Chapter 10

Relativistic Quantum Mechanics

In this Chapter we will address the issue that the laws of physics must be formulated in a form which is Lorentz–invariant, i.e., the description should not allow one to differentiate between frames of reference which are moving relative to each other with a constant uniform velocity $\vec{v}$. The transformations beween such frames according to the Theory of Special Relativity are described by Lorentz transformations. In case that $\vec{v}$ is oriented along the x_1–axis, i.e., $\vec{v} = v_1 \hat{x}_1$, these transformations are

$$x_1\prime = \frac{x_1 - v_1 t}{\sqrt{1 - \left(\frac{v_1}{c}\right)^2}} \, , \; t' = \frac{t - \frac{v_1}{c^2} x_1}{\sqrt{1 - \left(\frac{v_1}{c}\right)^2}} \, , \; x_2' = x_2 \, ; \; x_3' = x_3 \tag{10.1}$$

which connect space time coordinates (x_1, x_2, x_3, t) in one frame with space time coordinates (x_1', x_2', x_3', t') in another frame. Here c denotes the velocity of light. We will introduce below Lorentz-invariant differential equations which take the place of the Schrödinger equation of a particle of mass m and charge q in an electromagnetic field [c.f. (refeq:ham2, 8.45)] described by an electrical potential $V(\vec{r}, t)$ and a vector potential $\vec{A}(\vec{r}, t)$

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \left[\frac{1}{2m} \left(\frac{\hbar}{i} \nabla - \frac{q}{c} \vec{A}(\vec{r}, t) \right)^2 + qV(\vec{r}, t) \right] \psi(\vec{r}, t) \tag{10.2}$$

The replacement of (10.2) by Lorentz–invariant equations will have two surprising and extremely important consequences: some of the equations need to be formulated in a representation for which the wave functions $\psi(\vec{r}, t)$ are vectors of dimension larger one, the components representing the spin attribute of particles and also representing together with a particle its anti-particle. We will find that actually several Lorentz–invariant equations which replace (10.2) will result, any of these equations being specific for certain classes of particles, e.g., spin–0 particles, spin–$\frac{1}{2}$ particles, etc. As mentioned, some of the equations describe a particle together with its anti-particle. It is not possible to uncouple the equations to describe only a single type particle without affecting negatively the Lorentz invariance of the equations. Furthermore, the equations need to be interpreted as actually describing many–particle-systems: the equivalence of mass and energy in relativistic formulations of physics allows that energy converts into particles such that any particle described will have 'companions' which assume at least a virtual existence.

Obviously, it will be necessary to begin this Chapter with an investigation of the group of Lorentz transformations and their representation in the space of position $\vec{r}$ and time t. The representation in Sect. 10.1 will be extended in Sect. 10.4 to cover fields, i.e., wave functions $\psi(\vec{r}, t)$ and vectors with functions $\psi(\vec{r}, t)$ as components. This will provide us with a general set of Lorentz–invariant equations which for various particles take the place of the Schrödinger equation. Before introducing these general Lorentz–invariant field equations we will provide in Sects. 10.5, 10.7 a heuristic derivation of the two most widely used and best known Lorentz–invariant field equations, namely the Klein–Gordon (Sect. 10.5) and the Dirac (Sect. 10.7) equation.

10.1 Natural Representation of the Lorentz Group

In this Section we consider the natural representation of the Lorentz group $\mathcal{L}$, i.e. the group of Lorentz transformations (10.1). Rather than starting from (10.1), however, we will provide a more basic definition of the transformations. We will find that this definition will lead us back to the transformation law (10.1), but in a setting of representation theory methods as applied in Secti. 5 to the groups SO(3) and SU(2) of rotation transformations of space coordinates and of spin.
The elements $L \in \mathcal{L}$ act on 4–dimensional vectors of position– and time–coordinates. We will denote these vectors as follows

$$x^\mu \stackrel{\text{def}}{=} (x^0, x^1, x^2, x^3) \tag{10.3}$$

where $x^0 = ct$ describes the time coordinate and $(x^1, x^2, x^3) = \vec{r}$ describes the space coordinates. Note that the components of x^μ all have the same dimension, namely that of length. We will, henceforth, assume new units for time such that the velocity of light c becomes $c = 1$. This choice implies $dim(\text{time}) = dim(\text{length})$.

Minkowski Space

Historically, the Lorentz transformations were formulated in a space in which the time component of x^μ was chosen as a purely imaginary number and the space components real. This space is called the Minkowski space. The reason for this choice is that the transformations (10.1) leave the quantity

$$s^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{10.4}$$

invariant, i.e., for the transformed space-time–cordinates $x'^\mu = (x'^0, x'^1, x'^2, x'^3)$ holds

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x'^0)^2 - (x'^1)^2 - (x'^2)^2 - (x'^3)^2 \,. \tag{10.5}$$

One can interprete the quantity $\sqrt{-s^2}$ as a distance in a 4–dimensional Euclidean space if one chooses the time component purely imaginary. In such a space Lorentz transformations correspond to 4-dimensional rotations. Rather than following this avenue we will introduce Lorentz transformations within a setting which does not require real and imaginary coordinates.

The Group of Lorentz Transformations $\mathcal{L}$ = O(3,1)

The Lorentz transformations L describe the relationship between space-time coordinates x^μ of two reference frames which move relative to each other with uniform fixed velocity $\vec{v}$ and which might be reoriented relative to each other by a rotation around a common origin. Denoting by x^μ the coordinates in one reference frame and by x'^{μ} the coordinates in the other reference frame, the Lorentz transformations constitute a linear transformation which we denote by

$$x'^{\mu} \;=\; \sum_{\nu=0}^{3} L^{\mu}{}_{\nu} x_{\nu} \; . \tag{10.6}$$

Here $L^{\mu}{}_{\nu}$ are the elements of a 4×4–matrix representing the Lorentz transformation. The upper index closer to 'L' denotes the first index of the matrix and the lower index ν further away from 'L' denotes the second index. [A more conventional notation would be $L_{\mu\nu}$, however, the latter notation will be used for different quantities further below.] The following possibilities exist for the positioning of the indices $\mu, \nu \;=\; 0, 1, 2, 3$:

$$\text{4-vector: } x^{\mu},\; x_{\mu}\;; \quad 4 \times 4 \text{ tensor: } A^{\mu}{}_{\nu},\; A_{\mu}{}^{\nu},\; A^{\mu\nu},\; A_{\mu\nu}\; . \tag{10.7}$$

The reason for the notation is two-fold. First, the notation in (10.6) allows us to introduce the so-called *summation conventon*: any time the *same* index appears in an upper **and** a lower position, summation over that index is assumed without explicitly noting it, i.e.,

$$\underbrace{y_{\mu} x^{\mu}}_{\text{new}} \;=\; \underbrace{\sum_{\mu=0}^{3} y_{\mu} x^{\mu}}_{\text{old}}\;;\; \underbrace{A^{\mu}{}_{\nu} x^{\nu}}_{\text{new}} \;=\; \underbrace{\sum_{\nu=0}^{3} A^{\mu}{}_{\nu} x^{\nu}}_{\text{old}}\;;\; \underbrace{A^{\mu}{}_{\nu} B^{\nu}{}_{\rho}}_{\text{new}} \;=\; \underbrace{\sum_{\nu=0}^{3} A^{\mu}{}_{\nu} B^{\nu}{}_{\rho}}_{\text{old}}\; . \tag{10.8}$$

The summation convention allows us to write (10.6) $x'^{\mu} \;=\; L^{\mu}{}_{\nu} x^{\nu}$. The second reason is that upper and lower positions allow us to accomodate the expression (10.5) into scalar products. This will be explained further below.
The Lorentz transformations are non-singular 4×4–matrices with real coefficients, i.e., $L \in \mathrm{GL}(4, \mathbb{R})$, the latter set constituting a group. The Lorentz transformations form the subgroup of all matrices which leave the expression (10.5) invariant. This condition can be written

$$x^{\mu} g_{\mu\nu} x^{\nu} \;=\; x'^{\mu} g_{\mu\nu} x'^{\nu} \tag{10.9}$$

where

$$(\, g_{\mu\nu} \,) \;=\; \begin{pmatrix} 1 & 0 & 0 & \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \;=\; \mathbf{g}\; . \tag{10.10}$$

Combining condition (10.9) and (10.6) yields

$$L^{\mu}{}_{\rho} g_{\mu\nu} L^{\nu}{}_{\sigma}\, x^{\rho} x^{\sigma} \;=\; g_{\rho\sigma}\, x^{\rho} x^{\sigma}\; . \tag{10.11}$$

Since this holds for any x^{μ} it must be true

$$L^{\mu}{}_{\rho} g_{\mu\nu} L^{\nu}{}_{\sigma} \;=\; g_{\rho\sigma}\; . \tag{10.12}$$

This condition specifies the key property of Lorentz transformations. We will exploit this property below to determine the general form of the Lorentz transformations. The subset of $\mathrm{GL}(4, \mathbb{R})$, the

elements of which satisfy this condition, is called O(3,1). This set is identical with the set of all Lorentz transformations $\mathcal{L}$. We want to show now $\mathcal{L} = \mathrm{O}(3{,}1) \subset \mathrm{GL}(4, \mathbb{R})$ is a group.
To simplify the following proof of the key group properties we like to adopt the conventional matrix notation for $L^{\mu}{}_{\nu}$

$$\mathbf{L} \;=\; (\, L^{\mu}{}_{\nu} \,) \;=\; \begin{pmatrix} L^0{}_0 & L^0{}_1 & L^0{}_2 & L^0{}_3 \\ L^1{}_0 & L^1{}_1 & L^1{}_2 & L^1{}_3 \\ L^2{}_0 & L^2{}_1 & L^2{}_2 & L^3{}_3 \\ L^3{}_0 & L^3{}_1 & L^3{}_2 & L^3{}_3 \end{pmatrix} . \tag{10.13}$$

Using the definition (10.10) of $\mathbf{g}$ one can rewrite the invariance property (10.12)

$$\mathbf{L}^T \mathbf{g} \mathbf{L} \;=\; \mathbf{g} \,. \tag{10.14}$$

From this one can obtain using

$$\mathbf{g}^2 \;=\; \mathbb{1} \tag{10.15}$$

$(\mathbf{g}\mathbf{L}^T\mathbf{g})\mathbf{L} \;=\; \mathbb{1}$ and, hence, the inverse of $\mathbf{L}$

$$\mathbf{L}^{-1} \;=\; \mathbf{g}\,\mathbf{L}^T\,\mathbf{g} \;=\; \begin{pmatrix} L^0{}_0 & -L^1{}_0 & -L^2{}_0 & -L^3{}_0 \\ -L^0{}_1 & L^1{}_1 & L^2{}_1 & L^3{}_1 \\ -L^0{}_2 & L^1{}_2 & L^2{}_2 & L^3{}_2 \\ -L^0{}_3 & L^1{}_3 & L^2{}_3 & L^3{}_3 \end{pmatrix} . \tag{10.16}$$

The corresponding expression for $(\mathbf{L}^T)^{-1}$ is obviously

$$(\mathbf{L}^T)^{-1} \;=\; (\mathbf{L}^{-1})^T \;=\; \mathbf{g}\,\mathbf{L}\,\mathbf{g} \,. \tag{10.17}$$

To demonstrate the group property of O(3,1), i.e., of

$$\mathrm{O}(3,1) \;=\; \{\, \mathbf{L}, \mathbf{L} \,\in\, \mathrm{GL}(4,\mathbb{R}),\, \mathbf{L}^T\mathbf{g}\mathbf{L} \;=\; \mathbf{g} \,\} \,, \tag{10.18}$$

we note first that the identity matrix $\mathbb{1}$ is an element of O(3,1) since it satisfies (10.14). We consider then $\mathbf{L}_1, \mathbf{L}_2 \in \mathrm{O}(3{,}1)$. For $\mathbf{L}_3 \;=\; \mathbf{L}_1\mathbf{L}_2$ holds

$$\mathbf{L}_3^T\,\mathbf{g}\,\mathbf{L}_3 \;=\; \mathbf{L}_2^T\mathbf{L}_1^T\,\mathbf{g}\,\mathbf{L}_1\mathbf{L}_2 \;=\; \mathbf{L}_2^T\,(\mathbf{L}_1^T\,\mathbf{g}\mathbf{L}_1)\,\mathbf{L}_2 \;=\; \mathbf{L}_2^T\,\mathbf{g}\,\mathbf{L}_2 \;=\; \mathbf{g} \,, \tag{10.19}$$

i.e., $\mathbf{L}_3 \in \mathrm{O}(3{,}1)$. One can also show that if $\mathbf{L} \in \mathrm{O}(3{,}1)$, the associated inverse obeys (10.14), i.e., $\mathbf{L}^{-1} \in \mathrm{O}(3{,}1)$. In fact, employing expressions (10.16, 10.17) one obtains

$$(\mathbf{L}^{-1})^T\,\mathbf{g}\,\mathbf{L}^{-1} \;=\; \mathbf{g}\mathbf{L}\mathbf{g}\mathbf{g}\mathbf{g}\mathbf{L}^T\mathbf{g} \;=\; \mathbf{g}\mathbf{L}\mathbf{g}\mathbf{L}^T\mathbf{g} \,. \tag{10.20}$$

Multiplying (10.14) from the right by $\mathbf{g}\mathbf{L}^T$ and using (10.15) one can derive $\mathbf{L}^T\mathbf{g}\mathbf{L}\mathbf{g}\mathbf{L}^T \;=\; \mathbf{L}^T$ and multiplying this from the left by by $\mathbf{g}(\mathbf{L}^T)^{-1}$ yields

$$\mathbf{L}\,\mathbf{g}\,\mathbf{L}^T \;=\; \mathbf{g} \tag{10.21}$$

Using this result to simplify the r.h.s. of (10.20) results in the desired property

$$(\mathbf{L}^{-1})^T\,\mathbf{g}\,\mathbf{L}^{-1} \;=\; \mathbf{g} \,, \tag{10.22}$$

i.e., property (10.14) holds for the inverse of $\mathbf{L}$. This stipulates that O(3,1) is, in fact, a group.

Classification of Lorentz Transformations

We like to classify now the elements of $\mathcal{L} = \mathrm{O}(3,1)$. For this purpose we consider first the value of $\det L$. A statement on this value can be made on account of property (10.14). Using $\det AB = \det A \det B$ and $\det A^T = \det A$ yields $(\det L)^2 = 1$ or

$$\det L = \pm 1 . \tag{10.23}$$

One can classify Lorentz transformations according to the value of the determinant into two distinct classes.
A second class property follows from (10.14) which we employ in the formulation (10.12). Considering in (10.12) the case $\rho = 0$, $\sigma = 0$ yields

$$\left(L^0{}_0\right)^2 - \left(L^1{}_0\right)^2 - \left(L^2{}_0\right)^2 - \left(L^3{}_0\right)^2 = 1 . \tag{10.24}$$

or since $(L^1{}_0)^2 + (L^2{}_0)^2 + (L^3{}_0)^2 \geq 0$ it holds $(L^0{}_0)^2 \geq 1$. From this we can conclude

$$L^0{}_0 \geq 1 \quad \text{or} \qquad L^0{}_0 \leq -1 , \tag{10.25}$$

i.e., there exist two other distinct classes. Properties (10.23) and (10.25) can be stated as follows: The set of all Lorentz transformations $\mathcal{L}$ is given as the union

$$\mathcal{L} = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow \cup \mathcal{L}_-^\uparrow \cup \mathcal{L}_-^\downarrow \tag{10.26}$$

where $\mathcal{L}_+^\uparrow$, $\mathcal{L}_+^\downarrow$, $\mathcal{L}_-^\uparrow$, $\mathcal{L}_-^\downarrow$ are disjunct sets defined as follows

$$\mathcal{L}_+^\uparrow = \{ L, L \in \mathrm{O}(3,1), \det L = 1, L^0{}_0 \geq 1\} ; \tag{10.27}$$
$$\mathcal{L}_+^\downarrow = \{ L, L \in \mathrm{O}(3,1), \det L = 1, L^0{}_0 \leq -1\} ; \tag{10.28}$$
$$\mathcal{L}_-^\uparrow = \{ L, L \in \mathrm{O}(3,1), \det L = -1, L^0{}_0 \geq 1\} ; \tag{10.29}$$
$$\mathcal{L}_-^\downarrow = \{ L, L \in \mathrm{O}(3,1), \det L = -1, L^0{}_0 \leq -1\} . \tag{10.30}$$

It holds $\mathbf{g} \in \mathcal{L}$ and $-1\!\!1 \in \mathcal{L}$ as one can readily verify testing for property (10.14). One can also verify that one can write

$$\mathcal{L}_-^\uparrow = \mathbf{g}\mathcal{L}_+^\uparrow = \mathcal{L}_+^\uparrow\mathbf{g} ; \tag{10.31}$$
$$\mathcal{L}_+^\downarrow = -\mathcal{L}_+^\uparrow ; \tag{10.32}$$
$$\mathcal{L}_-^\downarrow = -\mathbf{g}\mathcal{L}_+^\uparrow = -\mathcal{L}_+^\uparrow\mathbf{g} \tag{10.33}$$

where we used the definition $a\mathcal{M} = \{M_1, \exists M_2, M_2 \in \mathcal{M}, M_1 = a\, M_2\}$. The above shows that the set of proper Lorentz transformations $\mathcal{L}_+^\uparrow$ allows one to generate all Lorentz transformations, except for the trivial factors $\mathbf{g}$ and $-1\!\!1$. It is, hence, entirely suitable to investigate first only Lorentz transformations in $\mathcal{L}_+^\uparrow$.
We start our investigation by demonstrating that $\mathcal{L}_+^\uparrow$ forms a group. Obviously, $\mathcal{L}_+^\uparrow$ contains $1\!\!1$. We can also demonstrate that for $A, B \in \mathcal{L}_+^\uparrow$ holds $C = AB \in \mathcal{L}_+^\uparrow$. For this purpose we consider the value of $C^0{}_0 = A^0{}_\mu B^\mu{}_0 = \sum_{j=1}^3 A^0{}_j B^j{}_0 + A^0{}_0 B^0{}_0$. Schwartz's inequality yields

$$\left(\sum_{j=1}^{3} A^0{}_j B^j{}_0 \right)^2 \leq \sum_{j=1}^{3} \left(A^0{}_j\right)^2 \sum_{j=1}^{3} \left(B^j{}_0\right)^2 . \tag{10.34}$$

From (10.12) follows $(B^0{}_0)^2 - \sum_{j=1}^3 (B^j{}_0)^2 = 1$ or $\sum_{j=1}^3 (B^j{}_0)^2 = (B^0{}_0)^2 - 1$. Similarly, one can conclude from (10.21) $\sum_{j=1}^3 (A^0{}_j)^2 = (A^0{}_0)^2 - 1$. (10.34) provides then the estimate

$$\left(\sum_{j=1}^{3} A^0{}_j B^j{}_0 \right)^2 \leq \left[(A^0{}_0)^2 - 1 \right] \left[(B^0{}_0)^2 - 1 \right] < (A^0{}_0)^2 \, (B^0{}_0)^2 \, . \tag{10.35}$$

One can conclude, therefore, $| \sum_{j=1}^3 A^0{}_j B^j{}_0 | < A^0{}_0 B^0{}_0$. Since $A^0{}_0 \geq 1$ and $B^0{}_0 \geq 1$, obviously $A^0{}_0 B^0{}_0 \geq 1$. Using the above expression for $C^0{}_0$ one can state $C^0{}_0 > 0$. In fact, since the group property of O(3,1) ascertains $\mathbf{C}^T \mathbf{g} \mathbf{C} = \mathbf{g}$ it must hold $C^0{}_0 \geq 1$.
The next group property of $\mathcal{L}_+^\uparrow$ to be demonstrated is the existence of the inverse. For the inverse of any $\mathbf{L} \in \mathcal{L}_+^\uparrow$ holds (10.16). This relationship shows $(\mathbf{L}^{-1})^0{}_0 = L^0{}_0$, from which one can conclude $\mathbf{L}^{-1} \in \mathcal{L}_+^\uparrow$. We also note that the identity operator $1\!\!1$ has elements

$$1\!\!1^\mu{}_\nu = \delta^\mu{}_\nu \tag{10.36}$$

where we defined[1]

$$\delta^\mu{}_\nu = \begin{cases} 1 & \text{for } \mu = \nu \\ 0 & \text{for } \mu \neq \nu \end{cases} \tag{10.37}$$

It holds, $1\!\!1^0{}_0 = \geq 1$ and, hence, $1\!\!1 \in \mathcal{L}_+^\uparrow$. Since the associative property holds for matrix multiplication we have verified that $\mathcal{L}_+^\uparrow$ is indeed a subgroup of SO(3,1).
$\mathcal{L}_+^\uparrow$ is called the subgroup of *proper, orthochronous Lorentz transformations*. In the following we will consider solely this subgroup of SO(3,1).

Infinitesimal Lorentz transformations

The transformations in $\mathcal{L}_+^\uparrow$ have the property that they are continously connected to the identity $1\!\!1$, i.e., these transformations can be parametrized such that a continuous variation of the parameters connects any element of $\mathcal{L}_+^\uparrow$ with $1\!\!1$. This property will be exploited now in that we consider first transformations in a small neighborhood of $1\!\!1$ which we parametrize by infinitesimal parameters. We will then employ the Lie group properties to generate all transformations in $\mathcal{L}_+^\uparrow$.
Accordingly, we consider transformations

$$L^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon^\mu{}_\nu \, ; \quad \epsilon^\mu{}_\nu \quad \text{small} \, . \tag{10.38}$$

For these transformations, obviously, holds $L^0{}_0 > 0$ and the value of the determinant is close to unity, i.e., if we enforce (10.14) actually $L^0{}_0 \geq 1$ and $\det \mathbf{L} = 1$ must hold. Property (10.14) implies

$$\left(1\!\!1 + \epsilon^T \right) \mathbf{g} \left(1\!\!1 + \epsilon \right) = \mathbf{g} \tag{10.39}$$

where we have employed the matrix form ϵ defined as in (10.13). To order $O(\epsilon^2)$ holds

$$\epsilon^T \mathbf{g} + \mathbf{g} \epsilon = 0 \, . \tag{10.40}$$

[1] It should be noted that according to our present definition holds $\delta_{\mu\nu} = g_{\mu\rho} \delta^\rho{}_\nu$ and, accordingly, $\delta_{00} = 1$ and $\delta_{11} = \delta_{22} = \delta_{33} = -1$.

Using (10.15) one can conclude

$$\epsilon^T = -\mathbf{g}\,\epsilon\,\mathbf{g} \tag{10.41}$$

which reads explicitly

$$\begin{pmatrix} \epsilon^0{}_0 & \epsilon^1{}_0 & \epsilon^2{}_0 & \epsilon^3{}_0 \\ \epsilon^0{}_1 & \epsilon^1{}_1 & \epsilon^2{}_1 & \epsilon^3{}_1 \\ \epsilon^0{}_2 & \epsilon^1{}_2 & \epsilon^2{}_2 & \epsilon^3{}_2 \\ \epsilon^0{}_3 & \epsilon^1{}_3 & \epsilon^2{}_3 & \epsilon^3{}_3 \end{pmatrix} = \begin{pmatrix} -\epsilon^0{}_0 & \epsilon^0{}_1 & \epsilon^0{}_2 & \epsilon^0{}_3 \\ \epsilon^1{}_0 & -\epsilon^1{}_1 & -\epsilon^1{}_2 & -\epsilon^1{}_3 \\ \epsilon^2{}_0 & -\epsilon^2{}_1 & -\epsilon^2{}_2 & -\epsilon^2{}_3 \\ \epsilon^3{}_0 & -\epsilon^3{}_1 & -\epsilon^3{}_2 & -\epsilon^3{}_3 \end{pmatrix} . \tag{10.42}$$

This relationship implies

$$\begin{aligned} \epsilon^\mu{}_\mu &= 0 \\ \epsilon^0{}_j &= \epsilon^j{}_0 \,, \qquad j = 1,2,3 \\ \epsilon^j{}_k &= -\,\epsilon^k{}_j \,, \quad j,k = 1,2,3 \end{aligned} \tag{10.43}$$

Inspection shows that the matrix ϵ has 6 independent elements and can be written

$$\epsilon(\vartheta_1,\vartheta_2,\vartheta_3,w_1,w_2,w_3) = \begin{pmatrix} 0 & -w_1 & -w_2 & -w_3 \\ -w_1 & 0 & -\vartheta_3 & \vartheta_2 \\ -w_2 & \vartheta_3 & 0 & -\vartheta_1 \\ -w_3 & -\vartheta_2 & \vartheta_1 & 0 \end{pmatrix} . \tag{10.44}$$

This result allows us now to define six generators for the Lorentz transformations($k = 1,2,3$)

$$\mathbf{J}_k = \epsilon(\vartheta_k = 1, \text{other five parameters zero}) \tag{10.45}$$

$$\mathbf{K}_k = \epsilon(w_k = 1, \text{other five parameters zero}) \,. \tag{10.46}$$

The generators are explicitly

$$\mathbf{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} ;\ \mathbf{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} ;\ \mathbf{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{10.47}$$

$$\mathbf{K}_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ;\ \mathbf{K}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ;\ \mathbf{K}_3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{10.48}$$

These commutators obey the following commutation relationships

$$\begin{aligned} [\,\mathbf{J}_k,\,\mathbf{J}_\ell\,] &= \epsilon_{k\ell m}\,\mathbf{J}_m \\ [\,\mathbf{K}_k,\,\mathbf{K}_\ell\,] &= -\,\epsilon_{k\ell m}\,\mathbf{J}_m \\ [\,\mathbf{J}_k,\,\mathbf{K}_\ell\,] &= \epsilon_{k\ell m}\,\mathbf{K}_m \,. \end{aligned} \tag{10.49}$$

The operators also obey

$$\vec{\mathbf{J}}\cdot\vec{\mathbf{K}} = \mathbf{J}_1\mathbf{J}_1 + \mathbf{J}_2\mathbf{J}_2 + \mathbf{J}_3\mathbf{J}_3 = 0 \tag{10.50}$$

as can be readily verified.

Exercise 7.1:
Demonstrate the commutation relationships (10.49, 10.50).

The commutation relationships (10.49) define the Lie algebra associated with the Lie group $\mathcal{L}_+^\uparrow$. The commutation relationships imply that the algebra of the generators $\mathbf{J}_k$, $\mathbf{K}_k$, $k = 1, 2, 3$ is closed. Following the treatment of the rotation group SO(3) one can express the elements of $\mathcal{L}_+^\uparrow$ through the exponential operators

$$\mathbf{L}(\vec{\vartheta}, \vec{w}) = \exp\left(\vec{\vartheta} \cdot \vec{\mathbf{J}} + \vec{w} \cdot \vec{\mathbf{K}}\right) \quad ; \quad \vec{\vartheta}, \vec{w} \in \mathbb{R}^3 \tag{10.51}$$

where we have defined $\vec{\vartheta} \cdot \vec{\mathbf{J}} = \sum_{k=1}^3 \vartheta_k \mathbf{J}_k$ and $\vec{w} \cdot \vec{\mathbf{K}} = \sum_{k=1}^3 w_k \mathbf{K}_k$. One can readily show, following the algebra in Chapter 5, and using the relationship

$$\mathbf{J}_k = \begin{pmatrix} 0 & 0 \\ 0 & L_k \end{pmatrix} \tag{10.52}$$

where the 3×3–matrices L_k are the generators of SO(3) defined in Chapter 5, that the transformations (10.51) for $\vec{w} = 0$ correspond to rotations of the spatial coordinates, i.e.,

$$\mathbf{L}(\vec{\vartheta}, \vec{w} = 0) = \begin{pmatrix} 0 & 0 \\ 0 & R(\vec{\vartheta}) \end{pmatrix} . \tag{10.53}$$

Here $R(\vec{\vartheta})$ are the 3×3–rotation matrices constructed in Chapter 5. For the parameters ϑ_k of the Lorentz transformations holds obviously

$$\vartheta_k \in [0, 2\pi[\quad , \; k = 1, 2, 3 \tag{10.54}$$

which, however, constitutes an overcomplete parametrization of the rotations (see Chapter 5).
We consider now the Lorentz transformations for $\vec{\vartheta} = 0$ which are referred to as 'boosts'. A boost in the x_1–direction is $L = \exp(w_1 \mathbf{K}_1)$. To determine the explicit form of this transformation we evaluate the exponential operator by Taylor expansion. In analogy to equation (5.35) it issufficient to consider in the present case the 2×2–matrix

$$L' = \exp\left(w_1 \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\right) = \sum_{n=0}^{\infty} \frac{w_1^n}{n!} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}^n \tag{10.55}$$

since

$$\exp\left(w_1 \mathbf{K}_1\right) = \exp\begin{pmatrix} \multicolumn{2}{c}{\multirow{2}{*}{L'}} & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{10.56}$$

Using the idempotence property

$$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1\!\!1 \tag{10.57}$$

one can carry out the Taylor expansion above:

$$L' = \sum_{n=0}^{\infty} \frac{w_1^{2n}}{(2n)!} \, 1\!\!1 \; + \; \sum_{n=0}^{\infty} \frac{w_1^{2n+1}}{(2n+1)!} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \tag{10.58}$$

$$= \cosh w_1 \, 1\!\!1 \; + \; \sinh w_1 \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \cosh w_1 & -\sinh w_1 \\ -\sinh w_1 & \cosh w_1 \end{pmatrix} .$$

The conventional form (10.1) of the Lorentz transformations is obtained through the parameter change

$$v_1 = \frac{\sinh w_1}{\cosh w_1} = \tanh w_1 \tag{10.59}$$

Using $\cosh^2 w_1 - \sinh^2 w_1 = 1$ one can identify $\sinh w_1 = \sqrt{\cosh^2 w_1 - 1}$ and $\cosh w_1 = \sqrt{\sinh^2 w_1 + 1}$. Correspondingly, one obtains from (10.59)

$$v_1 = \frac{\sqrt{\cosh^2 w_1 - 1}}{\cosh w_1} = \frac{\sinh w_1}{\sqrt{\sinh^2 w_1 + 1}} . \tag{10.60}$$

These two equations yield

$$\cosh w_1 = 1/\sqrt{1 - v_1^2} \; ; \quad \sinh w_1 = v_1/\sqrt{1 - v_1^2} \; , \tag{10.61}$$

and (10.56, 10.59) can be written

$$\exp(w_1 \mathbf{K}_1) = \begin{pmatrix} \frac{1}{\sqrt{1-v_1^2}} & \frac{-v_1}{\sqrt{1-v_1^2}} & 0 & 0 \\ \frac{-v_1}{\sqrt{1-v_1^2}} & \frac{1}{\sqrt{1-v_1^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{10.62}$$

According to (10.3, 10.6, 10.51) the explicit transformation for space–time–coordinates is then

$$x_1' = \frac{x_1 - v_1 t}{\sqrt{1 - v_1^2}} \; , \; t' = \frac{t - v_1 x_1}{\sqrt{1 - v_1^2}} \; , \; x_2' = x_2 \; , \; x_3' = x_3 \tag{10.63}$$

which agrees with (10.1).

The range of the parameters w_k can now be specified. v_k defined in (10.59) for the case $k = 1$ corresponds to the relative velocity of two frames of reference. We expect that v_k can only assume values less than the velocity of light c which in the present units is $c = 1$. Accordingly, we can state $v_k \in]-1, 1[$. This property is, in fact, consistent with (10.59). From (10.59) follows, however, for w_k

$$w_k \in]-\infty, \infty[\; . \tag{10.64}$$

We note that the range of w_k-values is not a compact set even though the range of v_k-values is compact. This property of the w_k-values contrasts with the property of the parameters ϑ_k specifying rotational angles which assume only values in a compact range.

10.2 Scalars, 4–Vectors and Tensors

In this Section we define quantities according to their behaviour under Lorentz transformations. Such quantities appear in the description of physical systems and statements about transformation properties are often extremely helpful and usually provide important physical insight. We have encountered examples in connection with rotational transformations, namely, scalars like $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$, vectors like $\vec{r} = (x_1, x_2, x_3)^T$, spherical harmonics $Y_{\ell m}(\hat{r})$, total angular momentum states of composite systems like $\mathcal{Y}_{\ell m}(\ell_1, \ell_2|\hat{r}_1, \hat{r}_2)$ and, finally, tensor operators T_{km}. Some of these quantities were actually defined with respect to representations of the rotation group in function spaces, not in the so-called natural representation associated with the 3–dimensional Euclidean space $\mathbb{E}^3$.

Presently, we have not yet defined representations of Lorentz transformations beyond the 'natural' representation acting in the 4–dimensional space of position– and time–coordinates. Hence, our definition of quantities with special properties under Lorentz transformations presently is confined to the natural representation. Nevertheless, we will encounter an impressive example of physical properties.

Scalars The quantities with the simplest transformation behaviour are so-called scalars $f \in \mathbb{R}$ which are invariant under transformations, i.e.,

$$f' = f . \tag{10.65}$$

An example is s^2 defined in (10.4), another example is the rest mass m of a particle. However, not any physical property $f \in \mathbb{R}$ is a scalar. Counterexamples are the energy, the charge density, the z–component x_3 of a particle, the square of the electric field $|\vec{E}(\vec{r}, t)|^2$ or the scalar product $\vec{r}_1 \cdot \vec{r}_2$ of two particle positions. We will see below how true scalars under Lorentz transformations can be constructed.

4-Vectors The quantities with the transformation behaviour like that of the position–time vector x^μ defined in (10.3) are the so-called 4–vectors a^μ. These quantites always come as four components $(a^0, a^1, a^2, a^3)^T$ and transform according to

$$a'^\mu = L^\mu{}_\nu a^\nu . \tag{10.66}$$

Examples of 4-vectors beside x^μ are the momentum 4-vector

$$p^\mu = (E, \vec{p}) , \; E = \frac{m}{\sqrt{1 - \vec{v}^2}} , \; \vec{p} = \frac{m\vec{v}}{\sqrt{1 - \vec{v}^2}} \tag{10.67}$$

the transformation behaviour of which we will demonstrate further below. A third 4-vector is the so-called current vector

$$J^\mu = (\rho, \vec{J}) \tag{10.68}$$

where $\rho(\vec{r}, t)$ and $\vec{J}(\vec{r}, t)$ are the charge density and the current density, respectively, of a system of charges. Another example is the potential 4-vector

$$A^\mu = (V, \vec{A}) \tag{10.69}$$

where $V(\vec{r}, t)$ and $\vec{A}(\vec{r}, t)$ are the electrical and the vector potential of an electromagnetic field. The 4-vector character of J^μ and of A^μ will be demonstrated further below.

Scalar Product 4-vectors allow one to construct scalar quantities. If a^μ and b^μ are 4-vectors then

$$a^\mu g_{\mu\nu} b^\nu \tag{10.70}$$

is a scalar. This property follows from (10.66) together with (10.12)

$$a'^\mu g_{\mu\nu} b'^\nu \;=\; L^\mu{}_\rho g_{\mu\nu} L^\nu{}_\sigma a^\rho b^\sigma \;=\; a^\rho g_{\rho\sigma} b^\sigma \tag{10.71}$$

Contravariant and Covariant 4-Vectors It is convenient to define a second class of 4-vectors. The respective vectors a_μ are associated with the 4-vectors a^μ, the relationship being

$$a_\mu \;=\; g_{\mu\nu} a^\nu \;=\; (a^0, -a^1, -a^2, -a^3) \tag{10.72}$$

where a^ν is a vector with transformation behaviour as stated in (10.66). One calls 4-vectors a_μ *covariant* and 4-vectors a^μ *contravariant*. Covariant 4-vectors transform like

$$a'_\mu \;=\; g_{\mu\nu} L^\nu{}_\rho g^{\rho\sigma} a_\sigma \tag{10.73}$$

where we defined

$$g^{\mu\nu} \;=\; g_{\mu\nu} \,. \tag{10.74}$$

We like to point out that from definition (10.72) of the covariant 4-vector follows $a^\mu \;=\; g^{\mu\nu} a_\nu$. In fact, one can employ the tensors $g^{\mu\nu}$ and $g_{\mu\nu}$ to raise and lower indices of $L^\mu{}_\nu$ as well. We do not establish here the consistency of the ensuing notation. In any case one can express (10.73)

$$a'_\mu \;=\; L_\mu{}^\sigma a_\sigma \,. \tag{10.75}$$

Note that according to (10.17) $L_\mu{}^\sigma$ is the transformation inverse to $L^\sigma{}_\mu$. In fact, one can express $[(\mathbf{L}^{-1})^T]^\mu{}_\nu \;=\; (L^{-1})^\nu{}_\mu$ and, accordingly, (10.17) can be written

$$(L^{-1})^\nu{}_\mu \;=\; L_\mu{}^\nu \,. \tag{10.76}$$

The 4-Vector ∂_μ An important example of a covariant 4-vector is the differential operator

$$\partial_\mu \;=\; \frac{\partial}{\partial x^\mu} \;=\; \left(\frac{\partial}{\partial t}, \nabla \right) \tag{10.77}$$

The transformed differential operator will be denoted by

$$\partial'_\mu \stackrel{\text{def}}{=} \frac{\partial}{\partial x'^\mu} \,. \tag{10.78}$$

To prove the 4-vector property of ∂_μ we will show that $g^{\mu\nu}\partial_\nu$ transforms like a contravariant 4-vector, i.e., $g^{\mu\nu}\partial'_\nu \;=\; L^\mu{}_\rho g^{\rho\sigma}\partial_\sigma$. We start from $x'^\mu \;=\; L^\mu{}_\nu x^\nu$. Multiplication (and summation) of $x'^\mu \;=\; L^\mu{}_\nu x^\nu$ by $L^\rho{}_\sigma g_{\rho\mu}$ yields, using (10.12), $g_{\sigma\nu} x^\nu \;=\; L^\rho{}_\sigma g_{\rho\mu} x'^\mu$ and $g^{\mu\sigma} g_{\sigma\nu} \;=\; \delta^\mu{}_\nu$,

$$x^\nu \;=\; g^{\nu\sigma} L^\rho{}_\sigma g_{\rho\mu} x'^\mu \,. \tag{10.79}$$

This is the inverse Lorentz transformation consistent with (10.16). We have duplicated the expression for the inverse of $L^\mu{}_\nu$ to obtain the correct notation in terms of covariant, i.e., lower, and contravariant, i.e., upper, indices. (10.79) allows us to determine the connection between ∂_μ and ∂'_μ. Using the chain rule of differential calculus we obtain

$$\partial'_\mu \;=\; \sum_{\nu=0}^{3} \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} \;=\; g^{\nu\sigma} L^\rho{}_\sigma g_{\rho\mu} \partial_\nu \;=\; L_\mu{}^\nu \, \partial_\nu \tag{10.80}$$

Multiplication by $g^{\lambda\mu}$ (and summation over μ) together with $g^{\lambda\mu} g_{\rho\mu} \;=\; \delta^\lambda{}_\rho$ yields

$$g^{\lambda\mu} \partial'_\mu \;=\; L^\lambda{}_\sigma g^{\sigma\nu} \partial_\nu \;, \tag{10.81}$$

i.e., ∂_μ does indeed transform like a covariant vector.

d'Alembert Operator We want to construct now a scalar differential operator. For this purpose we define first the contravariant differential operator

$$\partial^\mu \;=\; g^{\mu\nu} \partial_\nu \;=\; \left(\frac{\partial}{\partial t}, -\nabla \right) \;. \tag{10.82}$$

Then the operator

$$\partial_\mu \partial^\mu \;=\; \partial_t^2 \;-\; \nabla^2 \tag{10.83}$$

is a scalar under Lorentz transformations. In fact, this operator is equal to the d'Alembert operator which is known to be Lorentz-invariant.

Proof that p^μ is a 4-vector We will demonstrate now that the momentum 4-vector p^μ defined in (10.67) transforms like (10.66). For this purpose we consider the scalar differential

$$(d\tau)^2 \;=\; dx^\mu dx_\mu \;=\; (dt)^2 \;-\; (d\vec{r})^2 \tag{10.84}$$

It holds

$$\left(\frac{d\tau}{dt} \right)^2 \;=\; 1 \;-\; (\vec{v})^2 \tag{10.85}$$

from which follows

$$\frac{d}{d\tau} \;=\; \frac{1}{\sqrt{1 \;-\; \vec{v}^{\,2}}} \frac{d}{dt} \;. \tag{10.86}$$

One can write

$$p^0 \;=\; E \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^{\,2}}} \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^{\,2}}} \frac{dt}{dt} \;. \tag{10.87}$$

The remaining components of p^μ can be written, e.g.,

$$p^1 \;=\; \frac{m \, v^1}{\sqrt{1 \;-\; \vec{v}^{\,2}}} \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^{\,2}}} \frac{dx^1}{dt} \;. \tag{10.88}$$

One can express then the momentum vector

$$p^\mu \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^{\,2}}} \frac{dx^\mu}{dt} \;=\; m \frac{d}{d\tau} x^\mu \;. \tag{10.89}$$

The operator $m\frac{d}{d\tau}$ transforms like a scalar. Since x^μ transforms like a contravariant 4-vector, the r.h.s. of (10.89) alltogether transforms like a contravariant 4-vector, and, hence, p^μ on the l.h.s. of (10.89) must be a 4-vector.
The momentum 4-vector allows us to construct a scalar quantity, namely

$$p^\mu p_\mu \;=\; p^\mu g_{\mu\nu} p^\nu \;=\; E^2 \;-\; \vec{p}^{\,2} \tag{10.90}$$

Evaluation of the r.h.s. yields according to (10.67)

$$E^2 \;-\; \vec{p}^{\,2} \;=\; \frac{m^2}{1 \;-\; \vec{v}^{\,2}} \;-\; \frac{m^2 \vec{v}^{\,2}}{1 \;-\; \vec{v}^{\,2}} \;=\; m^2 \tag{10.91}$$

or

$$p^\mu p_\mu \;=\; m^2 \tag{10.92}$$

which, in fact, is a scalar. We like to rewrite the last result

$$E^2 \;=\; \vec{p}^{\,2} \;+\; m^2 \tag{10.93}$$

or

$$E \;=\; \pm\sqrt{\vec{p}^{\,2} \;+\; m^2}\;. \tag{10.94}$$

In the non-relativistic limit the rest energy m is the dominant contribution to E. Expansion in $\frac{1}{m}$ should then be rapidly convergent. One obtains

$$E \;=\; \pm m \pm \frac{\vec{p}^{\,2}}{2m} \mp \frac{(\vec{p}^{\,2})^2}{4m^3} \;+\; O\left(\frac{(\vec{p}^{\,2})^3}{4m^5}\right)\;. \tag{10.95}$$

This obviously describes the energy of a free particle with rest energy $\pm m$, kinetic energy $\pm\frac{\vec{p}^{\,2}}{2m}$ and relativistic corrections.

10.3 Relativistic Electrodynamics

In the following we summarize the Lorentz-invariant formulation of electrodynamics and demonstrate its connection to the conventional formulation as provided in Sect. 8.

Lorentz Gauge In our previous description of the electrodynamic field we had introduced the scalar and vector potential $V(\vec{r},t)$ and $\vec{A}(\vec{r},t)$, respectively, and had chosen the so-called Coulomb gauge (8.12), i.e., $\nabla \cdot \vec{A} \;=\; 0$, for these potentials. This gauge is not Lorentz-invariant and we will adopt here another gauge, namely,

$$\partial_t V(\vec{r},t) \;+\; \nabla \cdot \vec{A}(\vec{r},t) \;=\; 0\;. \tag{10.96}$$

The Lorentz-invariance of this gauge, the so-called *Lorentz gauge*, can be demonstrated readily using the 4-vector notation (10.69) for the electrodynamic potential and the 4-vector derivative (10.77) which allow one to express (10.96) in the form

$$\partial_\mu A^\mu \;=\; 0\;. \tag{10.97}$$

We have proven already that ∂_μ is a contravariant 4-vector. If we can show that A^μ defined in (10.69) is, in fact, a contravariant 4-vector then the l.h.s. in (10.97) and, equivalently, in (10.96) is a scalar and, hence, Lorentz-invariant. We will demonstrate now the 4-vector property of A^μ.

Transformation Properties of J^μ and A^μ

The charge density $\rho(\vec{r},t)$ and current density $\vec{J}(\vec{r},t)$ are known to obey the continuity property

$$\partial_t \rho(\vec{r},t) \;+\; \nabla\cdot\vec{J}(\vec{r},t) \;=\; 0 \tag{10.98}$$

which reflects the principle of charge conservation. This principle should hold in any frame of reference. Equation (10.98) can be written, using (10.77) and (10.68),

$$\partial_\mu J^\mu(x^\mu) \;=\; 0\;. \tag{10.99}$$

Since this equation must be true in any frame of reference the right hand side must vanish in all frames, i.e., must be a scalar. Consequently, also the l.h.s. of (10.99) must be a scalar. Since ∂_μ transforms like a covariant 4-vector, it follows that J^μ, in fact, has to transform like a contravariant 4-vector.

We want to derive now the differential equations which determine the 4-potential A^μ in the Lorentz gauge (10.97) and, thereby, prove that A^μ is, in fact, a 4-vector. The respective equation for $A^0 = V$ can be obtained from Eq. (8.13). Using $\nabla\cdot\partial_t\vec{A}(\vec{r},t) = \partial_t\nabla\cdot\vec{A}(\vec{r},t)$ together with (10.96), i.e., $\nabla\cdot\vec{A}(\vec{r},t) = -\partial_t V(\vec{r},t)$, one obtains

$$\partial_t^2 V(\vec{r},t) \;-\; \nabla^2 V(\vec{r},t) \;=\; 4\pi\rho(\vec{r},t)\;. \tag{10.100}$$

Similarly, one obtains for $\vec{A}(\vec{r},t)$ from (8.17) using the identity (8.18) and, according to (10.96), $\nabla\cdot\vec{A}(\vec{r},t) = -\partial_t V(\vec{r},t)$

$$\partial_t^2 \vec{A}(\vec{r},t) \;-\; \nabla^2\vec{A}(\vec{r},t) \;=\; 4\,\pi\,\vec{J}(\vec{r},t)\;. \tag{10.101}$$

Combining equations (10.100, 10.101), using (10.83) and (10.69), yields

$$\partial_\mu\partial^\mu\, A^\nu(x^\sigma) \;=\; 4\,\pi\, J^\nu(x^\sigma)\;. \tag{10.102}$$

In this equation the r.h.s. transforms like a 4-vector. The l.h.s. must transform likewise. Since $\partial_\mu\partial^\mu$ transforms like a scalar one can conclude that $A^\nu(x^\sigma)$ must transform like a 4-vector.

The Field Tensor

The electric and magnetic fields can be collected into an anti-symmetric 4×4 tensor

$$F^{\mu\nu} \;=\; \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}\;. \tag{10.103}$$

Alternatively, this can be stated

$$F^{k0} \;=\; -F^{0k} \;=\; E^k\;, \quad F^{mn} \;=\; -\epsilon^{mn\ell}B^\ell\;, \quad k,\ell,m,n \;=\; 1,2,3 \tag{10.104}$$

where $\epsilon^{mn\ell} = \epsilon_{mn\ell}$ is the totally anti-symmetric three-dimensional tensor defined in (5.32).

One can readily verify, using (8.6) and (8.9), that $F^{\mu\nu}$ can be expressed through the potential A^μ in (10.69) and ∂^μ in (10.82) as follows

$$F^{\mu\nu} \;=\; \partial^\mu A^\nu \;-\; \partial^\nu A^\mu\;. \tag{10.105}$$

The relationships (10.103, 10.104) establishe the transformation behaviour of $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$. In a new frame of reference holds

$$F'^{\mu\nu} \;=\; L^{\mu}{}_{\alpha}\, L^{\nu}{}_{\beta}\, F^{\alpha\beta} \tag{10.106}$$

In case that the Lorentz transformation $L^{\mu}{}_{\nu}$ is given by (10.62) or, equivalently, by (10.63), one obtains

$$F'^{\mu\nu} \;=\; \begin{pmatrix} 0 & -E_x & -\frac{E_y - v_1 B_z}{\sqrt{1-v_1^2}} & -\frac{E_z + v_1 B_y}{\sqrt{1-v_1^2}} \\ E_x & 0 & -\frac{B_z - v_1 E_y}{\sqrt{1-v_1^2}} & \frac{B_y + v_1 E_z}{\sqrt{1-v_1^2}} \\ \frac{E_y - v_1 B_z}{\sqrt{1-v_1^2}} & \frac{B_z - v_1 E_y}{\sqrt{1-v_1^2}} & 0 & -B_x \\ \frac{E_z + v_1 B_y}{\sqrt{1-v_1^2}} & -\frac{B_y + v_1 E_z}{\sqrt{1-v_1^2}} & B_x & 0 \end{pmatrix} \tag{10.107}$$

Comparision with

$$F'^{\mu\nu} \;=\; \begin{pmatrix} 0 & -E'_x & -E'_y & -E'_z \\ E'_x & 0 & -B'_z & B'_y \\ E'_y & B'_z & 0 & -B'_x \\ E'_z & -B'_y & B'_x & 0 \end{pmatrix} \tag{10.108}$$

yields then the expressions for the transformed fields $\vec{E}'$ and $\vec{B}'$. The results can be put into the more general form

$$\vec{E}'_{\parallel} \;=\; \vec{E}_{\parallel}\,, \quad \vec{E}'_{\perp} \;=\; \frac{\vec{E}_{\perp} \,+\, \vec{v}\times\vec{B}}{\sqrt{1\,-\,\vec{v}^{\,2}}} \tag{10.109}$$

$$\vec{B}'_{\parallel} \;=\; \vec{B}_{\parallel}\,, \quad \vec{B}'_{\perp} \;=\; \frac{\vec{B}_{\perp} \,-\, \vec{v}\times\vec{E}}{\sqrt{1\,-\,\vec{v}^{\,2}}} \tag{10.110}$$

where $\vec{E}_{\parallel}$, $\vec{B}_{\parallel}$ and $\vec{E}_{\perp}$, $\vec{B}_{\perp}$ are, respectively, the components of the fields parallel and perpendicular to the velocity $\vec{v}$ which determines the Lorentz transformation. These equations show that under Lorentz transformations electric and magnetic fields convert into one another.

Maxwell Equations in Lorentz-Invariant Form

One can express the Maxwell equations in terms of the tensor $F^{\mu\nu}$ in Lorentz-invariant form. Noting

$$\partial_\mu F^{\mu\nu} \;=\; \partial_\mu\partial^\mu A^\nu \;-\; \partial_\mu\partial^\nu A^\mu \;=\; \partial_\mu\partial^\mu A^\nu \;-\; \partial^\nu\partial_\mu A^\mu \;=\; \partial_\mu\partial^\mu A^\nu\,, \tag{10.111}$$

where we used (10.105) and (10.97), one can conclude from (10.102)

$$\partial_\mu F^{\mu\nu} \;=\; 4\pi\, J^\nu\,. \tag{10.112}$$

One can readily prove that this equation is equivalent to the two inhomogeneous Maxwell equations (8.1, 8.2). From the definition (10.105) of the tensor $F^{\mu\nu}$ one can conclude the property

$$\partial^\sigma F^{\mu\nu} \;+\; \partial^\mu F^{\nu\sigma} \;+\; \partial^\nu F^{\sigma\mu} \;=\; 0 \tag{10.113}$$

which can be shown to be equivalent to the two homogeneous Maxwell equations (8.3, 8.4).

Lorentz Force

One important property of the electromagnetic field is the Lorentz force acting on charged particles moving through the field. We want to express this force through the tensor $F^{\mu\nu}$. It holds for a particle with 4-momentum p^μ as given by (10.67) and charge q

$$\frac{dp^\mu}{d\tau} \;=\; \frac{q}{m}\, p_\nu\, F^{\mu\nu} \tag{10.114}$$

where $d/d\tau$ is given by (10.86). We want to demonstrate now that this equation is equivalent to the equation of motion (8.5) where $\vec{p} = m\vec{v}/\sqrt{1-v^2}$.
To avoid confusion we will employ in the following for the energy of the particle the notation $\mathcal{E} = m/\sqrt{1-v^2}$ [see (10.87)] and retain the definition $\vec{E}$ for the electric field. The $\mu = 0$ component of (10.114) reads then, using (10.104),

$$\frac{d\mathcal{E}}{d\tau} \;=\; \frac{q}{m}\, \vec{p}\cdot\vec{E} \tag{10.115}$$

or with (10.86)

$$\frac{d\mathcal{E}}{dt} \;=\; \frac{q}{\mathcal{E}}\, \vec{p}\cdot\vec{E}\;. \tag{10.116}$$

From this one can conclude, employing (10.93),

$$\frac{1}{2}\frac{d\mathcal{E}^2}{dt} \;=\; \frac{1}{2}\frac{d\vec{p}^{\,2}}{dt} \;=\; q\,\vec{p}\cdot\vec{E} \tag{10.117}$$

This equation follows, however, also from the equation of motion (8.5) taking the scalar product with $\vec{p}$

$$\vec{p}\cdot\frac{d\vec{p}}{dt} \;=\; q\vec{p}\cdot\vec{E} \tag{10.118}$$

where we exploited the fact that according to $\vec{p} = m\vec{v}/\sqrt{1-v^2}$ holds $\vec{p} \parallel \vec{v}$.
For the spatial components, e.g., for $\mu = 1$, (10.114) reads using (10.103)

$$\frac{dp_x}{d\tau} \;=\; \frac{q}{m}\,(\,\mathcal{E}E_x \;+\; p_y B_z \;-\; p_z B_y\,)\;. \tag{10.119}$$

Employing again (10.86) and (10.67), i.e., $\mathcal{E} = m/\sqrt{1-v^2}$, yields

$$\frac{dp_x}{dt} \;=\; q\,\Big[\,E_x \;+\; (\vec{v}\times\vec{B})_x\Big] \tag{10.120}$$

which is the x-component of the equation of motion (8.5). We have, hence, demonstrated that (10.114) is, in fact, equivalent to (8.5). The term on the r.h.s. of (10.120) is referred to as the Lorentz force. Equation (10.114), hence, provides an alternative description of the action of the Lorentz force.

10.4 Function Space Representation of Lorentz Group

In the following it will be required to decribe the transformation of wave functions under Lorentz transformations. In this section we will investigate the transformation properties of *scalar* functions $\psi(x^\mu)$, $\psi \in \mathbb{C}_\infty(4)$. For such functions holds in the transformed frame

$$\psi'(L^\mu{}_\nu x^\nu) \;=\; \psi(x^\mu) \tag{10.121}$$

which states that the function values $\psi'(x'^\mu)$ at each point x'^μ in the new frame are identical to the function values $\psi(x^\mu)$ in the old frame *taken at the same space–time point* x^μ, i.e., taken at the pairs of points $(x'^\mu = L^\mu{}_\nu x^\nu, x^\mu)$. We need to emphasize that (10.121) covers solely the transformation behaviour of *scalar* functions. Functions which represent 4-vectorsor other non-scalar entities, e.g., the charge-current density in case of Sect. 10.3 or the bi-spinor wave function of electron-positron pairs in Sect. 10.7, obey a different transformation law.
We like to express now $\psi'(x'^\mu)$ in terms of the old coordinates x^μ. For this purpose one replaces x^μ in (10.121) by $(L^{-1})^\mu{}_\nu x^\nu$ and obtains

$$\psi'(x^\mu) \;=\; \psi((L^{-1})^\mu{}_\nu x^\nu)\;. \tag{10.122}$$

This result gives rise to the definition of the function space representation $\rho(L^\mu{}_\nu)$ of the Lorentz group

$$(\rho(L^\mu{}_\nu)\psi)(x^\mu) \;\stackrel{\text{def}}{=}\; \psi((L^{-1})^\mu{}_\nu x^\nu)\;. \tag{10.123}$$

This definition corresponds closely to the function space representation (5.42) of SO(3). In analogy to the situation for SO(3) we seek an expression for $\rho(L^\mu{}_\nu)$ in terms of an exponential operator and transformation parameters $\vec{\vartheta}$, $\vec{w}$, i.e., we seek an expression which corresponds to (10.51) for the natural representation of the Lorentz group. The resulting expression should be a generalization of the function space representation (5.48) of SO(3), in as far as SO(3,1) is a generalization (rotation + boosts) of the group SO(3). We will denote the intended representation by

$$\mathcal{L}(\vec{\vartheta}, \vec{w}) \;\stackrel{\text{def}}{=}\; \rho(L^\mu{}_\nu(\vec{\vartheta}, \vec{w})) \;=\; \rho\left(e^{\vec{\vartheta}\cdot\vec{J}+\vec{w}\cdot\vec{K}}\right) \tag{10.124}$$

which we present in the form

$$\mathcal{L}(\vec{\vartheta}, \vec{w}) \;=\; \exp\left(\vec{\vartheta}\cdot\vec{\mathcal{J}} \;+\; \vec{w}\cdot\vec{\mathcal{K}}\right)\;. \tag{10.125}$$

In this expression $\vec{\mathcal{J}} = (\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3)$ and $\vec{\mathcal{K}} = (\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3)$ are the generators of $\mathcal{L}(\vec{\vartheta}, \vec{w})$ which correspond to the generators J_k and K_k in (10.47), and which can be constructed following the procedure adopted for the function space representation of SO(3). However, in the present case we exclude the factor '$-i$' [cf. (5.48) and (10.125)]. Accordingly, one can evaluate $\mathcal{J}_k$ as follows

$$\mathcal{J}_k \;=\; \lim_{\vartheta_k\to 0}\frac{1}{\vartheta_1}\left[\rho\left(e^{\vartheta_k J_k}\right) \;-\; 1\!\!1\right] \tag{10.126}$$

and $\mathcal{K}_k$

$$\mathcal{K}_k \;=\; \lim_{w_k\to 0}\frac{1}{w_1}\left[\rho\left(e^{w_k K_k}\right) \;-\; 1\!\!1\right]\;. \tag{10.127}$$

One obtains

$$\begin{aligned} \mathcal{J}_1 &= x^3\partial_2 - x^2\partial_3 ; & \mathcal{K}_1 &= x^0\partial_1 + x^1\partial_0 \\ \mathcal{J}_2 &= x^1\partial_3 - x^3\partial_1 ; & \mathcal{K}_2 &= x^0\partial_2 + x^2\partial_0 \\ \mathcal{J}_3 &= x^2\partial_1 - x^1\partial_2 ; & \mathcal{K}_3 &= x^0\partial_3 + x^3\partial_0 \end{aligned} \tag{10.128}$$

which we like to demonstrate for $\mathcal{J}_1$ and $\mathcal{K}_1$.
In order to evaluate (10.126) for $\mathcal{J}_1$ we consider first

$$\left(e^{\vartheta_1 J_1}\right)^{-1} = e^{-\vartheta_1 J_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\vartheta_1 & \sin\vartheta_1 \\ 0 & 0 & -\sin\vartheta_1 & \cos\vartheta_1 \end{pmatrix} \tag{10.129}$$

which yields for small ϑ_1

$$\begin{aligned} \rho\left(e^{\vartheta_1 J_1}\right)\psi(x^\mu) &= \psi(x^0, x^1, \cos\vartheta_1\, x^2 + \sin\vartheta_1\, x^3, -\sin\vartheta_1\, x^2 + \cos\vartheta_1\, x^3) \\ &= \psi(x^\mu) + \vartheta_1\,(x^3\partial_2 - x^2\partial_3)\,\psi(x^\mu) + O(\vartheta_1^2) . \end{aligned} \tag{10.130}$$

This result, obviously, reproduces the expression for $\mathcal{J}_1$ in (10.128).
One can determine similarly $\mathcal{K}_1$ starting from

$$\left(e^{w_1 K_1}\right)^{-1} = e^{-w_1 K_1} = \begin{pmatrix} \cosh w_1 & \sinh w_1 & 0 & 0 \\ \sinh w_1 & \cosh w_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{10.131}$$

This yields for small w_1

$$\begin{aligned} \rho\left(e^{w_1 K_1}\right)\psi(x^\mu) &= \psi(\cosh w_1\, x^0 + \sinh w_1\, x^1, \sinh w_1\, x^0 + \cosh w_1\, x^1, x^2, x^3) \\ &= \psi(x^\mu) + w_1\,(x^1\partial_0 + x^0\partial_1)\,\psi(x^\mu) + O(w_1^2) \end{aligned} \tag{10.132}$$

and, obviously, the expression for $\mathcal{K}_1$ in (10.126).
The generators $\vec{\mathcal{J}}$, $\vec{\mathcal{K}}$ obey the same Lie algebra (10.49) as the generators of the natural representation, i.e.

$$\begin{aligned} [\,\mathcal{J}_k, \mathcal{J}_\ell\,] &= \epsilon_{k\ell m}\,\mathcal{J}_m \\ [\,\mathcal{K}_k, \mathcal{K}_\ell\,] &= -\,\epsilon_{k\ell m}\,\mathcal{J}_m \\ [\,\mathcal{J}_k, \mathcal{K}_\ell\,] &= \epsilon_{k\ell m}\,\mathcal{K}_m . \end{aligned} \tag{10.133}$$

We demonstrate this for three cases, namely $[\mathcal{J}_1, \mathcal{J}_2] = \mathcal{J}_3$, $[\mathcal{K}_1, \mathcal{K}_2] = -\mathcal{J}_3$, and $[\mathcal{J}_1, \mathcal{K}_2] = \mathcal{K}_3$:

$$\begin{aligned} [\,\mathcal{J}_1, \mathcal{J}_2\,] &= [x^3\partial_2 - x^2\partial_3, x^1\partial_3 - x^3\partial_1] \\ &= [x^3\partial_2, x^1\partial_3] - [x^2\partial_3, x^3\partial_1] \\ &= -x^1\partial_2 + x^2\partial_1 = \mathcal{J}_3 , \end{aligned} \tag{10.134}$$

$$
\begin{aligned}
[\,\mathcal{K}_1,\,\mathcal{K}_2\,] &= [x^0\partial_1 \,+\, x^1\partial_0,\, x^0\partial_2 \,+\, x^2\partial_0] \\
&= [x^0\partial_1,\, x^2\partial_0] \,-\, [x^1\partial_0,\, x^0\partial_2] \\
&= -x^2\partial_1 \,+\, x^1\partial_2 \;=\; -\,\mathcal{J}_3\;,
\end{aligned} \tag{10.135}
$$

$$
\begin{aligned}
[\,\mathcal{J}_1,\,\mathcal{K}_2\,] &= [x^3\partial_2 \,-\, x^2\partial_3,\, x^0\partial_2 \,+\, x^2\partial_0] \\
&= [x^3\partial_2,\, x^2\partial_0] \,-\, [x^2\partial_3,\, x^0\partial_2] \\
&= x^3\partial_0 \,+\, x^0\partial_3 \;=\; \mathcal{K}_3\;.
\end{aligned} \tag{10.136}
$$

One-Dimensional Function Space Representation

The exponential operator (10.125) in the case of a one-dimensional transformation of the type

$$\mathcal{L}(w^3) \;=\; \exp\left(\,w^3\mathcal{K}_3\,\right)\;, \tag{10.137}$$

where $\mathcal{K}_3$ is given in (10.128), can be simplified considerably. For this purpose one expresses $\mathcal{K}_3$ in terms of hyperbolic coordinates R, Ω which are connected with x^0, x^3 as follows

$$x^0 \;=\; R\cosh\Omega\;, \quad x^3 \;=\; R\sinh\Omega \tag{10.138}$$

a relationship which can also be stated

$$R \;=\; \begin{cases} +\sqrt{(x^0)^2 \,-\, (x^3)^2} & \text{if } x^0 \,\geq\, 0 \\ -\sqrt{(x^0)^2 \,-\, (x^3)^2} & \text{if } x^0 \,<\, 0 \end{cases} \tag{10.139}$$

and

$$\tanh\Omega \;=\; \frac{x^3}{x^0}\;, \quad \coth\Omega \;=\; \frac{x^0}{x^3}\;. \tag{10.140}$$

The transformation to hyperbolic coordinates closely resembles the transformation to radial coordinates for the generators of SO(3) in the function space representation [cf. Eqs. (5.85-5.87)]. In both cases the radial coordinate is the quantity conserved under the transformations, i.e., $\sqrt{x_1^2 + x_2^2 + x_3^2}$ in the case of SO(3) and $\sqrt{(x^0)^2 - (x^3)^2}$ in case of transformation (10.137).

In the following we consider solely the case $x^0 \geq 0$. The relationships (10.139, 10.140) allow one to express the derivatives ∂_0, ∂_3 in terms of $\frac{\partial}{\partial R}, \frac{\partial}{\partial \Omega}$. We note

$$\frac{\partial R}{\partial x^0} \;=\; \frac{x^0}{R}\;, \quad \frac{\partial R}{\partial x^3} \;=\; -\frac{x^0}{R} \tag{10.141}$$

and

$$
\begin{aligned}
\frac{\partial \Omega}{\partial x^3} &= \frac{\partial \Omega}{\partial \tanh\Omega}\frac{\partial \tanh\Omega}{\partial x^3} \;=\; \cosh^2\Omega\,\frac{1}{x^0} \\
\frac{\partial \Omega}{\partial x^0} &= \frac{\partial \Omega}{\partial \coth\Omega}\frac{\partial \coth\Omega}{\partial x^0} \;=\; -\sinh^2\Omega\,\frac{1}{x^3}\;.
\end{aligned} \tag{10.142}
$$

The chain rule yields then

$$
\begin{aligned}
\partial_0 &= \frac{\partial R}{\partial x^0}\frac{\partial}{\partial R} \,+\, \frac{\partial \Omega}{\partial x^0}\frac{\partial}{\partial \Omega} \;=\; \frac{x^0}{R}\frac{\partial}{\partial R} \,-\, \sinh^2\Omega\,\frac{1}{x^3}\frac{\partial}{\partial \Omega} \\
\partial_3 &= \frac{\partial R}{\partial x^3}\frac{\partial}{\partial R} \,+\, \frac{\partial \Omega}{\partial x^3}\frac{\partial}{\partial \Omega} \;=\; -\frac{x^3}{R}\frac{\partial}{\partial R} \,+\, \cosh^2\Omega\,\frac{1}{x^0}\frac{\partial}{\partial \Omega}\;.
\end{aligned} \tag{10.143}
$$

Inserting these results into the definition of $\mathcal{K}_3$ in (10.128) yields

$$\mathcal{K}_3 \;=\; x^0\partial_3 \;+\; x^3\partial_0 \;=\; \frac{\partial}{\partial\Omega}\,. \tag{10.144}$$

The action of the exponential operator (10.137) on a function $f(\Omega) \in \mathbb{C}_\infty(1)$ is then that of a shift operator

$$\mathcal{L}(w^3)\, f(\Omega) \;=\; \exp\left(w^3 \frac{\partial}{\partial\Omega} \right) f(\Omega) \;=\; f(\Omega \,+\, w^3)\,. \tag{10.145}$$

10.5 Klein–Gordon Equation

In the following Sections we will provide a heuristic derivation of the two most widely used quantum mechanical descriptions in the relativistic regime, namely the Klein–Gordon and the Dirac equations. We will provide a 'derivation' of these two equations which stem from the historical development of relativistic quantum mechanics. The historic route to these two equations, however, is not very insightful, but certainly is short and, therefore, extremely useful. Further below we will provide a more systematic, representation theoretic treatment.

Free Particle Case

A quantum mechanical description of a relativistic free particle results from applying the *correspondence principle*, which allows one to replace classical observables by quantum mechanical operators acting on wave functions. In the position representation the correspondence principle states

$$\begin{aligned} E \;&\Longrightarrow\; \hat{E} = -\frac{\hbar}{i}\partial_t \\ \vec{p} \;&\Longrightarrow\; \hat{\vec{p}} = \frac{\hbar}{i}\nabla \end{aligned} \tag{10.146}$$

which, in 4-vector notation reads

$$p_\mu \Longrightarrow \hat{p}_\mu \;=\; i\hbar(\partial_t, \nabla) \;=\; i\hbar\partial_\mu\,; \quad p^\mu \Longrightarrow \hat{p}^\mu \;=\; i(\partial_t, -\nabla) \;=\; i\hbar\partial^\mu\,. \tag{10.147}$$

Applying the correspondence principle to (10.92) one obtains the wave equation

$$-\,\hbar^2\partial^\mu\partial_\mu\,\psi(x^\nu) \;=\; m^2\,\psi(x^\nu) \tag{10.148}$$

or

$$\left(\hbar^2\,\partial^\mu\partial_\mu \;+\; m^2\right)\,\psi(x^\nu) \;=\; 0\,. \tag{10.149}$$

where $\psi(x^\mu)$ is a *scalar*, complex-valued function. The latter property implies that upon change of reference frame $\psi(x^\mu)$ transforms according to (10.121, 10.122). The partial differential equation (10.151) is called the *Klein-Gordon equation*.
In the following we will employ so-called *natural units* $\hbar \,=\, c \,=\, 1$. In these units the quantities energy, momentum, mass, $(\text{length})^{-1}$, and $(\text{time})^{-1}$ all have the same dimension. In natural units the Klein–Gordon equation (10.151) reads

$$\left(\partial_\mu\partial^\mu \;+\; m^2\right)\,\psi(x^\mu) \;=\; 0 \tag{10.150}$$

or

$$\left(\partial_t^2 \; - \; \nabla^2 \; + \; m^2 \right) \, \psi(x^\mu) \; = \; 0 \, . \tag{10.151}$$

One can notice immeadiately that (10.150) is invariant under Lorentz transformations. This follows from the fact that $\partial_\mu \partial^\mu$ and m^2 are scalars, and that (as postulated) $\psi(x^\mu)$ is a scalar. Under Lorentz transformations the free particle Klein–Gordon equation (10.150) becomes

$$\left(\partial'_\mu \partial'^\mu + \; m^2 \right) \, \psi'(x'^\mu) \; = \; 0 \tag{10.152}$$

which has the same form as the Klein–Gordon equation in the original frame.

Current 4-Vector Associated with the Klein-Gordon Equation

As is well-known the Schrödinger equation of a free particle

$$i\partial_t \psi(\vec{r}, t) \; = \; -\frac{1}{2m} \nabla^2 \, \psi(\vec{r}, t) \tag{10.153}$$

is associated with a conservation law for particle probability

$$\partial_t \rho_S(\vec{r}, t) \; + \; \nabla \cdot \vec{j}_S(\vec{r}, t) \; = \; 0 \tag{10.154}$$

where

$$\rho_S(\vec{r}, t) \; = \; \psi^*(\vec{r}, t) \, \psi(\vec{r}, t) \tag{10.155}$$

describes the *positive definite* probability to detect a particle at position $\vec{r}$ at time t and where

$$\vec{j}_S(\vec{r}, t) \; = \; \frac{1}{2mi} \left[\psi^*(\vec{r}, t) \nabla \psi(\vec{r}, t) \; - \; \psi(\vec{r}, t) \nabla \psi^*(\vec{r}, t) \right] \tag{10.156}$$

describes the current density connected with motion of the particle probability distribution. To derive this conservation law one rewrites the Schrödinger equation in the form $(i\partial_t - \frac{1}{2m}\nabla^2)\psi \; = \; 0$ and considers

$$\mathrm{Im} \left[\psi^* \left(i\partial_t \; - \; \frac{1}{2m} \nabla^2 \right) \psi \right] \; = \; 0 \tag{10.157}$$

which is equivalent to (10.154).

In order to obtain the conservation law connected with the Klein–Gordon equation (10.150) one considers

$$\mathrm{Im} \left[\psi^* \left(\partial_\mu \partial^\mu \; + \; m^2 \right) \psi \right] \; = \; 0 \tag{10.158}$$

which yields

$$\begin{aligned} \psi^* \partial_t^2 \psi \; - \; \psi \partial_t^2 \psi^* \; - \; \psi^* \nabla^2 \psi \; + \; \psi \nabla^2 \psi^* \qquad &= \\ \partial_t \left(\psi^* \partial_t \psi \; - \; \psi \partial_t \psi^* \right) \; + \; \nabla \cdot \left(\psi \nabla \psi^* \; - \; \psi^* \nabla \psi \right) \quad &= \; 0 \end{aligned} \tag{10.159}$$

which corresponds to

$$\partial_t \rho_{KG}(\vec{r}, t) \; + \; \nabla \cdot \vec{j}_{KG}(\vec{r}, t) \; = \; 0 \tag{10.160}$$

where

$$\begin{aligned} \rho_{KG}(\vec{r}, t) \quad &= \; \frac{i}{2m} \left(\psi^*(\vec{r}, t) \partial_t \psi(\vec{r}, t) \; - \; \psi(\vec{r}, t) \partial_t \psi^*(\vec{r}, t) \right) \\ \vec{j}_{KG}(\vec{r}, t) \quad &= \; \frac{1}{2mi} \left(\psi^*(\vec{r}, t) \nabla \psi(\vec{r}, t) \; - \; \psi(\vec{r}, t) \nabla \psi^*(\vec{r}, t) \right) \, . \end{aligned} \tag{10.161}$$

This conservation law differs in one important aspect from that of the Schrödinger equation (10.154), namely, in that the expression for ρ_{KG} is *not* positive definite. When the Klein-Gordon equation had been initially suggested this lack of positive definiteness worried physicists to a degree that the Klein–Gordon equation was rejected and the search for a Lorentz–invariant quantum mechanical wave equation continued. Today, the Klein-Gordon equation is considered as a suitable equation to describe spin–0 particles, for example pions. The proper interpretation of $\rho_{KG}(\vec{r},t)$, it had been realized later, is actually that of a charge density, not of particle probability.

Solution of the Free Particle Klein–Gordon Equation

Solutions of the free particle Klein–Gordon equation are

$$\psi(x^\mu) \;=\; N\, e^{-ip_\mu x^\mu} \;=\; N\, e^{i(\vec{p}_0\cdot\vec{r}\,-\,E_o t)} \,. \tag{10.162}$$

Inserting this into the Klein–Gordon equation (10.151) yields

$$\left(E_o^2 \;-\; \vec{p}_0^{\,2} \;-\; m^2 \right) \psi(\vec{r},t) \;=\; 0 \tag{10.163}$$

which results in the expected [see (10.93] dispersion relationship connecting E_0, $\vec{p}_0$, m

$$E_0^2 \;=\; m^2 \;+\; \vec{p}_o^{\,2} \,. \tag{10.164}$$

The corresponding energy is

$$E_o(\vec{p}_o,\,\pm) \;=\; \pm\sqrt{m^2 \;+\; \vec{p}_o^{\,2}} \tag{10.165}$$

This result together with (10.162) shows that the solutions of the free particle Klein-Gordon equation (10.150) are actually determined by $\vec{p}_o$ and by the choice of sign $\pm$. We denote this by summarizing the solutions as follows

$$\begin{aligned} \left(\partial_\mu\partial^\mu \;+\; m^2 \right) \psi_o(\vec{p},\lambda|x^\mu) \qquad\quad &= \; 0 \\ \psi_o(\vec{p},\lambda|x^\mu) \;=\; N_{\lambda,p}\, e^{i(\vec{p}\cdot\vec{r}\,-\,\lambda E_o(\vec{p})t)} \quad E_o(\vec{p}) \; &= \; \sqrt{m^2 \;+\; \vec{p}_o^{\,2}}\;,\lambda \;=\pm \end{aligned} \tag{10.166}$$

The spectrum of the Klein–Gordon equation (10.150) is a continuum of positive energies $E \,\geq\, m$, corresponding to $\lambda \,=\, +$, and of negative energies $E \,\leq\, -m$, corresponding to $\lambda \,=\, -$. The density $\rho_{KG}(\vec{p},\lambda)$ associated with the corresponding wave functions $\psi_o(\vec{p},\lambda|x^\mu)$ according to (10.161) and (10.166) is

$$\rho_{KG}(\vec{p},\lambda) \;=\; \lambda\,\frac{E_o(\vec{p})}{m}\,\psi_o^*(\vec{p},\lambda|x^\mu)\psi_o(\vec{p},\lambda|x^\mu) \tag{10.167}$$

which is positive for $\lambda \,=\, +$ and negative for $\lambda \,=\, -$. The proper interpretation of the two cases is that the Klein–Gordon equation describes particles as well as anti-particles; the *anti-particles* carry a charge opposite to that of the associated *particles*, and the density $\rho_{KG}(\vec{p},\lambda)$ actually describes charge density rather than probability.

Generating a Solution Through Lorentz Transformation

A particle at rest, i.e., with $\vec{p} \;=\; 0$, according to (**??**) is decribed by the $\vec{r}$–independent wave function

$$\psi_o(\vec{p}=0,\lambda|x^\mu) \;=\; N\, e^{-i\lambda m t}\;, \quad \lambda \;=\; \pm\,. \tag{10.168}$$

We want to demonstrate now that the wave functions for $\vec{p} \neq 0$ in (**??**) can be obtained through appropriate Lorentz transformation of (10.168). For this purpose we consider the wave function for a particle moving with momentum velocity v in the direction of the x^3–axis. Such wave function should be generated by applying the Lorentz transformation in the function space representation (10.145) choosing $\frac{p}{m} = \text{sinh}w^3$. This yields, in fact, for the wave function (10.168), using (10.138) to replace $t = x^0$ by hyperbolic coordinates R, Ω,

$$\begin{aligned} \mathcal{L}(w^3)\psi_o(\vec{p}=0,\lambda|x^\mu) &= \exp\left(w^3 \frac{\partial}{\partial\Omega} \right) N\, e^{-i\lambda m R \cosh\Omega} \\ &= N\, e^{-i\lambda m R \cosh(\Omega + w^3)} \,. \end{aligned} \tag{10.169}$$

The addition theorem of hyperbolic functions $\cosh(\Omega + w^3) = \cosh\Omega \cosh w^3 + \sinh\Omega \sinh w^3$ allows us to rewrite the exponent on the r.h.s. of (10.169)

$$-i\lambda\,(\,m\cosh w^3\,)\,(\,R\cosh\Omega\,) \;-\; i\lambda\,(\,m\sinh w^3\,)\,(\,R\sinh\Omega\,)\,. \tag{10.170}$$

The coordinate transformation (10.138) and the relationships (10.61) yield for this expression

$$-i\lambda \frac{m}{\sqrt{1 - v^2}}\, x^0 \;-\; i\lambda \frac{m\,v}{\sqrt{1 - v^2}}\, x^3 \,. \tag{10.171}$$

One can interpret then for $\lambda = +$, i.e., for positive energy solutions,

$$p \;=\; -\,mv/\sqrt{1 - v^2} \tag{10.172}$$

as the momentum of the particle relative to the moving frame and

$$\frac{m}{\sqrt{1 - v^2}} \;=\; \sqrt{\frac{m^2}{1 - v^2}} \;=\; \sqrt{m^2 + \frac{m^2 v^2}{1 - v^2}} \;=\; \sqrt{m^2 + p^2} \;=\; E_o(p) \tag{10.173}$$

as the energy [c.f. (10.166)] of the particle. In case of $\lambda = +$ one obtains finally

$$\mathcal{L}(w^3)\psi_o(\vec{p}=0,\lambda=+|x^\mu) \;=\; N\, e^{i(px^3 \,-\, E_o(p)x^0} \tag{10.174}$$

which agrees with the expression given in (10.166). In case of $\lambda = -$, i.e., for negative energy solutions, one has to interprete

$$p \;=\; mv/\sqrt{1 - v^2} \tag{10.175}$$

as the momentum of the particle and one obtains

$$\mathcal{L}(w^3)\psi_o(\vec{p}=0,\lambda=-|x^\mu) \;=\; N\, e^{i(px^3 \,+\, E_o(p)x^0} \,. \tag{10.176}$$

10.6 Klein–Gordon Equation for Particles in an Electromagnetic Field

We consider now the quantum mechanical wave equation for a spin–0 particle moving in an electromagnetic field described by the 4-vector potential

$$A^\mu(x^\mu) \;=\; (V(\vec{r},t), \vec{A}(\vec{r},t))\,; \quad A_\mu(x^\mu) \;=\; (V(\vec{r},t), -\vec{A}(\vec{r},t)) \tag{10.177}$$

	free classical particle	classical particle in field $(V, \vec{A})$	free quantum particle	quantum particle in field $(V, \vec{A})$
energy	E	$E - qV$	$i\partial_t$	$i\partial_t - qV$
momentum	$\vec{p}$	$\vec{p} - q\vec{A}$	$\hat{\vec{p}} = -i\nabla$	$\hat{\vec{p}} - q\vec{A} = \hat{\vec{\pi}}$
4-vector	p^μ	$p^\mu - qA^\mu$	$i\partial_\mu$	$i\partial_\mu - qA_\mu = \pi_\mu$

Table 10.1:
Coupling of a particle of charge q to an electromagnetic field described by the 4-vector potential $A^\mu = (V, \vec{A})$ or $A_\mu = (V, -\vec{A})$. According to the so-called minimum coupling principle the presence of the field is accounted for by altering energy, momenta for classical particles and the respective operators for quantum mechanical particles in the manner shown. See also Eq. (10.147).

To obtain the appropriate wave equation we follow the derivation of the free particle Klein–Gordon equation above and apply again the correspondence principle to (10.93), albeit in a form, which couples a particle of charge q to an electromagnetic field described through the potential $A_\mu(x^\nu)$. According to the principle of minimal coupling [see (10.69)] one replaces the quantum mechanical operators, i.e., $i\partial_t$ and $-i\nabla$ in (10.150), according to the rules shown in Table 10.1. For this purpose one writes the Klein-Gordon equation (10.150)

$$\left(-g^{\mu\nu}(i\partial_\mu)(i\partial_\nu) + m^2 \right) \psi(x^\mu) = 0 . \tag{10.178}$$

According to the replacements in Table 10.1 this becomes

$$g^{\mu\nu}(i\partial_\mu - qA_\mu)(i\partial_\nu - A_\nu)\,\psi(x^\mu) = m^2\,\psi(x^\mu) \tag{10.179}$$

which can also be written

$$g^{\mu\nu}\pi_\mu\pi_\nu - ; m^2\,)\,\psi(x^\mu) = 0 . \tag{10.180}$$

In terms of space-time derivatives this reads

$$(i\partial_t - qV(\vec{r},t))^2\,\psi(\vec{r},t) = \left[\left(-i\nabla - q\vec{A}(\vec{r},t) \right)^2 + m^2 \right] \psi(\vec{r},t) . \tag{10.181}$$

Non-Relativistic Limit of Free Particle Klein–Gordon Equation

In order to consider further the interpretation of the positive and negative energy solutions of the Klein–Gordon equation one can consider the non-relativistic limit. For this purpose we split-off a factor $\exp(-imt)$ which describes the oscillations of the wave function due to the rest energy, and focus on the remaining part of the wave function, i.e., we define

$$\psi(\vec{r},t) = e^{-imt}\,\Psi(\vec{r},t) , \tag{10.182}$$

and seek an equation for $\Psi(\vec{r},t)$. We will also assume, in keeping withnthe non-relativistic limit, that the mass m of the particle, i.e., it's rest energy, is much larger than all other energy terms, in particular, larger than $|i\partial_t \Psi / \Psi|$ and alrger than qV, i.e.,

$$|\frac{i\partial_t \Psi}{\Psi}| << m \, , \quad |q V| << m \, . \tag{10.183}$$

The term on the l.h.s. of (10.181) can then be approximated as follows:

$$\begin{aligned} (i\partial_t - qV)^2 e^{-imt}\Psi &= (i\partial_t - qV)\left(m e^{-imt}\Psi + e^{-imt} i\partial_t\Psi - qVe^{-imt}\Psi\right) \\ &= m^2 e^{-imt}\Psi + m e^{-imt} i\partial_t \Psi - qVe^{-imt}\Psi \\ &\quad + m e^{-imt} i\partial_t\Psi - e^{-imt}\partial^2\Psi - qVe^{-imt} i\partial_t\Psi \\ &\quad - m e^{-imt} qV\Psi - e^{-imt} i\partial_t qV\Psi + q^2V^2 e^{-imt}\Psi \\ &\approx m^2 e^{-imt}\Psi - 2mqVe^{-imt}\Psi - 2me^{-imt} i\partial_t \Psi \end{aligned} \tag{10.184}$$

where we neglected all terms which did not contain factors m. The approximation is justified on the ground of the inequalities (10.183). The Klein-Gordon equation (10.181) reads then

$$i\,\partial_t\,\Psi(\vec{r},t) = \left[\frac{[\hat{\vec{p}} - q\vec{A}(\vec{r},t)]^2}{2m} + qV(\vec{r},t)\right]\Psi(\vec{r},t) \tag{10.185}$$

This is, however, identical to the Schrödinger equation (10.2) of a non-relativistic spin-0 particle moving in an electromagnetic field.

Pionic Atoms

To apply the Klein–Gordon equation (10.181) to a physical system we consider pionic atoms, i.e., atoms in which one or more electrons are replaced by π^- mesons. This application demonstrates that the Klein–Gordon equation describes spin zero particles, e.g., spin-0 mesons.
To 'manufacture' pionic atoms, π^- mesons are generated through inelastic proton–proton scattering

$$p + p \longrightarrow p + p + \pi^- + \pi^+ \, , \tag{10.186}$$

then are slowed down, filtered out of the beam and finally fall as slow pions onto elements for which a pionic variant is to be studied. The process of π^- meson capture involves the so-called Auger effect, the binding of a negative charge (typically an electron) while at the same time a lower shell electron is being emitted

$$\pi^- + atom \longrightarrow (atom - e^- + \pi^-) + e^- \, . \tag{10.187}$$

We want to investigate in the following a description of a *stationary state* of a pionic atom involving a nucleus with charge $+Ze$ and a π^- meson. A stationary state of the Klein–Gordon equation is described by a wave function

$$\psi(x^\mu) = \varphi(\vec{r})\, e^{-i\epsilon t} \, . \tag{10.188}$$

Inserting this into (10.181) yields (we assume now that the Klein–Gordon equation describes a particle with mass m_π and charge $-e$) for $qV(\vec{r},t) = -\frac{Ze^2}{r}$ and $\vec{A}(\vec{r},t) \equiv 0$

$$\left[\left(\epsilon + \frac{Ze^2}{r}\right)^2 + \nabla^2 - m_\pi^2\right]\varphi(\vec{r}) = 0 \, . \tag{10.189}$$

Because of the radial symmetry of the Coulomb potential we express this equation in terms of spherical coordinates r, θ, ϕ. The Laplacian is

$$\nabla^2 \;=\; \frac{1}{r}\partial_r^2 r \;+\; \frac{1}{r^2 \sin^2\theta}\partial_\theta \sin\theta \partial_\theta \;+\; \frac{1}{r^2 \sin^2\theta}\partial_\phi^2 \;=\; \frac{1}{r}\partial_r^2 r \;-\; \frac{\hat{L}^2}{r^2} \;. \tag{10.190}$$

With this expression and after expanding $(\epsilon + \frac{Ze^2}{r})^2$ one obtains

$$\left(\frac{d^2}{dr^2} \;-\; \frac{\hat{L}^2 \;-\; Z^2 e^4}{r^2} \;+\; \frac{2\epsilon Z e^2}{r} \;+\; \epsilon^2 \;-\; m_\pi^2 \right) r\,\phi(\vec{r}) \;=\; 0 \;. \tag{10.191}$$

The operator $\hat{L}^2$ in this equation suggests to choose a solution of the type

$$\varphi(\vec{r}\,) \;=\; \frac{R_\ell(r)}{r}\, Y_{\ell m}(\theta, \phi) \tag{10.192}$$

where the functions $Y_{\ell m}(\theta, \phi)$ are spherical harmonics, i.e., the eigenfunctions of the operator $\hat{L}^2$ in (10.191)

$$\hat{L}^2\, Y_{\ell m}(\theta, \phi) \;=\; \ell\,(\ell \,+\, 1)\, Y_{\ell m}(\theta, \phi) \;. \tag{10.193}$$

(10.192) leads then to the ordinary differential equation

$$\left(\frac{d^2}{dr^2} \;-\; \frac{\ell(\ell+1) \;-\; Z^2 e^4}{r^2} \;+\; \frac{2\epsilon Z e^2}{r} \;+\; \epsilon^2 \;-\; m_\pi^2 \right) R_\ell(r) \;=\; 0 \;. \tag{10.194}$$

Bound state solutions can be obtained readily noticing that this equation is essentially identical to that posed by the Coulomb problem (potential $-\frac{Ze^2}{r}$) for the Schrödinger equation

$$\left(\frac{d^2}{dr^2} \;-\; \frac{\ell(\ell+1)}{r^2} \;+\; \frac{2 m_\pi Z e^2}{r} \;+\; 2 m_\pi E \right) R_\ell(r) \;=\; 0 \tag{10.195}$$

The latter problem leads to the well-known spectrum

$$E_n \;=\; -\,\frac{m_\pi\,(Ze^2)^2}{2n^2} \;;n \;=\; 1, 2, \ldots \;;\ell \;=\; 0, 1, \ldots n-1 \;. \tag{10.196}$$

In this expression the number n' defined through

$$n' \;=\; n \;-\; \ell \;-1 \tag{10.197}$$

counts the number of nodes of the wave function, i.e., this quantity definitely must be an integer. The similarity of (10.194) and (10.195) can be made complete if one determines λ such that

$$\lambda(\ell)\,(\lambda(\ell) \,+\, 1) \;=\; \ell\,(\ell \,+\, 1) \;-\; Z^2 e^4 \;. \tag{10.198}$$

The suitable choice is

$$\lambda(\ell) \;=\; -\frac{1}{2} \;+\; \sqrt{(\ell \,+\, \frac{1}{2})^2 \;-\; Z^2 e^4} \tag{10.199}$$

and one can write (10.194)

$$\left(\frac{d^2}{dr^2} - \frac{\lambda(\ell)\,(\,\lambda(\ell)\,+\,1\,)}{r^2} + \frac{2\epsilon Z e^2}{r} + \epsilon^2 - m_\pi^2 \right) R_\ell(r) \;=\; 0 \,. \tag{10.200}$$

The bound state solutions of this equation should correspond to ϵ values which can be obtained from (10.196) if one makes the replacement

$$E \longrightarrow \frac{\epsilon^2 - m_\pi^2}{2m_\pi} \,, \qquad \ell \longrightarrow \lambda(\ell) \,, \qquad e^2 \longrightarrow e^2 \, \frac{\epsilon}{m_\pi} \,. \tag{10.201}$$

One obtains

$$\frac{\epsilon^2 - m_\pi^2}{2m_\pi} \;=\; -\frac{m_\pi\, Z^2 e^4 \frac{\epsilon^2}{m_\pi^2}}{2\,(n' + \lambda(\ell) + 1)^2} \,. \tag{10.202}$$

Solving this for ϵ (choosing the root which renders $\epsilon \le m_\pi$, i.e., which corresponds to a bound state) yields

$$\epsilon \;=\; \frac{m_\pi}{\sqrt{1 + \frac{Z^2 e^4}{(\,n' + \lambda(\ell) + 1\,)^2}}} \quad ; \; n' \;=\; 0, 1, \ldots \; ; \; \ell \;=\; 0, 1, \ldots \,. \tag{10.203}$$

Using (10.197, 10.199) and defining $E_{\rm KG} \;=\; \epsilon$ results in the spectrum

$$E_{\rm KG}(n, \ell, m) \;=\; \frac{m_\pi}{\sqrt{1 + \frac{Z^2 e^4}{(\,n - \ell - \frac{1}{2} + \sqrt{(\ell + \frac{1}{2})^2 - Z^2 e^4}\,)^2}}}$$

$$\begin{aligned} n &= 1, 2, \ldots \\ \ell &= 0, 1, \ldots, n-1 \\ m &= -\ell, -\ell+1, \ldots, +\ell \end{aligned} \tag{10.204}$$

In order to compare this result with the spectrum of the non-relativistic hydrogen-like atom we expand in terms of the fine structure constant e^2 to order $O(\epsilon^8)$. Introducing $\alpha \;=\; Z^2 e^4$ and $\beta \;=\; \ell + \frac{1}{2}$ (10.204) reads

$$\frac{1}{\sqrt{1 + \frac{\alpha}{(n - \beta + \sqrt{\beta^2 - \alpha})^2}}} \tag{10.205}$$

and one obtains the series of approximations

$$\begin{aligned} &\frac{1}{\sqrt{1 + \frac{\alpha}{(n - \beta + \sqrt{\beta^2 - \alpha})^2}}} \\ &\approx \quad \frac{1}{\sqrt{1 + \frac{\alpha}{(n - \frac{\alpha}{2\beta} + O(\alpha^2))^2}}} \\ &\approx \quad \frac{1}{\sqrt{1 + \frac{\alpha}{n^2 - \frac{\alpha}{\beta} n + O(\alpha^2)}}} \end{aligned}$$

$$\approx \frac{1}{\sqrt{1 + \frac{\alpha}{n^2} + \frac{\alpha^2}{\beta n^3} + O(\alpha^3)}}$$
$$\approx \frac{1}{1 + \frac{\alpha}{2n^2} + \frac{\alpha^2}{2\beta n^3} - \frac{\alpha^2}{8n^4} + O(\alpha^3)}$$
$$\approx 1 - \frac{\alpha}{2n^2} - \frac{\alpha^2}{2\beta n^3} + \frac{\alpha^2}{8n^4} + \frac{\alpha^2}{4n^4} + O(\alpha^3)\,. \qquad (10.206)$$

From this results for (10.204)

$$E_{\rm KG}(n,\ell,m) \approx m - \frac{mZ^2e^4}{2n^2} - \frac{mZ^4e^8}{2n^3}\left[\frac{1}{\ell+\frac{1}{2}} - \frac{3}{4n}\right] + O(Z^6e^{12})\,. \qquad (10.207)$$

Here the first term represents the rest energy, the second term the non-relativistic energy, and the third term gives the leading relativistic correction. The latter term agrees with observations of pionic atoms, however, it does not agree with observations of the hydrogen spectrum. The latter spectrum shows, for example, a splitting of the six $n = 2, \ell = 1$ states into groups of two and four degenerate states. In order to describe electron spectra one must employ the Lorentz-invariant wave equation for spin-$\frac{1}{2}$ particles, i.e., the Dirac equation introduced below.
It must be pointed out here that ϵ does not denote energy, but in the present case rather the negative of the energy. Also, the π^- meson is a pseudoscalar particle, i.e., the wave function changes sign under reflection.

10.7 The Dirac Equation

Historically, the Klein–Gordon equation had been rejected since it did not yield a positive-definite probability density, a feature which is connected with the 2nd order time derivative in this equation. This derivative, in turn, arises because the Klein–Gordon equation, through the correspondence principle, is related to the equation $E^2 = m^2 + \vec{p}^2$ of the classical theory which involves a term E^2. A more satisfactory Lorentz–invariant wave equation, i.e., one with a positive-definite density, would have only a first order time derivative. However, because of the equivalence of space and time coordinates in the Minkowski space such equation necessarily can only have then first order derivatives with respect to spatial coordinates. It should feature then a differential operator of the type $\mathcal{D} = i\gamma^\mu\partial_\mu$.

Heuristic Derivation Starting from the Klein-Gordon Equation

An obvious starting point for a Lorentz-invariant wave equation with only a first order time derivative is $E = \pm\sqrt{m^2 + \vec{p}^2}$. Application of the correspondence principle (10.146) leads to the wave equation

$$i\partial_t \Psi(\vec{r},t) = \pm\sqrt{m^2 - \nabla^2}\,\Psi(\vec{r},t)\,. \qquad (10.208)$$

These two equation can be combined

$$\left(i\partial_t + \sqrt{m^2 - \nabla^2}\right)\left(i\partial_t - \sqrt{m^2 - \nabla^2}\right)\Psi(\vec{r},t) \qquad (10.209)$$

which, in fact, is identical to the two equations (10.208). Equations (10.208, 10.209), however, are unsatisfactory since expansion of the square root operator involves all powers of the Laplace operator, but not an operator $i\vec{\gamma}\cdot\nabla$ as suggested by the principle of relativity (equivalence of space and time). Many attempts were made by theoretical physicists to 'linearize' the square root operator in (10.208, 10.209), but for a long time to no avail. Finally, Dirac succeeded. His solution to the problem involved an ingenious step, namely, the realization that the linearization can be carried out only if one assumes a 4-dimensional representation of the coefficients γ^μ.
Initially, it was assumed that the 4-dimensional space introduced by Dirac could be linked to 4-vectors, i.e., quantities with the transformation law (10.66). However, this was not so. Instead, the 4-dimensionsional representation discovered by Dirac involved new physical properties, spin-$\frac{1}{2}$ and anti-particles. The discovery by Dirac, achieved through a beautiful mathematical theory, strengthens the believe of many theoretical physicists today that the properties of physical matter ultimately derive from a, yet to be discovered, beautiful mathematical theory and that, therefore, one route to important discoveries in physics is the creation of new mathematical descriptions of nature, these descriptions ultimately merging with the true theory of matter.

Properties of the Dirac Matrices

Let us now trace Dirac's steps in achieving the linearization of the 'square root operator' in (10.208). Starting point is to boldly factorize, according to (10.209), the operator of the Klein–Gordon equation

$$\partial_\mu\partial^\mu + m^2 = -(P + m)(P - m) \tag{10.210}$$

where

$$P = i\gamma^\mu\partial_\mu\,. \tag{10.211}$$

Obviously, this would lead to the two wave equations $(P - m)\Psi = 0$ and $(P + m)\Psi = 0$ which have a first order time derivative and, therefore, are associated with a positive-definite particle density. We seek to identify the coefficients γ^μ. Inserting (10.211) into (10.210) yields

$$\begin{aligned} & -g^{\mu\nu}\partial_\mu\partial_\nu - m^2 = (i\gamma^\mu\partial_\mu + m)(i\gamma^\mu\partial_\mu - m) \\ = \; & -\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu - m^2 = -\frac{1}{2}(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + \gamma^\nu\gamma^\mu\partial_\nu\partial_\mu) - m^2 \\ = \; & -\frac{1}{2}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)\partial_\mu\partial_\nu - m^2 \end{aligned} \tag{10.212}$$

where we have changed 'dummy' summation indices, exploited $\partial_\mu\partial_\nu = \partial_\nu\partial_\mu$, but did not commute the, so far, unspecified algebraic objects γ^μ and γ^ν. Comparing the left-most and the right-most side of the equations above one can conclude the following property of γ_μ

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = [\gamma^\mu, \gamma^\nu]_+ = 2\,g^{\mu\nu} \tag{10.213}$$

We want to determine now the simplest algebraic realization of γ^μ. It turns out that no 4-vector of real or complex coefficients can satisfy these conditions. In fact, the quantities γ^0, γ^1, γ^2, γ^3 can only be realized by $d\times d$–matrices requiring that the wave function $\Psi(x^\mu)$ is actually a d–dimensional vector of functions $\psi_1(x^\mu)$, $\psi_2(x^\mu)$, $\ldots$ $\psi_d(x^\mu)$.
For $\mu = \nu$ condition (10.213) reads

$$(\gamma^\mu)^2 = \begin{cases} 1 & \mu = 0 \\ -1 & \mu = 1,2,3 \end{cases}\;. \tag{10.214}$$

From this follows that γ^0 has real eigenvalues ± 1 and γ^j, $j = 1,2,3$ has imaginary eigenvalues $\pm i$. Accordingly, one can impose the condition

$$\gamma^0 \quad \text{is hermitian} \qquad ; \gamma^j \ , \ j = 1,2,3 \quad \text{are anti-hermitian} \ . \tag{10.215}$$

For $\mu \neq \nu$ (10.213) reads

$$\gamma^\mu \gamma^\nu \ = \ -\gamma^\nu \gamma^\mu \ , \tag{10.216}$$

i.e., the γ^μ are anti-commuting. From this one can conclude for the determinants of γ^μ

$$\det(\gamma^\mu \gamma^\nu) \ = \ \det(-\gamma^\nu \gamma^\mu) \ = \ (-1)^d \det(\gamma^\nu \gamma^\mu) \ = \ (-1)^d \det(\gamma^\mu \gamma^\nu) \ . \tag{10.217}$$

Obviously, as long as $\det(\gamma^\mu) \neq 0$ the dimension d of the square matrices γ^μ must be even so that $(-1)^d = 1$.
For $d = 2$ there exist only three anti-commuting matrices, namely the Pauli matrices $\sigma^1, \sigma^2, \sigma^3$ for which, in fact, holds

$$\left(\sigma^j\right)^2 \ = \ 1\!\!1 \ ; \quad \sigma^j \sigma^k \ = \ -\sigma^k \sigma^j \quad \text{for } j \neq k \ . \tag{10.218}$$

The Pauli matrices allow one, however, to construct four matrices γ^μ for the next possible dimension $d = 4$. A proper choice is

$$\gamma^0 \ = \ \begin{pmatrix} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{pmatrix} \quad ; \quad \gamma^j \ = \ \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} \ , \tag{10.219}$$

Using property (10.218) of the Pauli matrices one can readily prove that condition (10.213) is satisfied. We will argue further below that the choice f γ^μ, except for similarity trasnformations, is unique.

The Dirac Equation

Altogether we have shown that the Klein–Gordon equation can be factorized formally

$$(\, i\gamma^\mu \partial_\mu \ + \ m \,)\, (\, i\gamma^\mu \partial_\mu \ - \ m \,)\ \Psi(x^\mu) \ = 0 \tag{10.220}$$

where $\Psi(x^\mu)$ represents a 4-dimensional wave function, rather than a scalar wave function. From this equation one can conclude that also the following should hold

$$(\, i\gamma^\mu \partial_\mu \ - \ m \,)\ \Psi(x^\mu) \ = 0 \tag{10.221}$$

which is the celebrated *Dirac equation.*

The Adjoint Dirac Equation

The adjoint equation is

$$\Psi^\dagger(x^\mu) \left(i(\gamma^\mu)^\dagger \overleftarrow{\partial_\mu} \ + m \right) \ = 0 \tag{10.222}$$

where we have defined $\Psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$ and where $\overleftarrow{\partial_\mu}$ denotes the differential operator ∂^μ operating to the left side, rather than to the right side. One can readily show using the hermitian property of the Pauli matrices $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^j)^\dagger = -\gamma^j$ for $j = 1, 2, 3$ which, in fact, is implied by (10.215). This property can also be written

$$(\gamma^\mu)^\dagger = \gamma^0 \gamma^\mu \gamma^0 . \tag{10.223}$$

Inserting this into (10.222) and multiplication from the right by γ^0 yields the *adjoint Dirac equation*

$$\Psi^\dagger(x^\mu)\, \gamma^0 \left(i\gamma^\mu \overleftarrow{\partial_\mu} + m \right) = 0 . \tag{10.224}$$

Similarity Transformations of the Dirac Equation - Chiral Representation

The Dirac equation can be subject to any similarity transformation defined through a non-singular 4×4–matrix S. Defining a new representation of the wave function $\tilde{\Psi}(x^\mu)$

$$\tilde{\Psi}(x^\mu) = S\, \Psi(x^\mu) \tag{10.225}$$

leads to the 'new' Dirac equation

$$(i\tilde{\gamma}^\mu \partial_\mu - m)\, \tilde{\Psi}(x^\mu) = 0 \tag{10.226}$$

where

$$\tilde{\gamma}^\mu = S\, \gamma^\mu S^{-1} \tag{10.227}$$

A representation often adopted beside the one given by (10.222, 10.219) is the socalled *chiral representation* defined through

$$\tilde{\Psi}(x^\mu) = S\, \Psi(x^\mu)\, ; \quad S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1\!\!1 & 1\!\!1 \\ 1\!\!1 & -1\!\!1 \end{pmatrix} \tag{10.228}$$

and

$$\tilde{\gamma}^0 = \begin{pmatrix} 0 & 1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix} \quad ; \quad \tilde{\gamma}^j = \begin{pmatrix} 0 & -\sigma^j \\ \sigma^j & 0 \end{pmatrix} , \; j = 1, 2, 3 . \tag{10.229}$$

The similarity transformation (10.227) leaves the algebra of the Dirac matrices unaffected and commutation property (10.213) still holds, i.e.,

$$[\tilde{\gamma}^\mu, \tilde{\gamma}^\nu]_+ = 2\, g^{\mu\nu} . \tag{10.230}$$

Exercise 10.7.1: Derive (refeq:Dirac-intro20a) from (10.213), (10.227).

Schrödinger Form of the Dirac Equation

Another form in which the Dirac equation is used often results from multiplying (10.221) from the left by γ^0

$$\left(i\, \partial_t + i\, \hat{\vec{\alpha}} \cdot \nabla - \hat{\beta}\, m \right) \Psi(\vec{r}, t) = 0 \tag{10.231}$$

where $\hat{\vec{\alpha}}$ has the three components α_j, $j = 1,2,3$ and

$$\hat{\beta} \;=\; \begin{pmatrix} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{pmatrix} \;;\quad \hat{\alpha}_j \;=\; \begin{pmatrix} 0 & \sigma^j \\ \sigma^j & 0 \end{pmatrix} \;,\; j = 1,2,3 \;. \tag{10.232}$$

This form of the Dirac equation is called the Schrödinger form since it can be written in analogy to the time-dependent Schrödinger equation

$$i\partial_t \Psi(x^\mu) \;=\; \mathcal{H}_o \Psi(x^\mu)\,;\quad \mathcal{H}_o \;=\; \hat{\vec{\alpha}}\cdot\hat{\vec{p}} \;+\; \hat{\beta}\, m \;. \tag{10.233}$$

The eigenstates and eigenvalues of $\mathcal{H}$ correspond to the stationary states and energies of the particles described by the Dirac equation.

Clifford Algebra and Dirac Matrices

The matrices defined through

$$d_j \;=\; i\gamma^j \;, j \;=\; 1,2,3\,;\quad d_4 \;=\; \gamma^0 \tag{10.234}$$

satisfy the anti-commutation property

$$d_j d_k \;+\; d_k d_j \;=\; \begin{cases} 2 & \text{for } j \;=\; k \\ 0 & \text{for } j \;\neq\; k \end{cases} \tag{10.235}$$

as can be readily verified from (10.213). The associative algebra generated by $d_1 \ldots d_4$ is called a *Clifford algebra* C_4. The three Pauli matrices also obey the property (10.235) and, hence, form a Clifford algebra C_3. The representations of Clifford algebras C_m are well established. For example, in case of C_4, a representation of the d_j's is

$$\begin{aligned} d_1 \;&=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;;\; d_2 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ d_3 \;&=\; \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;;\; d_4 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \end{aligned} \tag{10.236}$$

where '$\otimes$' denotes the Kronecker product between matrices, i.e., the matrix elements of $C \;=\; A \otimes B$ are $C_{jk,\ell m} \;=\; A_{j\ell} B_{km}$.
The Clifford algebra C_4 entails a subgroup G_4 of elements

$$\pm\, d_{j_1} d_{j_2} \cdots d_{j_s} \;,\qquad j_1 < j_2 < \cdots < j_s \; s \leq 4 \tag{10.237}$$

which are the ordered products of the operators $\pm 1\!\!1$ and d_1, d_2, d_3, d_4. Obviously, any product of the d_j's can be brought to the form (10.237) by means of the property (10.235). There are (including the different signs) 32 elements in G_4 which we define as follows

$$\begin{aligned} &\Gamma_{\pm 1} \;=\; \pm 1\!\!1 \\ &\Gamma_{\pm 2} \;=\; \pm d_1 \,,\; \Gamma_{\pm 3} \;=\; \pm d_2 \,,\; \Gamma_{\pm 4} \;=\; \pm d_3 \,,\; \Gamma_{\pm 5} \;=\; \pm d_4 \\ &\Gamma_{\pm 6} \;=\; \pm d_1 d_2 \,,\; \Gamma_{\pm 7} \;=\; \pm d_1 d_3 \,,\; \Gamma_{\pm 8} \;=\; \pm d_1 d_4 \,,\; \Gamma_{\pm 9} \;=\; \pm d_2 d_3 \\ &\Gamma_{\pm 10} \;=\; \pm d_2 d_4 \,,\; \Gamma_{\pm 11} \;=\; \pm d_3 d_4 \\ &\Gamma_{\pm 12} \;=\; \pm d_1 d_2 d_3 \,,\; \Gamma_{\pm 13} \;=\; \pm d_1 d_2 d_4 \,,\; \Gamma_{\pm 14} \;=\; \pm d_1 d_3 d_4 \,,\; \Gamma_{\pm 15} \;=\; \pm d_2 d_3 d_4 \\ &\Gamma_{\pm 16} \;=\; \pm d_1 d_2 d_3 d_4 \end{aligned} \tag{10.238}$$

These elements form a group since obviously any product of two Γ_r's can be expressed in terms of a third Γ_r. The representations of this group are given by a set of 32 4×4–matrices which are equivalent with respect to similarity transformations. Since the Γ_j are hermitian the similarity transformations are actually given in terms of unitary transformations. One can conclude then that also any set of 4×4–matrices obeying (10.235) can differ only with respect to unitary similarity transformations. This property extends then to 4×4–matrices which obey (10.213), i.e., to Dirac matrices.
To complete the proof in this section the reader may consult Miller 'Symmetry Groups and their Application' Chapter 9.6 and R.H.Good, Rev.Mod.Phys. 27, (1955), page 187. The reader may also want to establish the unitary transformation which connects the Dirac matrices in the form (10.236) with the Dirac representation (10.219).

Exercise 7.2:
Demonstrate the anti-commutation relationships (10.218) of the Pauli matrices σ^j.
Exercise 7.3:
Demonstrate the anti-commutation relationships (10.218) of the Dirac matrices γ^μ.
Exercise 7.4:
Show that from (10.214) follows that γ^0 has real eigenvalues ± 1 and can be represented by a hermitian matrix, and γ^j, $j = 1, 2, 3$ has imaginary eigenvalues $\pm i$ and can be represented by an anti-hermitian matrix.

10.8 Lorentz Invariance of the Dirac Equation

We want to show now that the Dirac equation is invariant under Lorentz transformations, i.e., the form of the Dirac equation is identical in equivalent frames of reference, i.e., in frames connected by Lorentz transformations. The latter transformations imply that coordinates transform according to (10.6), i.e., $x' = L^\mu{}_\nu x^\nu$, and derivatives according to (10.80). Multiplication and summation of (10.80) by $L^\mu{}_\rho$ and using (10.76) yields $\partial_\rho = L^\nu{}_\rho \partial'_\nu$, a result one could have also obtained by applying the chain rule to (10.6). We can, therefore, transform coordinates and derivatives of the Dirac equation. However, we do not know yet how to transform the 4-dimensional wave function Ψ and the Dirac matrices γ^μ.

Lorentz Transformation of the Bispinor State

Actually, we will approach the proof of the Lorentz invariance of the Dirac equation by testing if there exists a transformation of the bispinor wave function Ψ and of the Dirac matrices γ^μ which together with the transformations of coordinates and derivatives leaves the form of the Dirac equation invariant, i.e., in a moving frame should hold

$$\left(i\gamma'^\mu \partial'_\mu \,-\, m \right) \Psi'(x'^\mu) \,=0\,. \tag{10.239}$$

Form invariance implies that the matrices γ'^μ should have the same properties as γ^μ, namely, (10.213, 10.215). Except for a similarity transformation, these properties determine the matrices γ'^μ uniquely, i.e., it must hold $\gamma'^\mu = \gamma^\mu$. Hence, in a moving frame holds

$$\left(i\gamma^\mu \partial'_\mu \,-\, m \right) \Psi'(x'^\mu) \,=0\,. \tag{10.240}$$

Infinitesimal Bispinor State Transformation

We want to show now that a suitable transformation of $\Psi(x^\mu)$ does, in fact, exist. The transformation is assumed to be linear and of the form

$$\Psi'(x'^\mu) \;=\; \mathcal{S}(L^\mu{}_\nu)\,\Psi(x^\mu) \quad ;\; x'^\mu \;=\; L^\mu{}_\nu x^\nu \tag{10.241}$$

where $\mathcal{S}(L^\mu{}_\nu)$ is a non-singular 4×4–matrix, the coefficients of which depend on the matrix $L^\mu{}_\nu$ defining the Lorentz transformation in such a way that $\mathcal{S}(L^\mu{}_\nu) \;=\; 1\!\!1$ for $L^\mu{}_\nu \;=\; \delta^\mu{}_\nu$ holds. Obviously, the transformation (10.241) implies a similarity transformation $\mathcal{S}\gamma^\mu\mathcal{S}^{-1}$. One can, hence, state that the Dirac equation (10.221) upon Lorentz transformation yields

$$\big(\, i\mathcal{S}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^{-1}(L^\eta{}_\xi)\,L^\nu{}_\mu\,\partial'_\nu \;-\; m\,\big)\;\Psi'(x'^\mu) \;=\; 0\,. \tag{10.242}$$

The form invariance of the Dirac equation under this transformation implies then the condition

$$\mathcal{S}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^{-1}(L^\eta{}_\xi)\,L^\nu{}_\mu \;=\; \gamma^\nu\,. \tag{10.243}$$

We want to determine now the 4×4–matrix $\mathcal{S}(L^\eta{}_\xi)$ which satisfies this condition.
The proper starting point for a constructiuon of $\mathcal{S}(L^\eta{}_\xi)$ is actually (10.243) in a form in which the Lorentz transformation in the form $L^{\mu\nu}$ is on the r.h.s. of the equation. For this purpose we exploit (10.12) in the form $L^\nu{}_\mu g_{\nu\sigma} L^\sigma{}_\rho \;=\; g_{\mu\rho} \;=\; g_{\rho\mu}$. Multiplication of (10.243) from the left by $L^\sigma{}_\rho g_{\sigma\nu}$ yields

$$\mathcal{S}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^{-1}(L^\eta{}_\xi)\,g_{\rho\mu} \;=\; L^\sigma{}_\rho g_{\sigma\nu}\gamma_\sigma\,. \tag{10.244}$$

from which, using $g_{\rho\mu}\gamma^\mu \;=\; \gamma_\rho$, follows

$$\mathcal{S}(L^\eta{}_\xi)\gamma_\rho\mathcal{S}^{-1}(L^\eta{}_\xi) \;=\; L^\sigma{}_\rho\gamma_\sigma\,. \tag{10.245}$$

One can finally conclude multiplying both sides by $g^{\rho\mu}$

$$\mathcal{S}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^{-1}(L^\eta{}_\xi) \;=\; L^{\nu\mu}\gamma_\nu\,. \tag{10.246}$$

The construction of $\mathcal{S}(L^\eta{}_\xi)$ will proceed using the avenue of infinitesimal transformations. We had introduced in (10.38) the infinitesimal Lorentz transformations in the form $L^\mu{}_\nu \;=\; \delta^\mu{}_\nu + \epsilon^\mu{}_\nu$ where the infinitesimal operator $\epsilon^\mu{}_\nu$ obeyed $\epsilon^T \;=\; -\,\mathbf{g}\epsilon\mathbf{g}$. Multiplication of this property by $\mathbf{g}$ from the right yields $(\epsilon\mathbf{g})^T \;=\; -\,\epsilon\mathbf{g}$, i.e., $\epsilon\mathbf{g}$ is an anti-symmetric matrix. The elements of $\epsilon\mathbf{g}$ are, however, $\epsilon^\mu{}_\rho g^{\rho\nu} \;=\; \epsilon^{\mu\nu}$ and, hence, in the expression of the infinitesimal transformation

$$L^{\mu\nu} \;=\; g^{\mu\nu} \,+\, \epsilon^{\mu\nu} \tag{10.247}$$

the infinitesimal matrix $\epsilon^{\mu\nu}$ is anti-symmetric.
The infinitesimal transformation $\mathcal{S}(L^\rho{}_\sigma)$ which corresponds to (10.247) can be expanded

$$\mathcal{S}(\epsilon^{\mu\nu}) \;=\; 1\!\!1 - \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu} \tag{10.248}$$

Here $\sigma_{\mu\nu}$ denote 4×4–matrices operating in the 4-dimensional space of the wave functions Ψ. $\mathcal{S}(\epsilon^{\mu\nu})$ should not change its value if one replaces in its argument $\epsilon^{\mu\nu}$ by $-\epsilon^{\nu\mu}$. It holds then

$$\mathcal{S}(\epsilon^{\mu\nu}) \;=\; 1\!\!1 - \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu} \;=\; \mathcal{S}(-\epsilon^{\nu\mu}) \;=\; 1\!\!1 + \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\nu\mu} \tag{10.249}$$

from which we can conclude $\sigma_{\mu\nu}\epsilon^{\mu\nu} = -\sigma_{\mu\nu}\epsilon^{\nu\mu} = -\sigma_{\nu\mu}\epsilon^{\mu\nu}$, i.e., it must hold

$$\sigma_{\mu\nu} = -\sigma_{\nu\mu} \,. \tag{10.250}$$

One can readily show expanding $\mathcal{S}\mathcal{S}^{-1} = \mathbb{1}$ to first order in $\epsilon^{\mu\nu}$ that for the inverse infinitesimal transformation holds

$$\mathcal{S}^{-1}(\epsilon^{\mu\nu}) = \mathbb{1} + \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu} \tag{10.251}$$

Inserting (10.248, 10.251) into (10.246) results then in a condition for the generators $\sigma_{\mu\nu}$

$$-\frac{i}{4}\left(\sigma_{\alpha\beta}\gamma^{\mu} - \gamma^{\mu}\sigma_{\alpha\beta}\right)\epsilon^{\alpha\beta} = \epsilon^{\nu\mu}\gamma_{\nu} \,. \tag{10.252}$$

Since six of the coefficients $\epsilon^{\alpha\beta}$ can be chosen independently, this condition can actually be expressed through six independent conditions. For this purpose one needs to express formally the r.h.s. of (10.252) also as a sum over both indices of $\epsilon^{\alpha\beta}$. Furthermore, the expression on the r.h.s., like the expression on the l.h.s., must be symmetric with respect to interchange of the indices α and β. For this purpose we express

$$\begin{aligned}\epsilon^{\nu\mu}\gamma_{\nu} &= \frac{1}{2}\epsilon^{\alpha\mu}\gamma_{\alpha} + \frac{1}{2}\epsilon^{\beta\mu}\gamma_{\beta} = \frac{1}{2}\epsilon^{\alpha\beta}\,\delta^{\mu}{}_{\beta}\gamma_{\alpha} + \frac{1}{2}\epsilon^{\alpha\beta}\,\delta^{\mu}{}_{\alpha}\gamma_{\beta} \\ &= \frac{1}{2}\epsilon^{\alpha\beta}\left(\delta^{\mu}{}_{\beta}\gamma_{\alpha} - \delta^{\mu}{}_{\alpha}\gamma_{\beta}\right) \,. \end{aligned}\tag{10.253}$$

Comparing this with the l.h.s. of (10.252) results in the condition for each α, β

$$[\,\sigma_{\alpha\beta}, \gamma^{\mu}\,]_{-} = 2i\left(\delta^{\mu}{}_{\beta}\gamma_{\alpha} - \delta^{\mu}{}_{\alpha}\gamma_{\beta}\right) \,. \tag{10.254}$$

The proper $\sigma_{\alpha\beta}$ must be anti-symmetric in the indices α, β and operate in the same space as the Dirac matrices. In fact, a solution of condition (10.254) is

$$\sigma_{\alpha\beta} = \frac{i}{2}\,[\,\gamma_{\alpha}, \gamma_{\beta}\,]_{-} \tag{10.255}$$

which can be demonstrated using the properties (10.213, 10.216) of the Dirac matrices.

Exercise 7.5:
Show that the $\sigma_{\alpha\beta}$ defined through (10.255) satisfy condition (10.254).

Algebra of Generators of Bispinor Transformation

We want to construct the bispinor Lorentz transformation by exponentiating the generators $\sigma_{\mu\nu}$. For this purpose we need to verify that the algebra of the generators involving addition and multiplication is closed. For this purpose we inspect the properties of the generators in a particular representation, namely, the chiral representation introduced above in Eqs. (10.228, 10.229). In this representation the Dirac matrices $\tilde{\gamma}_{\mu} = (\tilde{\gamma}^0, -\vec{\tilde{\gamma}})$ are

$$\tilde{\gamma}_0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad ; \quad \tilde{\gamma}_j = \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} \,,\; j = 1, 2, 3 \,. \tag{10.256}$$

One can readily verify that the non-vanishing generators $\tilde{\sigma}_{\mu\nu}$ are given by (note $\tilde{\sigma}_{\mu\nu} = -\tilde{\sigma}_{\nu\mu}$, i.e. only six generators need to be determined)

$$\tilde{\sigma}_{0j} = \frac{i}{2}[\tilde{\gamma}_0, \tilde{\gamma}_j] = \begin{pmatrix} -i\sigma^j & 0 \\ 0 & i\sigma^j \end{pmatrix} \; ; \; \tilde{\sigma}_{jk} = [\tilde{\gamma}_j, \tilde{\gamma}_k] = \epsilon_{jk\ell} \begin{pmatrix} \sigma^\ell & 0 \\ 0 & \sigma^\ell \end{pmatrix} . \tag{10.257}$$

Obviously, the algebra of these generators is closed under addition and multiplication, since both operations convert block-diagonal operators

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \tag{10.258}$$

again into block-diagonal operators, and since the algebra of the Pauli matrices is closed.
We can finally note that the closedness of the algebra of the generators $\sigma_{\mu\nu}$ is not affected by similarity transformations and that, therefore, any representation of the generators, in particular, the representation (10.255) yields a closed algebra.

Finite Bispinor Transformation

The closedness of the algebra of the generators $\sigma_{\mu\nu}$ defined through (10.248) allows us to write the transformation $\mathcal{S}$ for any, i.e., not necessarily infinitesimal, $\epsilon^{\mu\nu}$ in the exponential form

$$\mathcal{S} = \exp\left(-\frac{i}{4}\, \sigma_{\mu\nu}\epsilon^{\mu\nu} \right) . \tag{10.259}$$

We had stated before that the transformation $\mathcal{S}$ is actually determined through the Lorentz transformation $L^\mu{}_\nu$. One should, therefore, be able to state $\mathcal{S}$ in terms of the same parameters $\vec{w}$ and $\vec{\vartheta}$ as the Lorentz transformation in (10.51). In fact, one can express the tensor $\epsilon^{\mu\nu}$ through $\vec{w}$ and $\vec{\vartheta}$ using $\epsilon^{\mu\nu} = \epsilon^\mu{}_\rho g^{\rho\nu}$ and the expression (10.44)

$$\epsilon^{\mu\nu} = \begin{pmatrix} 0 & -w_1 & -w_2 & -w_3 \\ w_1 & 0 & \vartheta_3 & -\vartheta_2 \\ w_2 & -\vartheta_3 & 0 & \vartheta_1 \\ w_3 & \vartheta_2 & -\vartheta_1 & 0 \end{pmatrix} \tag{10.260}$$

Inserting this into (10.259) yields the desired connection between the Lorentz transformation (10.51) and $\mathcal{S}$.
In order to construct an explicit expression of $\mathcal{S}$ in terms of $\vec{w}$ and $\vec{\vartheta}$ we employ again the chiral representation. In this representation holds

$$\begin{aligned} -\tfrac{i}{4}\tilde{\sigma}_{\mu\nu}\epsilon^{\mu\nu} &= -\frac{i}{2}\left(\tilde{\sigma}_{01}\epsilon^{01} + \tilde{\sigma}_{02}\epsilon^{02} + \tilde{\sigma}_{03}\epsilon^{03} + \tilde{\sigma}_{12}\epsilon^{12} + \tilde{\sigma}_{13}\epsilon^{13} + \tilde{\sigma}_{23}\epsilon^{23}\right) \\ &= \frac{1}{2}\begin{pmatrix} (\vec{w} - i\vec{\vartheta})\cdot\vec{\sigma} & 0 \\ 0 & -(\vec{w} + i\vec{\vartheta})\cdot\vec{\sigma} \end{pmatrix} . \end{aligned} \tag{10.261}$$

We note that this operator is block-diagonal. Since such operator does not change its block-diagonal form upon exponentiation the bispinor transformation (10.259) becomes in the chiral representation

$$\tilde{\mathcal{S}}(\vec{w}, \vec{\vartheta}) = \begin{pmatrix} e^{\frac{1}{2}(\vec{w} - i\vec{\vartheta})\cdot\vec{\sigma}} & 0 \\ 0 & e^{-\frac{1}{2}(\vec{w} + i\vec{\vartheta})\cdot\vec{\sigma}} \end{pmatrix} \tag{10.262}$$

This expression allows one to transform according to (10.241) bispinor wave functions from one frame of reference into another frame of reference.

Current 4-Vector Associated with Dirac Equation

We like to derive now an expression for the current 4-vector j^μ associated with the Dirac equation which satisfies the conservation law

$$\partial_\mu j^\mu \;=\; 0\,. \tag{10.263}$$

Starting point are the Dirac equation in the form (10.221) and the adjoint Dirac equation (10.224). Multiplying (10.221) from the left by $\Psi^\dagger(x^\mu)\gamma^0$, (10.224) from the right by $\Psi(x^\mu)$, and addition yields

$$\Psi^\dagger(x^\mu)\gamma^0\left(i\gamma^\mu\partial_\mu \;+\; i\gamma^\mu\overleftarrow{\partial_\mu}\right)\Psi(x^\mu) \;=\; 0\,. \tag{10.264}$$

The last result can be written

$$\partial_\mu\,\Psi^\dagger(x^\nu)\gamma^0\gamma^\mu\Psi(x^\nu) \;=\; 0\,, \tag{10.265}$$

i.e., the conservation law (10.263) does hold, in fact, for

$$j^\mu(x^\mu) \;=\; (\rho,\,\vec{j}) \;=\; \Psi^\dagger(x^\mu)\gamma^0\gamma^\mu\Psi(x^\mu)\,. \tag{10.266}$$

The time-like component ρ of j^μ

$$\rho(x^\mu) \;=\; \Psi^\dagger(x^\mu)\Psi(x^\mu) \;=\; \sum_{s=1}^{4}|\psi_s(x^\mu)|^2 \tag{10.267}$$

has the desired property of being positive definite.
The conservation law (10.263) allows one to conclude that j^μ must transform like a contravariant 4-vector as the notation implies. The reason is that the r.h.s. of (10.263) obviously is a scalar under Lorentz transformations and that the left hand side must then also transform like a scalar. Since ∂_μ transforms like a covariant 4-vector, j^μ must transform like a contravariant 4-vector. This transformation behaviour can also be deduced from the transformation properties of the bispinor wave function $\Psi(x_\mu)$. For this purpose we prove first the relationship

$$\mathcal{S}^{-1} \;=\; \gamma^0\,\mathcal{S}^\dagger\gamma^0\,. \tag{10.268}$$

We will prove this property in the chiral representation. Obviously, the property applies then in any representation of $\mathcal{S}$.
For our proof we note first

$$\tilde{\mathcal{S}}^{-1}(\vec{w},\vec{\vartheta}) \;=\; \tilde{\mathcal{S}}(-\vec{w},-\vec{\vartheta}) \;=\; \begin{pmatrix} e^{-\frac{1}{2}(\vec{w}\,-\,i\vec{\vartheta})\cdot\vec{\sigma}} & 0 \\ 0 & e^{\frac{1}{2}(\vec{w}\,+\,i\vec{\vartheta})\cdot\vec{\sigma}} \end{pmatrix} \tag{10.269}$$

One can readily show that the same operator is obtained evaluating

$$\tilde{\gamma}^0\tilde{\mathcal{S}}^\dagger(\vec{w},\vec{\vartheta})\tilde{\gamma}^0 \;=\; \begin{pmatrix} 0 & 1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}\begin{pmatrix} e^{\frac{1}{2}(\vec{w}\,+\,i\vec{\vartheta})\cdot\vec{\sigma}} & 0 \\ 0 & e^{-\frac{1}{2}(\vec{w}\,-\,i\vec{\vartheta})\cdot\vec{\sigma}} \end{pmatrix}\begin{pmatrix} 0 & 1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}\,. \tag{10.270}$$

We conclude that (10.268) holds for the bispinor Lorentz transformation.
We will now determine the relationship between the flux

$$j'^\mu \;=\; \Psi'^\dagger(x'^\mu)\gamma^0\gamma^\mu\Psi'(x') \tag{10.271}$$

in a moving frame of reference and the flux j^μ in a frame at rest. Note that we have assumed in (10.271) that the Dirac matrices are independent of the frame of reference. One obtains using (10.268)

$$j'^\mu \;=\; \Psi^\dagger(x^\mu)\mathcal{S}^\dagger\gamma^0\gamma^\mu\mathcal{S}\Psi(x^\mu) \;=\; \Psi^\dagger(x^\mu)\gamma^0\mathcal{S}^{-1}\gamma^\mu\mathcal{S}\Psi(x^\mu)\;. \tag{10.272}$$

With $\mathcal{S}^{-1}(L^\eta{}_\xi) \;=\; \mathcal{S}((L^{-1})^\eta{}_\xi)$ one can restate (10.246)

$$\mathcal{S}^{-1}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^(L^\eta{}_\xi) \;=\; (L^{-1})^{\nu\mu}\gamma_\nu \;=\; (L^{-1})_\nu{}^\mu\gamma^\nu \;=\; L^\mu{}_\nu\gamma_\nu. \tag{10.273}$$

where we have employed (10.76). Combining this with (10.272) results in the expected transformation behaviour

$$j'^\mu \;=\; L^\mu{}_\nu j^\nu\;. \tag{10.274}$$

10.9 Solutions of the Free Particle Dirac Equation

We want to determine now the wave functions of free particles described by the Dirac equation. Like in non-relativistic quantum mechanics the free particle wave function plays a central role, not only as the most simple demonstration of the theory, but also as providing a basis in which the wave functions of interacting particle systems can be expanded and characterized. The solutions provide also a complete, orthonormal basis and allows one to quantize the Dirac field $\Psi(x^\mu)$ much like the classical electromagnetic field is quantized through creation and annihilation operators representing free electromagnetic waves of fixed momentum and frequency.

In case of non-relativistic quantum mechanics the free particle wave function has a single component $\psi(\vec{r},t)$ and is determined through the momentum $\vec{p} \in \mathbb{R}^3$. In relativistic quantum mechanics a Dirac particle can also be characterized through a momentum, however, the wave function has four components which invite further characterization of the free particle state. In the following we want to provide this characterization, specific for the Dirac free particle.

We will start from the Dirac equation in the Schrödinger form (10.231, 10.232, 10.233)

$$i\partial_t\Psi(x^\mu) \;=\; \mathcal{H}_o\Psi(x^\mu)\;. \tag{10.275}$$

The free particle wave function is an eigenfunction of $\mathcal{H}_o$, a property which leads to the energy–momentum (dispersion) relationship of the Dirac particle. The additional degrees of freedom described by the four components of the bispinor wave function require, as just mentioned, additional characterizations, i.e., the identification of observables and their quantum mechanical operators, of which the wave functions are eigenfunctions as well. As it turns out, only two degrees of freedom of the bispinor four degrees of freedom are independent [c.f. (10.282, 10.283)]. The independent degrees of freedom allow one to choose the states of the free Dirac particle as eigenstates of the 4-momentum operator $\hat{p}_\mu$ and of the helicity operator $\Gamma \sim \vec{\sigma}\cdot\hat{\vec{p}}/|\hat{\vec{p}}|$ introduced below. These operators, as is required for the mentioned property, commute with each other. The operators commute also with $\mathcal{H}_o$ in (10.233).

Like for the free particle wave functions of the non-relativistic Schrödinger and the Klein–Gordon equations one expects that the space–time dependence is governed by a factor $\exp[i(\vec{p}\cdot\vec{r} \,-\, \epsilon t)]$. As pointed out, the Dirac particles are described by 4-dimensional, bispinor wave functions and we need to determine corresponding components of the wave function. For this purpose we consider the following form of the free Dirac particle wave function

$$\Psi(x^\mu) \;=\; \begin{pmatrix} \phi(x^\mu) \\ \chi(x^\mu) \end{pmatrix} \;=\; \begin{pmatrix} \phi_o \\ \chi_o \end{pmatrix} e^{i(\vec{p}\cdot\vec{r} - \epsilon t)} \tag{10.276}$$

where $\vec{p}$ and ϵ together represent four real constants, later to be identified with momentum and energy, and ϕ_o, χ_o each represent a constant, two-dimensional spinor state. Inserting (10.276) into (10.231, 10.232) leads to the 4-dimensional eigenvalue problem

$$\begin{pmatrix} m & \vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -m \end{pmatrix} \begin{pmatrix} \phi_o \\ \chi_o \end{pmatrix} \;=\; \epsilon \begin{pmatrix} \phi_o \\ \chi_o \end{pmatrix} . \tag{10.277}$$

To solve this problem we write (10.277) explicitly

$$\begin{aligned} (\epsilon \,-\, m)\, 1\!\!1\, \phi_o \;-\; \vec{\sigma}\cdot\vec{p}\,\chi_o \quad &= \; 0 \\ -\,\vec{\sigma}\cdot\vec{p}\,\phi_o \;+\; (\epsilon \,+\, m)\, 1\!\!1\, \chi_o \quad &= \; 0 \,. \end{aligned} \tag{10.278}$$

Multiplication of the 1st equation by $(\epsilon + m)1\!\!1$ and of the second equation by $-\vec{\sigma}\cdot\vec{p}$ and subtraction of the results yields the 2-dimensional equation

$$\left[\, (\epsilon^2 \,-\, m^2)\, 1\!\!1 \;-\; (\vec{\sigma}\cdot\vec{p})^2 \,\right] \phi_o \;=\; 0 \,. \tag{10.279}$$

According to the property (5.234) of Pauli matrices holds $(\vec{\sigma}\cdot\vec{p})^2 \;=\; \vec{p}^{\,2} 1\!\!1$. One can, hence, conclude from (10.279) the well-known relativistic dispersion relationship

$$\epsilon^2 \;=\; m^2 \;+\; \vec{p}^{\,2} \tag{10.280}$$

which has a positive and a negative solution

$$\epsilon \;=\; \pm\, E(\vec{p}) \,, \quad E(\vec{p}) \;=\; \sqrt{m^2 \;+\; \vec{p}^{\,2}} \,. \tag{10.281}$$

Obviously, the Dirac equation, like the Klein–Gordon equation, reproduce the classical relativistic energy–momentum relationships (10.93, 10.94)
Equation (10.278) provides us with information about the components of the bispinor wave function (10.276), namely ϕ_o and χ_o are related as follows

$$\phi_0 \;=\; \frac{\vec{\sigma}\cdot\vec{p}}{\epsilon - m}\,\chi_0 \tag{10.282}$$

$$\chi_o \;=\; \frac{\vec{\sigma}\cdot\vec{p}}{\epsilon \,+\, m}\,\phi_0 \,, \tag{10.283}$$

where ϵ is defined in (10.281). These two relationships are consistent with each other. In fact, one finds using (5.234) and (10.280)

$$\frac{\vec{\sigma}\cdot\vec{p}}{\epsilon + m}\,\phi_0 \;=\; \frac{(\vec{\sigma}\cdot\vec{p})^2}{(\epsilon + m)\,(\epsilon + m)}\,\chi_0 \;=\; \frac{\vec{p}^{\,2}}{\epsilon^2 - m^2}\,\chi_0 \;=\; \chi_0 \,. \tag{10.284}$$

The relationships (10.282, 10.283) imply that the bispinor part of the wave function allows only two degrees of freedom to be chosen independently. We want to show now that these degrees of freedom correspond to a spin-like property, the socalled helicity of the particle.

For our further characterization we will deal with the positive and negative energy solutions [cf. (10.281)] separately. For the positive energy solution, i.e., the solution for $\epsilon = +E(\vec{p})$, we present ϕ_o through the normalized vector

$$\phi_o = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \mathbf{u} \in \mathbb{C}^2 , \; \mathbf{u}^\dagger \mathbf{u} = |u_1|^2 + |u_2|^2 = 1 . \tag{10.285}$$

The corresponding free Dirac particle is then described through the wave function

$$\Psi(\vec{p},+|x^\mu) = \mathcal{N}_+(\vec{p}) \begin{pmatrix} \mathbf{u} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u} \end{pmatrix} e^{i(\vec{p}\cdot\vec{r} - \epsilon t)} , \quad \epsilon = +E(\vec{p}) . \tag{10.286}$$

Here $\mathcal{N}_+(\vec{p})$ is a constant which will be chosen to satisfy the normalization condition

$$\Psi^\dagger(\vec{p},+)\,\gamma^0\,\Psi^\dagger(\vec{p},+) = 1 , \tag{10.287}$$

the form of which will be justified further below. Similarly, we present the negative energy solution, i.e., the solution for $\epsilon = -E(\vec{p})$, through χ_o given by

$$\chi_o = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \mathbf{u} \in \mathbb{C}^2 , \; \mathbf{u}^\dagger \mathbf{u} = |u_1|^2 + |u_2|^2 = 1 . \tag{10.288}$$

corresponding to the wave function

$$\Psi(\vec{p},-|x^\mu) = \mathcal{N}_-(\vec{p}) \begin{pmatrix} \frac{-\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u} \\ \mathbf{u} \end{pmatrix} e^{i(\vec{p}\cdot\vec{r} - \epsilon t)} , \quad \epsilon = -E(\vec{p}) . \tag{10.289}$$

Here $\mathcal{N}_-(\vec{p})$ is a constant which will be chosen to satisfy the normalization condition

$$\Psi^\dagger(\vec{p},+)\,\gamma^0\,\Psi^\dagger(\vec{p},+) = -1 , \tag{10.290}$$

which differs from the normalization condition (10.287) in the minus sign on the r.h.s. The form of this condition and of (10.287) will be justified now.
First, we demonstrate that the product $\Psi^\dagger(\vec{p},\pm)\gamma^0\Psi(\vec{p},\pm)$, i.e., the l.h.s. of (10.287, 10.290), is invariant under Lorentz transformations. One can see this as follows: Let $\Psi(\vec{p},\pm)$ denote the solution of a free particle moving with momentum $\vec{p}$ in the laboratory frame, and let $\Psi(0,\pm)$ denote the corresponding solution of a particle in its rest frame. The connection between the solutions, according to (10.241), is $\Psi(\vec{p},\pm) = S\,\Psi(0,\pm)$, where S is given by (10.259). Hence,

$$\begin{aligned} \Psi^\dagger(\vec{p},\pm)\,\gamma^0\,\Psi(\vec{p},\pm) &= \left(\Psi^\dagger(0,\pm)\,S^\dagger\right)\gamma^0\,S\,\Psi(0,\pm) \\ &= \Psi^\dagger(0,\pm)\,\left(\gamma^0\,S^{-1}\,\gamma^0\right)\gamma^0\,S\,\Psi(0,\pm) \\ &= \Psi^\dagger(0,\pm)\,\gamma^0\,\Psi(0,\pm). \end{aligned} \tag{10.291}$$

Note that we have used that, according to (10.268), $S^{-1} = \gamma^0 S^\dagger \gamma^0$ and, hence, $S^\dagger = \gamma^0 S^{-1} \gamma^0$.
We want to demonstrate now that the signs on the r.h.s. of (10.287, 10.290) should differ. For this purpose we consider first the positive energy solution. Employing (10.286) for $\vec{p} = 0$ yields, using γ^0 as given in (10.219) and $\mathbf{u}^\dagger\mathbf{u} = 1$ [c.f. (10.285)],

$$\Psi^\dagger(0,+)\,\gamma^0\,\Psi(0,+) = |\mathcal{N}_+(0)|^2\,(\mathbf{u}^\dagger, \quad 0)\,\gamma^0 \begin{pmatrix} \mathbf{u} \\ 0 \end{pmatrix} = |\mathcal{N}_+(0)|^2 . \tag{10.292}$$

The same calculation for the negative energy wave function as given in (10.289) yields

$$\Psi^\dagger(0,-)\,\gamma^0\,\Psi(0,-) \;=\; |\mathcal{N}_-(0)|^2\,(0, \quad \mathbf{u}^\dagger)\,\gamma^0 \begin{pmatrix} 0 \\ \mathbf{u} \end{pmatrix} \;=\; -\,|\mathcal{N}_-(0)|^2\;. \tag{10.293}$$

Obviously, this requires the choice of a negative side on the r.h.s. of (10.290) to assign a positive value to $|\mathcal{N}_-(0)|^2$. We can also conclude from our derivation

$$\mathcal{N}_\pm(0) \;=\; 1\;. \tag{10.294}$$

We want to determine now $\mathcal{N}_\pm(\vec{p})$ for arbitrary $\vec{p}$. We consider first the positive energy solution. Condition (10.287) written explicitly using (10.286) is

$$\mathcal{N}_+^2(\vec{p}) \left((\mathbf{u}^*)^T, \; \left[\frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})\,+\,m}\,\mathbf{u}^* \right]^T \right) \gamma^o \begin{pmatrix} \mathbf{u} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u} \end{pmatrix} \;=\; 1 \tag{10.295}$$

Evaluating the l.h.s. using γ^0 as given in (10.219) yields

$$\mathcal{N}_+^2(\vec{p}) \left[\mathbf{u}^\dagger\mathbf{u} \;-\; \mathbf{u}^\dagger\,\frac{(\vec{\sigma}\cdot\vec{p})^2}{(E(\vec{p})\,+\,m)^2}\,\mathbf{u} \right] \;=\; 1\;. \tag{10.296}$$

Replacing $(\vec{\sigma}\cdot\vec{p})^2$ by $\vec{p}^{\,2}$ [c.f. (5.234)] and using the normalization of $\mathbf{u}$ in (10.285) results in

$$\mathcal{N}_+^2(\vec{p}) \left[1 \;-\; \frac{\vec{p}^{\,2}}{(E(\vec{p})\,+\,m)^2} \right] \;=\; 1 \tag{10.297}$$

from which follows

$$\mathcal{N}_+(\vec{p}) \;=\; \sqrt{\frac{(\,m\,+\,E(\vec{p})\,)^2}{(\,m\,+\,E(\vec{p})\,)^2\;-\;\vec{p}^{\,2}}}\;. \tag{10.298}$$

Noting

$$(\,m\,+\,E(\vec{p})\,)^2\;-\;\vec{p}^{\,2} \;=\; m^2\;-\;\vec{p}^{\,2}\;+\;2mE(\vec{p})\;+\;E^2(\vec{p}) \;=\; 2(m\,+\,E(\vec{p}))\,m \tag{10.299}$$

the normalization coefficient (10.298) becomes

$$\mathcal{N}_+(\vec{p}) \;=\; \sqrt{\frac{m\,+\,E(\vec{p})}{2m}}\;. \tag{10.300}$$

This result completes the expression for the wave function (10.286).

Exercise 7.6: Show that the normalization condition

$$\mathcal{N}_+'^2(\vec{p}) \left((\mathbf{u}^*)^T, \; \left[\frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})\,+\,m}\,\mathbf{u}^* \right]^T \right) \begin{pmatrix} \mathbf{u} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u} \end{pmatrix} \;=\; 1 \tag{10.301}$$

yields the normalization coefficient

$$\mathcal{N}_+'(\vec{p}) \;=\; \sqrt{\frac{m\,+\,E(\vec{p})}{2\,E(\vec{p})}}\;. \tag{10.302}$$

We consider now the negative energy solution. Condition (10.290) written explicitly using (10.289) is

$$\mathcal{N}_-^2(\vec{p}) \left(\left[\frac{-\vec{\sigma}\cdot\vec{p}}{E(\vec{p})\,+\,m}\,\mathbf{u}^* \right]^T ,\; (\mathbf{u}^*)^T \right) \gamma^o \left(\begin{array}{c} \frac{-\vec{\sigma}\cdot\vec{p}}{E(\vec{p})\,+\,m}\,\mathbf{u} \\ \mathbf{u} \end{array} \right) \;=\; -1 \tag{10.303}$$

Evaluating the l.h.s. yields

$$\mathcal{N}_-^2(\vec{p}) \left[\mathbf{u}^\dagger \, \frac{(\vec{\sigma}\cdot\vec{p})^2}{(E(\vec{p})\,+\,m)^2}\,\mathbf{u} \;-\; \mathbf{u}^\dagger\mathbf{u} \right] \;=\; -1\;. \tag{10.304}$$

This condition is, however, identical to the condition (10.296) for the normalization constant $\mathcal{N}_+(\vec{p})$ of the positive energy solution. We can, hence, conclude

$$\mathcal{N}_-(\vec{p}) \;=\; \sqrt{\frac{m\,+\,E(\vec{p})}{2m}} \tag{10.305}$$

and, thereby, have completed the determination for the wave function (10.289).
The wave functions (10.286, 10.289, 10.300) have been constructed to satisfy the free particle Dirac equation (10.275). Inserting (10.286) into (10.275) yields

$$\mathcal{H}_o\,\Psi(\vec{p},\lambda|x^\mu) \;=\; \lambda\,E(\vec{p})\,\Psi(\vec{p},\lambda|x^\mu)\;, \tag{10.306}$$

i.e., the wave functions constructed represent eigenstates of $\mathcal{H}_o$. The wave functions are also eigenstates of the momentum operator $i\partial_\mu$, i.e.,

$$i\partial_\mu\Psi(\vec{p},\lambda|x^\mu) \;=\; p_\mu\,\Psi(\vec{p},\lambda|x^\mu) \tag{10.307}$$

where $p_\mu \;=\; (\epsilon, -\vec{p})$. This can be verified expressing the space–time factor of $\Psi(\vec{p},\lambda|x^\mu)$ in 4-vector notation, i.e., $\exp[i(\vec{p}\cdot\vec{r}\,-\,\epsilon t)] \;=\; \exp(ip_\mu x^\mu)$.

Helicity

The free Dirac particle wave functions (10.286, 10.289) are not completely specified, the two components of $\mathbf{u}$ indicate another degree of freedom which needs to be defined. This degree of freedom describes a spin–$\frac{1}{2}$ attribute. This attribute is the so-called helicity, defined as the component of the particle spin along the direction of motion. The corresponding operator which measures this observable is

$$\Lambda \;=\; \frac{1}{2}\,\sigma\cdot\frac{\vec{\hat{p}}}{|\vec{\hat{p}}|}\;. \tag{10.308}$$

Note that $\vec{\hat{p}}$ represents here an operator, not a constant vector. Rather than considering the observable (10.307) we investigate first the observable due to the simpler operator $\vec{\sigma}\cdot\vec{\hat{p}}$. We want to show that this operator commutes with $\mathcal{H}_o$ and $\vec{\hat{p}}$ to ascertain that the free particle wave function can be simultaneously an eigenvector of all three operators. The commutation property $[\vec{\sigma}\cdot\vec{\hat{p}}, \hat{p}_j] \;=\; 0\,, j = 1,2,3$ is fairly obvious. The property $[\vec{\sigma}\cdot\vec{\hat{p}}, \mathcal{H}_o] \;=\; 0$ follows from (10.233) and from the two identities

$$\left(\begin{array}{cc} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{array} \right) \left(\begin{array}{cc} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{array} \right) \cdot \vec{\hat{p}} \;-\; \left(\begin{array}{cc} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{array} \right) \cdot \vec{\hat{p}} \left(\begin{array}{cc} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{array} \right) \;=\; 0 \tag{10.309}$$

and

$$\begin{aligned} & \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \cdot \vec{p} \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{p} - \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{p} \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \cdot \vec{p} \\ & = \begin{pmatrix} 0 & (\vec{\sigma}\cdot\vec{p})^2 \\ (\vec{\sigma}\cdot\vec{p})^2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & (\vec{\sigma}\cdot\vec{p})^2 \\ (\vec{\sigma}\cdot\vec{p})^2 & 0 \end{pmatrix} = 0\,. \end{aligned} \tag{10.310}$$

We have shown altogether that the operators $\vec{p}$, $\mathcal{H}_o$ and $\vec{\sigma}\cdot\vec{p}$ commute with each other and, hence, can be simultaneously diagonal. States which are simultaneously eigenvectors of these three operators are also simulteneously eigenvectors of the three operators $\vec{p}$, $\mathcal{H}_o$ and Λ defined in (10.308) above. The condition that the wave functions (10.286) are eigenfunctions of Λ as well will specify now the vectors **u**.

Since helicity is defined relative to the direction of motion of a particle the characterization of **u** as an eigenvector of the helicity operator, in principle, is independent of the direction of motion of the particle. We consider first the simplest case that particles move along the x_3–direction, i.e., $\vec{p} = (0,0,p_3)$. In this case $\Lambda = \frac{1}{2}\sigma^3$.

We assume first particles with positive energy, i.e., $\epsilon = +E(\vec{p})$. According to the definition (5.224) of σ_3 the two **u** vectors $(1,0)^T$ and $(0,1)^T$ are eigenstates of $\frac{1}{2}\sigma^3$ with eigenvalues $\pm\frac{1}{2}$. Therefore, the wave functions which are eigenstates of the helicity operator, are

$$\begin{aligned} \Psi(p\hat{e}_3,+,+\tfrac{1}{2}|\vec{r},t) &= \mathcal{N}_p \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \frac{p}{m+E_p}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} e^{i(px^3 - E_p t)} \\ \Psi(p\hat{e}_3,+,-\tfrac{1}{2}|\vec{r},t) &= \mathcal{N}_p \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{-p}{m+E_p}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} e^{i(px^3 - E_p t)} \end{aligned} \tag{10.311}$$

where $\hat{e}_3$ denotes the unit vector in the x_3-direction and where

$$E_p = \sqrt{m^2 + p^2}\,; \quad \mathcal{N}_p = \sqrt{\frac{m + E_p}{2m}}\,. \tag{10.312}$$

We assume now particles with negative energy, i.e., $\epsilon = -E(\vec{p})$. The wave functions which are eigenfunctions of the helicity operator are in this case

$$\begin{aligned} \Psi(p\hat{e}_3,-,+\tfrac{1}{2}|\vec{r},t) &= \mathcal{N}_p \begin{pmatrix} \frac{-p}{m+E_p}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)} \\ \Psi(p\hat{e}_3,-,-\tfrac{1}{2}|\vec{r},t) &= \mathcal{N}_p \begin{pmatrix} \frac{p}{m+E_p}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)} \end{aligned} \tag{10.313}$$

where E_p and $\mathcal{N}_p$ are defined in (10.312).

To obtain free particle wave functions for arbitrary directions of $\vec{p}$ one can employ the wave functions (10.311, 10.313) except that the states $(1,\, 0)^T$ and $(0,\, 1)^T$ have to be replaced by eigenstates $\mathbf{u}_\pm(\vec{p})$ of the spin operator along the direction of $\vec{p}$. These eigenstates are obtained through a rotational transformation (5.222, 5.223) as follows

$$\mathbf{u}(\vec{p}, +\frac{1}{2}) \;=\; \exp\left(-\frac{i}{2}\,\vec{\vartheta}(\vec{p})\cdot\vec{\sigma}\right)\begin{pmatrix}1\\0\end{pmatrix}\,, \tag{10.314}$$

$$\mathbf{u}(\vec{p}, -\frac{1}{2}) \;=\; \exp\left(-\frac{i}{2}\,\vec{\vartheta}(\vec{p})\cdot\vec{\sigma}\right)\begin{pmatrix}0\\1\end{pmatrix} \tag{10.315}$$

where

$$\vec{\vartheta}(\vec{p}) \;=\; \frac{\hat{e}_3\times\vec{p}}{|\vec{p}|}\,\angle(\hat{e}_3,\,\vec{p}) \tag{10.316}$$

describes a rotation which aligns the x_3–axis with the direction of $\vec{p}$. [One can also express the rotation through Euler angles α, β, γ , in which case the transformation is given by (5.220).] The corresponding free particle wave functions are then

$$\Psi(\vec{p}, +, +\tfrac{1}{2}|\vec{r}, t) \;=\; \mathcal{N}_p\begin{pmatrix}\mathbf{u}(\vec{p}, +\frac{1}{2})\\ \frac{p}{m+E_p}\mathbf{u}(\vec{p}, +\frac{1}{2})\end{pmatrix} e^{i(\vec{p}\cdot\vec{r}\,-\,E_p t)} \tag{10.317}$$

$$\Psi(\vec{p}, +, -\tfrac{1}{2}|\vec{r}, t) \;=\; \mathcal{N}_p\begin{pmatrix}\mathbf{u}(\vec{p}, -\frac{1}{2})\\ \frac{-p}{m+E_p}\mathbf{u}(\vec{p}, -\frac{1}{2})\end{pmatrix} e^{i(\vec{p}\cdot\vec{r}\,-\,E_p t)} \tag{10.318}$$

$$\Psi(\vec{p}, -, +\tfrac{1}{2}|\vec{r}, t) \;=\; \mathcal{N}_{\surd}\begin{pmatrix}\frac{-p}{m+E_p}\mathbf{u}(\vec{p}, +\frac{1}{2})\\ \mathbf{u}(\vec{p}, +\frac{1}{2})\end{pmatrix} e^{i(\vec{p}\cdot\vec{r}\,+\,E_p t)} \tag{10.319}$$

$$\Psi(\vec{p}, -, -\tfrac{1}{2}|\vec{r}, t) \;=\; \mathcal{N}_p\begin{pmatrix}\frac{p}{m+E_p}\mathbf{u}(\vec{p}, -\frac{1}{2})\\ \mathbf{u}(\vec{p}, -\frac{1}{2})\end{pmatrix} e^{i(\vec{p}\cdot\vec{r}\,+\,E_p t)} \tag{10.320}$$

where E_p and $\mathcal{N}$ are again given by (10.312).

Generating Solutions Through Lorentz Transformation

The solutions (10.311, 10.312) can be obtained also by means of the Lorentz transformation (10.262) for the bispinor wave function and the transformation (10.123). For this purpose one starts from the solutions of the Dirac equation in the chiral representation (10.226, 10.229), denoted by $\tilde{\ }$, for an $\vec{r}$–independent wave function, i.e., a wave function which represents free particles at rest. The corresponding wave functions are determined through

$$\left(i\tilde{\gamma}^0\partial_t \;-\; m\right)\,\tilde{\Psi}(t) \;=\; 0\,. \tag{10.321}$$

and are

$$\tilde{\Psi}(p=0,\,+,\,\,\tfrac{1}{2}|t) \;=\; \tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix} e^{-imt}\,,$$

$$\tilde{\Psi}(p=0,\,+,\,-\tfrac{1}{2}|t) \;=\; \tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\1\end{pmatrix} e^{-imt}\,,$$

$$\tilde{\Psi}(p=0,\,-,\,\tfrac{1}{2}|t) \;=\; \tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix} e^{+imt}\,,$$

$$\tilde{\Psi}(p=0,\,-,\,-\tfrac{1}{2}|t) \;=\; \tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\-1\end{pmatrix} e^{+imt}\,. \tag{10.322}$$

The reader can readily verify that transformation of these solutions to the Dirac representationsas defined in (10.228) yields the corresponding solutions (10.311, 10.313) in the $p \to 0$ limit. This correspondence justifies the characterization $\pm, \pm\frac{1}{2}$ of the wave functions stated in (10.322). The solutions (10.322) can be written in spinor form

$$\frac{1}{\sqrt{2}}\begin{pmatrix}\phi_o\\ \chi_o\end{pmatrix} e^{\mp imt}\,, \quad \phi_o\,,\,\chi_o \in \left\{ \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\} \tag{10.323}$$

Transformation (10.262) for a boost in the x_3–direction, i.e., for $\vec{w} = (0,0,w_3)$, yields for the exponential space–time dependence according to (10.174, 10.176)

$$\mp imt \;\to i\,(\,p_3 x^3 \;\mp Et\,) \tag{10.324}$$

and for the bispinor part according to (10.262)

$$\begin{pmatrix}\phi_o\\ \chi_o\end{pmatrix} \to \begin{pmatrix} e^{\frac{1}{2}w_3\sigma^3} & 0 \\ 0 & e^{-\frac{1}{2}w_3\sigma^3}\end{pmatrix}\begin{pmatrix}\phi_o\\ \chi_o\end{pmatrix} = \begin{pmatrix} e^{\frac{1}{2}w_3\sigma^3}\,\phi_o \\ e^{-\frac{1}{2}w_3\sigma^3}\,\chi_o\end{pmatrix}\,. \tag{10.325}$$

One should note that ϕ_o, χ_o are eigenstates of σ^3 with eigenvalues ± 1. Applying (10.324, 10.325) to (10.323) should yield the solutions for non-vanishing momentum p in the x_3–direction. For the resulting wave functions in the chiral representation one can use then a notation corresponding to that adopted in (10.311)

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\,+,\,+\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} e^{\frac{1}{2}w_3}\begin{pmatrix}1\\0\end{pmatrix} \\ e^{-\frac{1}{2}w_3}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} e^{i(px^3\,-\,E_p t)}$$

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\,+,\,-\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-\frac{1}{2}w_3}\begin{pmatrix}0\\1\end{pmatrix} \\ e^{\frac{1}{2}w_3}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} e^{i(px^3\,-\,E_p t)}$$

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\, -,\, +\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} e^{\frac{1}{2}w_3} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ -e^{-\frac{1}{2}w_3} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)}$$

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\, -,\, -\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-\frac{1}{2}w_3} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ -\,e^{\frac{1}{2}w_3} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)} \tag{10.326}$$

where according to (10.61) $p(w_3) \;=\; m\,\mathrm{sinh}w_3$. Transformation to the Dirac representation by means of (10.228) yields

$$\Psi(p(w_3)\hat{e}_3,\, +,\, +\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} \cosh\frac{w_3}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \sinh\frac{w_3}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} e^{i(px^3 - E_p t)}$$

$$\Psi(p(w_3)\hat{e}_3,\, +,\, -\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} \cosh\frac{w_3}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ -\sinh\frac{w_3}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} e^{i(px^3 - E_p t)}$$

$$\Psi(p(w_3)\hat{e}_3,\, -,\, +\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} \sinh\frac{w_3}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \cosh\frac{w_3}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)}$$

$$\Psi(p(w_3)\hat{e}_3,\, -,\, -\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} -\sinh\frac{w_3}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \cosh\frac{w_3}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)} \tag{10.327}$$

Employing the hyperbolic function properties

$$\cosh\frac{x}{2} \;=\; \sqrt{\frac{\cosh x \,+\, 1}{2}}\,, \quad \sinh\frac{x}{2} \;=\; \sqrt{\frac{\cosh x \,-\, 1}{2}}\,, \tag{10.328}$$

the relationship (10.61) between the parameter w_3 and boost velocity v_3, and the expression (10.311) for E_p one obtains

$$\begin{aligned} \cosh\tfrac{w_3}{2} \;&=\; \frac{1}{\sqrt{2}}\sqrt{\sqrt{\frac{1}{1 \,-\, v_3^2}} \,+\, 1} \;=\; \frac{1}{\sqrt{2}}\sqrt{\sqrt{1 \,+\, \frac{v_3^2}{1 \,-\, v_3^2}} \,+\, 1} \\ &=\; \sqrt{\frac{\sqrt{m^2 \,+\, \frac{m^2 v_3^2}{1-v_3^2}} \,+\, m}{2\,m}} \;=\; \sqrt{\frac{E_p \,+\, m}{2\,m}} \end{aligned} \tag{10.329}$$

and similarly

$$\sinh\frac{w_3}{2} \;=\; \sqrt{\frac{E_p \,-\, m}{2\,m}} \;=\; \frac{p}{\sqrt{2m\,(E_p \,+\, m)}} \tag{10.330}$$

Inserting expressions (10.329, 10.330) into (10.327), indeed, reproduces the positive energy wave functions (10.311) as well as the negative energy solutions (10.313) for $-p$. The change of sign for the latter solutions had to be expected as it was already noted for the negative energy solutions of the Klein–Gordon equation (10.168–10.176).

Invariance of Dirac Equation Revisited

At this point we like to provide a variation of the derivation of (10.243), the essential property stating the Lorentz–invariance of the Dirac equation. Actually, we will derive this equation only for infinitesimal transformations, which however, is sufficient since (1) it must hold then for any finite transformation, and since (2) the calculations following (10.243) considered solely the limit of infinitesimal transformations anyway.

The reason why we provide another derivation of (10.243) is to familiarize ourselves with a formulation of Lorentz transformations of the bispinor wave finction $\Psi(x^\mu)$ which treats the spinor and the space-time part of the wave function on the same footing. Such description will be essential for the formal description of Lorentz invariant wave equations for arbitray spin further below.

In the new derivation we consider the particle described by the wave function transformed, but not the observer. This transformation, refered to as the active transformation, expresses the system in the old coordinates. The transformation is

$$\Psi'(x^\mu) \;=\; S(L^\eta{}_\xi)\,\rho(L^\eta{}_\xi)\,\Psi(x^\mu) \tag{10.331}$$

where $S(L^\eta{}_\xi)$ denotes again the transformation acting on the bispinor character of the wave function $\Psi(x^\mu)$ and where $\rho(L^\eta{}_\xi)$ denotes the transformation acting on the space-time character of the wave function $\Psi(x^\mu)$. $\rho(L^\eta{}_\xi)$ has been defined in (10.123) above and characterized there. Such transformation had been applied by us, of course, when we generated the solutions $\Psi(\vec{p}, \lambda, \Lambda|x^\mu)$ from the solutions describing particles at rest $\Psi(\vec{p} = 0, \lambda, \Lambda|t)$. We expect, in general, that if $\Psi(x^\mu)$ is a solution of the Dirac equation that $\Psi'(x^\mu)$ as given in (10.331) is a solution as well. Making this expectation a postulate allows one to derive the condition (10.243) and, thereby, the proper transformation $S(L^\eta{}_\xi)$.

To show this we rewrite the Dirac equation (10.221) using (10.331)

$$\left(\,i\,S(L^\eta{}_\xi)\gamma^\mu S^{-1}(L^\eta{}_\xi)\,\rho(L^\eta{}_\xi)\partial_\mu\rho^{-1}(L^\eta{}_\xi) \;-\; m\,\right)\,\Psi'(x^\mu) \;=\; 0 \tag{10.332}$$

Here we have made use of the fact that $S(L^\eta{}_\xi)$ commutes with ∂_μ and $\rho(L^\eta{}_\xi)$ commutes with γ^μ. The fact that any such $\Psi'(x^\mu)$ is a solution of the Dirac equation allows us to conclude

$$S(L^\eta{}_\xi)\gamma^\mu S^{-1}(L^\eta{}_\xi)\,\rho(L^\eta{}_\xi)\partial_\mu\rho^{-1}(L^\eta{}_\xi) \;=\; \gamma^\nu\partial_\nu \tag{10.333}$$

which is satisfied in case that the following conditions are met

$$\begin{aligned} \rho(L^\eta{}_\xi)\partial_\mu\rho^{-1}(L^\eta{}_\xi) \quad &= \; L^\nu{}_\mu\,\partial_\nu \;; \\ \mathcal{S}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^{-1}(L^\eta{}_\xi)\,L^\nu{}_\mu \quad &= \; \gamma^\nu \;. \end{aligned} \tag{10.334}$$

We will demonstrate now that the first condition is satisfied by $\rho(L^\eta{}_\xi)$. The second condition is identical to (10.243) and, of course, it is met by $\mathcal{S}(L^\eta{}_\xi)$ as given in the chiral representation by (10.262).
As mentioned already we will show condition (10.334) for infinitesimal Lorentz transformations $L^\eta{}_\xi$. We will proceed by employing the generators (10.128) to express $\rho(L^\eta{}_\xi)$ in its infinitesimal form and evaluate the expression

$$\left(1\!\!1 \,+\, \epsilon\vec{\vartheta}\cdot\vec{\mathcal{J}} \,+\, \epsilon\vec{w}\cdot\vec{\mathcal{K}} \right) \partial_\mu \left(1\!\!1 \,-\, \epsilon\vec{\vartheta}\cdot\vec{\mathcal{J}} \,-\, \epsilon\vec{w}\cdot\vec{\mathcal{K}} \right) = \partial_\mu \,+\, \epsilon M^\nu{}_\mu \partial_\nu \,+\, O(\epsilon^2) \tag{10.335}$$

The result will show that the matrix $M^\nu{}_\mu$ is identical to the generators of $L^\nu{}_\mu$ for the six choices $\vec{\vartheta} = (1,0,0)$, $\vec{w} = (0,0,0)$, $\vec{\vartheta} = (0,1,0)$, $\vec{w} = (0,0,0)$, $\ldots$, $\vartheta = (0,0,0)$, $\vec{w} = (0,0,1)$. Inspection of (10.335) shows that we need to demonstrate

$$[\mathcal{J}_\ell, \partial_\mu] \;=\; (J_\ell)^\nu{}_\mu \partial_\nu \;; \quad [\mathcal{K}_\ell, \partial_\mu] \;=\; (K_\ell)^\nu{}_\mu \partial_\nu \;. \tag{10.336}$$

We will proceed with this task considering all six cases:

$$[\mathcal{J}_1, \partial_\mu] \;=\; [x^3\partial_2 \,-\, x^2\partial_3, \partial_\mu] \;=\; \begin{cases} 0 & \mu = 0 \\ 0 & \mu = 1 \\ \partial_3 & \mu = 2 \\ -\partial_2 & \mu = 3 \end{cases} \tag{10.337}$$

$$[\mathcal{J}_2, \partial_\mu] \;=\; [x^1\partial_3 \,-\, x^3\partial_1, \partial_\mu] \;=\; \begin{cases} 0 & \mu = 0 \\ -\partial_3 & \mu = 1 \\ 0 & \mu = 2 \\ \partial_1 & \mu = 3 \end{cases} \tag{10.338}$$

$$[\mathcal{J}_3, \partial_\mu] \;=\; [x^2\partial_1 \,-\, x^1\partial_2, \partial_\mu] \;=\; \begin{cases} 0 & \mu = 0 \\ \partial_2 & \mu = 1 \\ -\partial_1 & \mu = 2 \\ 0 & \mu = 3 \end{cases} \tag{10.339}$$

$$[\mathcal{K}_1, \partial_\mu] \;=\; [x^0\partial_1 \,+\, x^1\partial_0, \partial_\mu] \;=\; \begin{cases} -\partial_1 & \mu = 0 \\ -\partial_0 & \mu = 1 \\ 0 & \mu = 2 \\ 0 & \mu = 3 \end{cases} \tag{10.340}$$

$$[\mathcal{K}_2, \partial_\mu] \;=\; [x^0\partial_2 \,+\, x^2\partial_0, \partial_\mu] \;=\; \begin{cases} -\partial_2 & \mu = 0 \\ 0 & \mu = 1 \\ -\partial_0 & \mu = 2 \\ 0 & \mu = 3 \end{cases} \tag{10.341}$$

$$[\mathcal{K}_3, \partial_\mu] \;=\; [x^0\partial_3 \,+\, x^3\partial_0, \partial_\mu] \;=\; \begin{cases} -\partial_3 & \mu = 0 \\ 0 & \mu = 1 \\ 0 & \mu = 2 \\ -\partial_0 & \mu = 3 \end{cases} \tag{10.342}$$

One can readily convince oneself that these results are consistent with (10.336). We have demonstrated, therefore, that any solution $\Psi(x^\mu)$ transformed according to (10.331) is again a solution of the Dirac equation, i.e., the Dirac equation is invariant under *active* Lorentz transformations.

10.10 Dirac Particles in Electromagnetic Field

We like to provide now a description for particles governed by the Dirac equation which includes the coupling to an electromagnetic field in the minimum coupling description. Following the respective procedure developed for the Klein-Gordon equation in Sect. 10.6 we assume that the field is described through the 4-vector potential A^μ and, accordingly, we replace in the Dirac equation the momentum operator $\hat{p}_\mu = i\partial_\mu$ by $i\partial_\mu - qA_\mu$ where q is the charge of the respective particles (see Table 10.1 in Sect 10.6 above). Equivalently, we replace the operator ∂_μ by $\partial_\mu + iqA_\mu$. The Dirac equation (10.221) reads then

$$[\, i\gamma^\mu(\partial_\mu + iqA_\mu) - m\,]\, \Psi(x^\nu) = 0 \tag{10.343}$$

One may also include the electromagnetic field in the Dirac equation given in the Schrödinger form (10.233) by replacing $i\partial_t$ by (see Table 10.1) $i\partial_t - qV$ and $\hat{\vec{p}}$ by

$$\hat{\vec{\pi}} = \hat{\vec{p}} - q\vec{A}\,. \tag{10.344}$$

The Dirac equation in the Schrödinger form reads then

$$i\partial_t \Psi(x^\mu) = \left(\hat{\vec{\alpha}} \cdot \hat{\vec{\pi}} + qV + \hat{\beta}\, m \right) \Psi(x^\mu) \tag{10.345}$$

where $\hat{\vec{\alpha}}$ and $\hat{\beta}$ are defined in (10.232).

Non-Relativistic Limit

We want to consider now the Dirac equation (10.345) in the so-called non-relativistic limit in which all energies are much smaller than m, e.g., for the scalar field V in (10.345) holds

$$|qV| << m\,. \tag{10.346}$$

For this purpose we choose the decomposition

$$\Psi(x^\mu) = \begin{pmatrix} \phi(x^\mu) \\ \chi(x^\mu) \end{pmatrix}. \tag{10.347}$$

Using the notation $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)^T$ one obtains then

$$i\partial_t \phi = \vec{\sigma} \cdot \hat{\vec{\pi}}\, \chi + qV\, \phi + m\phi \tag{10.348}$$

$$i\partial_t \chi = \vec{\sigma} \cdot \hat{\vec{\pi}}\, \phi + qV\, \chi - m\chi\,. \tag{10.349}$$

We want to focus on the stationary positive energy solution. This solution exhibits a time-dependence $\exp[-i(m+\epsilon)t]$ where for ϵ holds in the non-relativistic limit $|\epsilon| << m$. Accordingly, we define

$$\phi(x^\mu) = e^{-imt}\, \Phi(x^\mu) \tag{10.350}$$

$$\chi(x^\mu) = e^{-imt}\, \mathcal{X}(x^\mu) \tag{10.351}$$

and assume that for the time-derivative of Φ and $\mathcal{X}$ holds

$$\left| \frac{\partial_t \Phi}{\Phi} \right| << m \, , \quad \left| \frac{\partial_t \mathcal{X}}{\mathcal{X}} \right| << m \, . \tag{10.352}$$

Using (10.350, 10.351) in (10.348, 10.349) yields

$$\begin{aligned} i\partial_t \Phi &= \vec{\sigma} \cdot \hat{\vec{\pi}} \, \mathcal{X} + qV \, \Phi & (10.353) \\ i\partial_t \mathcal{X} &= \vec{\sigma} \cdot \hat{\vec{\pi}} \, \Phi + qV \, \mathcal{X} - 2m\mathcal{X} \, . & (10.354) \end{aligned}$$

The properties (10.346, 10.352) allow one to approximate (10.354)

$$0 \approx \vec{\sigma} \cdot \hat{\vec{\pi}} \, \Phi - 2m \, \mathcal{X} \, . \tag{10.355}$$

and, accordingly, one can replace $\mathcal{X}$ in (10.353) by

$$\mathcal{X} \approx \frac{\vec{\sigma} \cdot \hat{\vec{\pi}}}{2m} \, \Phi \tag{10.356}$$

to obtain a closed equation for Φ

$$i\partial_t \Phi \approx \frac{\left(\vec{\sigma} \cdot \hat{\vec{\pi}} \right)^2}{2m} \, \Phi + qV \, \Phi \, . \tag{10.357}$$

Equation (10.356), due to the m^{-1} factor, identifies $\mathcal{X}$ as the small component of the bi-spinor wave function which, henceforth, does not need to be considered anymore.
Equation (10.357) for Φ can be reformulated by expansion of $(\vec{\sigma} \cdot \hat{\vec{\pi}})^2$. For this purpose we employ the identity (5.230), derived in Sect. 5.7, which in the present case states

$$(\vec{\sigma} \cdot \hat{\vec{\pi}})^2 = \hat{\vec{\pi}}^2 + i \, \vec{\sigma} \cdot (\hat{\vec{\pi}} \times \hat{\vec{\pi}}) \, . \tag{10.358}$$

For the components of $\hat{\vec{\pi}} \times \hat{\vec{\pi}}$ holds

$$\left(\hat{\vec{\pi}} \times \hat{\vec{\pi}} \right)_\ell = \epsilon_{jk\ell} \, (\, \pi_j \pi_k - \pi_k \pi_j \,) = \epsilon_{jk\ell} \, [\pi_j, \, \pi_\ell] \, . \tag{10.359}$$

We want to evaluate the latter commutator. One obtains

$$\begin{aligned} [\pi_j, \pi_k] &= [\frac{1}{i}\partial_j + qA_j, \, \frac{1}{i}\partial_k + qA_k] \\ &= \underbrace{[\frac{1}{i}\partial_j, \, \frac{1}{i}\partial_k]}_{=0} + q \, [A_j, \, \frac{1}{i}\partial_k] + q \, [\frac{1}{i}\partial_j, \, A_k] + q^2 \underbrace{[A_j, \, A_k]}_{=0} \\ &= \frac{q}{i} \, [A_j, \, \partial_k] + \frac{q}{i} \, [\partial_j, \, A_k] \, . \end{aligned} \tag{10.360}$$

For an arbitrary function $f(\vec{r})$ holds

$$(\, [A_j, \, \partial_k] + [\partial_j, \, A_k] \,) \, f = (\, \partial_j A_k - A_k \partial_j + A_j \partial_k - \partial_k A_j \,) \, f \, . \tag{10.361}$$

Using

$$\partial_j A_k f \;=\; ((\partial_j A_k))\, f \;+\; A_k \partial_j\, f$$
$$\partial_k A_j f \;=\; ((\partial_k A_j))\, f \;+\; A_j \partial_k\, f$$

where $((\partial_j \cdots))$ denotes confinement of the differential operator to within the brackets $((\cdots))$, one obtains

$$(\, [A_j,\, \partial_k] \;+\; [\partial_j,\, A_k]\,)\, f \;=\; [\, ((\partial_j A_k)) \;-\; ((\partial_k A_j))\,]\, f \tag{10.362}$$

or, using (10.360) and $A_\mu \;=\; (V,\, -\vec{A})$,

$$[\pi_j, \pi_k] \;=\; \frac{q}{i}\, ((\, \partial_j A_k \;-\; \partial_k A_j\,)) \;=\; -\frac{q}{i} \left(\nabla \times \vec{A} \right)_\ell \epsilon_{jk\ell} \;=\; -\frac{q}{i} B_\ell\, \epsilon_{jk\ell} \tag{10.363}$$

where we employed $\vec{B}(\vec{r},t) \;=\; \nabla \times \vec{A}(\vec{r},t)$ [see (8.6)]. Equations (10.344, 10.358, 10.359, 10.363) allow us to write (10.357) in the final form

$$i\partial_t \Phi(\vec{r},t) \;\approx\; \left[\frac{[\hat{\vec{p}} \;-\; q\vec{A}(\vec{r},t)]^2}{2m} \;-\; \frac{q}{2m}\, \vec{\sigma} \cdot \vec{B}(\vec{r},t) \;+\; q\, V(\vec{r},t) \right] \Phi(\vec{r},t) \tag{10.364}$$

which is referred to as the *Pauli equation*.
Comparision of (10.364) governing a two-dimensional wave function $\Phi \,\in\, \mathbb{C}^2$ with the corresponding non-relativistic Schrödinger equation (10.2) governing a one-dimensional wave function $\psi \,\in\, \mathbb{C}$, reveals a stunning feature: the Pauli equation does justice to its two-dimensional character; while agreeing in all other respects with the non-relativistic Schrödinger equation (10.2) it introduces the extra term $q\vec{\sigma} \cdot \vec{B}\, \Phi$ which describes the well-known interaction of a spin-$\frac{1}{2}$ particle with a magnetic field $\vec{B}$. In other words, the spin-$\frac{1}{2}$ which emerged in the Lorentz-invariant theory as an algebraic necessity, does not leave the theory again when one takes the non-relativistic limit, but rather remains as a steady "guest" of non-relativistic physics with the proper interaction term.
Let us consider briefly the consequences of the interaction of a spin-$\frac{1}{2}$ with the magnetic field. For this purpose we disregard the spatial degrees of freedom and assume the Schrödinger equation

$$i\partial_t \Phi(t) \;=\; q\, \vec{\sigma} \cdot \vec{B}\, \Phi(t)\;. \tag{10.365}$$

The formal solution of this equation is

$$\Phi(t) \;=\; e^{-iqt\vec{B}\cdot\vec{\sigma}}\, \Phi(0)\;. \tag{10.366}$$

Comparision of this expression with (5.222, 5.223) shows that the propagator in (10.366) can be interpreted as a rotation around the field $\vec{B}$ by an angle qtB, i.e., the interaction $q\vec{\sigma} \cdot \vec{B}$ induces a precession of the spin-$\frac{1}{2}$ around the magnetic field.

Dirac Particle in Coulomb Field - Spectrum

We want to describe now the spectrum of a relativistic electron ($q \;=\; -e$) in the Coulomb field of a nucleus with charge Ze. The respective bispinor wave function $\Psi(x^\mu) \,\in\, \mathbb{C}^4$ is described as the stationary solution of the Dirac equation (10.343) for the vector potential

$$A_\mu \;=\; (-\frac{Ze^2}{r},\, 0,\, 0,\, 0)\;. \tag{10.367}$$

For the purpose of the solution we assume the *chiral representation*, i.e, we solve

$$[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,-\, m \,]\; \tilde{\Psi}(x^\mu) \;=0 \tag{10.368}$$

where $\tilde{\Psi}(x^\mu)$ and $\tilde{\gamma}^\mu$ are defined in (10.228) and in (10.229), respectively. Employing π_μ as defined in Table 10.1 one can write (10.368)

$$(\, \tilde{\gamma}^\mu \,\pi_\mu \,-\, m \,)\; \tilde{\Psi}(x^\mu) \;=0\;. \tag{10.369}$$

For our solution we will adopt presently a strategy which follows closely that for the spectrum of pionic atoms in Sect. 10.6. For this purpose we 'square' the Dirac equation, multiplying (10.369) from the left by $\gamma^\nu \pi_\nu \,+\, m$. This yields

$$\begin{aligned}&[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,+\, m \,]\,[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,-\, m \,]\; \tilde{\Psi}(x^\mu)\\ &\qquad =\; (\tilde{\gamma}^\mu \hat{\pi}_\mu \,\tilde{\gamma}^\nu \hat{\pi}_\nu \,-\, m^2)\, \tilde{\Psi}(x^\mu) \;=\; 0\;.\end{aligned} \tag{10.370}$$

Any solution of (10.368) is also a solution of (10.370), but the converse is not necessarily true. However, once a solution $\tilde{\Psi}(x^\mu)$ of (10.370) is obtained then

$$[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,+\, m \,]\; \tilde{\Psi}(x^\mu) \tag{10.371}$$

is a solution of (10.369). This follows from

$$\begin{aligned}&[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,+\, m \,]\,[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,-\, m \,]\\ &\quad =\; [\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,-\, m \,]\,[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,+\, m \,]\end{aligned} \tag{10.372}$$

according to which follows from (10.370)

$$[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,-\, m \,]\,[\, i\tilde{\gamma}^\mu(\partial_\mu \,+\, iqA_\mu) \,+\, m \,]\; \tilde{\Psi}(x^\mu) \;=\; 0 \tag{10.373}$$

such that we can conclude that (10.371), indeed, is a solution of (10.369).
Equation (10.370) resembles closely the Klein-Gordon equation (10.180), but differs from it in an essential way. The difference arises from the term $\tilde{\gamma}^\mu \hat{\pi}_\mu \,\tilde{\gamma}^\nu \hat{\pi}_\nu$ in (10.370) for which holds

$$\tilde{\gamma}^\mu \hat{\pi}_\mu \,\tilde{\gamma}^\nu \hat{\pi}_\nu \;=\; \sum_{\mu=0}^{3} \,(\tilde{\gamma}^\mu)^2\, \hat{\pi}_\mu^2 \;+\; \sum_{\substack{\mu,\nu=1\\ \mu\,\neq\nu}} \,\tilde{\gamma}^\mu \tilde{\gamma}^\nu \,\hat{\pi}_\mu \hat{\pi}_\nu\;. \tag{10.374}$$

The first term on the r.h.s. can be rewritten using, according to (10.230), $(\tilde{\gamma}^0)^2 \;=\; 1\!\!1$ and $(\tilde{\gamma}^j)^2 \;=\; -1\!\!1$, $j = 1, 2, 3$,

$$\sum_{\mu=0}^{3} \,(\tilde{\gamma}^\mu)^2\, \hat{\pi}_\mu^2 \;=\; \hat{\pi}_0^2 \;-\; \hat{\vec{\pi}}^{\,2}\;. \tag{10.375}$$

Following the algebra that connected Eqs. (5.231), (5.232) in Sect. 5.7 one can write the second term in (10.374), noting from (10.230) $\tilde{\gamma}^\mu \tilde{\gamma}^\nu \;=\; -\,\tilde{\gamma}^\nu \tilde{\gamma}^\mu$, $\mu \neq \nu$ and altering 'dummy' summation indices,

$$\sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}} \tilde{\gamma}^\mu \tilde{\gamma}^\nu \hat{\pi}_\mu \hat{\pi}_\nu = \frac{1}{2} \sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}} \left(\tilde{\gamma}^\mu \tilde{\gamma}^\nu \hat{\pi}_\mu \hat{\pi}_\nu \; + \; \tilde{\gamma}^\nu \tilde{\gamma}^\mu \hat{\pi}_\nu \hat{\pi}_\mu \right)$$
$$= \frac{1}{4} \sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}} \left(\tilde{\gamma}^\mu \tilde{\gamma}^\nu \hat{\pi}_\mu \hat{\pi}_\nu \; - \; \tilde{\gamma}^\nu \tilde{\gamma}^\mu \hat{\pi}_\mu \hat{\pi}_\nu \; + \; \tilde{\gamma}^\nu \tilde{\gamma}^\mu \hat{\pi}_\nu \hat{\pi}_\mu \; - \; \tilde{\gamma}^\mu \tilde{\gamma}^\nu \hat{\pi}_\nu \hat{\pi}_\mu \right)$$
$$= \frac{1}{4} \sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}} [\tilde{\gamma}^\mu, \tilde{\gamma}^\nu] \, [\hat{\pi}_\mu, \hat{\pi}_\nu] \tag{10.376}$$

This expression can be simplified due to the special form (10.367) of A_μ, i.e., due to $\vec{A} = 0$. Since

$$[\hat{\pi}_\mu, \hat{\pi}_\nu] \; = \; 0 \quad \text{for } \mu, \nu \; = \; 1, 2, 3 \tag{10.377}$$

which follows readily from the definition (10.344), it holds

$$\frac{1}{4} \sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}} [\tilde{\gamma}^\mu, \tilde{\gamma}^\nu] \, [\hat{\pi}_\mu, \hat{\pi}_\nu]$$
$$= \frac{1}{4} \sum_{j=1}^{3} [\tilde{\gamma}^0, \tilde{\gamma}^j] \, [\hat{\pi}_0, \hat{\pi}_j] \; + \; \frac{1}{4} \sum_{j=1}^{3} [\tilde{\gamma}^j, \tilde{\gamma}^0] \, [\hat{\pi}_j, \hat{\pi}_0]$$
$$= \frac{1}{2} \sum_{j=1}^{3} [\tilde{\gamma}^0, \tilde{\gamma}^j] \, [\hat{\pi}_0, \hat{\pi}_j] \, . \tag{10.378}$$

According to the definition (10.229), the commutators $[\tilde{\gamma}^0, \tilde{\gamma}^j]$ are

$$[\tilde{\gamma}^0, \tilde{\gamma}^j] \; = \; \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} 0 & -\sigma^j \\ \sigma^j & 0 \end{pmatrix} - \begin{pmatrix} 0 & -\sigma^j \\ \sigma^j & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$$
$$= \; 2 \begin{pmatrix} \sigma^j & 0 \\ 0 & -\sigma^j \end{pmatrix} \tag{10.379}$$

The commutators $[\hat{\pi}_0, \hat{\pi}_j]$ in (10.378) can be evaluated using (10.367) and the definition (10.344)

$$[\hat{\pi}_0, \hat{\pi}_j] \; = \; (-i\partial_t \, + \, qA_0, \, -i\partial_j] \; = \; - \left[\, (\partial_t \, + \, iqA_0) \, \partial_j \; - \; \partial_j \, (\partial_t \, + \, iqA_0) \, \right] f$$
$$= \; i \, ((\partial_j q A_0)) \, f \tag{10.380}$$

where $f \; = \; f(\vec{r}, t)$ is a suitable test function and where $((\cdots))$ denotes the range to which the derivative is limited. Altogether, one can summarize (10.376–10.380)

$$\sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}} \tilde{\gamma}^\mu \tilde{\gamma}^\nu \hat{\pi}_\mu \hat{\pi}_\nu \; = \; i \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix} \cdot ((\nabla q A_0)) \tag{10.381}$$

According to (10.367) holds

$$\nabla q A_0 \; = \; \hat{r} \, \frac{Ze^2}{r^2} \, . \tag{10.382}$$

where $\hat{r} = \vec{r}/|\vec{r}|$ is a unit vector. Combining this result with (10.381), (10.374), (10.375) the 'squared' Dirac equation (10.368) reads

$$\left[- \left(\partial_t - i\frac{Ze^2}{r} \right)^2 + \nabla^2 + i \begin{pmatrix} \vec{\sigma}\cdot\hat{r} & 0 \\ 0 & -\vec{\sigma}\cdot\hat{r} \end{pmatrix} \frac{Ze^2}{r^2} - m^2 \right] \Psi(x^\mu) = 0 \qquad (10.383)$$

We seek stationary solutions of this equation. Such solutions are of the form

$$\tilde{\Psi}(x^\mu) = \tilde{\Phi}(\vec{r})\, e^{-i\epsilon t} . \qquad (10.384)$$

ϵ can be interpreted as the energy of the stationary state and, hence, it is this quantity that we want to determine. Insertion of (10.384) into (10.383) yields the purely spatial four-dimensional differential equation

$$\left[\left(\epsilon + \frac{Ze^2}{r} \right)^2 + \nabla^2 + i \begin{pmatrix} \vec{\sigma}\cdot\hat{r} & 0 \\ 0 & -\vec{\sigma}\cdot\hat{r} \end{pmatrix} \frac{Ze^2}{r^2} - m^2 \right] \Phi(\vec{r}) = 0 . \qquad (10.385)$$

We split the wave function into two spin-$\frac{1}{2}$ components

$$\tilde{\Psi}(\vec{r}) = \begin{pmatrix} \tilde{\phi}_+(\vec{r}) \\ \tilde{\phi}_-(\vec{r}) \end{pmatrix} \qquad (10.386)$$

and obtain for the separate components $\phi_\pm(\vec{r})$

$$\left[\left(\epsilon + \frac{Ze^2}{r} \right)^2 + \nabla^2 \pm i\,\vec{\sigma}\cdot\hat{r}\,\frac{Ze^2}{r^2} - m^2 \right] \phi_\pm(\vec{r}) = 0 . \qquad (10.387)$$

The expression (10.189) for the Laplacian and expansion of the term $(\cdots)^2$ result in the two-dimensional equation

$$\left[\partial_r^2 - \frac{\hat{L}^2 - Z^2e^4 \mp i\vec{\sigma}\cdot\hat{r}\,Ze^2}{r^2} + \frac{2Ze^2\epsilon}{r} + \epsilon^2 - m^2 \right] r\,\phi_\pm(\vec{r}) = 0 . \qquad (10.388)$$

Except for the term $i\vec{\sigma}\cdot\hat{r}$ this equation is identical to that posed by the one-dimensional Klein-Gordon equation for pionic atoms (10.191) solved in Sect. 10.6. In the latter case, a solution of the form $\sim Y_{\ell m}(\hat{r})$ can be obtained. The term $i\vec{\sigma}\cdot\hat{r}$, however, is genuinely two-dimensional and, in fact, couples the orbital angular momentum of the electron to its spin-$\frac{1}{2}$. Accordingly, we express the solution of (10.388) in terms of states introduced in Sect. 6.5 which describe the coupling of orbital angular momentum and spin

$$\{ (\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}), \mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r})) , \quad j = \tfrac{1}{2}, \tfrac{3}{2} \ldots ; m = -j, -j+1, \ldots + j \} \qquad (10.389)$$

According to the results in Sect. 6.5 the operator $i\vec{\sigma}\cdot\hat{r}$ is block-diagonal in this basis such that only the states for identical j, m values are coupled, i.e., only the two states $\{\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}), \mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r})\}$ as given in (6.147, 6.148). We note that these states are also eigenstates of the angular momentum operator $\hat{L}^2$ [cf. (6.151]. We select, therefore, a specific pair of total spin-orbital angular momentum quantum numbers j, m and expand

$$\phi_\pm(\vec{r}) = \frac{h_\pm(r)}{r}\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) + \frac{g_\pm(r)}{r}\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{10.390}$$

Using

$$\vec{\sigma}\cdot\hat{r}\,\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) = -\,\mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{10.391}$$

derived in Sect. 6.5 [c.f. (6.186)], property (6.151), which states that the states $\mathcal{Y}_{jm}(j\pm\frac{1}{2},\frac{1}{2}|\hat{r})$ are eigenfunctions of $\hat{L}^2$, together with the orthonormality of these two states leads to the coupled differential equation

$$\left[\left(\partial_r^2 + \frac{2Ze^2\epsilon}{r} + \epsilon^2 - m\right)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \right. \tag{10.392}$$
$$\left. - \frac{1}{r^2}\begin{pmatrix}(j-\frac{1}{2})(j+\frac{1}{2}) - Z^2e^4 & \pm iZe^2\\ \pm i\,Ze^2 & (j+\frac{1}{2})(j+\frac{3}{2}) - Z^2e^4\end{pmatrix}\right]\begin{pmatrix}h_\pm(r)\\ g_\pm(r)\end{pmatrix} = 0\,.$$

We seek to bring (10.392) into diagonal form. Any similarity transformation leaves the first term in (10.392), involving the 2×2 unit matrix, unaltered. However, such transformation can be chosen as to diagonalize the second term. Since, in the present treatment, we want to determine solely the spectrum, not the wave functions, we require only the eigenvalues of the matrices

$$B_\pm = \begin{pmatrix}(j-\frac{1}{2})(j+\frac{1}{2}) - Z^2e^4 & \pm iZe^2\\ \pm i\,Ze^2 & (j+\frac{1}{2})(j+\frac{3}{2}) - Z^2e^4\end{pmatrix}, \tag{10.393}$$

but do not explicitly consider further the wavefunctions. Obviously, the eigenvalues are independent of m. The two eigenvalues of both matrices are identical and can be written in the form

$$\lambda_1(j)\,[\lambda_1(j) + 1] \qquad \text{and} \qquad \lambda_2(j)\,[\lambda_2(j) + 1] \tag{10.394}$$

where

$$\lambda_1(j) = \sqrt{(j+\tfrac{1}{2})^2 - Z^2e^4} \tag{10.395}$$

$$\lambda_2(j) = \sqrt{(j+\tfrac{1}{2})^2 - Z^2e^4} - 1 \tag{10.396}$$

Equation (10.392) reads then in the diagonal representation

$$\left(\partial_r^2 - \frac{\lambda_{1,2}(j)[\lambda_{1,2}(j)+1)}{r^2} + \frac{2\epsilon Ze^2}{r} + \epsilon^2 - m^2\right) f_{1,2}(r) = 0 \tag{10.397}$$

This equation is identical to the Klein-Gordon equation for pionic atoms written in the form (10.200), except for the slight difference in the expression of $\lambda_{1,2}(j)$ as given by (10.395, 10.396) and (10.199), namely, the missing additive term $-\frac{1}{2}$, the values of the argument of $\lambda_{1,2}(j)$ being $j = \frac{1}{2}, \frac{3}{2}, \ldots$ rather than $\ell = 0, 1, \ldots$ as in the case of pionic atoms, and except for the fact that we have two sets of values for $\lambda_{1,2}(j)$, namely, $\lambda_1(j)$ and $\lambda_2(j)$.. We can, hence, conclude that the spectrum of (10.397) is again given by eq. (10.203), albeit with some modifications. Using (10.395, 10.396) we obtain, accordingly,

$$\epsilon_1 = \frac{m}{\sqrt{1 + \frac{Z^2 e^4}{(n' + 1 + \sqrt{(j+\frac{1}{2})^2 - Z^2 e^4})^2}}} \; ; \tag{10.398}$$

$$\epsilon_2 = \frac{m}{\sqrt{1 + \frac{Z^2 e^4}{(n' \;\; + \;\; \sqrt{(j+\frac{1}{2})^2 - Z^2 e^4})^2}}} \; ; \tag{10.399}$$

$$n' = 0, 1, 2, \ldots \; , \quad j = \tfrac{1}{2}, \tfrac{3}{2}, \ldots \; , \quad m = -j, -j+1, \ldots, j$$

where ϵ_1 corresponds to $\lambda_1(j)$ as given in (10.395) and ϵ_2 corresponds to $\lambda_2(j)$ as given in (10.396). For a given value of n' the energies ϵ_1 and ϵ_2 for identical j-values correspond to mixtures of states with orbital angular momentum $\ell = j - \frac{1}{2}$ and $\ell = j + \frac{1}{2}$. The magnitude of the relativistic effect is determined by $Z^2 e^4$. Expanding the energies in terms of this parameter allows one to identify the relationship between the energies ϵ_1 and ϵ_2 and the non-relativistic spectrum. One obtains in case of (10.398, 10.399)

$$\epsilon_1 \approx m - \frac{m Z^2 e^4}{2\,(n' + j + \frac{3}{2})^2} + O(Z^4 e^8) \tag{10.400}$$

$$\epsilon_2 \approx m - \frac{m Z^2 e^4}{2\,(n' + j + \frac{1}{2})^2} + O(Z^4 e^8) \tag{10.401}$$

$$n' = 0, 1, 2, \ldots \; , \quad j = \tfrac{1}{2}, \tfrac{3}{2}, \ldots \; , \quad m = -j, -j+1, \ldots, j \; .$$

These expressions can be equated with the non-relativistic spectrum. Obviously, the second term on the r.h.s. of these equations describe the binding energy. In case of non-relativistic hydrogen-type atoms, including spin-$\frac{1}{2}$, the stationary states have binding energies

$$E = -\frac{m Z^2 e^4}{2 n^2} \, , \quad \begin{array}{ll} n = 1, 2, \ldots & \ell = 0, 1, \ldots, n-1 \\ m = -\ell, -\ell+1, \ldots, \ell & m_s = \pm\frac{1}{2} \end{array} \; . \tag{10.402}$$

In this expression n is the so-called main quantum number. It is given by $n = n' + \ell + 1$ where ℓ is the orbital angular momentum quantum number and $n' = 0, 1, \ldots$ counts the nodes of the wave function. One can equate (10.402) with (10.400) and (10.401) if one attributes to the respective states the angular momentum quantum numbers $\ell = j + \frac{1}{2}$ and $\ell = j - \frac{1}{2}$. One may also state this in the following way: (10.400) corresponds to a non-relativistic state with quantum numbers n, ℓ and spin-orbital angular momentum $j = \ell - \frac{1}{2}$; (10.401) corresponds to a non-relativistic state with quantum numbers n, ℓ and spin-orbital angular momentum $j = \ell + \frac{1}{2}$. These considerations are summarized in the following equations

$$E_{\rm D}(n, \ell, j = \ell - \tfrac{1}{2}, m) = \frac{m}{\sqrt{1 + \frac{Z^2 e^4}{(n - \ell + \sqrt{(\ell+1)^2 - Z^2 e^4})^2}}} \; ; \tag{10.403}$$

$$E_{\rm D}(n, \ell, j = \ell + \tfrac{1}{2}, m) = \frac{m}{\sqrt{1 + \frac{Z^2 e^4}{(n - \ell - 1 + \sqrt{\ell^2 - Z^2 e^4})^2}}} \; ; \tag{10.404}$$

$$n = 1, 2, \ldots \; ; \quad \ell = 0, 1, \ldots, n-1 \; ; \quad m = -j, -j+1, \ldots, j$$

spectr. notation	main quantum number n	orbital angular mom. ℓ	spin-orbital ang. mom. j	non-rel. binding energy / eV Eq. (10.402)	rel. binding energy / eV Eq. (10.405)
$1s_{\frac{1}{2}}$	1	0	$\frac{1}{2}$	-13.60583	-13.60601
$2s_{\frac{1}{2}}$	2	0	$\frac{1}{2}$	-3.40146	-3.40151
$2p_{\frac{1}{2}}$	2	1	$\frac{1}{2}$	$\Uparrow$	$\Uparrow$
$2p_{\frac{3}{2}}$	2	1	$\frac{3}{2}$	$\Uparrow$	- 3.40147
$3s_{\frac{1}{2}}$	3	0	$\frac{1}{2}$	-1.51176	-1.551178
$3p_{\frac{1}{2}}$	3	1	$\frac{1}{2}$	$\Uparrow$	$\Uparrow$
$3p_{\frac{3}{2}}$	3	1	$\frac{3}{2}$	$\Uparrow$	- 1.551177
$3d_{\frac{3}{2}}$	3	2	$\frac{3}{2}$	$\Uparrow$	$\Uparrow$
$3d_{\frac{5}{2}}$	3	2	$\frac{5}{2}$	$\Uparrow$	- 1.551176

Table 10.2:
Binding energies for the hydrogen ($Z = 1$) atom. Degeneracies are denoted by $\Uparrow$. The energies were evaluated with $m = 511.0041$ keV and $e^2 = 1/137.036$ by means of Eqs. (10.402, 10.405).

One can combine the expressions (10.403, 10.404) finally into the single formula

$$E_{\rm D}(n,\ell,j,m) \;=\; \frac{m}{\sqrt{1 \;+\; \frac{Z^2 e^4}{(\,n - j - \frac{1}{2} + \sqrt{(j+\frac{1}{2})^2 - \cdots}}}}$$

$$\begin{array}{l} n \;=\; 1, 2, \ldots \\ \ell \;=\; 0, 1, \ldots, n-1 \\ j \;=\; \left\{ \begin{array}{ll} \frac{1}{2} & \text{for } \ell \;=\; 0 \\ \ell \pm \frac{1}{2} & \text{otherwise} \end{array} \right. \\ m \;=\; -j, -j+1, \ldots, j \end{array} \qquad (10.405)$$

In order to demonstrate relativistic effects in the spectrum of the hydrogen atom we compare in Table 10.2 the non-relativistic [cf. (10.402)] and the relativistic [cf. (10.405)] spectrum of the hydrogen atom. The table entries demonstrate that the energies as given by the expression (10.405) in terms of the non-relativistic quantum numbers n, ℓ relate closely to the corresponding non-relativistic states, in fact, the non-relativistic and relativistic energies are hardly discernible. The reason is that the mean kinetic energy of the electron in the hydrogen atom, is in the range of 10 eV, i.e., much less than the rest mass of the electron (511 keV). However, in case of heavier nuclei the kinetic energy of bound electrons in the ground state scales with the nuclear charge Z like Z^2 such that in case $Z = 100$ the kinetic energy is of the order of the rest mass and relativistic effects become important. This is clearly demonstrated by the comparision of non-relativistic and relativistic spectra of a hydrogen-type atom with $Z = 100$ in Table 10.3.

spectr. notation	main quantum number n	orbital angular mom. ℓ	spin-orbital ang. mom. j	non-rel. binding energy / keV Eq. (10.402)	rel. binding energy / keV Eq. (10.405)
$1s_{\frac{1}{2}}$	1	0	$\frac{1}{2}$	-136.1	-161.6
$2s_{\frac{1}{2}}$	2	0	$\frac{1}{2}$	-34.0	-42.1
$2p_{\frac{1}{2}}$	2	1	$\frac{1}{2}$	⇑	⇑
$2p_{\frac{3}{2}}$	2	1	$\frac{3}{2}$	⇑	- 35.2
$3s_{\frac{1}{2}}$	3	0	$\frac{1}{2}$	-15.1	-17.9
$3p_{\frac{1}{2}}$	3	1	$\frac{1}{2}$	⇑	⇑
$3p_{\frac{3}{2}}$	3	1	$\frac{3}{2}$	⇑	- 15.8
$3d_{\frac{3}{2}}$	3	2	$\frac{3}{2}$	⇑	⇑
$3d_{\frac{5}{2}}$	3	2	$\frac{5}{2}$	⇑	- 15.3

Table 10.3:
Binding energies for the hydrogen-type ($Z = 100$) atom. Degeneracies are denoted by ⇑. The energies were evaluated with $m = 511.0041$ keV and $e^2 = 1/137.036$ by means of Eqs. (10.402, 10.405).

Of particular interest is the effect of spin-orbit coupling which removes, for example, the non-relativistic degeneracy for the six $2p$ states of the hydrogen atom: in the present, i.e., relativistic, case these six states are split into energetically different $2p_{\frac{1}{2}}$ and $2p_{\frac{3}{2}}$ states. The $2p_{\frac{1}{2}}$ states with $j = \frac{1}{2}$ involve two degenerate states corresponding to $\mathcal{Y}_{\frac{1}{2}m}(1, \frac{1}{2}|\hat{r})$ for $m = \pm\frac{1}{2}$, the $2p_{\frac{3}{2}}$ states with $j = \frac{3}{2}$ involve four degenerate states corresponding to $\mathcal{Y}_{\frac{3}{2}m}(1, \frac{1}{2}|\hat{r})$ for $m = \pm\frac{1}{2}, \pm\frac{3}{2}$.
In order to investigate further the deviation between relativistic and non-relativistic spectra of hydrogen-type atoms we expand the expression (10.405) to order $O(Z^4e^8)$. Introducing $\alpha = Z^2e^4$ and $\beta = j + \frac{1}{2}$ (10.405) reads

$$\frac{1}{\sqrt{1 + \frac{\alpha}{(n - \beta + \sqrt{\beta^2 - \alpha})^2}}} \tag{10.406}$$

The expansion (10.206) provides in the present case

$$E_{\rm D}(n, \ell, j, m) \;\approx\; m - \frac{mZ^2e^4}{2n^2} - \frac{mZ^4e^8}{2n^3}\left[\frac{1}{j + \frac{1}{2}} - \frac{3}{4n}\right] + O(Z^6e^{12})\,. \tag{10.407}$$

This expression allows one, for example, to estimate the difference between the energies of the states $2p_{\frac{3}{2}}$ and $2p_{\frac{1}{2}}$ (cf. Tables 10.2,10.3). It holds for $n = 2$ and $j = \frac{3}{2}, \frac{1}{2}$

$$E\,(2p_{\frac{3}{2}}) - E\,(2p_{\frac{1}{2}}) \;\approx\; -\frac{mZ^4e^8}{2 \cdot 2^3}\left[\frac{1}{2} - 1\right] \;=\; \frac{mZ^4e^8}{32}\,. \tag{10.408}$$

Radial Dirac Equation

We want to determine now the wave functions for the stationary states of a Dirac particle in a 4-vector potential

$$A_\mu \;=\; (V(r), 0, 0, 0) \tag{10.409}$$

where $V(r)$ is spherically symmetric. An example for such potential is the Coulomb potential $V(r) = -Ze^2/r$ considered further below. We assume for the wave function the stationary state form

$$\Psi(x^\mu) \;=\; e^{-i\epsilon t} \begin{pmatrix} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{pmatrix} , \tag{10.410}$$

where $\Phi(\vec{r})$, $\mathcal{X}(\vec{r}) \in \mathbb{C}^2$ describe the spatial and spin-$\frac{1}{2}$ degrees of freedom, but are time-independent. The Dirac equation reads then, according to (10.232, 10.345),

$$\vec{\sigma} \cdot \hat{\vec{p}}\, \mathcal{X} \;+\; m\, \Phi \;+\; V(r)\, \Phi \;=\; \epsilon\, \Phi \tag{10.411}$$

$$\vec{\sigma} \cdot \hat{\vec{p}}\, \Phi \;-\; m \mathcal{X} \;+\; V(r)\, \mathcal{X} \;=\; \epsilon\, \mathcal{X} \tag{10.412}$$

In this equation a coupling between the wave functions $\Phi(\vec{r})$ and $\mathcal{X}(\vec{r})$ arises due to the term $\vec{\sigma} \cdot \hat{\vec{p}}$. This term has been discussed in detail in Sect. 6.5 [see, in particular, pp. 168]: the term is a scalar (rank zero tensor) in the space of the spin-angular momentum states $\mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{r})$ introduced in Sect. 6.5, i.e., the term is block-diagonal in the space spanned by the states $\mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{r})$ and does not couple states with different j, m-values; $\vec{\sigma} \cdot \hat{\vec{p}}$ has odd parity and it holds [c.f. (6.197, 6.198)]

$$\vec{\sigma} \cdot \hat{\vec{p}}\, f(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;=\; i \left[\partial_r \;+\; \frac{j + \frac{3}{2}}{r} \right] f(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \tag{10.413}$$

$$\vec{\sigma} \cdot \hat{\vec{p}}\, g(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;=\; i \left[\partial_r \;+\; \frac{\frac{1}{2} - j}{r} \right] g(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \,. \tag{10.414}$$

These equations can be brought into a more symmetric form using

$$\partial_r \;+\; \frac{1}{r} \;=\; \frac{1}{r}\, \partial_r\, r$$

which allows one to write (10.413, 10.414)

$$\vec{\sigma} \cdot \hat{\vec{p}}\, r f(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;=\; i \left[\partial_r \;+\; \frac{j + \frac{1}{2}}{r} \right] r\, f(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \tag{10.415}$$

$$\vec{\sigma} \cdot \hat{\vec{p}}\, r g(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;=\; i \left[\partial_r \;-\; \frac{j + \frac{1}{2}}{r} \right] r\, g(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \,. \tag{10.416}$$

The differential equations (10.411, 10.412) are four-dimensional with $\vec{r}$-dependent wave functions. The arguments above allow one to eliminate the angular dependence by expanding $\Phi(\vec{r})$ and $\mathcal{X}(\vec{r})$ in terms of $\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r})$ and $\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r})$, i.e.,

$$\begin{pmatrix} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{pmatrix} = \begin{pmatrix} \frac{a(r)}{r}\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) + \frac{b(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ \frac{c(r)}{r}\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) + \frac{d(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \end{pmatrix} . \tag{10.417}$$

In general, such expansion must include states with all possible j, m values. Presently, we consider the case that only states for one specific j, m pair contribute. Inserting (10.417) into (10.411, 10.412), using (10.415, 10.416), the orthonormality property (6.157), and multiplying by r results in the following two independent pairs of coupled differential equations

$$\begin{aligned} i\left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] d(r) + [\ \ m + V(r) - \epsilon]\, a(r) &= 0 \\ i\left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] a(r) + [-m + V(r) - \epsilon]\, d(r) &= 0 \end{aligned} \tag{10.418}$$

and

$$\begin{aligned} i\left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] c(r) + [\ \ m + V(r) - \epsilon]\, b(r) &= 0 \\ i\left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] b(r) + [-m + V(r) - \epsilon]\, c(r) &= 0\,. \end{aligned} \tag{10.419}$$

Obviously, only $a(r), d(r)$ are coupled and $b(r), c(r)$ are coupled. Accordingly, there exist two independent solutions (10.417) of the form

$$\begin{pmatrix} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{pmatrix} = \begin{pmatrix} i\,\frac{f_1(r)}{r}\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \\ -\frac{g_1(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \end{pmatrix} \tag{10.420}$$

$$\begin{pmatrix} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{pmatrix} = \begin{pmatrix} i\,\frac{f_2(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ -\frac{g_2(r)}{r}\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \end{pmatrix} \tag{10.421}$$

where the factors i and -1 have been introduced for convenience. According to (10.418) holds for $f_1(r)$, $g_1(r)$

$$\begin{aligned} \left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] g_1(r) + [\epsilon - m - V(r)]\, f_1(r) &= 0 \\ \left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] f_1(r) - [\epsilon + m - V(r)]\, g_1(r) &= 0 \end{aligned} \tag{10.422}$$

and for $f_2(r)$, $g_2(r)$

$$\left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] g_2(r) + [\,\epsilon - m - V(r)\,] f_2(r) = 0$$
$$\left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] f_2(r) - [\,\epsilon + m - V(r)\,] g_2(r) = 0 \qquad (10.423)$$

Equations (10.422) and (10.423) are identical, except for the opposite sign of the term $(j+\frac{1}{2})$; the equations determine, together with the appropriate boundary conditions at $r = 0$ and $r \to \infty$, the radial wave functions for Dirac particles in the potential (10.409).

Dirac Particle in Coulomb Field - Wave Functions

We want to determine now the wave functions of the stationary states of hydrogen-type atoms which correspond to the energy levels (10.405). We assume the 4-vector potential of pure Coulomb type (10.367) which is spherically symmetric such that equations (10.422, 10.423) apply for $V(r) = -Ze^2/r$. Equation (10.422) determines solutions of the form (10.420). In the non-relativistic limit, Φ in (10.420) is the large component and $\mathcal{X}$ is the small component. Hence, (10.422) corresponds to states

$$\Psi(x^\mu) \approx \begin{pmatrix} i\,\frac{f_1(r)}{r}\,\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \\ 0 \end{pmatrix}, \qquad (10.424)$$

i.e., to states with angular momentum $\ell = j+\frac{1}{2}$. According to the discussion of the spectrum (10.405) of the relativistic hydrogen atom the corresponding states have quantum numbers $n = 1, 2, \ldots,\ \ell = 0, 1, \ldots, n-1$. Hence, (10.422) describes the states $2p_{\frac{1}{2}}$, $3p_{\frac{1}{2}}$, $3d_{\frac{3}{2}}$, etc. Similarly, (10.423), determining wave functions of the type (10.421), i.e., in the non-relativistic limit wave functions

$$\Psi(x^\mu) \approx \begin{pmatrix} i\,\frac{f_2(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ 0 \end{pmatrix}, \qquad (10.425)$$

covers states with angular momentum $\ell = j-\frac{1}{2}$ and, correspodingly the states $1s_{\frac{1}{2}}$, $2s_{\frac{1}{2}}$, $2p_{\frac{3}{2}}$, $3s_{\frac{1}{2}}$, $3p_{\frac{3}{2}}$, $3d_{\frac{5}{2}}$, etc.

We consider first the solution of (10.422). The solution of (10.422) follows in this case from the same procedure as that adopted for the radial wave function of the non-relativistic hydrogen-type atom. According to this procedure, one demonstrates first that the wave function at $r \to 0$ behaves as r^γ for some suitable γ, one demonstrates then that the wave functions for $r \to \infty$ behaves as $\exp(-\mu r)$ for some suitable μ, and obtains finally a polynomial function p(r) such that $r^\gamma \exp(-\mu r) p(r)$ solves (10.422); enforcing the polynomial to be of finite order leads to discrete eigenvalues ϵ, namely, the ones given in (10.405).

Behaviour at $r \to 0$

We consider first the behaviour of the solutions $f_1(r)$ and $g_1(r)$ of (10.422) near $r = 0$. We note that (10.422), for small r, can be written

$$\left[\partial_r - \frac{j+\frac{1}{2}}{r} \right] g_1(r) + \frac{Ze^2}{r} f_1(r) = 0$$

$$\left[\partial_r + \frac{j+\frac{1}{2}}{r} \right] f_1(r) - \frac{Ze^2}{r} g_1(r) = 0 . \tag{10.426}$$

Setting

$$f_1(r) \underset{\to 0}{\sim} a\, r^{\gamma} , \quad g_1(r) \underset{\to 0}{\sim} b\, r^{\gamma} \tag{10.427}$$

yields

$$\gamma\, b\, r^{\gamma-1} - (j+\tfrac{1}{2})\, b\, r^{\gamma-1} + Ze^2\, a\, r^{\gamma-1} = 0$$

$$\gamma\, a\, r^{\gamma-1} + (j+\tfrac{1}{2})\, a\, r^{\gamma-1} - Ze^2\, b\, r^{\gamma-1} = 0 . \tag{10.428}$$

or

$$\begin{pmatrix} \gamma + (j+\frac{1}{2}) & -Ze^2 \\ Ze^2 & \gamma - (j+\frac{1}{2}) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0 . \tag{10.429}$$

This equation poses an eigenvalue problem (eigenvalue $-\gamma$) for proper γ values. One obtains $\gamma = \pm\sqrt{(j+\frac{1}{2})^2 - Z^2 e^4}$. The assumed r-dependence in (10.427) makes only the positive solution possible. We have, hence, determined that the solutions $f_1(r)$ and $g_1(r)$, for small r, assume the r-dependence in (10.427) with

$$\gamma = \sqrt{(j+\frac{1}{2})^2 - Z^2 e^4} . \tag{10.430}$$

Note that the exponent in (10.427), in case $j+\frac{1}{2})^2 < Ze^2$, becomes imaginary. Such r-dependence would make the expectation value of the potential

$$\int r^2 dr\, \rho(\vec{r})\, \frac{1}{r} \tag{10.431}$$

infinite since, according to (10.266, 10.267, 10.420), for the particle density holds then

$$\rho(\vec{r}) \sim |r^{\gamma-1}|^2 = \frac{1}{r^2} . \tag{10.432}$$

Behaviour at $r \to \infty$

For very large r values (10.422) becomes

$$\partial_r\, g_1(r) = -(\epsilon - m,)\, f_1(r)$$

$$\partial_r\, f_1(r) = (\epsilon + m)\, g_1(r) \tag{10.433}$$

Iterating this equation once yields

$$\begin{aligned} \partial_r^2 \, g_1(r) &= (\, m^2 - \epsilon^2 \,) \, g_1(r) \\ \partial_r^2 \, f_1(r) &= (\, m^2 - \epsilon^2 \,) \, f_1(r) \end{aligned} \tag{10.434}$$

The solutions of these equations are $f_1, g_1 \sim \exp(\pm\sqrt{m^2 - \epsilon^2}\, r)$. Only the exponentially decaying solution is admissable and, hence, we conclude

$$f_1(r) \underset{\to\infty}{\sim} e^{-\mu r} \,, \quad g_1(r) \underset{\to\infty}{\sim} e^{-\mu r} \,, \quad \mu = \sqrt{m^2 - \epsilon^2} \tag{10.435}$$

For bound states holds $\epsilon < m$ and, hence, μ is real. Let us consider then for the solution of (10.434)

$$f_1(r) = \sqrt{m + \epsilon}\, a \, e^{-\mu r} \,, \quad g_1(r) = -\sqrt{m - \epsilon}\, a \, e^{-\mu r} \,. \tag{10.436}$$

Insertion into (10.434) results in

$$\begin{aligned} (m - \epsilon) \sqrt{m + \epsilon}\, a \; - \; (m - \epsilon) \sqrt{m + \epsilon}\, a &= 0 \\ (m + \epsilon) \sqrt{m - \epsilon}\, a \; - \; (m + \epsilon) \sqrt{m - \epsilon}\, a &= 0 \end{aligned} \tag{10.437}$$

which is obviously correct.

Solution of the Radial Dirac Equation for a Coulomb Potential

To solve (10.422) for the Coulomb potential $V(r) = -Ze^2/r$ We assume a form for the solution which is adopted to the asymptotic solution (10.436). Accordingly, we set

$$f_1(r) = \sqrt{m + \epsilon}\; e^{-\mu r}\, \tilde{f}_1(r) \tag{10.438}$$

$$g_1(r) = -\sqrt{m - \epsilon}\; e^{-\mu r}\, \tilde{g}_1(r) \tag{10.439}$$

where μ is given in (10.435). Equation (10.422) leads to

$$\begin{aligned} -\sqrt{m - \epsilon}\, [\partial_r - \frac{j + \frac{1}{2}}{r}]\, \tilde{g}_1 + \sqrt{m + \epsilon}\, \frac{Ze^2}{r}\, \tilde{f}_1 + & \\ (m - \epsilon) \sqrt{m + \epsilon}\, \tilde{g}_1 - (m - \epsilon) \sqrt{m + \epsilon}\, \tilde{f}_1 &= 0 \end{aligned} \tag{10.440}$$

$$\begin{aligned} \sqrt{m + \epsilon}\, [\partial_r + \frac{j + \frac{1}{2}}{r}]\, \tilde{f}_1 + \sqrt{m - \epsilon}\, \frac{Ze^2}{r}\, \tilde{g}_1 - & \\ (m + \epsilon) \sqrt{m - \epsilon} \tilde{f}_1 + (m + \epsilon) \sqrt{m - \epsilon} \tilde{g}_1 &= 0 \end{aligned} \tag{10.441}$$

The last two terms on the l.h.s. of both (10.440) and (10.441) correspond to (10.437) where they cancelled in case $\tilde{f}_1 = \tilde{g}_1 = a$. In the present case the functions $\tilde{f}_1$ and $\tilde{g}_1$ cannot be chosen identical due to the terms in the differential equations contributing for finite r. However, without loss of generality we can choose

$$\tilde{f}_1(r) = \phi_1(r) + \phi_2(r) \,, \quad \tilde{g}_1(r) = \phi_1(r) - \phi_2(r) \tag{10.442}$$

which leads to a partial cancellation of the asymptotically dominant terms. We also introduce the new variable

$$\rho = 2\mu\, r \,. \tag{10.443}$$

From this results after a little algebra

$$[\partial_\rho - \frac{j+\frac{1}{2}}{\rho}]\,(\phi_1 - \phi_2) - \sqrt{\frac{m+\epsilon}{m-\epsilon}}\,\frac{Ze^2}{\rho}\,(\phi_1 + \phi_2) + \phi_2 = 0 \tag{10.444}$$

$$[\partial_\rho + \frac{j+\frac{1}{2}}{\rho}]\,(\phi_1 + \phi_2) + \sqrt{\frac{m-\epsilon}{m+\epsilon}}\,\frac{Ze^2}{\rho}\,(\phi_1 - \phi_2) - \phi_2 = 0\,. \tag{10.445}$$

Addition and subtraction of these equations leads finally to the following two coupled differential equations for ϕ_1 and ϕ_2

$$\partial_\rho\phi_1 + \frac{j+\frac{1}{2}}{\rho}\,\phi_2 - \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_1 - \frac{mZe^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_2 = 0 \tag{10.446}$$

$$\partial_\rho\phi_2 + \frac{j+\frac{1}{2}}{\rho}\,\phi_1 + \frac{mZe^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_1 + \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_2 - \phi_2 = 0 \tag{10.447}$$

We seek solutions of (10.446 , 10.447) of the form

$$\phi_1(\rho) = \rho^\gamma \sum_{s=0}^{n'} \alpha_s\,\rho^2 \tag{10.448}$$

$$\phi_2(\rho) = \rho^\gamma \sum_{s=0}^{n'} \beta_s\,\rho^2 \tag{10.449}$$

for γ given in (10.430) which conform to the proper $r \to 0$ behaviour determined above [c.f. (10.426–10.430)]. Inserting (10.448, 10.449) into (10.446, 10.447) leads to

$$\sum_s \Big[(s+\gamma)\,\alpha_s + (j+\tfrac{1}{2})\,\beta_s - \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}}\,\alpha_s - \frac{mZe^2}{\sqrt{m^2-\epsilon^2}}\,\beta_s\Big]\,\rho^{s+\gamma-1} = 0 \tag{10.450}$$

$$\sum_s \Big[(s+\gamma)\,\beta_2 + (j+\tfrac{1}{2})\,\alpha_s + \frac{mZe^2}{\sqrt{m^2-\epsilon^2}}\,\alpha_s + \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}}\,\beta_s - \beta_{s-1}\Big] = 0 \tag{10.451}$$

From (10.450) follows

$$\frac{\alpha_s}{\beta_s} = \frac{\frac{mZe^2}{\sqrt{m^2-\epsilon^2}} - (j+\frac{1}{2})}{s + \gamma - \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}}}\,. \tag{10.452}$$

From (10.451) follows

$$\begin{aligned}\beta_{s-1} &= \left(s + \gamma + \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}}\right)\beta_s + \frac{\frac{m^2Z^2e^4}{m^2-\epsilon^2} - (j+\frac{1}{2})^2}{s + \gamma - \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}}}\,\beta_s \\ &= \frac{(s+\gamma)^2 + Z^2e^4 - (j+\frac{1}{2})^2}{s + \gamma - \frac{\epsilon Ze^2}{\sqrt{m^2-\epsilon^2}}}\,\beta_s\,.\end{aligned} \tag{10.453}$$

Using (10.430) one can write this

$$\beta_s \;=\; \frac{s \,+\, \gamma \,-\, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}}}{s\,(s \,+\, 2\gamma)}\,\beta_{s-1}\;. \tag{10.454}$$

Defining

$$s_o \;=\; \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}} \,-\, \gamma \tag{10.455}$$

one obtains

$$\begin{aligned}
\beta_s \;&=\; \frac{s - s_o}{s(s+2\gamma)}\,\beta_{s-1} \\
&=\; \frac{(s-1-s_o)(s-s_o)}{(s-1)s\,(s-1+2\gamma)(s+2\gamma)}\,\beta_{s-2} \\
&\;\;\vdots \\
&=\; \frac{(1-s_o)(2-s_o)\dots(s-s_o)}{s!\,(2\gamma+1)(2\gamma+2)\cdots(2\gamma+s)}\,\beta_0
\end{aligned} \tag{10.456}$$

From (10.452) follows

$$\alpha_s \;=\; \frac{j+\frac{1}{2} \,-\, \frac{mZe^2}{\sqrt{m^2-\epsilon^2}}}{s_o}\;\frac{(-s_o)(1-s_o)(2-s_o)\dots(s-s_o)}{s!\,(2\gamma+1)(2\gamma+2)\cdots(2\gamma+s)}\,\beta_0 \tag{10.457}$$

One can relate the polynomials $\phi_1(\rho)$ and $\phi_2(\rho)$ defined through (10.448, 10.449) and (10.456, 10.457) with the *confluent hypergeometric functions*

$$F(a,c;x) \;=\; 1 \,+\, \frac{a}{c}\,x \,+\, \frac{a(a+1)}{c(c+1)}\,\frac{x^2}{2!} \,+\, \dots \tag{10.458}$$

or, equivalently, with the associated *Laguerre polynomials*

$$L_n^{(\alpha)} \;=\; F(-n,\alpha+1,x)\;. \tag{10.459}$$

It holds

$$\phi_1(\rho) \;=\; \beta_0\,\frac{j+\frac{1}{2} \,-\, \frac{mZe^2}{\sqrt{m^2-\epsilon^2}}}{s_o}\,\rho^\gamma\,F(-s_o,2\gamma+1;\rho) \tag{10.460}$$

$$\phi_2(\rho) \;=\; \beta_0\,\rho^\gamma\,F(1-s_o,2\gamma+1;\rho)\;. \tag{10.461}$$

In order that the wave functions remain normalizable the power series (10.448, 10.449) must be of finite order. This requires that all coefficients α_s and β_s must vanish for $s \geq n'$ for some $n' \in \mathbb{N}$. The expressions (10.456) and (10.457) for β_s and α_s imply that s_o must then be an integer, i.e., $s_o = n'$. According to the definitions (10.430, 10.455) this confinement of s_o implies discrete values for ϵ, namely,

$$\epsilon(n') \;=\; \frac{m}{\sqrt{1 \,+\, \frac{Z^2e^4}{(n' + \sqrt{(j+\frac{1}{2})^2 - Z^2e^4})^2}}}\;, \qquad n' \;=\; 0,1,2,\dots \tag{10.462}$$

This expression agrees with the spectrum of relativistic hydrogen-type atoms derived above and given by (10.405). Comparision with (10.405) allows one to identify $n' = n - j + \frac{1}{2}$ which, in fact, is an integer. For example, for the states $2p_{\frac{1}{2}}$, $3p_{\frac{1}{2}}$, $3d_{\frac{3}{2}}$ holds $n' = 1, 2, 1$. We can, hence, conclude that the polynomials in (10.461 for ϵ values given by (10.405) and the ensuing s_o values (10.455) are finite.

Altogether we have determined the stationary states of the type (10.421) with radial wave functions $f_1(r)$, $g_1(r)$ determined by (10.438, 10.439), (10.442), and (10.460, 10.461). The coefficients β_0 in (10.460, 10.461) are to be chosen to satisfy a normalization condition and to assign an overall phase. Due to the form (10.410) of the stationary state wave function the density $\rho(x^\mu)$ of the states under consideration, given by expression (10.267), is time-independent. The normalization integral is then

$$\int_0^\infty r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi \left(|\Phi(\vec{r})|^2 + |\mathcal{X}(\vec{r})|^2 \right) = 1 \qquad (10.463)$$

where Φ and $\mathcal{X}$, as in (10.421), are two-dimensional vectors determined through the explicit form of the spin-orbital angular momentum states $\mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{r})$ in (6.147, 6.148). The orthonormality properties (6.157, 6.158) of the latter states absorb the angular integral in (10.463) and yield [note the $1/r$ factor in (10.421)]

$$\int_0^\infty dr \left(|f_1(r)|^2 + |g_1(r)|^2 \right) = 1 \qquad (10.464)$$

The evaluation of the integrals, which involve the *confluent hypergeometric functions* in (10.460, 10.461), can follow the procedure adopted for the wave functions of the non-relativistic hydrogen atom and will not be carried out here.

The wave functions (10.421) correspond to non-relativistic states with orbital angular momentum $\ell = j + \frac{1}{2}$. They are described through quantum numbers $n, j, \ell = j + \frac{1}{2}, m$. The complete wave function is given by the following set of formulas

$$\Psi(n, j, \ell = j + \tfrac{1}{2}, m|x^\mu) = e^{-i\epsilon t} \begin{pmatrix} iF_1(r)\, \mathcal{Y}_{j,m}(j + \frac{1}{2}, \frac{1}{2}|\hat{r}) \\ G_1(r)\, \mathcal{Y}_{j,m}(j - \frac{1}{2}, \frac{1}{2}|\hat{r}) \end{pmatrix} \qquad (10.465)$$

$$F_1(r) = F_-(\kappa|r)\,, \quad G_1(r) = F_+(\kappa|r)\,, \quad \kappa = j + \tfrac{1}{2} \qquad (10.466)$$

where[2]

$$F_\pm(\kappa|r) = \mp N\, (2\mu r)^{\gamma-1} e^{-\mu r} \left\{ \left[\frac{(n'+\gamma)m}{\epsilon} - \kappa \right] F(-n', 2\gamma+1; 2\mu r) \right. \\ \left. \pm\, n' F(1 - n', 2\gamma + 1; 2\mu r) \right\} \qquad (10.467)$$

$$N = \frac{(2\mu)^{\frac{3}{2}}}{\Gamma(2\gamma+1)} \sqrt{\frac{m \mp \epsilon)\Gamma(2\gamma + n' + 1)}{4m\frac{(n'+\gamma)m}{\epsilon}\left(\frac{(n'+\gamma)m}{\epsilon} - \kappa\right) n'!}} \qquad (10.468)$$

[2]This formula has been adapted from "Relativistic Quantum Mechanics" by W. Greiner, (Springer, Berlin, 1990), Sect. 9.6.

and

$$\begin{aligned}
\mu &= \sqrt{(m-\epsilon)(m+\epsilon)} \\
\gamma &= \sqrt{(j+\tfrac{1}{2})^2 - Z^2 e^4} \\
n' &= n - j - \tfrac{1}{2} \\
\epsilon &= \frac{m}{\sqrt{1+\frac{Z^2 e^4}{(n'+\gamma)^2}}} \,.
\end{aligned} \tag{10.469}$$

We want to consider now the stationary states of the type (10.421) which, in the non-relativistic limit, become

$$\Psi(x^\mu) \approx e^{-1\epsilon t} \begin{pmatrix} i\,\dfrac{f_2(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ 0 \end{pmatrix} . \tag{10.470}$$

Obviously, this wavefunction has an orbital angular momentum quantum number $\ell = j - \frac{1}{2}$ and, accordingly, describes the complementary set of states $1s_{\frac{1}{2}}$, $2s_{\frac{1}{2}}$, $2p_{\frac{3}{2}}$, $3s_{\frac{1}{2}}$, $3p_{\frac{3}{2}}$, $3d_{\frac{5}{2}}$, etc. not not covered by the wave functions given by (10.465–10.469). The radial wave functions $f_2(r)$ and $g_2(r)$ in (10.421) are governed by the radial Dirac equation (10.423) which differs from the radial Dirac equation for $f_1(r)$ and $g_1(r)$ solely by the sign of the terms $(j+\frac{1}{2})/r$. One can verify, tracing all steps which lead from (10.422) to (10.469) that the following wave functions result

$$\Psi(n,j,\ell=j-\tfrac{1}{2},m|x^\mu) = e^{-i\epsilon t} \begin{pmatrix} iF_2(r)\,\mathcal{Y}_{j,m}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ G_2(r)\,\mathcal{Y}_{j,m}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \end{pmatrix} \tag{10.471}$$

$$F_2(r) = F_-(\kappa|r)\,, \quad G_2(r) = F_+(\kappa|r)\,, \quad \kappa = -j - \tfrac{1}{2} \tag{10.472}$$

where $F_\pm(\kappa|r)$ are as given in (10.467–10.469). We have, hence, obtained closed expressions for the wave functions of all the stationary bound states of relativistic hydrogen-type atoms.