^{Tc/2} dx_0 \, e^{i(p-p')x/\hbar} = VTc\, \delta_{p,p'} \,. \tag{12.140}$$

Note that the delta function in (12.139) is substituted by the Kronecker symbol in (12.140). Therefore, comparing (12.139) and (12.140) yields on formal grounds the following substitution rule

$$(2\pi\hbar)^4 \, \delta(0) = VTc \,, \tag{12.141}$$

which suggests an appropriate regularisation for the singular term $\delta(0)$. We now follow the calculation strategy that both the initial and the final states of the scattering process are still considered to be continuous, while the intermediate states are treated as discrete ones as in (12.138). Thus, we would then have to repeat the whole perturbative calculation for the Møller scattering and calculate how the scattering matrix element (12.45) and its absolute square (12.137) change from this modified point of view. This would yield the result (12.137) with a regularization by the formal substitution rule (12.141) together with the identification $(2\pi\hbar)^3 \to V$. With this we obtain for the transition rate per volume from (12.137) and (12.141):

$$\frac{\left|S_{fi}^{(2)}\right|^2}{VT} = \frac{\hbar^2 e^4}{\epsilon_0^2 c} \, (2\pi\hbar)^4 \, \delta(p_{f_1} + p_{f_2} - p_{i_1} - p_{i_2}) \, \frac{(Mc^2)^4}{V^4 E_{\mathbf{p}_{i_1}} E_{\mathbf{p}_{i_2}} E_{\mathbf{p}_{f_1}} E_{\mathbf{p}_{f_2}}} \left|M_{fi}^{(2)}\right|^2 . \tag{12.142}$$

This transition rate per volume is then to be integrated or summed up over all final states:

$$\frac{V}{(2\pi\hbar)^3} \int d^3p_{f_1} \, \frac{V}{(2\pi\hbar)^3} \int d^3p_{f_2} \sum_{s_{f_1}} \sum_{s_{f_2}} \tag{12.143}$$

and it is to be averaged over the polarizations of the initial states:

$$\frac{1}{4} \sum_{s_{i_1}} \sum_{s_{i_2}} \tag{12.144}$$

This yields the averaged transition rate per volume:

$$W = \frac{1}{4} \sum_{s_{i_1}} \sum_{s_{i_2}} \sum_{s_{f_1}} \sum_{s_{f_2}} \frac{V}{(2\pi\hbar)^3} \int d^3p_{f_1} \, \frac{V}{(2\pi\hbar)^3} \int d^3p_{f_2} \, \frac{|S_{fi}^{(2)}|^2}{VT} \,. \tag{12.145}$$

Inserting (12.142) into (12.145) as well as taking into account (12.48) then leads to

$$W = \frac{e^4 M^2 c^4}{4\pi^2 \epsilon_0^2 c V^2 E_{\mathbf{p}_{i_1}} E_{\mathbf{p}_{i_2}}} \int d^3p_{f_1} \int d^3p_{f_2} \, \delta(p_{f_1} + p_{f_2} - p_{i_1} - p_{i_2}) \, \frac{M^2 c^4}{E_{\mathbf{p}_{f_1}} E_{\mathbf{p}_{f_2}}} \, \overline{\left|M_{fi}^{(2)}\right|^2} , \tag{12.146}$$

where the polarisation average of the squared matrix element (12.48) was already calculated in (12.136). The two integrals over the outgoing momenta are of the following form:

$$I = \int \frac{d^3p_{f_1}}{2E_{\mathbf{p}_{f_1}}} \frac{d^3p_{f_2}}{2E_{\mathbf{p}_{f_2}}} \, \delta(p_{f_1} + p_{f_2} - p_{i_1} - p_{i_2}) \, f(\mathbf{p}_{f_1}, \mathbf{p}_{f_2}) \,. \tag{12.147}$$

In order to evaluate (12.147) we perform at first the following auxiliary calculation

$$\begin{aligned} \int_0^\infty dp^0 \delta(p^2 - M^2c^2) &= \int_0^\infty dp^0 \delta((p^0)^2 - \mathbf{p}^2 - M^2c^2) \\ &= \int_0^\infty dp^0 \left[\frac{c}{2E_{\mathbf{p}}} \, \delta\left(p^0 - \frac{E_{\mathbf{p}}}{c}\right) + \frac{c}{2E_{\mathbf{p}}} \, \delta\left(p^0 + \frac{E_{\mathbf{p}}}{c}\right)\right] = \frac{c}{2E_{\mathbf{p}}} \, . \end{aligned} \tag{12.148}$$

Note that we used here the distributional rule

$$\delta(g(x)) = \sum_i \frac{1}{|g'(x_i)|} \, \delta(x - x_i) \, , \qquad g(x_i) = 0 \tag{12.149}$$

for the function

$$g(p^0) = (p^0)^2 - \frac{E_{\mathbf{p}}^2}{c^2} = \left(p^0 - \frac{E_{\mathbf{p}}}{c}\right)\left(p^0 + \frac{E_{\mathbf{p}}}{c}\right) \, , \qquad g'\left(p^0 = \pm \frac{E_{\mathbf{p}}}{c}\right) = \pm 2 \, \frac{E_{\mathbf{p}}}{c} \, . \tag{12.150}$$

Inserting (12.148) into (12.147) leads to

$$\begin{aligned} I &= \frac{1}{c} \int \frac{d^3 p_{f_1}}{2E_{\mathbf{p}_{f_1}}} \int d^3 p_{f_2} \int_0^\infty dp_{f_2}^0 \, \delta(p_{f_2}^2 - M^2c^2)\delta(p_{f_1} + p_{f_2} - p_{i_1} - p_{i_2}) f(\mathbf{p}_{f_1}, \mathbf{p}_{f_2}) \\ &= \frac{1}{c} \int \frac{d^3 p_{f_1}}{2E_{\mathbf{p}_{f_1}}} \int d^4 p_{f_2} \Theta(p_{f_2}^0)\delta(p_{f_2}^2 - M^2c^2)\delta(p_{f_1} + p_{f_2} - p_{i_1} - p_{i_2}) f(\mathbf{p}_{f_1}, \mathbf{p}_{f_2}) \, . \end{aligned} \tag{12.151}$$

Now the four-dimensional p_{f_2}-integral can formally be evaluated and we obtain the intermediate result:

$$I = \frac{1}{c} \int \frac{d^3 p_{f_1}}{2E_{\mathbf{p}_{f_1}}} \, \Theta(p_{i_1}^0 + p_{i_2}^0 - p_{f_1}^0)\delta((p_{i_1} + p_{i_2} - p_{f_1})^2 - M^2c^2) f(\mathbf{p}_{f_1}, \mathbf{p}_{i_1} + \mathbf{p}_{i_2} - \mathbf{p}_{f_1}) \, . \tag{12.152}$$

In view of evaluating also the $\mathbf{p}_{f_1}$-integral we specialise for the center of mass reference frame, so that we can apply the considerations from the previous section. However, in contrast to (12.122), we cannot use the conservation of energy, as this is only established due the delta function in (12.152). Therefore, based on (12.114) and (12.117), we have to generalise the four-momentum vectors (12.122) accordingly:

$$p_{i_1} = \begin{pmatrix} E_{\mathbf{p}}/c \\ \mathbf{p} \end{pmatrix} \, , \quad p_{i_2} = \begin{pmatrix} E_{\mathbf{p}}/c \\ -\mathbf{p} \end{pmatrix} \, , \quad p_{f_1} = \begin{pmatrix} E_{\mathbf{p}'}/c \\ \mathbf{p}' \end{pmatrix} \, , \quad p_{f_2} = \begin{pmatrix} E_{\mathbf{p}'}/c \\ -\mathbf{p}' \end{pmatrix} \, . \tag{12.153}$$

From this we read off

$$p_{i_1}^0 + p_{i_2}^0 - p_{f_1}^0 = \frac{2E_{\mathbf{p}} - E_{\mathbf{p}'}}{c} \, , \tag{12.154}$$

$$\mathbf{p}_{i_1} + \mathbf{p}_{i_2} - \mathbf{p}_{f_1} = -\mathbf{p}' \, , \tag{12.155}$$

as well as

$$\begin{aligned} (p_{i_1} + p_{i_2} - p_{f_1})^2 &= (p_{i_1} + p_{i_2})^2 - 2(p_{i_1} + p_{i_2})p_{f_1} + p_{f_1}^2 \\ &= \left(\frac{2E_{\mathbf{p}}}{c}\right)^2 - 2\frac{2E_{\mathbf{p}}}{c}\frac{E_{\mathbf{p}'}}{c} + \frac{E_{\mathbf{p}'}^2}{c^2} - \mathbf{p}'^2 = \frac{4E_{\mathbf{p}}}{c^2}(E_{\mathbf{p}} - E_{\mathbf{p}'}) + M^2c^2 \, , \end{aligned} \tag{12.156}$$

where in the last step the relativistic energy-momentum dispersion

$$E_{\mathbf{p}'}^2 = \mathbf{p}'^2 c^2 + M^2 c^4 \tag{12.157}$$

was used. With this (12.152) reads using (12.149) in the center of mass reference frame

$$I = \frac{1}{c} \int \frac{d^3 p'}{2E_{\mathbf{p}'}} \Theta(2E_{\mathbf{p}} - E_{\mathbf{p}'}) \frac{c^2}{4E_{\mathbf{p}}} \delta(E_{\mathbf{p}} - E_{\mathbf{p}'}) f(\mathbf{p}', -\mathbf{p}') \,. \tag{12.158}$$

In view of a further evaluation of the $\mathbf{p}'$-integral, we introduce spherical coordinates for which we get

$$d^3 p' = |\mathbf{p}'|^2 \, d|\mathbf{p}'| d\Omega \,, \qquad d\Omega = \sin\theta d\theta d\phi \,. \tag{12.159}$$

Furthermore, due to a comparison of (12.136), (12.146), and (12.147), we identify $f(\mathbf{p}', -\mathbf{p}')$ with $F(|\mathbf{p}'|, \theta)$:

$$I = \frac{c}{8E_{\mathbf{p}}} \int_0^\infty d|\mathbf{p}'| \frac{|\mathbf{p}'|^2}{E_{\mathbf{p}'}} \int d\Omega \, \Theta(2E_{\mathbf{p}} - E_{\mathbf{p}'}) \delta(E_{\mathbf{p}} - E_{\mathbf{p}'}) F(|\mathbf{p}'|, \theta) \,. \tag{12.160}$$

Due to the relativistic energy-momentum dispersion (12.157) we obtain the following substitution:

$$2E_{\mathbf{p}'} \, dE_{\mathbf{p}'} = 2|\mathbf{p}'| \, d|\mathbf{p}'| c^2 \qquad \Longrightarrow \qquad d|\mathbf{p}'| = \frac{E_{\mathbf{p}'}}{|\mathbf{p}'| c^2} \, dE_{\mathbf{p}'} \,, \tag{12.161}$$

so that (12.160) goes over into

$$\begin{aligned} I &= \frac{c}{8E_{\mathbf{p}}} \int_0^\infty dE_{\mathbf{p}'} \frac{E_{\mathbf{p}'}}{|\mathbf{p}'| c^2} \frac{|\mathbf{p}'|^2}{E_{\mathbf{p}'}} \int d\Omega \, \Theta(2E_{\mathbf{p}} - E_{\mathbf{p}'}) \delta(E_{\mathbf{p}} - E_{\mathbf{p}'}) F(|\mathbf{p}'|, \theta) \qquad (12.162) \\ &= \frac{1}{8c^2 E_{\mathbf{p}}} \int d\Omega \int_0^\infty dE_{\mathbf{p}'} \sqrt{E_{\mathbf{p}'}^2 - M^2 c^4} \, \Theta(2E_{\mathbf{p}} - E_{\mathbf{p}'}) \delta(E_{\mathbf{p}} - E_{\mathbf{p}'}) F\left(\sqrt{E_{\mathbf{p}'}^2 / c^2 - M^2 c^2}, \theta\right) \,. \end{aligned}$$

Now the $E_{\mathbf{p}'}$ integral can be performed due to the delta function, yielding, finally, the conservation of energy $E_{\mathbf{p}'} = E_{\mathbf{p}}$:

$$I = \frac{\sqrt{E_{\mathbf{p}}^2 - M^2 c^4}}{8c^2 E_{\mathbf{p}}} \int d\Omega \, F\left(\sqrt{E_{\mathbf{p}}^2 / c^2 - M^2 c^2}, \theta\right) \,. \tag{12.163}$$

Based on the result (12.163) for the two integrals (12.147) in the center of mass reference frame, we now obtain for the averaged transition rate per volume (12.146) with identifying $E = E_{\mathbf{p}}$:

$$W = \frac{e^4}{\pi^2 \epsilon_0^2 c} \frac{M^4 c^8}{V^2 E^2} \frac{\sqrt{E^2 - M^2 c^4}}{8c^2 E} \int d\Omega \, \overline{|M_{fi}^{(2)}|^2} \,. \tag{12.164}$$

Checking the physical units of (12.164) by taking into account (12.136) yields, indeed, as expected: $[W] = 1/(\mathrm{s}\,\mathrm{m}^3)$.

12.7 Cross Section

In order to calculate the cross section we still need the number of incoming electrons per time unit and area. For this purpose, we consider again the normal order of the four current density operator (12.17), but this time for electrons being confined in a finite volume V. To this end we apply (12.138) and its Dirac adjoint, yielding instead of (12.17):

$$: \hat{j}^\mu(x) := c \sum_{\mathbf{p}_1} \sum_{\mathbf{p}_2} \sum_{s_1} \sum_{s_2} \sqrt{\frac{Mc^2}{VE_{\mathbf{p}_1}}} \sqrt{\frac{Mc^2}{VE_{\mathbf{p}_2}}} \tag{12.165}$$
$$\Big\{ e^{i(p_2-p_1)x/\hbar}\, \overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}_{\mathbf{p}_1,s_1} + e^{i(p_2+p_1)x/\hbar}\, \overline{u}(\mathbf{p}_2,s_2)\gamma^\mu v(\mathbf{p}_1,s_1)\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}^\dagger_{\mathbf{p}_1,s_1}$$
$$+ e^{-i(p_2+p_1)x/\hbar}\, \overline{v}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\hat{d}_{\mathbf{p}_2,s_2}\hat{b}_{\mathbf{p}_1,s_1} - e^{-i(p_2+p_1)x/\hbar}\, \overline{v}(\mathbf{p}_2,s_2)\gamma^\mu v(\mathbf{p}_1,s_1)\hat{d}^+_{\mathbf{p}_2,s_2}\hat{d}_{\mathbf{p}_1,s_1} \Big\} .$$

Evaluating the matrix element of (12.165) with respect to the initial state (12.13) leads to

$$\begin{aligned} \langle\psi_i| : \hat{j}^\mu(x) : |\psi_i\rangle &= c \sum_{\mathbf{p}_1} \sum_{\mathbf{p}_2} \sum_{s_1} \sum_{s_2} \sqrt{\frac{Mc^2}{VE_{\mathbf{p}_1}}} \sqrt{\frac{Mc^2}{VE_{\mathbf{p}_2}}} \\ &\quad \times e^{i(p_2-p_1)x/\hbar}\, \overline{u}(\mathbf{p}_2,s_2)\gamma^\mu u(\mathbf{p}_1,s_1)\, C(\mathbf{p}_1,\mathbf{p}_2;s_1,s_2)\,, \end{aligned} \tag{12.166}$$

where we have introduced as an abbreviation the vaccum expectation value

$$C(\mathbf{p}_1,\mathbf{p}_2;s_1,s_2) = \langle 0|\, \hat{b}_{\mathbf{p}_{i_2},s_{i_2}} \left(\hat{b}_{\mathbf{p}_{i_1},s_{i_1}} \hat{b}^\dagger_{2,s_2}\right) \left(\hat{b}_{\mathbf{p}_1,s_1} \hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}\right) \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\, |0\rangle\ . \tag{12.167}$$

Applying the anti-commutator algebra (10.407) we obtain from (12.167)

$$\begin{aligned} \cdots &= \langle 0|\, \hat{b}_{\mathbf{p}_{i_2},s_{i_2}} \left(-\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}_{\mathbf{p}_{i_1},s_{i_1}} + \delta_{\mathbf{p}_{i_1},\mathbf{p}_2}\delta_{s_{i_1},s_2}\right) \left(-\hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}\hat{b}_{\mathbf{p}_1,s_1} + \delta_{\mathbf{p}_{i_1},\mathbf{p}_1}\delta_{s_{i_1},s_1}\right) \hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\, |0\rangle \\ &= \langle 0|\, (\hat{b}_{\mathbf{p}_{i_2},s_{i_2}}\hat{b}^\dagger_{\mathbf{p}_2,s_2})\hat{b}_{\mathbf{p}_{i_1},s_{i_1}}\hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}(\hat{b}_{\mathbf{p}_1,s_1}\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}})\,|0\rangle - \delta_{\mathbf{p}_{i_1},\mathbf{p}_2}\delta_{s_{i_1},s_2}\,\langle 0|\,\hat{b}_{\mathbf{p}_{i_2},s_{i_2}}\hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}\hat{b}_{\mathbf{p}_1,s_1}\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\,|0\rangle \\ &\quad -\delta_{\mathbf{p}_{i_1},\mathbf{p}_1}\delta_{s_{i_1},s_1}\,\langle 0|\,\hat{b}_{\mathbf{p}_{i_2},s_{i_2}}\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}_{\mathbf{p}_{i_1},s_{i_1}}\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\,|0\rangle + \delta_{\mathbf{p}_{i_1},\mathbf{p}_1}\delta_{\mathbf{p}_{i_1},\mathbf{p}_2}\delta_{s_{i_1},s_2}\delta_{s_{i_1},s_2}\,\langle 0|\,\hat{b}_{\mathbf{p}_{i_2},s_{i_2}}\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\,|0\rangle\ . \end{aligned} \tag{12.168}$$

Since it is assumed that the initial momenta $\mathbf{p}_{i_1}, \mathbf{p}_{i_2}$ differ from each other, the respective fermionic operators $\hat{b}_{\mathbf{p}_{i_2},s_{i_2}}$, $\hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}$ and $\hat{b}_{\mathbf{p}_{i_1},s_{i_1}}$, $\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}$ anticommute, respectively. Therefore, the second and the third matrix element in (12.168) disappear, so we obtain

$$\begin{aligned} C(\mathbf{p}_1,\mathbf{p}_2;s_1,s_2) &= \langle 0|\Big(-\hat{b}^\dagger_{\mathbf{p}_2,s_2}\hat{b}_{\mathbf{p}_{i_2},s_{i_2}} + \delta_{\mathbf{p}_{i_2},\mathbf{p}_2}\delta_{s_{i_2},s_2}\Big)\,\hat{b}_{\mathbf{p}_{i_1},s_{i_1}}\hat{b}^\dagger_{\mathbf{p}_{i_1},s_{i_1}}\Big(-\hat{b}^\dagger_{\mathbf{p}_{i_2},s_{i_2}}\hat{b}_{\mathbf{p}_1,s_1} + \delta_{\mathbf{p}_{i_2},\mathbf{p}_1}\delta_{s_{i_2},s_1}\Big)\,|0\rangle \\ &+\delta_{\mathbf{p}_{i_1},\mathbf{p}_1}\delta_{\mathbf{p}_{i_2},\mathbf{p}_2}\delta_{s_{i_1},s_1}\delta_{s_{i_1},s_2}\,\langle 0|\,1 - \hat{b}^\dagger_{i_2,s_{i_2}}\hat{b}_{\mathbf{p}_{i_2},s_{i_2}}\,|0\rangle = \delta_{\mathbf{p}_1,\mathbf{p}_2}\delta_{s_1,s_2}\Big(\delta_{\mathbf{p}_1,\mathbf{p}_{i_1}}\delta_{s_1,s_{i_1}} + \delta_{\mathbf{p}_1,\mathbf{p}_{i_2}}\delta_{s_1,s_{i_2}}\Big)\ . \end{aligned} \tag{12.169}$$

Inserting the vaccum expectation value (12.169) into (12.166) leads to the matrix element

$$\langle\psi_i| : \hat{j}^\mu(x) : |\psi_i\rangle = c\,\frac{Mc^2}{VE_{\mathbf{p}_{i_1}}}\,\overline{u}(\mathbf{p}_{i_1},s_{i_1})\gamma^\mu u(\mathbf{p}_{i_1},s_{i_1}) + c\,\frac{Mc^2}{VE_{\mathbf{p}_{i_2}}}\,\overline{u}(\mathbf{p}_{i_2},s_{i_2})\gamma^\mu u(\mathbf{p}_{i_2},s_{i_2})\,. \tag{12.170}$$

Afterwards, we average this current density with respect to the polarizations of both incoming electrons:

$$J^\mu = \frac{1}{4}\sum_{s_{i_1}}\sum_{s_{i_2}} \langle\psi_i| : \hat{j}^\mu(x) : |\psi_i\rangle\ . \tag{12.171}$$

Substituting (12.170) into (12.171) we obtain

$$J^\mu = \frac{Mc^3}{2VE_{\mathbf{p}_{i_1}}} \sum_{s_{i_1}} \overline{u}_\alpha(\mathbf{p}_{i_1}, s_{i_1}) \gamma^\mu_{\alpha\beta} u_\beta(\mathbf{p}_{i_1}, s_{i_1}) + \frac{Mc^3}{2VE_{\mathbf{p}_{i_2}}} \sum_{s_{i_2}} \overline{u}_\alpha(\mathbf{p}_{i_2}, s_{i_2}) \gamma^\mu_{\alpha\beta} u_\beta(\mathbf{p}_{i_2}, s_{i_2}) \, . \quad (12.172)$$

The polarisation sums with respect to s_{i_1}, s_{i_2} were already calculated according to (10.438) and (10.447), yielding

$$J^\mu = \frac{Mc^3}{2VE_{\mathbf{p}_{i_1}}} \gamma^\mu_{\alpha\beta} \left(\frac{p_{i_1\nu}\gamma^\nu + Mc}{2Mc} \right)_{\alpha\beta} + \frac{Mc^3}{2VE_{\mathbf{p}_{i_2}}} \gamma^\mu_{\alpha\beta} \left(\frac{p_{i_2\nu}\gamma^\nu + Mc}{2Mc} \right)_{\alpha\beta} . \quad (12.173)$$

The sums with respect to the spinorial indices can be interpreted as traces:

$$J^\mu = \frac{c^2}{4VE_{\mathbf{p}_{i_1}}} \left\{ p_{i_1\nu} \operatorname{Tr}[\gamma^\mu\gamma^\nu] + Mc \operatorname{Tr}[\gamma^\mu] \right\} + \frac{c^2}{4VE_{\mathbf{p}_{i_2}}} \left\{ p_{i_2\nu} \operatorname{Tr}[\gamma^\mu\gamma^\nu] + Mc \operatorname{Tr}[\gamma^\mu] \right\} . \quad (12.174)$$

Due to the trace rules (12.62) and (12.63) the polarization averaged current density (12.174) reduces to

$$J^\mu = \frac{p^\mu_{i_1} c^2}{VE_{\mathbf{p}_{i_1}}} + \frac{p^\mu_{i_2} c^2}{VE_{\mathbf{p}_{i_2}}} \, . \quad (12.175)$$

In the center of mass reference frame (12.122) applies, so that the polarization averaged current density (12.175) vanishes:

$$J^\mu = 0 \, . \quad (12.176)$$

The relative current density, however, turns out to be

$$\Delta J = \frac{2|\mathbf{p}|c^2}{VE_{\mathbf{p}}} \quad (12.177)$$

and has, indeed, the correct physics unit $[\Delta J] = 1/(\mathrm{s\,m}^2)$. The cross section follows now from the quotient of the averaged transition rate per volume W and the averaged relative current density ΔJ per volume:

$$\sigma = \frac{W}{\Delta J / V} \, . \quad (12.178)$$

Substituting (12.164) and (12.177) into (12.178) yields the total cross section in the form of

$$\sigma = \int d\Omega \, \frac{d\sigma}{d\Omega} \, , \quad (12.179)$$

so that the differential cross section is definied by

$$\frac{d\sigma}{d\Omega} = \frac{e^4 M^4 c^4}{16\pi^2 \epsilon_0^2 E^2} \overline{\left| M^{(2)}_{fi} \right|^2} \, . \quad (12.180)$$

Inserting the polarisation-averaged matrix element (12.136) therein then yields

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 \hbar^2 c^2}{4E^2} \left[1 - \frac{8E^4 - 4M^2c^4E^2 - M^4c^8}{(E^2 - M^2c^4)^2} \frac{1}{\sin^2\theta} + \frac{4(2E^2 - M^2c^4)^2}{(E^2 - M^2c^4)^2} \frac{1}{\sin^4\theta} \right] . \quad (12.181)$$

Here we have introduced the Sommerfeld fine-structure constant

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c}\,, \tag{12.182}$$

which quantifies the strength of the electromagnetic interaction between elementary charged particles. It is a dimensionless quantity related to the elementary charge e, which denotes the strength of the coupling of an elementary charged particle with the electromagnetic field. As a dimensionless quantity, its numerical value is approximately given by

$$\alpha \approx \frac{1}{137}\,. \tag{12.183}$$

The result (12.181) predicts the differential cross section for the elastic scattering of two unpolarized electrons on the basis of quantum electrodynamics. It was first calculated in the ultra-relativistic regime by Christian Møller in 1932 based on some guesses and consistency requirements, not using quantunm electrodynamics. The full quantum electrodynamical calculation was provided only a few years later by Bethe and Fermi. Note that the indistinguishability of the two electrons involved in the scattering is represented by the forward-backward symmetry, i.e. the differential cross section is invariant with respect to the substitution $\theta \to \pi - \theta$. Within a classic experiment at the Laboratory of Nuclear Studies (Cornell University, Ithaca, New York) the Møller scattering formula (12.181) was checked in detail [Phys. Rev. **94**, 357 (1954)]. To this end the absolute differential electron-electron scattering cross section was measured for the incident electron energy in the laboratory frame varying in the interval from 0.6 to 1.2 Mev, which has to be compared with the rest energy of the electron of 0.513 MeV. The technique of measurement combined good resolution with large energy transfers between the particles, so this experiment allowed a sensitive test of the Møller scattering formula (12.181) in the relativistic regime. The results verified the theoretical predictions within a 7% experimental error.

In the ultra-relativistic limit $E \gg Mc^2$ the differential cross section (12.181) reduces to:

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{ur}} = \frac{\alpha^2\hbar^2c^2}{4E^2}\left(1 - \frac{8}{\sin^2\theta} + \frac{16}{\sin^4\theta}\right)\,. \tag{12.184}$$

With the help of the trigonometric formulae

$$\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) = \frac{1}{2}\sin\theta\,,\;\; \sin^2\left(\frac{\theta}{2}\right) = \frac{1}{2}\left(1-\cos\theta\right),\;\; \cos^2\left(\frac{\theta}{2}\right) = \frac{1}{2}\left(1+\cos\theta\right) \tag{12.185}$$

follows the trigonometric side calculation

$$\frac{1+\cos^4\left(\frac{\theta}{2}\right)}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{2}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)} + \frac{1+\sin^4\left(\frac{\theta}{2}\right)}{\cos^4\left(\frac{\theta}{2}\right)} = 2\left(1 - \frac{8}{\sin^2\theta} + \frac{16}{\sin^4\theta}\right)\,. \tag{12.186}$$

Inserting (12.186) into (12.184) leads to

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{ur}} = \frac{\alpha^2\hbar^2c^2}{8E^2}\left[\frac{1+\cos^4\left(\frac{\theta}{2}\right)}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{2}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)} + \frac{1-\sin^4\left(\frac{\theta}{2}\right)}{\cos^4\left(\frac{\theta}{2}\right)}\right]\,. \tag{12.187}$$

In the opposite non-relativistic limit $E = Mc^2 + \epsilon$ we obtain with $\epsilon \ll Mc^2$ from (12.181)

$$\left.\frac{d\sigma}{d\Omega}\right|_{\rm nr} = \frac{\alpha^2\hbar^2c^2}{16\epsilon^2}\left(\frac{4}{\sin^4\theta} - \frac{3}{\sin^2\theta}\right) . \tag{12.188}$$

With the trigonometric formulae (12.185) follows the trigonometric side calculation

$$\frac{1}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{1}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)} + \frac{1}{\cos^4\left(\frac{\theta}{2}\right)} = \frac{16}{\sin^4\theta}\left\{\frac{1}{4}(1+\cos\theta)^2 + \frac{1}{4}(1-\cos\theta)^2 \right.$$
$$\left. -\frac{1}{4}(1-\cos\theta)(1+\cos\theta)\right\} = \frac{16}{\sin^4\theta}\left(1 - \frac{3}{4}\sin^2\theta\right) = 4\left(\frac{4}{\sin^4\theta} - \frac{3}{\sin^2\theta}\right) . \tag{12.189}$$

With this (12.188) goes over into

$$\left.\frac{d\sigma}{d\Omega}\right|_{\rm nr} = \frac{\alpha^2\hbar^2c^2}{64\epsilon^2}\left[\frac{1}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{1}{\cos^4\left(\frac{\theta}{2}\right)} - \frac{1}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)}\right] . \tag{12.190}$$

With the non-relativistic dispersion relation $\epsilon = \mathbf{p}^2/(2M)$ it follows finally

$$\left.\frac{d\sigma}{d\Omega}\right|_{\rm nr} = \frac{\alpha^2\hbar^2c^2M^2}{16\mathbf{p}^4}\left[\frac{1}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{1}{\cos^4\left(\frac{\theta}{2}\right)} - \frac{1}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)}\right] . \tag{12.191}$$

The first term in (12.191) just corresponds to the cross section of the Rutherford scattering

$$\left.\frac{d\sigma}{d\Omega}\right|_{\rm R} = \frac{\alpha^2\hbar^2c^2M^2Z^2}{4\mathbf{p}^4}\,\frac{1}{\sin^4\left(\frac{\theta}{2}\right)} . \tag{12.192}$$

with the nuclear charge number $Z = 1$ and the mass M being substituted by the reduced mass $M/2$. This means that the forward peak of the non-relativistic Møller scattering at $\theta \approx 0$ agrees with the prediction of Rutherford prediction. Beyond that, however, there is another significant backward peak at $\theta = \pi$ that stems from interferences. Note that the latter must occur due to above mentioned forward-backward symmetry following from the indistinguishability of the electrons.

While formerly many particle colliders were designed specifically for electron-electron collisions, recently electron-positron colliders have become more common. Here one uses the so-called crossing symmetry, one of the useful tricks often used in quantum field theory to evaluate Feynman diagrams. Namely, from the Feynman rules follows directly that the unpolarized scattering matrix for any process involving a particle with momentum p in the initial state can be converted into the unpolarized scattering matrix for an otherwise identical process but with an antiparticle of momentum $-p$ in the final state. This implies that the Møller scattering between two electrons (12.1) goes over into the corresponding unpolarized cross section of the Bhabha scattering, i.e. the electron-positron scattering:

$$e^-\,e^+ \to e^-\,e^+ . \tag{12.193}$$

Applying this crossing symmetry to the unpolarized Møller cross section turns out to have the consequence that the unpolarized Bhabha cross section follows by interchaging the Mandelstam parameter s and u in (12.108), (12.109), and (12.112):

$$s \qquad \Longleftrightarrow \qquad u\,. \tag{12.194}$$

We refrain here from discussing the respective energy and angle dependence of the Bhabha differential cross section. Instead we refer to the above mentioned classic experiment at the Laboratory of Nuclear Studies, where the absolute differential positron-electron scattering cross section was checked in the energy interval from 0.6 to 1.0 Mev, which verified the Bhabha formula within the 10% experimental error. Furthermore, the ratio of the Møller and the Bhabha cross sections was also measured with somewhat increased accuracy, yielding a verification within about 8% experimental error.

In the last three decades Bhabha scattering has been used as a luminosity monitor in a number of e^-e^+ collider physics experiments. The accurate measurement of luminosity is necessary for accurate measurements of cross sections. Small-angle Bhabha scattering was used to measure the luminosity of the 1993 run of the Stanford Large Detector (SLD), with a relative uncertainty of less than 0.5%. Electron-positron colliders operating in the region of the low-lying hadronic resonances (about 1 GeV to 10 GeV), such as the Beijing Electron Synchrotron (BES) and the Belle and BaBar "B-factory" experiments, use large-angle Bhabha scattering as a luminosity monitor. To achieve the desired precision at the 0.1% level, the experimental measurements must be compared to a theoretical calculation including next-to-leading-order radiative corrections. The high-precision measurement of the total hadronic cross section at these low energies is, for instance, a crucial input into the theoretical calculation of the anomalous magnetic dipole moment of the muon, which is used to constrain supersymmetry and other models of physics beyond the Standard Model.