Theoretical Physics III

Advanced Quantum Mechanics

Prof. Hans Kroha

Winter Term 2004/05
University of Bonn

TEXed by

Alexander Karim & Sebastian Mai

Contents

Chapter 1

Relativistic Quantum Mechanics

1.1 Introductory remarks

1.1.1 Special relativity and Lorentz transformations

Einstein 1905:

> Postulate that the velocity of light $\left(c \approx 3 \cdot 10^8 \frac{m}{s}\right)$ is invariant in all inertial reference frames.

Mathematical formulation: Minkowski space-time

When we specify an instant of time t and a point (x, y, z) in space, we are defining a point in space-time. We denote the coordinates of such a point in space-time by (x^0, x^1, x^2, x^3), where

$$x^0 \equiv ct,\ x^1 \equiv x,\ x^2 \equiv y,\ x^3 \equiv z. \tag{1.1}$$

These coordinates are the components of a four-dimensional vector in space-time. To shorten the notation we label such a vector by *Greek* indices μ, ν, so that we get

$$x^\mu = (x^0, \vec{x}) = (x^0, x^1, x^2, x^3). \tag{1.2}$$

The space-time metric tensor (Minkwoski metric) is given by

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.3}$$

This metric defines Minkowski space. Later on we will see which consequences the Minowski metric implies. Further more we distinguish between two classes of vectors. Vectors like x^μ are called *contravariant* and those like x_ν are known as *covariant* vectors. Contravariant indices are placed as superscripts and covariant indices as subscripts. The Minkowski metric determines how to get a covariant vector of a contravariant one, and vice versa:

$$x^\mu = \sum_{\nu=0}^{3} g^{\mu\nu} x_\nu \qquad\qquad x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu \tag{1.4}$$

Using the Minkowski metric explicitly we obtain that a covariant vector x_μ (derived from a contravariant vector x^ν) has negative space components. To be more compactly we use the *Einstein summation convention*, i.e. Greek indices which appear twice (once as a contravariant index in superscript and once as a covariant index in subscript) are summed over.

$$x_\mu = g_{\mu\nu} x^\nu \tag{1.5}$$

So Greek indices $\mu, \nu = 0, 1, 2, 3$ denote the coordinates of a four-dimensional vector in Minkwoski space, while Roman indices $a, b = 1, 2, 3$ denote the coordinates of a three-dimensional vector in euclidian position space. To complete the structure of Minkowksi space, we define the scalar product of two four-vectors a^μ and b^μ:

$$a^\mu b_\mu = a_\mu b^\mu = a^0 b^0 - \vec{a} \cdot \vec{b}, \tag{1.6}$$

where $\vec{a} \cdot \vec{b}$ denotes the usual dot product between three-vectors. Now we can consider the geometrical structure of Minkowski space, i.e. we take a look at the norm of a four-vector a^μ:

$$a^\mu a_\mu = \left(a^0\right)^2 - |\vec{a}|^2 \tag{1.7}$$

Here we see that the norm is *not* positive-definite. Four-vectors can be classed into three types, depending on the sign of their norm:

$$a^\mu a_\mu = \begin{cases} < 0 & a^\mu \text{ is space-like} \\ = 0 & a^\mu \text{ is light-like} \\ > 0 & a^\mu \text{ is time-like} \end{cases} \tag{1.8}$$

The metric tensor defines how to calculate the scalar product and hence, the length of a vector in Minkowski space:

$$s^2 = a_\mu a^\mu = a^\mu g_{\mu\nu} a^\nu = (ct)^2 - |\vec{a}|^2 \tag{1.9}$$

Often another parametrization, the *eigentime* τ of a particle is used:

$$\tau = \frac{s}{c} = \sqrt{(ct)^2 - |\vec{x}|^2} \tag{1.10}$$

Since in the frame where the particle is at rest in the origin ($\vec{x} = const. \equiv 0$), i.e. in the particle's own frame, τ is by defintion the time coordinate of the particle (we will see that τ is a Lorentz invariant quantity).
With these definitions, Einstein's postulate can be formulated in a mathematical way.

> Transformations between reference frames S, S' moving with different velocities (= Lorentz transformations) leave the length s of any 4-vector *on the light cone* (i.e. $a^\mu a_\mu = 0$) in Minkowski space invariant.

Proof:
Because of general consistency arguments (homogeneity of Minkowski space), no point in Minkowski space should have more special properties than any other point. This means in particular:

Fundamental postulate of special relativity
The length of *any 4-vector* (with respect to the metric g) is invariant under transformations from one inertial reference frame to another.

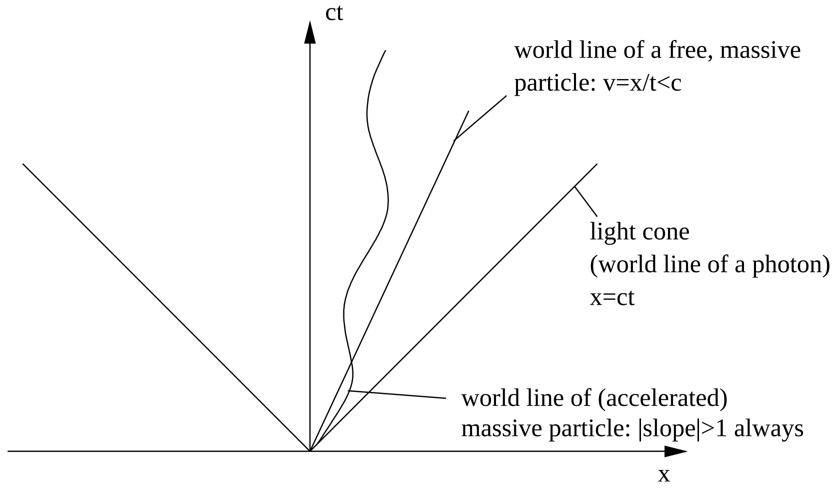

Figure 1.1: Classification of 4-vectors

The *Lorentz transformations* L describe the relationship between the coordinates x^μ of two reference frames which move relative to each other. Assume that the reference frame S' moves with velocity $-\vec{v}$ relative to the reference frame S ($\vec{v}||\widehat{x}$ without loss of generality). The Lorentz transformation L can be derived from the condition s^2 =invariant. One finds that only the primed coordinates x'^0 and x'^1 are changed while x^2 and x^3 are unchanged. If we set

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \quad \beta = \frac{v}{c} \tag{1.11}$$

the transformation L can be written as:

$$(L^\mu_\nu) = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.12}$$

It follows that

$$\det(L) = \pm 1. \tag{1.13}$$

The sign of the determinant leads to a classification of the Lorentz transformations.

$\det(L) = +1$	$L^0_0 \geq 1$	$\rightarrow$	proper orthochronous Lorentz transformation (no time reversal and no space inversion)
$\det(L) = -1$	$L^0_0 \geq 1$	$\rightarrow$	improper orthochronous Lorentz transformation (no time reversal)
$\det(L) = +1$	$L^0_0 \leq -1$	$\rightarrow$	proper nonorthochronous Lorentz transformation (no space inversion)
$\det(L) = -1$	$L^0_0 \leq -1$	$\rightarrow$	proper orthochronous Lorentz transformation

1.1.2 Relativistic generalization of quantum mechanics

The Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{x},t) = \left(-\frac{\hbar^2}{2m}\Delta + V(\vec{x})\right)\psi(\vec{x},t) \tag{1.14}$$

is obviously not relativistically covariant (form invariant) since it is of first order in the time variable x^0, but second order in the position variable $x^a, a = 1, 2, 3$. Hence a relativistic generalization of the wave equation is necessary. As will be seen in more detail in the course of the lecture, the combination of special relativity and quantum mechanics has two important consequences:

1. Relativistically, the mass - and hence the particle number - are not conserved any longer, but mass can be transformed into energy and vice versa, *if* there are interactions present. Therefore, any relativistic quantum theory must be a theory of variable particle number and obtains the character of a field theory (i.e. a theory with infinite number of degrees of freedom due to infinite particle number).

   $\rightarrow$ Number of particles as a new quantum number characterizing a quantumstate $\rightarrow$ *second quantization*

   $\rightarrow$ Particles can be created and destroyed by interaction $\rightarrow$ *particles and anti-particles*

   The non-conservation of the particle number recurs also in open many-particle systems at non-relativistic energy, for instance in superconductors where electrons form pairs and disappear in the superconducting condensate. Therefore, field theories for condensed matter systems use similar methods as relativistic field theories.

2. The motion of spin as an internal degree of freedom analogous (but not equivalent) to angular momentum follows necessarily from the combination of relativity and quantum mechanics.

In non-relativistic quantum mechanics we could to a large part use experimental results to obtain the Schrödinger equation (probabilistic nature of quantum mechanics, interference → wave functions, correspondence principle). In relativistic quantum mechanics this is much less the case, and we have to resort more and more to symmetry and consistency arguments, as will be seen below. This is a general feature (and strength) of modern theoretical physics.

1.2 Spin 0 bosons: Klein-Gordon equation

We seek a wave equation (at first for a free particle) which has *no internal degree of freedom* (spin 0). The state of the particle must be described by a *one-component* wave function (one-dimensional representation of the rotation group), i.e. it must also be a Lorentz scalar $\psi(x^\mu)$.

For a free particle the relativistic energy-momentum relation is

$$E = \sqrt{\vec{p}^2 c^2 + m^2 c^4}. \tag{1.15}$$

Plugging this into the Schrödinger equation (1.14) one would obtain

$$i\hbar \frac{\partial}{\partial t}\psi = \sqrt{-\hbar^2 c^2 \vec{\nabla}^2 + m^2 c^4}\psi, \tag{1.16}$$

where we have used

$$\vec{p} = \frac{\hbar}{i}\vec{\nabla}. \tag{1.17}$$

This form is problematic, because it involves gradients of the wave function of arbitrary order, as seen by expanding the square root. But this implies a *non-local* field theory, which would violate causality. Therefore we make the Ansatz that the equation above is quadratic in E:

$$-\hbar^2 \frac{\partial^2}{\partial t^2}\psi = \left(-\hbar^2 c^2 \vec{\nabla}^2 + m^2 c^4\right)\psi \tag{1.18}$$

This is the *Klein-Gordon equation*. Using the notation

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \qquad \partial^\mu = \frac{\partial}{\partial x \mu}, \tag{1.19}$$

where $x^\mu = (ct,\, x^1,\, x^2,\, x^3)$ and $x_\mu = (ct,\, x_1,\, x_2,\, x_3) = (ct, -x^1,\, -x^2,\, -x^3)$, we can rewrite the Klein-Gordon equation in a explicitly covariant form. Therefore we write

$$\left.\begin{array}{l} i\hbar\frac{\partial}{\partial(ct)} = i\hbar\frac{\partial}{\partial x^0} = i\hbar\partial_0 \\ i\hbar\frac{\partial}{\partial x^a} = -i\hbar\partial_a,\ a = 1,2,3 \end{array}\right\} p_\mu \longrightarrow i\hbar\partial_\mu = i\hbar \begin{pmatrix} \frac{\partial}{\partial(ct)} \\ -\vec{\nabla} \end{pmatrix} \tag{1.20}$$

and obtain the relativistically covariant wave equation:

$$-\hbar^2 \frac{\partial^2}{\partial(ct)^2}\psi(\vec{x},t) = (-\hbar^2\vec{\nabla}^2 + m^2c^2)\psi(\vec{x},t) \tag{1.21}$$

$$\Leftrightarrow \boxed{\left[\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right]\psi(x^\mu) = 0} \tag{1.22}$$

The operator

$$\Box := \partial_\mu\partial^\mu = \partial_0^2 - \vec{\nabla}^2 \tag{1.23}$$

is known as the *d'Alembert operator* which is a scalar under Lorentz transformation. The factor $\frac{mc}{\hbar}$ is the *inverse Compton wave length*.

1.2.1 Free solutions of the Klein-Gordon equation - Concept of antiparticles

The free Klein-Gordon equation has the plane-wave solutions

$$\psi(\vec{x},t) = \exp\left(-\frac{i}{\hbar}(Et - \vec{p}\vec{x})\right) = \exp\left(-\frac{i}{\hbar}p_\mu x^\mu\right), \tag{1.24}$$

with

$$E = \pm\sqrt{\vec{p}^2c^2 + (mc^2)^2}. \tag{1.25}$$

For a given momentum $\vec{p}$ there are always two solutions with equal, but opposite in sign energies. The solution with negative energy seems not to be meaningful for a *free* particle, since rest mass and kinetic energy are both non-negative. An arbitrary solution of the Klein-Gordon equation can be written as a 4-dimensional Fourierintegral ($\hbar k^\mu = p^\mu$)

$$\psi(x^\mu) = \int \frac{d^4k}{\sqrt{2\pi}^4}\,\delta^4\left(k_\mu k^\mu - \left(\frac{mc}{\hbar}\right)^2\right) A(k^\mu) e^{-ik_\mu x^\mu}. \tag{1.26}$$

Where $A(k^\mu)$ is arbitrary, i.e. the solutions with negative energy $k^0 < 0$ cannot be discarded.

Re-interpretation:
We choose always $E = +\sqrt{\vec{p}^2c^2 + (mc^2)^2} > 0$ and the solutions:

$$\psi^{(+)}(\vec{x},t) = e^{-\frac{i}{\hbar}(Et-\vec{p}\cdot\vec{x})} \tag{1.27}$$
$$\psi^{(-)}(\vec{x},t) = e^{-\frac{i}{\hbar}(-Et-\vec{p}\cdot\vec{x})} = e^{-\frac{i}{\hbar}(E(-t)-\vec{p}\cdot\vec{x})} \tag{1.28}$$

The solution with negative energy can be seen as that of a particle with *positive energy* propagating *backward in time*. A particle propagating backward in time is called *antiparticle*. The concept of particles and antiparticles will be developed further later on.

Remark:
The existence of $E < 0$ solutions is a consequence of second order in time derivatives. The wave equation must be of second order in $\partial/\partial\vec{x}$, since the kinetic term must be $\sim \mathcal{O}(\vec{p}^2)$ (more precisely: even order in $\vec{p}$), in order to fulfill space inversion symmetry (parity). The relativistic equation must therefore also be of second order in $\partial/\partial(ct)$. The relativistic formulation necessarily implies the existence of antiparticles.

1.2.2 Continuity equation: Violation of particle number conservation, conservation of energy

In order to derive a continuity equation describing the conservation of the quantum mechanical probability, one must again observe that the Klein-Gordon equation is of second order in time. Recall that the probability density for the Schrödinger equation is

$$\rho = \psi^*\psi \tag{1.29}$$

and the probability current is

$$\vec{\jmath} = \frac{\hbar}{2mi}\left(\psi^* \vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right). \tag{1.30}$$

They obey the continuity equation

$$\frac{\partial}{\partial t}\rho + \vec{\nabla}\cdot\vec{\jmath} = 0. \tag{1.31}$$

Now we determine the corresponding expressions for the Klein-Gordon equation. Multiplying the Klein-Gordon equation (1.22) from the left with ψ^* we get

$$\psi^*\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\psi = 0. \tag{1.32}$$

Taking the complex conjugate leads to

$$\psi\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\psi^* = 0. \tag{1.33}$$

By calculating the difference of these equations we finally get

$$\partial_\mu\left(\psi^*\partial_\mu\psi - \psi\partial^\mu\psi^*\right) = 0, \tag{1.34}$$

which can be written in a more familiar way as

$$\frac{\partial}{\partial t}\underbrace{\left[\frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial}{\partial t}\psi - \psi\frac{\partial}{\partial t}\psi^*\right)\right]}_{\rho} + \vec{\nabla}\cdot\underbrace{\frac{\hbar}{2mi}\left[\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right]}_{\vec{\jmath}} = 0. \tag{1.35}$$

While the current $\vec{\jmath}$ has the well known form, the probability density ρ identified from the continuity equation contains a time derivative, again a consequence of the second-order-in-time nature of the Klein-Gordon equation. As a consequence ρ is in general *not positive definite*! For this reason, the Klein-Gordon equation was at first rejected, since it did not seem to give positive probability. However, by Fourier transforming $\rho(\vec{x}, t)$ in time for plane waves, one obtains:

$$\boxed{\rho(\omega, \vec{x}) = \frac{\hbar\omega}{mc^2}\psi^*\psi} \tag{1.36}$$

This means that ρ must be interpreted as an *energy density* (in units of the rest mass), rather than a particle density. This is in accordance to the general observation that particle numbers are not conserved in relativistic theories, but rather the total energy.

1.3 Spin $\frac{1}{2}$ fermions: Dirac equation

1.3.1 Formulation of the Dirac equation and Dirac matrices

The Klein-Gordon equation was abandoned at first, because it seemed to yield negative probability density. The interpretation as an energy density came only later, after the concept of antiparticles had been introduced by Dirac (first for the case of spin $\frac{1}{2}$ fermions). In order to guarantee a positive definite probability density, we now seek a wave equation which is of first order in time and which is relativistically covariant (without spin in mind yet!). It must still obey the relativistic energy-momentum relation

$$\widehat{E} = \sqrt{\widehat{p}^2 c^2 + (mc^2)^2} \tag{1.37}$$

However,

1. the square root must be avoided (causility) and

2. the wave equation must be first order in the space derivatives as well, in order to be covariant.

Both requirements can be achieved by making the *Ansatz* that the squared total energy of a particle can be written as a *complete square* of a term involving the momentum only in first order. Hence this term must also be linear in the rest energy (Dirac). This means

$$E^2 = c^2 \vec{p}^2 + (mc^2)^2 \stackrel{!}{=} (c\alpha \cdot \vec{p} + \beta mc^2)^2, \tag{1.38}$$

with $\vec{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$, β (space independent). $\vec{\alpha}$, β must be determined such that the equality is fulfilled:

$$\begin{aligned} c^2(p_x^2 + p_y^2 + p_z^2) + m^2c^4 \quad = \quad & c^2(\alpha_x^2 p_x^2 + \alpha_y^2 p_y^2 + \alpha_z^2 p_z^2) + \beta^2 m^2 c^4 \\ & + c^2 p_x p_y(\alpha_x\alpha_y + \alpha_y\alpha_x) + c^2 p_y p_z(\alpha_y\alpha_z + \alpha_z\alpha_y) \\ & + c^2 p_z p_x(\alpha_z\alpha_x + \alpha_x\alpha_z) + mc^3[p_x(\alpha_x\beta + \beta\alpha_x) \\ & + p_y(\alpha_y\beta + \beta\alpha_y) + p_z(\alpha_z\beta + \beta\alpha_z)] \end{aligned} \tag{1.39}$$

It follows that $\alpha_x, \alpha_y, \alpha_z, \beta$ cannot be numbers, but must be matrices, the *Dirac matrices*, which obey the relations:

$$\alpha_i^2 = \beta^2 = \mathbb{1} \qquad i = x, y, z \tag{1.40}$$

$$\alpha_i\alpha_j + \alpha_j\alpha_i =: [\alpha_i, \alpha_j]_+ = 0 \qquad i, j = x, y, z \tag{1.41}$$

$$\alpha_i\beta + \beta\alpha_i =: [\alpha_i, \beta]_+ = 0 \qquad i = x, y, z \tag{1.42}$$

i.e. all α, β anticommute among each other, or in compact form:

Algebra of the Dirac matrices

$$\boxed{[M^\mu, M^\nu]_+ = 2\delta^{\mu\nu}\mathbb{1}} \tag{1.43}$$

with

$$M^\mu = \beta, \alpha_x\alpha_y\alpha_z. \tag{1.44}$$

Properties and determination of the Dirac matrices:

1. $\alpha_x, \alpha_y, \alpha_z, \beta$ are hermitean, in order for the Hamiltonoperator $H = c\vec{\alpha} \cdot \vec{p} + \beta mc^2$ to be hermitean.

2. All eigenvalues of M^μ are $\lambda = \pm 1$.

   Proof: (for $\mu = \nu$)

$$(M^\mu)^2 = \mathbb{1} \qquad \text{(projectors!)} \tag{1.45}$$

$$\text{in eigenbasis of } M^\mu \Rightarrow \lambda = \pm 1 \tag{1.46}$$

□

3. The trace of M^μ vanishes:

$$\mathrm{tr}(M^\mu) = 0, \mu = 0, 1, 2, 3 \tag{1.47}$$

Proof: (for $\mu \neq \nu$)

$$M^\mu M^\nu = -M^\nu M^\mu \tag{1.48}$$

$$\Leftrightarrow \underbrace{M^\mu M^\mu}_{\mathbb{1}} M^\nu = -M^\mu M^\nu M^\mu \tag{1.49}$$

$$\Rightarrow \mathrm{tr}(M^\nu) = -\mathrm{tr}(M^\mu M^\nu M^\mu) \tag{1.50}$$

$$= -\mathrm{tr}(M^\nu \underbrace{M^\mu M^\mu}_{=\mathbb{1}}) \tag{1.51}$$

$$= -\mathrm{tr}(M^\nu) \tag{1.52}$$

$$= 0 \tag{1.53}$$

□

4. The dimension d of the matrices $\alpha_x, \alpha_y, \alpha_z, \beta$ is *even*.

   Proof: (directly from 2. and 3.)

$$0 = \mathrm{tr}(M^\mu) = \sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} (\pm 1) \quad \Leftrightarrow \quad \text{d is even} \tag{1.54}$$

□

5. The $\alpha_x, \alpha_y, \alpha_z, \beta$ must be at least 4-dimensional ($d \geq 4$), since for $d = 2$ there are exactly three hermitean Pauli matrices obeying the Dirac algebra, and this set cannot be enlarged to include a fourth.

The Dirac matrices are not uniquely determined by their algebra, but can be chosen for $d = 4$ as

$$\boxed{\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}} \tag{1.55}$$

where $\vec{\sigma} = (\sigma_x,\, \sigma_y,\, \sigma_z)$ and

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.56}$$

are the Pauli matrices. This is the *standard representation* of the Dirac matrices $\vec{\alpha}, \beta$. Using the correspondence principle, we now have the *relativistic Dirac equation*:

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi(x^\mu) = (c\vec{\alpha}\cdot\vec{p}+\beta mc^2)\psi(x^\mu)} \tag{1.57}$$

with

$$\vec{p} = -i\hbar\vec{\nabla}. \tag{1.58}$$

Since $\alpha_x, \alpha_y, \alpha_z, \beta$ are 4-dimensional matrices, the $\psi(x^\mu)$ must be a 4-dimensional vector in an abstract representation space (in general: d-dimensional, d even).

- $\psi(x^\mu)$ is called *Dirac spinor*
- The fact that $\psi(x^\mu)$ is a d-dimensional object with *d even* implies that any relativistic particle described by the Dirac equation has a *two-fold* internal degree of freedom (i.e. comes in two flavors). This degree of freedom will be identified with particles and antiparticles.

The remaining two of the $2 \times 2 = 4 = d$ degrees of freedom will be identified with spin. The existence of particles/antiparticles as well as of spin follows from the relativistic formulation. Higher dimensional representation of the Dirac matrices are possible:

$$d = 2 \cdot (2S + 1) \qquad \text{even} \tag{1.59}$$

The factor 2 occurs because of the particle/antiparticle concept and S is the spin. This shows that the Dirac spinors are *not* 4-vectors.

1.3.2 Covariant form of the Dirac equation

The standard representation of the Dirac equation has non-trivial transformation properties under Lorentz transformations, since it represents the total energy of a particle, which is the 0-component of a Minkowski 4-vector. It is convenient to write the Dirac equation

$$i\hbar\frac{\partial}{\partial t}\psi = (c\vec{\alpha}\cdot\vec{p} + \beta mc^2)\psi \tag{1.60}$$

in explicitly covariant form. In order to give it a definite transformation behavior, we multiply by $\frac{1}{c}\beta$

$$(-i\hbar\beta\partial_0\psi + i\hbar\beta\alpha^i\partial_i\psi + mc)\psi = 0 \tag{1.61}$$

and define

$$\gamma^0 = \beta \tag{1.62}$$

$$\gamma^i = \beta\alpha^i, i = 1, 2, 3. \tag{1.63}$$

(Here α, β obey the Dirac algebra, but do not need to be determined explicitly)

$$\boxed{\left[-i\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right]\psi = 0} \tag{1.64}$$

Since $\frac{mc}{\hbar}$ is a Lorentz scalar (m is the *invariant* rest mass), so must

$$\gamma^\mu\partial_\mu = \gamma^\mu\frac{\partial}{\partial x^\mu} \tag{1.65}$$

be. Therefore $\gamma^\mu = \begin{pmatrix} \beta \\ \beta\vec{\alpha} \end{pmatrix}$ is a *contravariant 4-vector*, i.e. transforms under Lorentz transformation as $\gamma'^\mu = L^\mu_\nu\gamma^\nu$, and the equation above is explicitly Lorentz covariant.

Shorthand notation:

$$\gamma^\mu u_\mu = \gamma^0 u_0 - \vec{\gamma}\cdot\vec{u} =: \not{u} \qquad \text{(u slash)} \tag{1.66}$$

$$\boxed{\left(-i\not{\partial} + \frac{mc}{\hbar}\right)\psi = 0} \tag{1.67}$$

This is the Dirac equation in an explicitly covariant form.

Properties of the γ matrices (Dirac matrices):

$$\gamma^0 = \beta \quad \text{is hermitean and} \quad (\gamma^0)^2 = \mathbb{1} \tag{1.68}$$

$$\gamma^i,\ i = 1,2,3 \quad \text{are antihermitean and} \quad (\gamma^i)^2 = -\mathbb{1} \tag{1.69}$$

Proof:

$$(\gamma^i)^\dagger = (\beta\alpha^i)^\dagger = \alpha^i\beta = -\beta\alpha^i = -\gamma^i \tag{1.70}$$

$$(\gamma^i)^2 = \beta\alpha^i\beta\alpha^i = -\beta\beta\alpha^i\alpha^i = -\mathbb{1} \tag{1.71}$$

□

It follows that the Dirac γ matrices obey the algebra:

$$\boxed{[\gamma^\mu, \gamma^\nu]_+ = 2g^{\mu\nu}\mathbb{1}} \qquad \text{Dirac algebra} \tag{1.72}$$

In the standard representation of the $\vec{a}, \beta$ matrices the γ matrices read

$$\boxed{\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad i = 1,2,3} \tag{1.73}$$

An equivalent representation is obtained by the transformation

$$\widetilde{\gamma}^\mu = A\gamma^\mu A^{-1}, \qquad \text{(A non-singular, arbitrary)} \tag{1.74}$$

since this leaves the Dirac algebra valid. Hence the components of ψ are representation dependent and are not simple 4-vectors.

1.3.3 Continuity equation for Dirac spinors

Dirac spinor:

$$\psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_4 \end{pmatrix}, \qquad \psi^\dagger = (\psi_1^*, \ldots, \psi_4^*) \tag{1.75}$$

We want to derive a continuity equation for the density

$$\rho = \psi^\dagger \cdot \psi = \sum_{i=1}^{4} \psi_i^* \psi_i, \tag{1.76}$$

which involves

$$\frac{\partial}{\partial t}(\psi^\dagger \psi) = \left(\frac{\partial}{\partial t}\psi^\dagger\right)\psi + \psi^\dagger \left(\frac{\partial}{\partial t}\psi\right). \tag{1.77}$$

At first we consider the Dirac equation

$$i\hbar \left(\frac{\partial}{\partial t}\psi\right) = (-i\hbar c\vec{\alpha}\cdot\vec{\nabla} + \beta mc^2)\psi \tag{1.78}$$

and multiply by $\psi^\dagger$ from the left side and obtain

$$i\hbar\psi^\dagger \left(\frac{\partial}{\partial t}\psi\right) = (-i\hbar c\psi^\dagger\vec{\alpha}\cdot\vec{\nabla} + \beta mc^2\psi^\dagger)\psi. \tag{1.79}$$

Now we consider the Hermitean conjugate of the Dirac equation and multiply by ψ from the right side:

$$-i\hbar \left(\frac{\partial}{\partial t}\psi^\dagger\right)\psi = i\hbar\left(\vec{\nabla}\psi^\dagger\right)\cdot c\vec{\alpha}\psi + mc^2\psi^\dagger\beta\psi \tag{1.80}$$

Calculating the difference of these equations one gets:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) &= -i\hbar\left[\psi^\dagger(c\vec{\alpha})\cdot(\vec{\nabla}\psi) + (\vec{\nabla}\psi^\dagger)\cdot(c\vec{\alpha})\psi\right] && (1.81)\\ &= -\vec{\nabla}\cdot\left(\psi^\dagger(c\vec{\alpha})\psi\right) && (1.82)\end{aligned}$$

This leads to the continuity equation we aimed at:

$$\frac{\partial}{\partial t}\rho + \vec{\nabla}\cdot\vec{\jmath} = 0 \qquad \text{or} \qquad \partial_\mu j^\mu = 0, \tag{1.83}$$

where

$$(j^\mu) = \begin{pmatrix} c\rho \\ \vec{\jmath} \end{pmatrix}, \quad (\partial_\mu) = \begin{pmatrix} \frac{\partial}{\partial(ct)} \\ \frac{\partial}{\partial\vec{x}} \end{pmatrix}. \tag{1.84}$$

With

$\rho = \psi^\dagger\psi$	Dirac density	(1.85)
$\vec{\jmath} = \psi^\dagger(c\vec{\alpha})\psi = \vec{v}\rho$	Dirac current density	(1.86)
$\vec{v} := c\vec{\alpha}$	velocity operator in Dirac theory	(1.87)

Since the Dirac equation holds in any inertial reference frame (relativistic E-p relation), so does the continuity equation. Hence, j^μ is indeed a contravriant 4-vector, i.e. transforms according to

$$j'^\mu = L^\mu{}_\nu j^\nu. \tag{1.88}$$

The discussion up to now has shown that the relativistic formulation of quantum mechanics necessarily implies a fundamental reformulation of the wave equation.

Two possibilities:

1. The wave equation is second order in time, implying
   - positive and negative energy solutions $\longrightarrow$ particle/ antiparticle
   - non-conservation of particle number, conservation of energy $\longrightarrow$ bosons
2. The wave equation is first order in time, but multicomponent wave function, implying
   - spin, particles/ antiparticles
   - conservation of particle number $\longrightarrow$ fermions

1.3.4 Lorentz covariance of the Dirac equation: Lorentz transformation of the Dirac spinors ψ

We have derived the Dirac equation as a representation of the relativistic energy-momentum relation

$$E^2 = p^2c^2 + (mc^2)^2 \tag{1.89}$$

using the correspondence principle and postulating an equation linear in $E/c = i\hbar\,\partial/\partial(ct)$ and $\vec{p} = -i\hbar\vec{\nabla}$. The $E - \vec{p}$ relation is just an expression of the squared length of the energy-momentum 4-vector

$$p_\mu p^\mu = \left(\frac{E}{c}\right)^2 - \vec{p}^2 = (mc)^2, \tag{1.90}$$

which is invariant under Lorentz transformations. Therefore, the Dirac equation is by construction covariant (*shape invariant*) under Lorentz transformations.

We now calculate what this implies for the transformation of the Dirac spinor ψ under Lorentz transformation:

Although ψ is a 4-component object, it is not a 4-vector in Minkowski space. This is obvious, because there exist higher, even-dimensional representations of the γ matrices (see above). Rather, ψ is a spinor in an abstract representation space, as will be seen later.
Therefore, the transformation behaviour of ψ under Lorentz transformation is not a priori obvious.
Without restricting the generality we consider only special, orthochrone Lorentz transformations (no space inversion, no time inversion, without translations from inertial reference frame I to reference frame I'):

$$x^{\mu\prime} = L^{\mu}{}_{\nu} x^{\nu} \qquad \underline{x}' = \underline{\underline{L}}\,\underline{x} \tag{1.91}$$

(The matrix $\underline{\underline{\mathrm{L}}}$ is the representation of L in Minkowski space.)

The Dirac ψ transforms under L according to a linear transformation S(L):

$$\begin{aligned} \psi'(\underline{x}') &= S(L)\psi(\underline{x}) && (1.92)\\ &= S(L)\psi(\underline{\underline{L}}^{-1}(\underline{x}')) && (1.93) \end{aligned}$$

S(L) is the representation of L in the 4-dimensional spinor space (i.e. S(L) is a 4×4 matrix in spinor space).

Lorentz covariance of the Dirac equation:

$$\left(-i\gamma^{\mu}\partial_{\mu} + \frac{mc}{\hbar}\right)\psi(\underline{x}) = 0 \qquad \text{in frame } I \tag{1.94}$$

$$\left(-i\gamma^{\mu}\partial'_{\mu} + \frac{mc}{\hbar}\right)\psi'(\underline{x}') = 0 \qquad \text{in frame } I' \tag{1.95}$$

The above occurring partial derivatives are shorten by

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}} \qquad \text{and} \qquad \partial'_{\mu} = \frac{\partial}{\partial x^{\mu\prime}}. \tag{1.96}$$

The γ matrices can be *chosen* to be *Lorentz invariant*. This can be seen as follows:

γ^{μ} are determined by their algebra only up to a linear transformation,

$$\widetilde{\gamma}^{\mu} = A\gamma^{\mu}A^{-1} \qquad \text{(see above).} \tag{1.97}$$

Suppose γ^μ transforms under Lorentz transformation according to

$$\gamma^{\mu\prime} = T(L)\gamma^\mu T^{-1}(L). \tag{1.98}$$

Then we can always choose in the reference frame a new basis representation $\widetilde{\gamma}'$ for the matrices, such that:

$$\widetilde{\gamma}' = \gamma^\mu \qquad \Rightarrow \qquad A = T^{-1}(L) \tag{1.99}$$

□

Now we consider the behavior of the covariant derivative under Lorentz transformation.

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial x^{\nu\prime}}{\partial x^\mu}\frac{\partial}{\partial x^{\nu\prime}} = L^\nu{}_\mu \partial'_\nu \tag{1.100}$$

Which can be easily performed since:

$$x^{\nu\prime} = L^\nu{}_\mu x^\mu \tag{1.101}$$

$$\frac{\partial x^{\nu\prime}}{\partial x^\mu} = L^\nu{}_\mu \tag{1.102}$$

In (1.71) we obtained between ψ and ψ' the relation

$$S^{-1}\psi'(\underline{x}') = \psi(\underline{x}). \tag{1.103}$$

Plugging this into the Dirac equation (1.73) one finds that

$$\left(-i\gamma^\mu L^\nu{}_\mu \partial'_\nu + \frac{mc}{\hbar}\right) S^{-1}\psi'(\underline{x}') = 0. \tag{1.104}$$

By multiplying with S from the left side we get

$$-iSL^\nu{}_\mu \gamma^\mu S^{-1}\partial'_\nu \psi'(\underline{x}') + \frac{mc}{\hbar}\psi'(\underline{x}') = 0. \tag{1.105}$$

Lorentz covariance, comparing equation (1.83) with (1.74), we have the condition for $S(L)$:

$$SL^\nu{}_\mu \gamma^\mu S^{-1} = \gamma^\nu \tag{1.106}$$

or

$$\boxed{S^{-1}(L)\gamma^\nu S(L) = L^\nu{}_\mu \gamma^\mu} \tag{1.107}$$

Determining the representation $S(L)$ explicitly

We consider infinitesimal Lorentz transformations first. Representation in Minkwoski space:

$$L^{\nu}{}_{\mu} = \delta^{\nu}{}_{\mu} + \Delta\omega^{\nu}{}_{\mu}, \tag{1.108}$$

where $\Delta\omega^{\nu}{}_{\mu}$ is infinitesimal.

Diagonal elements:

$$\begin{aligned} L^{0}{}_{0} &= 1 && (1.109) \\ L^{a}{}_{b} &= \cos(\phi) \qquad \text{rotation} && (1.110) \\ L^{\nu}{}_{\nu} &= \cosh(\phi) \qquad \text{Lorentz boost} && (1.111) \\ \Rightarrow L^{\nu}{}_{\nu} &= 1 + O(\phi^2) && (1.112) \end{aligned}$$

Off-diagonal elements:

$$\begin{aligned} L^{\nu}{}_{\mu} \sim \left\{ \begin{array}{c} \sin(\phi) \\ \sinh(\phi) \end{array} \right\} &= O(\phi), \qquad \nu \neq \mu && (1.113) \\ \Rightarrow \Delta\omega^{\nu}{}_{\nu} &= O + O(\phi^2) && (1.114) \end{aligned}$$

$(\Delta\omega^{\nu}{}_{\mu})$ has at most 6 independent, non-zero elements to linear order in ϕ. Each one generates an independent Lorentz transformation, 3 rotations, 3 Lorentz boosts in 3 spatial dimensions.

Expansion of $S(L)$ in powers of $\Delta\omega^{\nu}{}_{\mu}$:

$$\begin{aligned} S &= \mathbb{1} + \tau && (1.115) \\ S^{-1} &= \mathbb{1} - \tau \qquad \text{with } \tau \text{ infinitesimal} && (1.116) \end{aligned}$$

Plugging this into equation (1.85) we find

$$\begin{aligned} (\mathbb{1} - \tau)\gamma^{\mu}(1 + \tau) &= \gamma^{\mu} + \gamma^{\mu}\tau - \tau\gamma^{\mu} + O(\tau^2) && (1.117) \\ &= \gamma^{\mu} + \Delta\omega^{\mu}{}_{\nu}\gamma^{\nu} && (1.118) \\ [\gamma^{\mu}, \tau] &= \Delta\omega^{\mu}{}_{\nu}\gamma^{\nu} && (1.119) \end{aligned}$$

This determines τ up to an additive matrix proportional to $\mathbb{1}$.

Proof:

If there were two solutions τ_1, τ_2, then

$$[\gamma^\mu, \tau_1 - \tau_2] = 0. \tag{1.120}$$

This yields

$$\tau_1 - \tau_2 = \alpha \cdot \mathbb{1} \qquad \text{with}\, \alpha \in \mathbb{R} \tag{1.121}$$

□

Unique determination of τ: norm invariance

- $S(L)$ must have the scalar product

$$\psi^\dagger(\underline{x})\psi(\underline{x}) = \sum_{\alpha=1}^{4} \psi^*_\alpha(\underline{x})\psi_\alpha(\underline{x}) \tag{1.122}$$

  in spinor space, which is *invariant.* (This implies that the density $\rho = \psi^\dagger(\underline{x})\psi(\underline{x})$ transforms only through the Lorentz contraction of the coordinates $\underline{x}$, and hence that ρ has the correct behavior under Lorentz transformations.)

  Thus we find

$$\det(S) = 1, \tag{1.123}$$

  up to an irrelevant phase factor.

$$1 = \det(S) = \det(\mathbb{1} + \tau) = \det(\mathbb{1}) + \mathrm{tr}(\tau) \tag{1.124}$$

$$= 1 + \mathrm{tr}(\tau) + O(\tau^2) \tag{1.125}$$

$$\mathrm{tr}(\tau) = 0 \tag{1.126}$$

The equations have the solutions

$$\tau = \frac{1}{8}\Delta\omega^\mu{}_{\nu'} g^{\nu'\nu}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu) = \frac{1}{8}\Delta\omega^\mu{}_{\nu'} g^{\nu'\nu}\,[\gamma_\mu, \gamma_\nu] \tag{1.127}$$

Representation of Lorentz transformation in Minkwoski and in Dirac space (infinitesimal transformation)

$$\begin{aligned} L^{\nu}{}_{\mu} &= \delta^{\nu}{}_{\mu} + \Delta\omega^{\nu}{}_{\mu} & (1.128) \\ S(L) &= \mathbb{1} + \frac{1}{8}\Delta\omega^{\mu}{}_{\nu'} g^{\nu'\nu} \left[\gamma_{\mu}, \gamma_{\nu}\right] & (1.129) \end{aligned}$$

Finite Lorentz transformations

- Transformations in Minkwoski space (4-vectors)

  A finite Lorentz transformation, generated by the real tensor $(\omega^{\nu}{}_{\mu})$ can be realized by N-times applying the infinitesimal transformation

$$\delta^{\nu}{}_{\mu} + \frac{\eta}{N}\omega^{\nu}{}_{\mu}, \qquad N \to \infty \tag{1.130}$$

  (where η is a free parameter which will be determined below).

$$L^{\nu}{}_{\mu} = \left[\lim_{N\to\infty} \left(\underline{\underline{1}} + \frac{\eta}{N}\underline{\underline{\omega}}\right)^{N}\right]^{\nu}{}_{\mu} = \left(e^{\eta\underline{\underline{\omega}}}\right)^{\nu}{}_{\mu} \tag{1.131}$$

  Example:

  Lorentz boost to a coordinate system moving in x direction with velocity $-\beta = -v/c$

$$\underline{\underline{\omega}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} =: (\tau^{01})_x \qquad 1^{01} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.132}$$

Therefore we get

$$
\begin{aligned}
(L^{\nu}{}_{\mu}) &= 1+\sum_{k=1}^{\infty}\frac{1}{k!}(\eta\tau_{x1})^k && (1.133)\\
&= 1-1^{01}+\sum_{k=0}^{\infty}\frac{1}{(2k)!}\eta^{2k}1^{01} && (1.134)\\
&\quad +\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\eta^{2ki+1}(\tau^{01})_x\\
(L^{\nu}{}_{\mu}) &= 1-1^{01}+\cosh(\eta)1^{01}+\sinh(\eta)(\tau^{01})_x && (1.135)\\
&= \begin{pmatrix} \cosh(\eta) & \sinh(\eta) & 0 & 0\\ \sinh(\eta) & \cosh(\eta) & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{pmatrix} && (1.136)
\end{aligned}
$$

From this result we see that the parameter η of the transformation is related to v/c by

$$
\begin{aligned}
\tanh(\eta) &= \frac{v}{c}=\beta, && (1.137)\\
\cosh(\eta) &= \gamma=\frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}, && (1.138)\\
\sinh(\eta) &= \beta\gamma. && (1.139)
\end{aligned}
$$

- Lorentz transformation of a Dirac spinor

$$
\begin{aligned}
S(L) &= \lim_{N\to\infty}[\mathbb{1}+\frac{\eta}{N}\frac{1}{8}\underbrace{(\omega^{\mu}{}_{\nu'}g^{\nu'\nu}\left[\gamma_\mu,\gamma_\nu\right])}_{(*)}]^N && (1.140)\\
&= \exp\left(\eta\cdot\frac{1}{8}(\omega^{\mu}{}_{\nu'}g^{\nu'\nu}\left[\gamma_\mu,\gamma_\nu\right])\right) && (1.141)
\end{aligned}
$$

[(∗) This product is a 4×4 matrix in spinor space and a scalar in Minkowksi space]

Example:

Lorentz boost by $-v/c$ in x direction

$$\omega^{\mu}{}_{\nu'} g^{\nu'\nu} [\gamma_\mu, \gamma_\nu] = \left[\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \right]^{\mu\nu} \times [\gamma_\mu, \gamma_\nu] \tag{1.142}$$

$$= 4\alpha \tag{1.143}$$

Where we have used

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.144}$$

and

$$[\gamma_\mu, \gamma_\nu] = [\gamma_\mu, \gamma_\nu]_+ - 2\gamma_\nu\gamma_\mu \tag{1.145}$$

$$= 2g_{\mu\nu}\mathbb{1} - 2\gamma_\nu\gamma_\mu \tag{1.146}$$

$$[\gamma_0, \gamma_1] = -2\gamma_1\gamma_0 \tag{1.147}$$

$$= -2 \begin{pmatrix} 0 & -\sigma_1 \\ \sigma_1 & 0 \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{1.148}$$

$$= -2 \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix} \tag{1.149}$$

$$= -2\alpha_1 \tag{1.150}$$

The transformation $S(L)$ then becomes here

$$S(L) = \exp\left(\frac{\eta}{2}\alpha_1\right) \tag{1.151}$$

$$= \sum_{k=0}^{\infty} \frac{1}{(2k)!} \left(\frac{\eta}{2}\right)^{2k} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \left(\frac{\eta}{2}\right)^{2k+1} \alpha_1 \tag{1.152}$$

$$S(L) = \cosh(\eta/2)\mathbb{1} + \sinh(\eta/2)\alpha_1 \tag{1.153}$$

$$S(L) = \begin{pmatrix} \cosh(\eta/2) & 0 & 0 & \sinh(\eta/2) \\ 0 & \cosh(\eta/2) & \sinh(\eta/2) & 0 \\ \sinh(\eta/2) & 0 & 0 & \cosh(\eta/2) \end{pmatrix} \tag{1.154}$$

This is a Lorentz boost in x direction by $\tanh(\eta) = v/c$.

- Lorentz transformation of bilinear forms of Dirac spinors

We investigate the behavior of bilinear forms like

$$j^\mu = c \underbrace{\psi^\dagger \gamma^0}_{\bar{\psi}} \gamma^\mu \psi = c\psi^\dagger \begin{pmatrix} \mathbb{1} \\ \vec{\alpha} \end{pmatrix} \psi \tag{1.155}$$

uner Lorentz transformation. To this end we need the relation between $S^\dagger$ and S^{-1}:

$$S^\dagger \gamma^0 = b\gamma^0 S^{-1} \tag{1.156}$$

with

$$b = \begin{cases} +1, & L^{00} \geq 1 \quad \text{(orthochronous)} \\ -1, & L^{00} \leq -1 \quad \text{(time reversal)} \end{cases} \tag{1.157}$$

(See F. Schwabl, Advanced quantum mechanics, p. 144 for the proof)

We introduce the shorthand notation

$$\bar{\psi} := \psi^\dagger \gamma^0, \tag{1.158}$$

since this combination appears regularly in bilinear forms like j^μ above.

Lorentz transformation

$$\psi' = S\psi \tag{1.159}$$

$$\bar{\psi}' = \psi^\dagger S^\dagger \gamma^0 = b\psi^\dagger \gamma^0 S^{-1} = b\bar{\psi} S^{-1} \tag{1.160}$$

Hence, we have the transformations:

- *Vector* (4-current):

$$j^\mu = c\bar{\psi}\gamma^\mu\psi \tag{1.161}$$

$$j^{\mu\prime} = cb\bar{\psi} \underbrace{S^{-1}\gamma\mu S}_{=L^\mu{}_\nu \gamma^\nu} \psi = cbL^\mu{}_\nu \bar{\psi}\gamma^\nu\psi = bL^\mu{}_\nu j^\nu \tag{1.162}$$

– *Lorentz scalar or pseudoscalar time reversal*

$$\bar{\psi}\psi = \psi^{\dagger}\gamma^{0}\psi \tag{1.163}$$

$$\bar{\psi}'\psi' = b\bar{\psi}S^{-1}S\psi = b\bar{\psi}\psi \tag{1.164}$$

1.3.5 Non-relativistic limit of the Dirac equation: Coupling to the electromagnetic field and the existence of spin

Standard representation:

$$i\hbar\frac{\partial}{\partial t}\psi = [c\vec{\alpha}\cdot\vec{p} + \beta mc^2]\psi \tag{1.165}$$

In order to "probe" the nature of the particles described by the Dirac equation, we need to couple the particles to some external field, especially the electromagnetic field. As in non-quantum mechanics, the coupling to the electromagnetic field arises from the postulate of *local* U(1) gauge invariance of the Dirac equation.

Phase transformation:

$$\psi(\underline{x}) \longmapsto \psi'(\underline{x}) = e^{-i\theta(\underline{x})}\psi(\underline{x}) \tag{1.166}$$

$$\frac{\partial}{\partial(ct)} \longmapsto \frac{\partial}{\partial(ct)} + i\frac{\partial\theta}{\partial(ct)} =: \mathcal{D}_t \tag{1.167}$$

$$\frac{\partial}{\partial\vec{x}} \longmapsto \frac{\partial}{\partial\vec{x}} + i\frac{\partial\theta}{\partial\vec{x}} =: \mathcal{D}_{\vec{x}} \tag{1.168}$$

or

$$p_\mu = i\hbar\frac{\partial}{\partial x^\mu} \longmapsto p_\mu - \hbar\frac{\partial\theta}{\partial x^\mu} = \Pi_\mu \tag{1.169}$$

Where the 4-vector x is denoted as

$$\underline{x} = \begin{pmatrix} ct \\ \vec{x} \end{pmatrix}. \tag{1.170}$$

Generalizing the gradient field $(\partial\theta/\partial x^\mu)$ to an arbitrary field, the electromagnetic 4-vector potential is introduced:

Kinetic 4-momentum:

$$\Pi_\mu = \left(p_\mu - \frac{q}{c}A_\mu\right) \tag{1.171}$$

Where

$$\frac{q}{c}A_\mu = \hbar\frac{\partial\theta}{\partial x^\mu} \tag{1.172}$$

and q is the charge of the particle.

Contravariant representation:

$$\begin{aligned}
\Pi^\mu &= g^{\mu\nu}\Pi_\nu = p^\mu - \frac{q}{c}A^\mu && (1.173)\\
c\Pi^0 &= i\hbar\frac{\partial}{\partial t} - q\Phi && (1.174)\\
\vec{\Pi} &= -i\hbar\frac{\partial}{\partial \vec{x}} - \frac{q}{c}\vec{A} \qquad \text{(kinetic momentum)} && (1.175)
\end{aligned}$$

Dirac equation in an electromagnetic field:

$$i\hbar\frac{\partial}{\partial t}\psi = [c\vec{\alpha}\cdot\underbrace{\left(\vec{p}-\frac{q}{c}\vec{A}\right)}_{\vec{\Pi}} + \beta mc^2 + q\Phi]\psi \tag{1.176}$$

We consider the non-relativistic limit $v \ll c$ for $\Phi = 0$ and obtain the Dirac equation

$$E\psi = \left[c\vec{\alpha}\cdot\vec{\Pi} + \beta mc^2\right]\psi. \tag{1.177}$$

Stationary solutions:

$$\psi(t) = \psi e^{-\frac{i}{\hbar}Et} \tag{1.178}$$

Defining components

$$\psi(\underline{x}) = \begin{pmatrix} \chi(\underline{x}) \\ \Phi(\underline{x}) \end{pmatrix}, \tag{1.179}$$

where χ, Φ are 2-component spinors.

$$\begin{aligned}
\begin{pmatrix} E - mc^2 & -c\vec{\sigma}\cdot\vec{\Pi} \\ -c\vec{\sigma}\cdot\vec{\Pi} & E + mc^2 \end{pmatrix}\begin{pmatrix} \chi \\ \Phi \end{pmatrix} &= 0 && (1.180)\\
\underbrace{(E - mc^2)}_{E_s}\chi - c\vec{\sigma}\cdot\vec{\Pi} &= 0 && (1.181)\\
\underbrace{(E + mc^2)}_{E_s+2mc^2}\Phi - c\vec{\sigma}\cdot\vec{\Pi}\chi &= 0 && (1.182)
\end{aligned}$$

With E_s the energy eigenvalue of the Schrödinger equation (shifted by mc^2 with respect to E).

Non-relativistic case, $v \ll c$ and small field $\left|\frac{q}{c}\vec{A}\right| \ll |\vec{p}|$:

$$E_s \approx \frac{\vec{p}^2}{2m} \ll mc^2 \tag{1.183}$$

$$|\vec{\Pi}| \approx m|\vec{v}| \ll mc \tag{1.184}$$

$$\Phi = \left(\frac{c\vec{\sigma}\cdot\vec{\Pi}}{E+mc^2}\right)\chi \tag{1.185}$$

$$E + mc^2 = E_s + 2mc^2 \approx 2mc^2 \tag{1.186}$$

$$c \cdot |\vec{\Pi}| \approx mc|\vec{v}| \tag{1.187}$$

$$\left|\frac{\Phi_i}{\chi_i}\right| \cong \frac{1}{2}\frac{v}{c} \ll 1, \qquad i = 1, 2 \tag{1.188}$$

$$\Phi \cong \frac{1}{2mc}(\vec{\sigma}\cdot\vec{\Pi})\chi \tag{1.189}$$

In the *non-relativistic limit*, χ and Φ are called the large and small components, respectively.

Eliminate the small component from the Dirac equation:

(1.148) in (1.140):

$$E_s\chi = \frac{1}{2m}(\vec{\sigma}\cdot\vec{P}i)(\vec{\sigma}\cdot\vec{P}i)\chi \tag{1.190}$$

Note that the product in the brackets is a scalar product in position space but the product of the terms in brackets is matrix product. by replacing

$$E_s \longrightarrow i\hbar\frac{\partial}{\partial t} \tag{1.191}$$

we finally obtain the *Pauli equation*:

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi = \frac{1}{2m}(\vec{\sigma}\cdot\vec{\Pi})(\vec{\sigma}\cdot\vec{\Pi})\psi} \tag{1.192}$$

Using the identity

$$(\vec{\sigma}\cdot\vec{u})(\vec{\sigma}\cdot\vec{v}) = \vec{u}\cdot\vec{v} + i\vec{\sigma}\cdot[\vec{u}\times\vec{v}] \tag{1.193}$$

(where $\vec{u}$ and $\vec{v}$ are arbitrary vectors) and

$$\begin{aligned}
[\vec{\Pi} \times \vec{\Pi}] &= \left(\vec{p} - \frac{q}{c}\vec{A}\right) \times \left(\vec{p} - \frac{q}{c}\vec{A}\right) && (1.194)\\
&= \underbrace{[\vec{p} \times \vec{p}]}_{=0} - \frac{q}{c}[\vec{A} \times \vec{p}] - \frac{q}{c}[\vec{p} \times \vec{A}] + \frac{q^2}{c^2}\underbrace{[\vec{A} \times \vec{A}]}_{=0} && (1.195)\\
&= -\frac{q}{c}[\vec{A} \times \vec{p}] - \frac{q}{c}[\vec{p} \times \vec{A}] + \frac{q}{c}[\vec{A} \times \vec{p}] && (1.196)\\
&= \frac{i\hbar q}{c}[\vec{\nabla} \times \vec{A}] && (1.197)\\
&= \frac{i\hbar q}{c}\vec{B} && (1.198)
\end{aligned}$$

one obtains:

$$i\hbar\frac{\partial}{\partial t}\chi = \Big[\frac{1}{2m}(\vec{p} - \frac{q}{c}\vec{A})^2 - \underbrace{\frac{q\hbar}{mc}}_{=2\mu_B} \cdot \frac{1}{2}\vec{\sigma} \cdot \vec{B}\Big]\chi \qquad (1.199)$$

This equation describes a 2-component spinor, whose 2 internal degrees of freedom couple to the magnetic field like an angular momentum $\frac{1}{2}\vec{\sigma}$ with a magnetic moment $\frac{q\hbar}{mc} = 2\mu_B = g\mu_B$.
The relativistic theory predicts spin $\frac{1}{2}$ with a Landé factor $g = 2$.

$$\mu_B = \frac{q\hbar}{mc} \qquad \text{Bohr magneton} \qquad (1.200)$$

1.3.6 Solutions of the Dirac equation for free particles

We set $\boxed{\hbar = c = 1}$ from now on. Consider the free Dirac equation:

$$(-i\gamma^\mu \partial_\mu + m)\psi = 0 \qquad (1.201)$$

The rest mass m is then the only scale in the problem. A length or energy scale is reconstructed from it in *a unique way* by multiplying with appropriate factors of $\hbar, c$.

$$\begin{aligned}
m &\longrightarrow \frac{mc}{\hbar} = \frac{1}{\lambda_{\text{Compton}}} && \text{inverse Compton wave length} && (1.202)\\
&\longrightarrow mc^2 = E_0 && \text{rest energy} && (1.203)
\end{aligned}$$

1. **Particles at rest:** $\vec{p}\psi = 0$

   Dirac equation:

$$(-i\gamma^0\partial_0 + m)\psi = 0 \tag{1.204}$$

   Putting in the γ matrix explicitly one obtains:

$$\left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}(-i\partial_0) + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} m\right]\psi = 0 \tag{1.205}$$

   This equation has normalized solutions:

$$\begin{aligned} \psi^{(+)}(\underline{x}) &= u_r(E=m, \vec{p}=0)e^{-imt} \qquad r = 1,2 && (1.206) \\ \psi^{(-)}(\underline{x}) &= v_r(E=m, \vec{p}=0)e^{+imt} \qquad m = \text{rest energy} && (1.207) \\ u_1(m,\vec{0}) &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad u_2(m,\vec{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} && (1.208) \\ v_1(m,\vec{0}) &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad v_2(m,\vec{0}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} && (1.209) \\ (p^\mu) &= \begin{pmatrix} m \\ \vec{0} \end{pmatrix} && (1.210) \end{aligned}$$

   There exist solutions with $E > 0(u_1, u_2)$ and solutions with $E < 0(v_1, v_2)$.

2. **Solutions with finite momentum $\vec{p}$ and total energy E**

   These solutions can be obtained from the 4 solutions at rest by applying a Lorentz transformation to an inertial frame with velocity $-\vec{v}$. Without loss of generality we choose $\vec{x}\|\vec{p}$: $p_x = p$
   The relation between p and v is obtained from the Lorentz transformation

of the 4-momentum:

$$\begin{aligned}
\begin{pmatrix} E \\ p \end{pmatrix} &= \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} m \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma m \\ \beta\gamma m \end{pmatrix} && (1.211)\\
\gamma &= \frac{1}{\sqrt{1-\beta^2}} && (1.212)\\
\beta &= \frac{v}{c} = v && (1.213)\\
p_x &= \beta\gamma m && (1.214)
\end{aligned}$$

The factor γm is the *relativistic mass enhancement*. For the Lorentz transformation in Minkowski space,

$$\left(L_x{}^\mu{}_\nu\right) = \begin{pmatrix} \cosh\eta & \sinh(\eta) & 0 & 0 \\ \sinh(\eta) & \cosh(\eta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.215}$$

The Dirac spinors transform with

$$S(L_x) = \cosh(\eta/2)\mathbb{1} + \sinh(\eta/2)\alpha_1, \tag{1.216}$$

i.e.

$$\begin{aligned}
u_1'(E,\vec{p}) &= \begin{pmatrix} \cosh(\eta/2) \\ 0 \\ 0 \\ \sinh(\eta/2) \end{pmatrix}, \quad u_2'(E,\vec{p}) = \begin{pmatrix} 0 \\ \cosh(\eta/2) \\ \sinh(\eta/2) \\ 0 \end{pmatrix} && (1.217)\\
v_1'(E,\vec{p}) &= \begin{pmatrix} 0 \\ \sinh(\eta/2) \\ \cosh(\eta/2) \\ 0 \end{pmatrix}, \quad v_2'(E,\vec{p}) = \begin{pmatrix} \sinh(\eta/2) \\ 0 \\ 0 \\ \cosh(\eta/2) \end{pmatrix} && (1.218)
\end{aligned}$$

$u_i', v_i', i = 1, 2$ can be expressed explicitly in terms of the energy E and the momentum p of the particle using

$$\cosh(\eta) = \gamma = \frac{E}{m} \qquad \sinh(\eta) = \beta\gamma = \frac{p_x}{m} \quad (E > 0) \tag{1.219}$$

and the theorems for the 1/2 arguments:

$$
\begin{aligned}
\cosh(\eta/2) &= \sqrt{\frac{1}{2}(\cosh(\eta)+1)} = \sqrt{\frac{1}{2m}(E+m)} && (1.220)\\
\sinh(\eta/2) &= \mathrm{sgn}(\eta)\sqrt{\frac{1}{2}(\cosh(\eta)-1)} = \mathrm{sgn}(p_x)\sqrt{\frac{1}{2m}(E_m)} && (1.221)\\
&= \mathrm{sgn}(p_x)\sqrt{\frac{1}{2m}\frac{E^2-m^2}{E+m}} = p_x\sqrt{\frac{1}{2m(E+m)}} && (1.222)\\
u_1(E,p_x) &= \begin{pmatrix} \sqrt{\frac{1}{2m}(E+m)}\chi_1 \\ p_x\sqrt{\frac{1}{2m(E+m)}}\chi_2 \end{pmatrix} && (1.223)\\
u_2(E,p_x) &= \begin{pmatrix} \sqrt{\frac{1}{2m}(E+m)}\chi_2 \\ p_x\sqrt{\frac{1}{2m(E+m)}}\chi_1 \end{pmatrix} && (1.224)\\
v_1(E,p_x) &= \begin{pmatrix} p_x\sqrt{\frac{1}{2m(E+m)}}\chi_2 \\ \sqrt{\frac{1}{2m}(E+m)}\chi_1 \end{pmatrix} && (1.225)\\
v_2(E,p_x) &= \begin{pmatrix} p_x\sqrt{\frac{1}{2m(E+m)}}\chi_1 \\ \sqrt{\frac{1}{2m}(E+m)}\chi_2 \end{pmatrix} && (1.226)
\end{aligned}
$$

with

$$
\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \qquad (1.227)
$$

As expected from the analysis of the non-relativistic limit for $v \ll c$

$$
\begin{aligned}
\text{the upper components of } u_1, u_2 \text{ are} &\approx 1,\\
\text{the lower components of } u_1, u_2 \text{ are} &\approx \frac{v}{2c} \ll 1,\\
\text{the upper components of } v_1, v_2 \text{ are} &\approx \frac{v}{2c} \ll 1,\\
\text{the lower components of } v_1, v_2 \text{ are} &\approx 1.
\end{aligned}
$$

The solutions for arbitrary direction of $\vec{p}$ are obtained by replacing

$$
\begin{aligned}
p_x\chi_2 &\longrightarrow \vec{\sigma}\cdot\vec{p}\chi_1 && (1.228)\\
p_x\chi_1 &\longrightarrow \vec{\sigma}\cdot\vec{p}\chi_2. && (1.229)
\end{aligned}
$$

This can be shown by applying a rotation in 3-dimensional position space to the above solutions for $\vec{p}\|\hat{x}$, or by direct solution of the Dirac equation. The space-time dependent phase factors of the solutions are Lorentz invariant and read:

$$e^{-imt'} = \underbrace{e^{-ip^{0\prime}x_0{}'}}_{\text{in rest frame}} = \underbrace{e^{-ip^\mu x_\mu}}_{\text{in moving frame}} \tag{1.230}$$

$$= e^{-i(Et-\vec{p}\vec{x})} \tag{1.231}$$

$$e^{+imt'} = e^{+ip^{0\prime}x_0{}'} = e^{+ip^\mu x_\mu} \tag{1.232}$$

$$= e^{-i(-Et+\vec{p}\vec{x})} \tag{1.233}$$

Hence, the free solutions of the Dirac equation with momentum $\vec{p}$ read:

$$\psi^{(+)}_{\vec{p},r}(\underline{x}) = u_r(E,\vec{p})e^{-i(Et-\vec{p}\vec{x})} \tag{1.234}$$

$$\psi^{(-)}_{\vec{p},r}(\underline{x}) = v_r(E,\vec{p})e^{-i(-Et+\vec{p}\vec{x})} \tag{1.235}$$

with the relativistic dispersion $E = +\sqrt{p^2+m^2}$. (Since by construction $H_D^2 = -i\hbar\frac{\partial}{\partial t}$, where H_D is the Dirac Hamilton operator.)

Orthogonality relations of the u_r, v_r

$$\bar{u}_r(\underline{k})u_s(\underline{k}) = \delta_{rs} \qquad r,s=1,2 \tag{1.236}$$

$$\vec{v}_r(\underline{k})v_s(\underline{k}) = -\delta_{rs} \tag{1.237}$$

$$\bar{u}_r(\underline{k})v_s(\underline{k}) = 0 \tag{1.238}$$

$$\bar{v}_r(\underline{k})u_s(\underline{k}) = 0 \tag{1.239}$$

1.3.7 Angular momentum and spin

We have derived the generates of the Lorentz group both in Minkowski and in spin space. Therefore, we can now identify the generates of 3-dimensional rotations, which are, by definition the components of angular momentum. The rotation in spinor space gives an additional contribution, identified with spin.

$$L^\mu{}_\nu = \delta^\mu{}_\nu + \Delta\omega^\mu{}_\nu \qquad \text{Minowski space} \tag{1.240}$$

$$S(L) = \mathbb{1} + \frac{1}{8}\Delta\omega^{\mu\nu}[\gamma_\mu,\gamma_\nu] \qquad \text{Spinor space} \tag{1.241}$$

A Dirac spinor is a variable both in Minkowski and in Spinor space. Rotation by $\Delta\vec{\varphi}, |\Delta\vec{\varphi}| =$ rotation angle.

$$\Delta\omega^{ij} = -\varepsilon^{ijk}\Delta\varphi^k \qquad (\Delta\omega^0{}_\mu = \Delta\omega^\mu{}_0 = 0) \tag{1.242}$$

Define:

$$\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu] \tag{1.243}$$

$$\sigma^{ij} = \sigma_{ij} = \varepsilon^{ijk}\Sigma^k \tag{1.244}$$

One can show that

$$\Sigma^k = \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \tag{1.245}$$

where σ^k are the Pauli matrices. Plugging the commutator relation above in equation (1.204) one obtains:

$$S(L) = \mathbb{1} - \frac{i}{4}\Delta\omega^{\mu\nu}\sigma_{\mu\nu} \tag{1.246}$$

And finally:

$$\psi'(\underline{x}') = L\left\{\psi(\underline{x})\right\} = S\psi(\underline{x}) = S\psi\left(L^{-1}\underline{x}'\right) \tag{1.247}$$

1.3.8 Physical interpretation of the solutions of the Dirac equation: The antiparticle concept

As seen in 1.2.6, the Dirac-Hamiltonian has positive as well as negative energy eigenvalues E, because β has the eigenvalues ± 1.
Free particles with a *negative* energy spectrum which is *not bounded from below* are a problem because of two reasons:

1. kinetic as well as rest energy should be positive
2. there is no stable ground state; i.e. a particle could go to infinitely low energies via scattering processes and the system would not be stable.

In this section we show at first that the $E < 0$ solutions cannot simply be discarded as unphysical, because they necessarily appear in physical systems in connection with the $E > 0$ solutions. We then give a physical interpretation of the $E < 0$ solutions, which remedies the paradox.

Wave packets and "Zitterbewegung"

The general wave packet is given by the expression:

$$\psi(\underline{x}) = \int \frac{d^4p}{(2\pi)^4}\,\delta(p_0^2 - E^2) \sum_{s=1,2} \left[(2\pi)2m\,\widetilde{b}(\underline{p}, s)W_s(\underline{p})\,e^{-ip_\mu x^\mu}\right] \tag{1.248}$$

wherein $\frac{d^4p}{(2\pi)^4}\,\delta(p_0^2 - E^2) = \frac{d^3p}{E}$ is a Lorentz-invariant measure and the Deltafunction guarantees, that ψ is a solution of the Dirac equation. The expansion coefficients are given by $2m\,\widetilde{b}(\underline{p}, s)$ and like always in this context is $E = \sqrt{|\vec{p}|^2 + m^2}$ (relativistic dispersion).

Furthermore we identify:

$$W_s(\underline{p}) = \begin{cases} u_s(\underline{p}), & p_0 > 0 \\ v_s(\underline{p}), & p_0 < 0. \end{cases} \tag{1.249}$$

Using $\delta(p_0^2 - E^2) = \frac{1}{2p_0}\left[\delta(p_0 - E) + \delta(p_0 + E)\right]$ we get

$$\psi(\underline{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{m}{E} \sum_{s=1,2} \left[b(\underline{p}, s)u_s(\underline{p})e^{-ip_\mu x^\mu} + d^*(\underline{p}, s)v_s(\underline{p})e^{+ip_\mu x^\mu}\right] \tag{1.250}$$

with

$$b(\underline{p}, s) = 2\pi\widetilde{b}(E, \vec{p}, s) \tag{1.251}$$

$$d^*(\underline{p}, s) = 2\pi\widetilde{b}(-E, \vec{p}, s). \tag{1.252}$$

a) **Superposition of states with $E > 0$ only**

For wave packets containing only $E > 0$ solutions no unexpected behavior appears: $\quad d^* \equiv 0, \quad p_0 = E.$

$$\psi^{(+)}(\underline{x}) = \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}}\frac{m}{E} \sum_{s=1,2} b(\underline{p}, s)u_s(\underline{p})e^{-ip_\mu x^\mu} \tag{1.253}$$

In this case the total current is given by

$$\vec{J}^{(+)} = \int \frac{d^3x}{(2\pi)^3} \vec{j}^{(+)}(\underline{x}) \tag{1.254}$$

$$= c\int \frac{d^3x}{(2\pi)^{\frac{3}{2}}} \psi^{(+)\dagger}(\underline{x})\vec{\alpha}\psi^{(+)}(\underline{x}) \tag{1.255}$$

$$= c\sum_s \int \frac{d^3p}{(2\pi)^3}\frac{m}{E}\left|b(\underline{p},s)\right|^2 \frac{\vec{p}}{E} \tag{1.256}$$

$$= \left\langle \frac{\vec{p}}{E} \right\rangle = \langle \vec{v}_G \rangle \tag{1.257}$$

which is the averaged group velocity over all states in the packet. The group velocity itself is given by $\vec{v}_G = \frac{\partial E}{\partial \vec{p}} = \frac{\partial\sqrt{|\vec{p}|^2+m^2}}{\partial \vec{p}} = \frac{\vec{p}}{E}$. In (1.256) we have used the orthogonality relations for u, v

b) **Wave packets containing $E > 0$ and $E < 0$ solutions**

Wave packets which are superpositions of $E > 0$ and $E < 0$ solutions of the Dirac equation cannot be avoided generally. In this case unexpected behavior appears, as the following example shows.

Time evolution of a free and localized wave packet which contains at $t = 0$ only $E > 0$ components and its width is characterized by a parameter d (total width: $4d$):

$$\psi(t=0,\vec{x}) = \frac{1}{(2\pi d^2)^{\frac{3}{4}}} e^{i\vec{x}\vec{k} - \frac{x^2}{(2d)^2}} w \tag{1.258}$$

wherein w is a pure $E > 0$ spinor, e.g. $w = \begin{pmatrix} \chi_1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

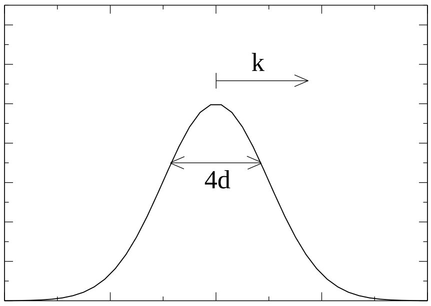

Figure 1.2: Gaussian wave packet

Decomposition of the Gaussian wave packets into plane waves in order to describe its time dependence using the Fourier transformation:

$$e^{i\vec{x}\vec{k}-\frac{x^2}{(2d^2)}} = \int \frac{d^3p}{(2\pi)^3}\left(\frac{4\pi d^2}{2\pi}\right)^{\frac{3}{2}} e^{-d^2\left(\vec{p}-\vec{k}\right)^2} \tag{1.259}$$

which is a Gaussian wave packet in $\vec{p}$-space shifted by $\vec{k}$.
Some further calculation delivers

$$b(\underline{p},s) = 2^{\frac{3}{2}}d^3e^{-d^2\left(\vec{p}-\vec{k}\right)^2}u_s^\dagger(vp)w \neq 0 \tag{1.260}$$

$$d^*(\underline{p},s) = 2^{\frac{3}{2}}d^3e^{-d^2\left(\vec{p}-\vec{k}\right)^2}v_s^\dagger(\underline{p})w \neq 0 \tag{1.261}$$

what shows that the general wave packet $\psi(t,\vec{x})$ contains both $E > 0$ and $E < 0$ components.

The $E < 0$ components are important for a wave packet with initially only $E > 0$ components, if the wave packet is localized much stronger than the particle's Compton-wavelength χ_C:

$$d \ll \chi_C = \frac{\hbar}{mc} \overset{\hbar,c=1}{=} \frac{1}{m}. \tag{1.262}$$

As we presupposed is $w = \binom{\chi_1}{0}$. Using the free particle spinors u_s, v_s we get the ratio

$$\frac{d^*(\underline{p})}{b(\underline{p})} = \frac{|\vec{p}-\vec{k}|}{E+m} \qquad \left(|\vec{k}| \ll 1\right). \tag{1.263}$$

Finally we use this ratio to show that the components with negative energy can be neglected for less strong localization of the particle but have to be considered for strong localization.

- weak particle localization: Wave packet extension $d \gg \frac{1}{m}$:

$$|\vec{p} - \vec{k}| \lesssim d^{-1} \ll m \qquad \Longleftrightarrow \qquad \frac{d^*}{b} \ll 1. \tag{1.264}$$

- strong particle localization: Wave packet extension $d \ll \frac{1}{m}$:

$$|\vec{p} - \vec{k}| \approx d^{-1} \gg m \quad \Longrightarrow \quad |\vec{p} - \vec{k}| \approx E \quad \Longleftrightarrow \quad \frac{d^*}{b} \approx 1 \tag{1.265}$$

Zitterbewegung:

In Physical expectation values such as $\langle \vec{x} \rangle$ the contributions with energies equal in size but opposite in sign lead to interference terms with oscillatory time dependence.
(It should be mentioned that in the modulus squared of terms with only $E > 0$ the time dependence cancels out.)

$$\langle \vec{x} \rangle = \int d^3x \, \psi^\dagger(\underline{x}) \vec{x} \psi(\underline{x}) \tag{1.266}$$

$$\frac{d}{dt} \langle \vec{x} \rangle = \frac{d}{dt} \int d^3x \, \psi^\dagger(0, \vec{x}) e^{+iHt} \vec{x} e^{-iHt} \psi(0, \vec{x}) \tag{1.267}$$

$$= \int d^3x \, \psi^\dagger(t, \vec{x}) i \underbrace{[H, \vec{x}]}_{-ic\vec{\alpha}} \psi(t, \vec{x}) \tag{1.268}$$

$$= \int d^3x \, \psi^\dagger(\underline{x}) c\vec{\alpha} \psi(\underline{x}) \quad = \quad \vec{J}(t) \tag{1.269}$$

$$J^i(t) = \int \frac{d^3p}{(2\pi)^3} \frac{m}{E} \left\{ \frac{p^i}{E} \sum_s \left[|b(\underline{p}, s)|^2 + |d(\underline{p}, s)|^2 \right] \right. \tag{1.270}$$

$$+ i \sum_{s,s'} \left[b^*(\underline{p}, s) d^*(\underline{p}, s) e^{2iEt} \overline{u}_s(\underline{p}) \sigma^{i0} v_{s'}(\underline{p}) \right. \tag{1.271}$$

$$\left. \left. - b(\underline{p}, s) d(\underline{p}, s) e^{-2iEt} \overline{v}_{s'}(\underline{p}) \sigma^{i0} u_s(\underline{p}) \right] \right\} \tag{1.272}$$

In this equation we see the time independent first part and a time dependent second contribution to the total current. The second part contains oscillations (Zitterbewegung) with frequencies of $2E > \frac{2mc^2}{\hbar} = 2 \times 10^{21} s^{-1}$ for electrons.

The Klein paradoxon: Dirac particle at a potential well (scattering problem in 1 dimension)

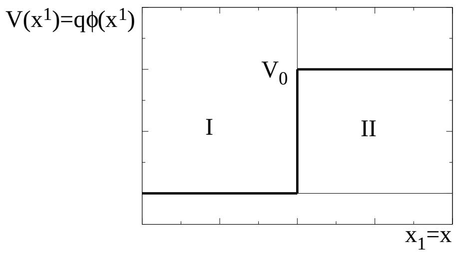

Figure 1.3: 1-dimensional scattering wall

Since $V(x^1) = const$ in the regions I and II respectively, there are "plane wave solutions" with energy E.

Region I:

We use unnormalized wave functions: Normalization factor $\sqrt{\frac{E+m}{2m}}$ is omitted!

Incident wave from left with $E > 0$:

$$\psi_{\text{in}}(\underline{x}) = e^{-iEt} e^{ipx} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \frac{p}{E+m} \end{pmatrix} \tag{1.273}$$

Ansatz for the reflected wave:

$$\psi_{\text{refl}}(\underline{x}) = e^{-iEt} \left\{ a e^{-ipx} \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ \frac{-p}{E+m} \end{pmatrix}}_{u_1} + b e^{-ipx} \underbrace{\begin{pmatrix} 0 \\ 1 \\ \frac{-p}{E+m} \\ 0 \end{pmatrix}}_{u_2} \right\} \tag{1.274}$$

with

$$E = +\sqrt{p^2 + m^2} > m, \qquad p = +\sqrt{E^2 - m^2}. \tag{1.275}$$

According to the interpretation *(Pauli equation)* the upper two components and the lower two components, respectively, represent the spin. The second term of

ψ_{refl} describes a free particle with opposite spin where we will find below that the coefficient b vanishes.

Region II:

Transmitted wave:

$$\psi_{\text{trans}}(\underline{x}) = e^{iEt}\left\{ce^{iqx}\begin{pmatrix}1\\0\\0\\\frac{q}{E-V_0+m}\end{pmatrix} + de^{iqx}\begin{pmatrix}0\\1\\\frac{q}{E-V_0+m}\\0\end{pmatrix}\right\} \tag{1.276}$$

with

$$q = \begin{cases} +\sqrt{(E-V_0)^2 - m^2}, & |E-V_0| \geq m \\ i\sqrt{m^2 - (E-V_0)^2}, & |E-V_0| \leq m \end{cases}. \tag{1.277}$$

We should remark that the normalized solution has no singularity at $E - V_0 + m$ since $q = 0$.

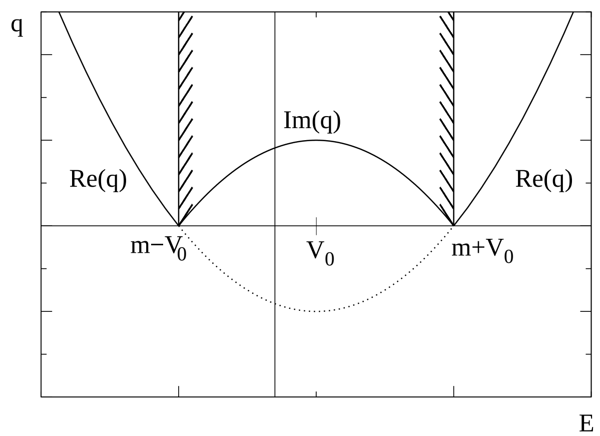

Figure 1.4: Transmitted wave solution

With the given information we find...

- real, i.e. propagating solutions for

$$E \geq m+V_0 > 0 \quad \text{or} \quad E \leq -m+V_0 \qquad (\text{if } E \geq m \,:\, V_0 \leq 2m) \tag{1.278}$$

- exponentially decaying solution for

$$-m + V_0 < E < m + V_0 \tag{1.279}$$

Since the Dirac equation is of first order in $\frac{\partial}{\partial x^\mu}$, the wavefunction must be continuous, not its derivative:

$$\psi_I(0) \stackrel{!}{=} \psi_{II}(0) \tag{1.280}$$

The conditions for the coefficients a, b, c, d result from the various components of ψ:

(i) $1 + a = c$

(ii) $b = d$

(iii) $-b\frac{p}{E+m} = d\frac{q}{E-V_0+m} \quad \stackrel{2.}{\Longleftrightarrow} \quad b = d = 0$ unless $V_0 = 0 \quad , \quad p = q$ (trivial)

(iv) $(1 - a) = rc \quad , \quad r = \frac{q}{p}\frac{E+m}{E-V_0+m}$

Therefore the spin is not flipped during the scattering process, because the coefficients b and d vanish.

Discussion of the solutions:

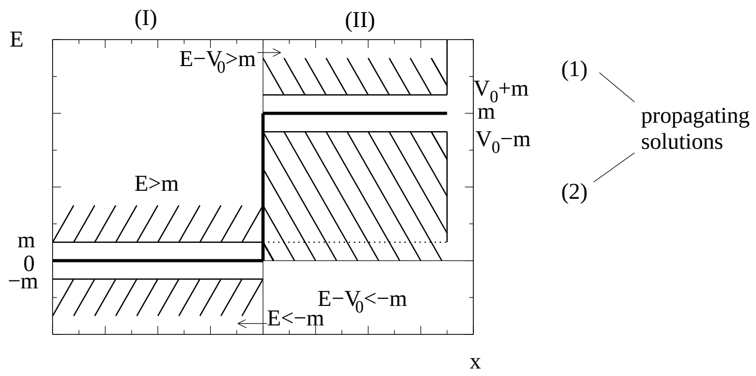

Figure 1.5: Complete solution in both regions

Propagating solutions exist, if the absolute value of the energy, measured relative to the potential level, is greater or equal to the rest energy m:

- region I : $\quad |E| > m$
- region II : $\quad |E - V_0| > m.$

We consider regular and propagating incident waves $E > m$ in region I and $V_0 > 0$:

Propagating transmitted solutions exist for

$$\begin{array}{llll} 1) & E - V_0 \geq m & \text{for any } \; V_0 > 0 & \text{regular, } E - V_0 > 0 \\ 2) & \left.\begin{array}{l} E - V_0 \leq -m \\ E \geq m \end{array}\right\} & \begin{array}{l} m \leq E \leq V_0 - m \\ \text{for } \quad E_0 > 2m \end{array} & \text{irregular, } E - V_0 < 0. \end{array}$$

Therefore we conclude, that for $E \geq m$ and $V_0 \geq 2m$ exists:

- a propagating transmitted solution in the well with *negative* energy compared to the potential V_0
- a propagating reflected solution with *positive* energy.

Transmitted and reflected currents:
The transmitted and the reflected current are orthogonal:

$$j_{\text{trans}} \perp j_{\text{refl}} \tag{1.281}$$

The continuity of ψ at $x = 0$ delivers:

$$c = \frac{2}{1+r} \qquad \text{and} \qquad a = \frac{1-r}{1+r}. \tag{1.282}$$

Furthermore we use the known expression

$$\vec{j} = c\psi^\dagger \vec{\alpha} \psi \qquad \text{with} \qquad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \tag{1.283}$$

where, in this case, c stands for the speed of light!

The information on the direction of $\vec{j}$ is encoded in the components of ψ and we get:

$$j_{\text{y}} = j_{\text{z}}, \quad j_{\text{x}} \neq 0. \tag{1.284}$$

$$\frac{j_{\text{trans}}}{j_{\text{in}}} = \frac{4r}{(1+r)^2} \quad , \qquad \frac{j_{\text{refl}}}{j_{\text{in}}} = \left(\frac{1-r}{1+r}\right)^2 \tag{1.285}$$

Therefore the current is conserved:

$$\frac{j_{\text{trans}}}{j_{\text{in}}} + \frac{j_{\text{refl}}}{j_{\text{in}}} = 1. \tag{1.286}$$

However, for $q, p > 0,\ m < E < V_0 - m,$ i.e. $\mathrm{V_0 > 2m,\ E - V_0 + m < 0,\ r < 0}$:

$$\frac{j_{\text{refl}}}{j_{\text{in}}} > 1, \qquad \frac{j_{\text{trans}}}{j_{\text{in}}} < 0 \tag{1.287}$$

For the group velocity of the transmitted wave in x-direction we find

$$\mathrm{v}_{\text{trans,x}} = \frac{dE}{dq} = \frac{d}{dq}\left(\sqrt{q^2+m^2}\right) \tag{1.288}$$

$$= \frac{2q}{+\sqrt{q^2+m^2}} = \frac{2q}{E - V_0} = -\frac{2q}{|E - V_0|} \tag{1.289}$$

what shows, that $\mathrm{v}_{\text{trans}}$ is opposite to q!

Interpretation:
A stationary solution in the parameter range $E > m$, $V_0 > 2m$ and an incident wave onto the well necessarily contains a particle current j_{trans} coming *out of* the well, or an antiparticle current going into the well.
The well does not provide any energy for particle-antiparticle pair production (since it is time independent and, hence, energy conserved).
Rather, the process corresponds to a partial *conversion* of particles into antiparticles at the well boundary, without change of the energy.

Hole theory: Interpretation of the $E < 0$ solutions in terms of antiparticles

We now have to find a way to solve the problems with the $E < 0$ solutions of the Dirac equation which are:

- $E < 0$ solutions are *necessary* as a part of the Dirac equation's complete set of eigenstates to describe the physical wave packet.
- Due to the $E < 0$ states matter would not be stable.

A *preliminary solution* first was given by Dirac. He supposed all the $E < 0$ states to be occupied by one electron each. In this sense a decay of $E > 0$ electrons into $E < 0$ states is prevented by the Pauli principle.

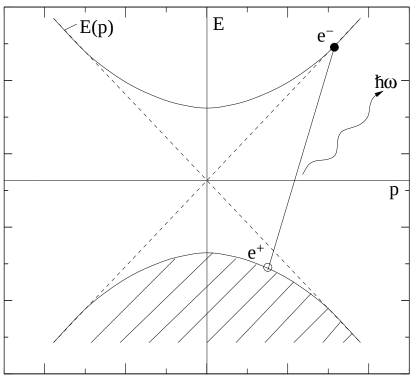

Figure 1.6: Excitation of a particle-antiparticle pair

Altogether the Vacuum is to be seen as an infinite sea of $E < 0$ particles (electrons), the *Dirac sea.*

Unfortunately the Dirac sea is not observable. Although it consists of an infinite number of charched particles, it does not create any force, because of translation invariance (infinite and homogeneous).

However: Problem of infinite mass

Consequently the Dirac sea can be observed only indirectly through its excitations.

Antiparticle concept
A *missing* electron in the Dirac sea with $E < 0,\ \vec{p},$ charge q is equivalent to a real particle with $E > 0,\ -\vec{p},$ charge $-q$; spin $-\sigma$, which is called an antiparticle. In case the particle is an electron, its antiparticle is the *positron.*

For the creation of particle-antiparticle pairs different ways a conceivable:

1. Time dependent fields $\hbar\omega > 2mc^2$

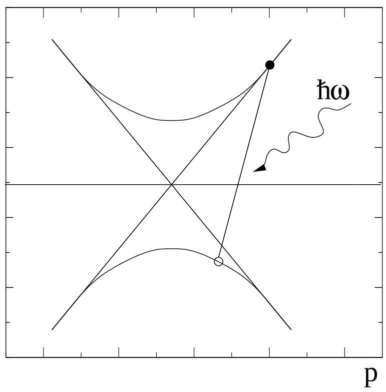

Figure 1.7: Creation of particle-antiparticle pairs

2. adiabatic creation of $E < 0$ *bound* states: *nuclear collisions*

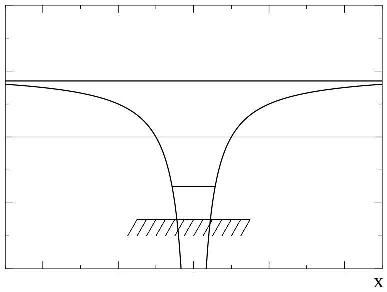

Figure 1.8: Adiabatic creation of bound states

Correspondence $E < 0$ solution - antiparticle
The Dirac sea equals the vacuum state and therefore $\psi^{(\pm)} \equiv 0$. In detail:

- $\psi^{(+)}(\underline{x})$: *occupied* $E > 0$, $\vec{p}$, σ state: particle $E > 0$, $\vec{p}$, σ, e
- $\psi^{(-)}(\underline{x})$: *unoccupied* $E < 0$, $\vec{p}$, σ state: antiparticle $-E > 0$, $-\vec{p}$, $-\sigma$, $-e$

Example: The Klein paradoxon

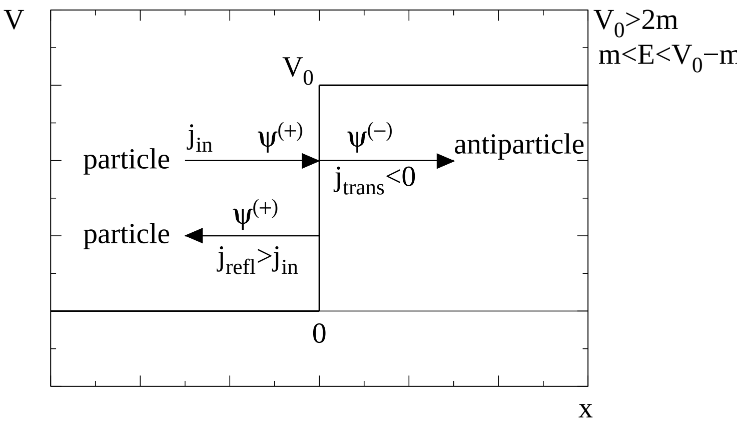

Figure 1.9: Klein paradoxon

Finally we give a summary of the problems behind the idea of the Dirac sea:

- Non-observability: translation invariance
- Infinite mass problem
- Many particle problem: The concept of particle-antiparticle pairs makes the Dirac sea idea inherently incomplete:

  Creation of many particles

  – Locality of the theory implies that the particle can be localized in the Dirac sea in an arbitrarily small length scale Δx.

  – In detail:

$$\Delta p \geq \hbar \, \Delta x^{-1} \tag{1.290}$$

$$\Delta E \approx c \, \Delta p \geq \frac{\hbar c}{\Delta x} \tag{1.291}$$

$$\Longrightarrow \quad \text{if} \quad \Delta x < \frac{1}{2} \frac{\hbar}{mc} = \frac{1}{2} \chi_{\mathrm{C}} \tag{1.292}$$

$$\Delta E \geq 2mc^2 \tag{1.293}$$

  which leads to particle-antiparticle production (In equation 1.217 we used the Compton wavelength χ_{C}).
  I.e. the locality of the relativistic theory necessarily implies that one must formulate it as a many-particle theory, which leads to the *field theory*.

1.3.9 Discrete symmetries of the Dirac equation

Charge conjugation C

The re-interpretation of $E < 0$ solutions in terms of antiparticles with opposite quantum numbers and charge suggests that there is a symmetry of the Dirac equation (i.e. a transformation which leaves the Dirac equation shape invariant) which transforms the $E < 0$ solutions for charge $e = -e_0 < 0$ into $E > 0$ solutions for charge $-e = e_0 > 0$.
This transformation will be called *charge conjugation C*.

For the construction of the transformation we start with the Dirac equation for the possible charges:

$$\begin{aligned} \text{for charge e:} \qquad [\gamma^\mu (i\partial_\mu - eA_\mu) - m]\,\psi &= 0 \qquad (1.294) \\ \text{for charge -e:} \qquad [\gamma^\mu (i\partial_\mu + eA_\mu) - m]\,\psi_\mathrm{C} &= 0 \qquad (1.295) \end{aligned}$$

We construct the transformation C, which transforms ψ into ψ_C: $\quad \psi_\mathrm{C} = \mathrm{C}\,\psi$.

The relative sign change between $i\partial_\mu$ and eA_μ is obtained by taking the complex conjugate of (1.245):

$$\widetilde{\mathrm{C}}\,[-\gamma^{\mu *}(i\partial_\mu + eA_\mu) - m]\,\widetilde{\mathrm{C}}^{-1}\widetilde{\mathrm{C}}\,\psi^* = 0. \qquad (1.296)$$

However this also introduces relative sign between the kinetic and the rest energy terms, and also changes $\gamma^\mu \longrightarrow \gamma^{\mu *}$.
In order to revert these changes, we look for a transformation $\widetilde{\mathrm{C}}$ such that

$$\widetilde{\mathrm{C}}\,\gamma^{\mu *}\,\widetilde{\mathrm{C}}^{-1} = -\gamma^\mu \qquad (1.297)$$

Note that such a transformation is allowed, since the γ^μ are determined by their algebra only up to similarity transformations.

Considering the γ^μ in the standard represantation,

$$\gamma^0 = \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad i = 1, 2, 3\,, \qquad (1.298)$$

(1.248) indicates that $\widetilde{\mathrm{C}}$

- interchanges the upper and lower components an
- transforms $\sigma^i \longrightarrow \sigma^{i*} \quad (i = 1, 2, 3)$.

This is performed by the matrix

$$\begin{aligned} \widetilde{\mathrm{C}} &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = i\gamma^2 \qquad \text{in standard representation} \qquad (1.299) \\ \widetilde{\mathrm{C}}^{-1} &= \widetilde{\mathrm{C}}\,. \qquad (1.300) \end{aligned}$$

Hence, we have

$$\boxed{\mathrm{C}\,\psi = \widetilde{\mathrm{C}}\,\psi^* = i\gamma^2\,\psi^*} \quad \text{(standard representation)} \tag{1.301}$$

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\mathbb{1} \tag{1.302}$$

$$(i\gamma^2)\gamma^{\mu *}(i\gamma^2) = [\underbrace{(-i\overbrace{\gamma^{2*}}^{-\gamma^2})}_{=i\gamma^2}\gamma^\mu(-i\gamma^{2*})]^* \tag{1.303}$$

$$= -[\gamma^2\gamma^\mu\gamma^2]^* \tag{1.304}$$

$$= -[\gamma^2(2g^{\mu 2}\mathbb{1} - \gamma^2\gamma^\mu)]^* \tag{1.305}$$

$$= [2\delta^{\mu 2}\gamma^2 - \gamma^\mu]^* \tag{1.306}$$

$$= \begin{cases} -\gamma^\mu & , \quad \mu \neq 2 \\ 2\gamma^{2*} - \gamma^{2*} = -\gamma^2 & , \quad \mu = 2 \end{cases} \tag{1.307}$$

where, in (1.256), we made use of $(\gamma^2)^2 = -\mathbb{1}$.

Altogether we conclude that C

- interchanges the upper and lower components of ψ, $\widetilde{\mathrm{C}}$
- transforms $E \longrightarrow -E$ in the exponential factor of ψ, $(\;\;)^*$
- inverts the momentum $\vec{p} \longrightarrow -\vec{p}$
- flips the spin ($\widetilde{\mathrm{C}}$)
- transforms from a solution for charge e to charge $-e$ (by construction).

Example: free Dirac spinor

$$\psi = \begin{pmatrix} 1 \\ 0 \\ \frac{p}{\sqrt{E+m}} \\ 0 \end{pmatrix} e^{-i(Et-pz)} \tag{1.308}$$

$$\mathrm{C}\,\psi = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{p}{\sqrt{E+m}} \\ 0 \end{pmatrix} e^{+i(Et-pz)} \tag{1.309}$$

$$= \begin{pmatrix} 0 \\ \frac{-p}{\sqrt{-E-m}} \\ 0 \\ 1 \end{pmatrix} e^{-i(-Et-(-p)z)} \tag{1.310}$$

where (1.258) is an $E > 0$ solution with $\vec{p}\,||\,\widehat{z}$, $p_z = p$ and spin $\uparrow$
and (1.259) is an $E < 0$ solution with $\vec{p}\,||\,\widehat{z}$, $p_z = -p$ and spin $\downarrow$.

Consequently the charge conjugation not only inverts the charge, but *completely* transforms a particle solution into an antiparticle solution and vice versa.

Since any spinor can be Fourier decomposed in terms of free spinors, this remains true for any solution for an arbitrary potential A^μ.

Time reversal (inversion of motion) T

The time reversal transformation T is defined by

$$\mathrm{T}: \quad t \longmapsto -t \tag{1.311}$$

i.e. reversal of the motion in all quantities:

$$\begin{aligned}
\mathrm{T}: \quad \vec{p} &\longmapsto -\vec{p} \quad \text{(momentum)} && (1.312)\\
\vec{L} = [\vec{r} \times \vec{p}] &\longmapsto -\vec{L} \quad \text{(angular momentum)} && (1.313)\\
\vec{S} &\longmapsto -\vec{S} \quad \text{(spin analogous to } \vec{L}) && (1.314)\\
\Phi(\vec{x}, t) &\longmapsto \Phi(\vec{x}, -t) = \Phi(\vec{x}, t) && (1.315)\\
\vec{A}(\vec{x}, t) &\longmapsto \vec{A}(\vec{x}, -t) = -\vec{A}(\vec{x}, t) && (1.316)\\
e &\longmapsto e \quad \text{(same for q in operators)} && (1.317)
\end{aligned}$$

$$\overset{(1.264),(1.265)}{\Longrightarrow} A^\mu(\vec{x}, t) \longmapsto A_\mu(\vec{x}, t) \qquad \text{(el.-magn. 4-potential)} \tag{1.318}$$

We construct the time reversal transformation from the form invariance of the Dirac equation:

$$[\gamma^\mu (i\partial_\mu - eA_\mu) - m]\, \psi = 0 \tag{1.319}$$

Time reversed Dirac equation:

$$[\gamma^\mu (-i\partial_\mu - eg^{\mu\nu} A_\nu) - m]\, \psi_{\mathrm{T}} = 0 \tag{1.320}$$

with $\psi_{\mathrm{T}} = \mathrm{T}\,\psi$ the time reversed spinor.
One can chose:

$$\boxed{\psi_{\mathrm{T}} = \mathrm{T}\,\psi = i\Sigma^2 \psi^*} \tag{1.321}$$

$$\text{with} \quad i\Sigma^2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = i \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}. \tag{1.322}$$

Note: For scalar particles is $\mathrm{C} \equiv \mathrm{T}$!

Parity (space inversion) P

$$\begin{aligned} \mathrm{P}: \quad \vec{x} &\longmapsto -\vec{x} \qquad t \longmapsto t\,,\ E \longmapsto E && (1.323)\\ \vec{p} &\longmapsto -\vec{p} && (1.324)\\ \vec{L} &\longmapsto +\vec{L} && (1.325)\\ \vec{S} &\longmapsto +\vec{S} && (1.326)\\ A^\mu &\longmapsto A_\mu \qquad \begin{cases} \Phi \longmapsto \Phi \\ \vec{A} \longmapsto -\vec{A} \end{cases} && (1.327) \end{aligned}$$

Parity transformation of the Dirac spinor
The phase factor $e^{-ip_\mu x^\mu}$ remains invariant, since

$$\begin{aligned} \mathrm{P}: \quad p_0 \longmapsto p_0 \quad &, \quad x^0 \longmapsto x^0 && (1.328)\\ p_i \longmapsto -p_i \quad &, \quad x^i \longmapsto -x^i \qquad , \ i = 1,2,3. && (1.329) \end{aligned}$$

From the parity symmetry of the Dirac equation one obtains, in a way analogous to C, T:

$$\boxed{\mathrm{P}\,\psi(\vec{x},t) = \gamma^0\,\psi(-\vec{x},t)} \qquad \text{with} \quad \gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \tag{1.330}$$

1.3.10 Separation of Spinor components into spin and particle-antiparticle degrees of freedom: The Foldy-Wouthuysen transformation

Problem: Physical meaning of the Dirac spinor components

Having derived the spin operator $\vec{\Sigma} = \frac{\hbar}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}$ and having introduced the particle- antiparticle degree of freedom, we are now able to discuss the meaning of the 4-spinor degrees of freedom for the general case, i.e. not only in the non-relativistic limits. This will ultimately lead to the systematic treatment of the relativistic corrections to the Schrödinger dynamics.

For particles *at rest* the Dirac spinors are in the standard representation:

$$\psi^{(+)}_{\vec{p}=0\,,\uparrow} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{-imt} \;,\quad \psi^{(+)}_{\vec{p}=0\,,\downarrow} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} e^{-imt} \quad ,E = m > 0 \tag{1.331}$$

$$\psi^{(-)}_{\vec{p}=0\,,\uparrow} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{+imt} \;,\quad \psi^{(-)}_{\vec{p}=0\,,\downarrow} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{+imt} \quad ,E = -m < 0 \tag{1.332}$$

These are eigenstates of the spin operator $\Sigma_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ as indicated, i.e. the upper two components describe the spin of particles, the lower two components the spin of antiparticles, in accordance to the Pauli equation. However, for free, moving particles one obtains after a Lorentz boost in the standard representation:

- For motion in z-direction (|| spin quantization axis $\widehat{z}$)

$$\psi^{(+)}_{\vec{p}||\widehat{z},\uparrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ 0 \end{pmatrix} e^{-ip_\mu x^\mu} \tag{1.333}$$

E $>$ 0

$$\psi^{(+)}_{\vec{p}||\widehat{z},\downarrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{p_z}{E+m} \end{pmatrix} e^{-ip_\mu x^\mu} \tag{1.334}$$

$$\psi^{(-)}_{\vec{p}||\widehat{z},\uparrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_z}{E+m} \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{+ip_\mu x^\mu} \tag{1.335}$$

E $<$ 0

$$\psi^{(-)}_{\vec{p}||\widehat{z},\downarrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ \frac{p_z}{E+m} \\ 0 \\ 1 \end{pmatrix} e^{+ip_\mu x^\mu} \tag{1.336}$$

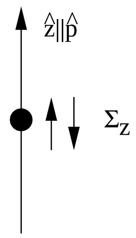

Figure 1.10: Motion in z-direction

These are eigenstates of the spin Σ_z with eigenvalues $s = \pm\frac{1}{2}\hbar$, as for particles at rest.

- For motion in x-direction ($\perp$ spin quantization axis $\widehat{z}$)

$$\psi^{(+)}_{\vec{p}\|\widehat{x},\uparrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \frac{p_z}{E+m} \end{pmatrix} e^{-ip_\mu x^\mu} \tag{1.337}$$

E $>$ 0

$$\psi^{(+)}_{\vec{p}\|\widehat{x},\downarrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ 1 \\ \frac{p_z}{E+m} \\ 0 \end{pmatrix} e^{-ip_\mu x^\mu} \tag{1.338}$$

$$\psi^{(-)}_{\vec{p}\|\widehat{x},\uparrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ \frac{p_z}{E+m} \\ 1 \\ 0 \end{pmatrix} e^{+ip_\mu x^\mu} \tag{1.339}$$

E $<$ 0

$$\psi^{(-)}_{\vec{p}\|\widehat{x},\downarrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_z}{E+m} \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{+ip_\mu x^\mu} \tag{1.340}$$

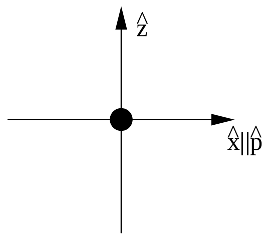

Figure 1.11: Motion in x-direction

The mixing of upper and lower components occurs because in the standard representation the Dirac equation has off-diagonal blocks:

$$\mathbb{1}\, i\frac{\partial}{\partial t}\psi = \begin{pmatrix} m + e\Phi & \vec{\alpha}\cdot\vec{\pi} \\ \vec{\alpha}\cdot\vec{\pi} & -m + e\Phi \end{pmatrix}\psi = \mathrm{H}\,\psi. \tag{1.341}$$

Since the representation of the γ^μ matrices is unique only up to an arbitrary similarity transformation

$$\widetilde{\gamma}^\mu = A\,\gamma^\mu\,A^{-1} \tag{1.342}$$

it is possible to find a representation such that $E < 0$ solutions have only upper components and $E < 0$ solutions have only lower components, i.e. a seperation of particle and antiparticle solutions in upper and lower components is possible.

The unitarity transformation separating particle and antiparticle components is defined such that it brings the Dirac equation to block diagonal form.
In general, this transformation is difficult and time dependent, since it requires transforming the exact, time dependent, solutions ψ.

Definition:

- Operator with only diagonal blocks: *even operator*
- Operator with only off-diagonal blocks: *odd operator*

These spinors are not eigenstates of Σ_z and have a *reduced* σ_z expectation value:

$$\langle \sigma_z \rangle = \pm \frac{m}{E}\frac{\hbar}{2} \quad ! \tag{1.343}$$

For velocity $|\vec{v}| \to c$, i.e. $E = \sqrt{p^2 + m^2} \to \infty$:

$$\langle \sigma_z \rangle_{p_x \to \infty} = 0 \qquad \text{similarly:} \quad \langle \sigma_y \rangle_{p_x \to \infty} = 0. \tag{1.344}$$

The spin polarization of particles moving at the speed of light, $|\vec{v}| \to c$, is either parallel or antiparallel to the direction of motion, what leads to *helicity*.

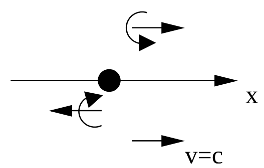

Figure 1.12: Helicity

From this example we see that, in the standard representation of the γ^μ matrices, for general, *moving* particles the spin and particle-antiparticle degrees of freedom are *mixed* among the 4-spinor components (i.e. $\psi^{(+)}$ has both upper and lower components, and has, in general, no definite z-component of spin).
Only in the non-relativistic limit the upper two components correspond to particles (spin $\uparrow$, $\downarrow$) and the lower two components to antiparticles (spin $\uparrow$, $\downarrow$).

Let us now have a closer look on equation (1.289):
For the normalization we calculate:

$$\psi^{(+)\dagger}_{\vec{p}\|\widehat{x},\uparrow} \cdot \psi^{(+)}_{\vec{p}\|\widehat{x},\uparrow} = \frac{E+m}{2m} \begin{pmatrix} 1 & 0 & 0 & \frac{p_x}{E+m} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \frac{p_x}{E+m} \end{pmatrix} \tag{1.345}$$

$$= \frac{E+m}{2m} \frac{(E^2+m^2+2Em) + \overbrace{p^2}^{E^2-m^2}}{(E+m)^2} \tag{1.346}$$

$$= \frac{E+m}{2m} \frac{2E^2+2Em}{(E+m)^2} = \frac{E}{m} \tag{1.347}$$

which cancels with the invariant measure in $\int d^4x$ for total laprobability.
Therefore we get:

$$\langle \sigma_z \rangle = \overbrace{\frac{m}{E}}^{\text{normalization}} \psi^{(+)\dagger}_{\vec{p}\|\widehat{x},\uparrow} \Sigma_z \psi^{(+)}_{\vec{p}\|\widehat{x},\uparrow} \tag{1.348}$$

$$= \frac{m}{E} \frac{E+m}{2m} \frac{(E^2+m^2+2Em) - p^2}{(E+m)^2} \tag{1.349}$$

$$= \frac{m}{E} \frac{E+m}{2m} \frac{2m^2+2Em}{(E+m)^2} = \frac{m}{E} \,. \tag{1.350}$$

Finally the *general spinor* is given by:

$$\psi^{(+)}_{\vec{p},\uparrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \end{pmatrix} e^{-ip_\mu x^\mu} \tag{1.351}$$

E > 0

$$\psi^{(+)}_{\vec{p},\downarrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ 1 \\ \frac{p_x+ip_y}{E+m} \\ \frac{p_z}{E+m} \end{pmatrix} e^{-ip_\mu x^\mu} \tag{1.352}$$

$$\psi^{(-)}_{\vec{p},\uparrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \\ 1 \\ 0 \end{pmatrix} e^{+ip_\mu x^\mu} \tag{1.353}$$

E < 0

$$\psi^{(-)}_{\vec{p},\downarrow}(\underline{x}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_x+ip_y}{E+m} \\ \frac{p_z}{E+m} \\ 0 \\ 1 \end{pmatrix} e^{+ip_\mu x^\mu} \tag{1.354}$$

The Foldy-Wouthuysen transformation

The separating transformation can be found systematically in powers of $\frac{E-m}{m}$, which, hence, leads to systematic relativistic corrections.

In general:

$A^{-1} =: e^{-i\mathrm{S}}$, S hermitean, time dependent; e^{-iS} decouples the upper and lower components.

$$\psi' = e^{+i\mathrm{S}}\,\psi \tag{1.355}$$

$$i\frac{\partial}{\partial t}\underbrace{e^{-i\mathrm{S}}e^{+i\mathrm{S}}}_{\mathbb{1}}\psi = \mathrm{H}\,\underbrace{e^{-i\mathrm{S}}e^{+i\mathrm{S}}}_{\mathbb{1}}\psi \qquad \text{Dirac equation} \tag{1.356}$$

$$\underbrace{i\frac{\partial}{\partial t}\left(e^{-i\mathrm{S}}\,\psi'\right)}_{i\left(\frac{\partial}{\partial t}e^{-i\mathrm{S}}\right)\psi' + ie^{-i\mathrm{S}}\frac{\partial}{\partial t}\psi'} = \mathrm{H}\,e^{-i\mathrm{S}}\,\psi' \qquad \Big| e^{+i\mathrm{S}}\,\cdot \tag{1.357}$$

$$i\frac{\partial}{\partial t}\psi' = \underbrace{e^{+i\mathrm{S}}\left(\mathrm{H} - i\frac{\partial}{\partial t}\right)e^{-i\mathrm{S}}}_{\mathrm{H}'}\,\psi'. \tag{1.358}$$

wherein H' is block-diagonal.

a) The Foldy-Wouthuysen transformation for free particles

$$\mathrm{H} = \vec{\alpha}\cdot\vec{p} + \beta m = \begin{pmatrix} \mathbb{1}m & \vec{\alpha}\cdot\vec{p} \\ \vec{\alpha}\cdot\vec{p} & -\mathbb{1}m \end{pmatrix} = \beta m + \underbrace{\mathcal{O}}_{\text{odd}} \tag{1.359}$$

For $\vec{B} = \begin{pmatrix} B_x \\ 0 \\ B_z \end{pmatrix}$ in the x-z plane analogous to diagonalizing $\mathrm{H} = \sigma_x B_x + \sigma_z B_z$

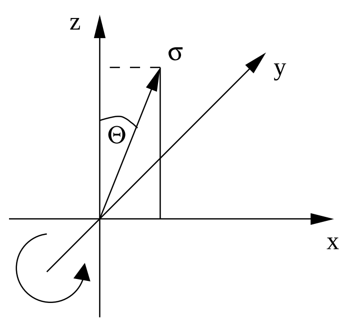

Figure 1.13: Rotation by an angle ϑ_0 about the y-axis

$(\rightarrow \quad e^{\frac{i}{2}\sigma_y\vartheta_0} = e^{-\frac{1}{2}\sigma_z\sigma_x\vartheta_0})$.

For the transformation we chose the following *Ansatz*:

$$\begin{aligned} e^{\pm i\mathrm{S}} &= e^{\pm \overbrace{\beta}^{\sigma_z} \overbrace{(\vec{\alpha}\cdot\widehat{p}\,)}^{\sigma_x} \vartheta(\vec{p})} = \mathbb{1}\cos\vartheta \pm \beta(\vec{\alpha}\cdot\vec{p}\,)\sin\vartheta & (1.360)\\ &= e^{\pm\beta\,\mathcal{O}\frac{\vartheta}{|\vec{p}|}} \quad \text{(time independent)} & (1.361)\\ \text{with} \quad \widehat{p} &= \frac{\vec{p}}{|\vec{p}|} & (1.362) \end{aligned}$$

where in (1.308) we used a Taylor expansion. Furthermore:

$$\begin{aligned} \mathrm{H}' &= e^{i\mathrm{S}}\,\mathrm{H}\,e^{-i\mathrm{S}} & (1.363)\\ &= e^{\beta(\vec{\alpha}\cdot\widehat{p})\vartheta}(\vec{\alpha}\cdot\widehat{p}+\beta m)(\mathbb{1}\cos\vartheta - \beta(\vec{\alpha}\cdot\widehat{p}\,)\sin\vartheta) & (1.364)\\ &= e^{\beta(\vec{\alpha}\cdot\widehat{p})\vartheta}\underbrace{(\mathbb{1}\cos\vartheta + \beta(\vec{\alpha}\cdot\widehat{p}\,)\sin\vartheta)}_{e^{\beta(\vec{\alpha}\cdot\widehat{p})\vartheta}}(\vec{\alpha}\cdot\widehat{p}+\beta m) & (1.365)\\ &= (\cos 2\vartheta + \beta(\vec{\alpha}\cdot\widehat{p}\,)\sin 2\vartheta)(\vec{\alpha}\cdot\widehat{p}+\beta m) & (1.366)\\ &= \vec{\alpha}\cdot\vec{p}\underbrace{\left(\cos 2\vartheta - \frac{m}{|\vec{p}|}\right)}_{\overset{!}{=}0} + \beta m\left(\cos 2\vartheta + \frac{|\vec{p}|}{m}\sin 2\vartheta\right) & (1.367)\\ \Longrightarrow \tan 2\vartheta &= \frac{|\vec{p}|}{m} & (1.368)\\ \sin 2\vartheta &= \frac{p}{\sqrt{m^2+p^2}} = \frac{p}{E} \quad , \quad \cos 2\vartheta = \frac{m}{\sqrt{m^2+p^2}} = \frac{m}{E} & (1.369) \end{aligned}$$

in (1.313) we used the anti-commutator $[\alpha\,,\beta]_+ = 0$.

We now put (1.317) into (1.315) to get the final result for H′:

$$\mathrm{H}' = \beta m\left(\frac{m}{E} + \frac{|\vec{p}|^2}{mE}\right) = \beta\frac{1}{E}\left(|\vec{p}|^2 + m^2\right) \tag{1.370}$$

which is block diagonal and wherein E is the energy eigenvalue and $\vec{p}$ is an operator.

For small momenta: $|\vec{p}| \ll m$

$$i\mathrm{S} = \beta\,\mathcal{O}\frac{\vartheta}{|\vec{p}|} \approx \beta\,\mathcal{O}\frac{1}{2m} \tag{1.371}$$

$$e^{\pm\beta(\vec{\alpha}\cdot\widehat{p})\vartheta} = \sum_{k=0}^{\infty}\frac{1}{(2k)!}\left[\beta(\vec{\alpha}\cdot\widehat{p})\right]^{2k} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\left[\beta(\vec{\alpha}\cdot\widehat{p})\right]^{2k+1} \qquad (1.372)$$

$$(\vec{\alpha}\cdot\widehat{p})^2 = \alpha^i p^i \alpha^j p^j = \underbrace{\frac{1}{2}\{\alpha^i, \alpha^j\}}_{\delta^{ij}} p^i p^j = |\vec{p}|^2 \qquad (1.373)$$

Foldy-Wouthuysen transformation in an electromagnetic field

$$\mathrm{H} = \vec{\alpha}\cdot\left(\vec{p} - e\vec{A}\right) + \beta m + e\phi = \beta m + \varepsilon + \mathcal{O} \qquad (1.374)$$

$$\text{with} \quad \varepsilon = \mathbb{1} e\phi \qquad \beta\varepsilon = \varepsilon \qquad (1.375)$$

$$\mathcal{O} = \vec{\alpha}(\vec{p} - e\vec{A}) \qquad \beta\mathcal{O} = -\mathcal{O}\beta \qquad (1.376)$$

where in (1.322) the βm-term is the $\mathcal{O}(m)$-term explicitly written in order to generate an $\frac{1}{m}$ expansion.

As in the field-free case we expect that in the non-relativistic limit the proper Foldy-Wouthuysen transformation is given by:

$$\boxed{i\mathrm{S} = \frac{\beta}{2m}\mathcal{O} \qquad \text{now with} \quad \mathcal{O} = \vec{\alpha}\cdot\left(\vec{p} - e\vec{A}\right).} \qquad (1.377)$$

The relativistic case can then be treated by repeating this transformation sufficiently often, leading to an expansion in $\frac{\left|\vec{p}-e\vec{A}\right|}{m}$.

$$\mathrm{H}' = e^{i\mathrm{S}}\left(\mathrm{H} - i\frac{\partial}{\partial t}\right)e^{-i\mathrm{S}} \qquad (1.378)$$

wherein $\frac{\partial}{\partial t}$ only acts on $e^{-i\mathrm{S}}$.

With the Baker-Campbell-Hausdorff identity:

$$e^A B e^{-A} = B + [A\,, B] + \frac{1}{2}[A\,, [A\,, B]] + \ldots + \frac{1}{k!}\underbrace{[A\,, [A\,, \ldots, [A\,, B]\ldots]]}_{\text{k commutators}} + \ldots \qquad (1.379)$$

one finds:

$$
\begin{aligned}
\mathrm{H}' = \mathrm{H} \quad &+ \quad i[\mathrm{S}\,,\mathrm{H}] + \frac{i^2}{2}[\mathrm{S}\,,[\mathrm{S}\,,\mathrm{H}]] && (1.380)\\
&+ \quad \frac{i^3}{6}[\mathrm{S}\,,[\mathrm{S}\,,[\mathrm{S}\,,\mathrm{H}]]] + \frac{i^4}{24}\overbrace{[\mathrm{S}\,,[\mathrm{S}\,,[\mathrm{S}\,,[\mathrm{S}\,,\mathrm{H}]]]]}^{\text{contains } \mathcal{O}(\frac{1}{m^3}) \text{ terms}} && (1.381)\\
&- \quad \dot{\mathrm{S}} - \frac{i}{2}[\mathrm{S}\,,\dot{\mathrm{S}}] - \frac{i^2}{6}[\mathrm{S}\,,[\mathrm{S}\,,\dot{\mathrm{S}}]] + \ldots && (1.382)
\end{aligned}
$$

With the commutation relations for β and $\vec{\alpha}$ the term of first order in $\frac{1}{m}$ is given by:

$$
i[\mathrm{S}\,,\mathrm{H}] = -\mathcal{O} + \frac{\beta}{2m}[\mathcal{O}\,,\varepsilon] + \frac{1}{m}\beta\mathcal{O}^2 \qquad (1.383)
$$

wherein the first $\mathcal{O}$ cancels the odd term to $\mathcal{O}\left(\frac{1}{m^0}\right)$.

Computing all commutators, disregarding the time dependence of the fields ($\dot{\mathrm{S}} = -i\frac{\beta}{2m}\dot{\mathcal{O}} \equiv 0$), we obtain:

$$
\begin{aligned}
\mathrm{H}' &= \beta m + \overbrace{\beta\left(\frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^2}\right) + \varepsilon}^{\varepsilon'} && (1.384)\\
&\quad \underbrace{-\frac{1}{8m^2}[\mathcal{O}\,,[\mathcal{O}\,,\varepsilon]] + \frac{\beta}{2m}[\mathcal{O}\,,\varepsilon] - \frac{\mathcal{O}^3}{3m^2}}_{\mathcal{O}'}\\
&= \beta m + \varepsilon' + \mathcal{O}' && (1.385)
\end{aligned}
$$

with $\mathcal{O}' \sim \mathcal{O}\left(\frac{1}{m}\right)$ only.

The odd terms can be determined to $\mathcal{O}\left(\frac{1}{m}\right)$ by another Foldy-Wouthuysen transformation, defined by

$$
i\mathrm{S} = \frac{\beta}{2m}\mathcal{O}' \quad \sim \quad \mathcal{O}\left(\frac{1}{m^2}\right) \,. \qquad (1.386)
$$

$$
\begin{aligned}
\mathrm{H}'' &= \beta m + \varepsilon' + \frac{\beta}{2m}[\mathcal{O}'\,,\varepsilon'] + \big(\text{even and odd terms of } \mathcal{O}\left(\frac{1}{m^4}\right)\big) && (1.387)\\
&= \beta m + \varepsilon' + \mathcal{O}''. && (1.388)
\end{aligned}
$$

A third Foldy-Wouthuysen transformation $i\mathrm{S} = \frac{\beta}{2m}\mathcal{O}'' \sim \mathcal{O}\left(\frac{1}{m^3}\right)$ blockdiagonalizes H to $\mathcal{O}\left(\frac{1}{m^3}\right)$:

$$\mathrm{H}''' = \beta\left(m + \frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3}\right) + \varepsilon - \frac{1}{8m^2}\left[\mathcal{O}\,,[\mathcal{O}\,,\varepsilon] + i\dot{\mathcal{O}}\right] + \mathcal{O}\left(\frac{1}{m^4}\right) \tag{1.389}$$

with only even powers of $\mathcal{O} \longrightarrow \mathrm{H}'''$ is even up to $\mathcal{O}\left(\frac{1}{m^3}\right)$.

Evaluating the powers of $\mathcal{O}$ in H''' one obtains finally

$$\begin{aligned}\mathrm{H}''' \;=\; & \beta\left(m + \frac{\left(\vec{p} - e\vec{A}\right)^2}{2m} - \frac{1}{8m^3}\left[\left(\vec{p} - e\vec{A}\right)^2 - e\vec{\Sigma}\cdot\vec{B}\right]^2\right) + e\phi \\ & -\frac{e}{2m}\beta\,\vec{\Sigma}\cdot\vec{B} - \frac{e}{8m^2}\vec{\Sigma}\cdot\left[\vec{\nabla}\times\vec{E}\right] \\ & -\frac{e}{4m^2}\vec{\Sigma}\cdot\left[\vec{E}\times\left(\vec{p} - e\vec{A}\right)\right] - \frac{e}{8m^2}\vec{\nabla}\cdot\vec{E}\,.\end{aligned} \tag{1.390}$$

In H''' the upper and lower components are decoupled.

For particle solutions ($E > 0$): $\psi' = \binom{\chi}{0}$:

$$\begin{aligned} i\frac{\partial\varphi}{\partial t} \;=\; & \Bigg\{m + e\phi + \frac{1}{2m}\left(\vec{p} - e\vec{A}\right)^2 - \frac{e}{2m}\vec{\sigma}\cdot\vec{B} \\ & -\frac{(|\vec{p}\,|^2)^2}{8m^3} - \frac{e}{4m^2}\vec{\sigma}\cdot\left[\vec{E}\times\left(\vec{p} - e\vec{A}\right)\right] - \frac{e}{8m^2}\vec{\nabla}\cdot\vec{E}\Bigg\}\varphi\end{aligned} \tag{1.391}$$

wherein up to the third term in order of their appearance one recognizes the particle's rest energy, its kinetic energy and furthermore the spin and Landé factor.
This Hamiltonian also contains the *relativistic corrections*:

1. **Relativistic mass correction** $\mathrm{H}_1 = -\frac{(|\vec{p}\,|^2)^2}{8m^3}$
   From expansion:

$$\frac{|\vec{p}\,|^2}{2\sqrt{m^2 + p^2}} = \frac{|\vec{p}\,|^2}{2m} - \frac{(|\vec{p}\,|^2)^2}{8m^3} + \dots \tag{1.392}$$

2. **Spin-orbit coupling**

   for central symmetric potential:

$$E = -\vec{\nabla}\phi(r) = -\frac{1}{r}\frac{\partial\phi}{\partial r}\vec{r} \quad , \quad \vec{A} = 0 \tag{1.393}$$

$$\vec{\sigma}\cdot\left[-\frac{1}{r}\frac{\partial\phi}{\partial r}\vec{r}\times\vec{p}\right] = -\frac{1}{r}\frac{\partial\phi}{\partial r}\left(\vec{\sigma}\cdot\vec{L}\right) \tag{1.394}$$

$$\mathrm{H}_2 = \frac{e}{4m^2}\frac{1}{r}\frac{\partial\phi}{\partial r}\left(\vec{\sigma}\cdot\vec{L}\right) \tag{1.395}$$

   what means that the Spin-orbit coupling is strongest for $\frac{\partial\phi}{\partial r}$ large, i.e. in heavy elements or near surfaces of metals.

3. **Darwin term**

$$\mathrm{H}_3 = -\frac{e}{8m^2}\vec{\nabla}\cdot\vec{E} \tag{1.396}$$

   H_3 is sensitive to local variations of the electric field, and can, therefore, be interpreted as due to the *Zitterbewegung* with an amplitude $\sim \lambda_{\mathrm{Compton}} = \frac{\hbar}{mc}$.

1.4 Further representations of the Dirac equation

1.4.1 Massless fermions (neutrinos)

For massless particles ($m = 0$) the Dirac equation reads

$$\gamma^\mu p_\mu \psi = 0 \qquad \text{or} \tag{1.397}$$

$$\mathbb{1}p^0 \psi = c\vec{\alpha}\cdot\vec{p}\psi \tag{1.398}$$

It is seen that for $m = 0$ only three matrices obeying the Dirac equation appear, $\{\alpha^i, \alpha^j\} = 2i\,\varepsilon^{ijk}\alpha^k$. Therefore a *2-dimensional* representation of the Dirac equation in terms of Pauli matrices is possible.
In the following we investigate what these two components correspond to physically, and why in the massless case the number of degrees of freedom is reduced from four to two.

From $E = \sqrt{p^2c^2 + (mc^2)^2}$ for free particles, it follows thatfor $m = 0$ the group velocity is ($E > 0$ solutions):

$$\mathrm{v}_{\vec{p}} = \frac{\partial E}{\partial \vec{p}} \overset{m=0}{=} \frac{\partial}{\partial \vec{p}}(|\vec{p}|c) = c\frac{\vec{p}}{|\vec{p}|} \,, \tag{1.399}$$

i.e. massless particles move at the speed of light.

But, as we have seen in section 1.3.10.1, particles moving at the speed of light $\left(\frac{|\vec{p}|}{E} = 1\right)$ cannot have a non-zero expectation value of spin components perpendicular to $\vec{p}$: $\langle \sigma_\perp \rangle = 0$ for $m = 0$.
It should be mentioned that the perpendicular components can fluctuate about their vanishing expectation value: $\langle \vec{\sigma}_\perp^2 \rangle > 0$ in general, just like a σ_z eigenstate has fluctuating σ_x, σ_y components.

As will be seen, the reduction of degrees of freedom from four to two corresponds to the absence of perpendicular (transversal) spin components in the massless case.
Because of the longitudinal[1] nature of spin for $m = 0$, it is useful to consider the projection of the spin operator $\vec{\Sigma}$ along the direction of $\vec{p}$:

$$\boxed{\widehat{\mathrm{h}}(\vec{p}) = \vec{\Sigma} \cdot \frac{\vec{p}}{|\vec{p}|} \qquad \text{helicity operator}} \tag{1.400}$$

$\widehat{\mathrm{h}}(\vec{p})$ obeys

$$\left(\vec{\Sigma} \cdot \frac{\vec{p}}{|\vec{p}|^2}\right)^2 = \sum_{i,j=1}^{3} \Sigma_i \Sigma_j \frac{p_i p_j}{|\vec{p}|^2} \tag{1.401}$$

$$= \sum_{i=1}^{3} \underbrace{\Sigma_i^2}_{=\mathbb{1}} \frac{p_i^2}{|\vec{p}|^2} = \mathbb{1}, \tag{1.402}$$

because for plane waves:

$$p_i p_j = \begin{cases} 0 & , \quad i \neq j \\ p_i^2 & , \quad i = j \end{cases} \tag{1.403}$$

[1]**Note:** For general, non-plane wave states the longitudinality holds for such components of their Fourier decomposition into momentum eigenstate.

I.e. $\widehat{\mathrm{h}}(\vec{p})$ has the eigenvalues *helicity* $h = \pm 1$.

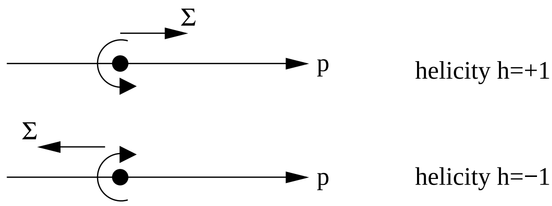

Figure 1.14: Eigenvalues of the helicity operator

As easily be verified $\left(\vec{\Sigma} \cdot \vec{p}\right)$ commutes with the Dirac Hamiltonian H:

$$\Big[\vec{\Sigma} \cdot \vec{p}, \underbrace{(c\vec{\alpha} \cdot \vec{p} + mc2)}_{\mathrm{H}}\Big] = 0 \tag{1.404}$$

so that $\left(\vec{\Sigma} \cdot \vec{p}\right)$ and H can be diagonalized simultaneously.
Therefore we seek a representation of the massless Dirac equation in terms of $\left(\vec{\Sigma} \cdot \vec{p}\right)$:

$$\gamma^\mu p_\mu \psi = 0 \qquad \text{or} \tag{1.405}$$

$$\vec{\gamma} \cdot \vec{p}\psi = \gamma^0 p^0 \psi \qquad \text{with} \quad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \tag{1.406}$$

$$\longrightarrow \vec{\Sigma} \cdot \vec{p} \qquad \text{and} \quad \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \tag{1.407}$$

Therefore $\vec{\gamma}$ is transformed into $\vec{\Sigma}$ by multiplying with $\begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$ which can be represented as

$$\begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} = \gamma^5\gamma^0 \quad , \quad \boxed{\gamma^5 \equiv \underbrace{i\gamma^0\gamma^1\gamma^2\gamma^3}_{\text{in general}} = \underbrace{\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}}_{\text{standard representation}}} \tag{1.408}$$

Multiplying the Dirac equation, $m = 0$, with $\gamma^5\gamma^0$ from left delivers finally:

$$\boxed{\vec{\Sigma} \cdot \vec{p}\psi = \gamma^5 p^0 \psi} \tag{1.409}$$

Since $\left[\vec{\Sigma}\cdot\vec{p},\gamma^5\right]=0$ these operators can be diagonalized simultaneously, which leads to the following

Definition:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \qquad \text{is called the } \textit{chirality operator.} \tag{1.410}$$

Equation (1.355) is brought to block diagonal form by applying the unitary transformation

$$\psi^{\text{ch}} = U\psi \qquad \text{with} \quad U = \frac{1}{\sqrt{2}}\big(\mathbb{1}+\gamma^5\big) \tag{1.411}$$

$$\gamma^{\mu\,\text{ch}} = U\psi U^{-1} \quad \text{etc.,} \tag{1.412}$$

leading to *two independent 2-component* equations, the *Weyl equations*:

$$\big(p_0 - \vec{\sigma}\cdot\vec{p}\big)\psi_1^{\text{ch}} = 0 \tag{1.413}$$

$$\big(p_0 + \vec{\sigma}\cdot\vec{p}\big)\psi_2^{\text{ch}} = 0 \tag{1.414}$$

with the 2-component helicity operator $\vec{\sigma}\cdot\frac{\vec{p}}{|\vec{p}|}$.
$\psi^{\text{ch}} = \frac{1}{\sqrt{2}}\big(\mathbb{1}+\gamma^5\big)$ is called the *chiral representation.*

The Weyl equations are, each one by itself, not parity invariant ($\vec{\sigma}\cdot\vec{p}\overset{\text{P}}{\longmapsto}-\vec{\sigma}\cdot\vec{p}$), but transform mutually into each other. The consequence is, that if a particle is described by one of the Weyl equations, parity is not an allowed transformation, i.e. the particle occurs either with helicity $h=+1$ or $h=-1$ in nature, but cannot occur with either one. For this reason, the Weyl equations were historically not considered further at first.
However, experiments showed that the only known massless fermion, the *neutrino*, does indeed occur in nature only with helicity $h=-1$ and are called *lefthanded.* With the same background *antineutrinos* with $h=+1$ are called *righthanded.* This shows that neutrinos are indeed described by the simplest possible, 2-dimensional representation of massless fermions.

Outlook:
Experiments by *Wu et al.* showed that parity is *not conserved* by the weak interaction (SU(2) gauge interaction), i.e. the weak interaction couples the two Weyl

equations. It is believed that this interaction also generates the small, but finite mass[2] of neutrinos via vacuum fluctuations.

[2]observed first time in 2002; *Kamiokande*, Japan

1.4.2 Majorana representation

Definition: Majorana representation of the γ matrices:

- γ^0 imaginary, antisymmetric
- γ^k imaginary, symmetric

Example:

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \qquad \gamma^2 = i\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{1.415}$$

$$\gamma^1 = i\begin{pmatrix} 0 & \sigma^1 \\ \sigma^1 & 0 \end{pmatrix}, \qquad \sigma^3 = i\begin{pmatrix} 0 & \sigma^3 \\ \sigma^3 & 0 \end{pmatrix} \tag{1.416}$$

It follows that in Majorana representation the free Dirac equation is real:

$$\left[i\gamma^\mu \partial_\mu - m\right]\psi = 0\,. \tag{1.417}$$

The charge conjugation is in Majorana representation:

$$\psi_{\mathrm{C}} = \psi^* \quad , \quad \text{since} \tag{1.418}$$

$$\left[\gamma^\mu(i\partial_\mu - eA_\mu) - m\right]\psi = 0 \qquad |\,^* \tag{1.419}$$

$$\left[-\gamma^\mu(-i\partial_\mu - eA_\mu) - m\right]\psi^* = 0 \tag{1.420}$$

$$\left[\gamma^\mu(i\partial_\mu + eA_\mu) - m\right]\psi^* = 0 \qquad \psi^* = \psi_{\mathrm{C}} \tag{1.421}$$

While the Majorana representation is completely general, it is especially useful for neutral fermions:
ψ can be chosen real ($\psi^* = \psi$) in this case and is called *Majorana spinor*. Therefore a Majorana spinor has half as many degrees of freedom than a Dirac spinor.