Chapter 2

Relativistic Wave Mechanics

Goal: construct wave equation that is compatible with special relativity, i.e., Lorentz transformations between inertial systems.

'Compatible' or 'covariant': equations have the same form for all observers, in all inertial systems.

Related to symmetry: rotations, Lorentz transformations as 'symmetry group' that leaves certain quantities unchanged. Length of a vector, Minkowski distance. Technically: representations of these symmetry groups.

2.1 Relativistic classical mechanics

Relativistic mechanics: action for point particle

$$\begin{aligned} S &= -mc\int \mathrm{d}\tau - e\int \mathrm{d}x^\mu A_\mu(x) \\ &= -mc\int \mathrm{d}t\sqrt{c^2-\dot{\mathbf{x}}^2} - e\int \mathrm{d}t\,(\phi(x) - \dot{\mathbf{x}}\cdot\mathbf{A}(x)) \end{aligned} \tag{2.1}$$

with proper time (*Eigenzeit*) $\mathrm{d}\tau = \mathrm{d}t/\gamma$ and vector potential $A_\mu = (\phi/c, -\mathbf{A})$, evaluated at the world point x with coordinates $x^\mu = (ct, \mathbf{x})$.

Discussion: 4-momentum

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = \frac{mc\dot{\mathbf{x}}}{\sqrt{c^2-\dot{\mathbf{x}}^2}} + e\mathbf{A} \tag{2.2}$$

differs from 'kinetic momentum' $m\dot{\mathbf{x}}$ for two reasons.

Energy

$$E = \dot{\mathbf{x}}\mathbf{p} - L = \frac{mc^3}{\sqrt{c^2-\dot{\mathbf{x}}^2}} + e\phi \tag{2.3}$$

combine into a ("covariant") 4-vector

$$p_\mu = (E/c, -\mathbf{p}) = m\frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} + eA_\mu \tag{2.4}$$

(note again that $1/\mathrm{d}\tau = \gamma/\mathrm{d}t = c/(\mathrm{d}t\,\sqrt{c^2-\dot{\mathbf{x}}^2})$.

Minkowski norm or dispersion relation

$$(p_\mu - eA_\mu)(p^\mu - eA^\mu) = \frac{(E-e\phi)^2}{c^2} - (\mathbf{p} - e\mathbf{A})^2 = (mc)^2 \tag{2.5}$$

Equation of motion:

$$\frac{\mathrm{d}p_\mu}{\mathrm{d}\tau} = eF_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \tag{2.6}$$

with Faraday tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.7}$$

covariant version of the Coulomb-Lorentz force

2.2 Klein-Gordon equation

Cook book recipe

First the usual recipe, as in the Schrödinger theory. From the classical dispersion relation

$$E - e\phi = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} \tag{2.8}$$

right hand side is kinetic energy $\frac{1}{2}m\dot{\mathbf{x}}^2$ with $m\dot{\mathbf{x}} = \mathbf{p} - e\mathbf{A}$
make the replacements

$$E \mapsto \mathrm{i}\hbar\partial_t, \qquad \mathbf{p} \mapsto -\mathrm{i}\hbar\nabla \tag{2.9}$$

and get the Schrödinger equation (in an electromagnetic field)

$$\mathrm{i}\hbar\partial_t\psi = -\frac{\hbar^2}{2m}[\nabla - \mathrm{i}(e/\hbar)\mathbf{A}]^2\psi + e\phi\psi \tag{2.10}$$

This does not work at the relativistic level, of course, because t and $\mathbf{x}$ do not enter in a symmetric way. But we can start from the relativistic energy-momentum relation (dispersion relation) Eq.(2.5). The replacement rules (2.9) are equivalent to

$$p_\mu \mapsto \mathrm{i}\hbar\partial_\mu \tag{2.11}$$

Note that the same sign is sufficient here because the covariant derivative is $\partial_\mu = \partial/\partial x^\mu = ((1/c)\partial_t, \nabla)$. One gets

$$\begin{aligned}&[\mathrm{i}\partial_\mu - (e/\hbar)A_\mu][\mathrm{i}\partial^\mu - (e/\hbar)A^\mu]\psi \\ &= -\frac{\hbar^2}{c^2}(\partial_t + \mathrm{i}(e/\hbar)\phi)^2\psi + \hbar^2(\nabla - \mathrm{i}(e/\hbar)\mathbf{A})^2\psi = (mc)^2\psi\end{aligned} \tag{2.12}$$

or, in the free field case

$$\left[\frac{1}{c^2}\partial_t^2 - \nabla^2 + \frac{m^2c^2}{\hbar^2}\right]\psi = 0 \tag{2.13}$$

with D'Alembert operator $c^{-2}\partial_t - \nabla^2 = \partial_\mu\partial^\mu$, the relativistic generalization of the Laplace operator. This is the Klein-Gordon equation.

World formula

Action for the Klein-Gordon equation (without field)

$$S = \int \mathrm{d}t\, \mathrm{d}^3x \left[\frac{\hbar^2}{c^2} \partial_t \psi^* \partial_t \psi - \hbar^2 (\nabla \psi^*) \cdot (\nabla \psi) - (mc)^2 \psi^* \psi \right] \tag{2.14}$$

is covariant under rotation, phase factor change. Also under Lorentz transformation since it involves the scalar $\partial_\mu \psi^* \partial^\mu \psi$. No need to change the field under Lorentz transformation (a 'scalar field').

A more precise formulation: change of coordinates $x' = \Lambda x$ with 4×4 matrix Λ. The scalar function $\psi(x)$ is represented in the new coordinates by a new function $\psi'(x')$ that takes the same value at points that are mapped onto each other by the Lorentz transformation

$$\psi'(x') = \psi(x) = \psi(\Lambda^{-1} x'), \qquad \text{or } \psi' = \psi \circ \Lambda^{-1} \tag{2.15}$$

Compare to a vector field and coordinate rotation R:

$$E_i'(\mathbf{x}') = R_{ij} E_j(\mathbf{x}) = R_{ij} E_j(R^{-1} \mathbf{x}'), \qquad \text{or } \mathbf{E}' = R(\mathbf{E} \circ R^{-1}) \tag{2.16}$$

where the rotation matrix R acts also on the components of the field.

Euler-Lagrange equations: here for a complex field ψ. Exercises: work out separately for real and imaginary part of ψ, like components of a generalized coordinate vector. Here: differentiate directly with respect to ψ or ψ^* as if they were independent. With Lagrangian density $\mathcal{L}$:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial \psi^*} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi^*)} && (2.17) \\ &= -(mc)^2 \psi - \hbar^2 \partial_\mu \partial^\mu \psi && (2.18) \end{aligned}$$

hence we recover again the Klein-Gordon equation.

Current conservation

Conserved current for Klein-Gordon field:

$$j^\mu = (nc, \mathbf{j}) = \frac{\mathrm{i}\hbar}{2m} \left(\psi^* \partial^\mu \psi - \partial^\mu \psi^* \psi \right) \tag{2.19}$$

easy to check that is conserved:

$$\partial_\mu j^\mu = \partial_t n + \nabla \cdot \mathbf{j} = 0 \tag{2.20}$$

from the Klein-Gordon equation (2.18) and its complex conjugate.

However, the density n cannot be interpreted in terms of a positive definite probability density:

$$n = \frac{\mathrm{i}\hbar}{2mc^2} \left(\psi^* \partial_t \psi - \partial_t \psi^* \psi \right) = \frac{E}{mc^2} \psi^* \psi \gtrless 0 \tag{2.21}$$

In the last step, we have assumed a stationary state $\psi(\mathbf{x}, t) = \psi(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}Et/\hbar}$. This is negative if we take for E the negative branch of the dispersion relation:

$$E = \pm \sqrt{\mathbf{p}^2 c^2 + (mc^2)^2} \tag{2.22}$$

In our days, this negative density is not really a problem: we would simply interpret these “negative energy solutions” as “antiparticles” that count with a minus sign in the total number of particles.

The fact that the Klein-Gordon equation has solutions whose energy is negative and arbitrary large, was troubling physicists in the 1920s. One of the reasons is that one could imagine a stable ground state only as a state of lowest energy. From the thermodynamics viewpoint, a Hamiltonian with energy eigenvalues that extend to $-\infty$ seem to make problems as well. This motivated Dirac to derive his equation, but it turned out that it suffers from the same problem.

Klein-Gordon propagator

Solution of (rewrite $mc/\hbar$ as m; unit: 1/length)

$$\left(\partial_\mu \partial^\mu + m^2\right) G(x) = \delta(x) \tag{2.23}$$

Solution with frequency Fourier transform

$$G(x) = \frac{1}{4\pi|\mathbf{x}|} \int \frac{\mathrm{d}\omega}{2\pi} \exp[-\mathrm{i}\omega t + \mathrm{i}|\mathbf{x}|\sqrt{\omega^2 - m^2}] \tag{2.24}$$

using solution for Helmholtz equation for $k = \sqrt{\omega^2 - m^2}$. General result is complicated. Is nonzero also in/outside the light cone, different from (massless) photon.

Special case $\omega = 0$, relevant for static source (instead of ‘event flash’): Yukawa potential

$$G(\mathbf{x}; \omega \to 0) = \frac{1}{4\pi|\mathbf{x}|} \mathrm{e}^{-m|\mathbf{x}|} \tag{2.25}$$

Hence a static source coupled to the Klein-Gordon field creates a nonzero field amplitude only at a distance $|\mathbf{x}| \sim 1/m$. This field has the form of a ‘screened’ Coulomb potential. For the electron mass ($0.5\,\mathrm{MeV}$), the range of this potential corresponds to $\sim 10^{-12}\,\mathrm{m}$. But the electron is not described by the scalar Klein-Gordon field, as we shall see. The particles called π-mesons or ‘pions’ (π^0, π^+, π^-) have masses around $140\,\mathrm{MeV}$ and their quantum numbers are compatible with the Klein-Gordon equation. The range is of the order of a few $\sim 10^{-15}\,\mathrm{m}$ which is of the size of the nucleus. Indeed, the pions are thought to be the ‘photons’ (i.e., the particles that carry the interaction) of the ‘strong force’ that binds together protons and neutrons in the nucleus.

For a source term oscillating at a frequency $\omega \approx m$, write $\omega = m + E$ and one gets the non-relativistic limit:

$$G(x) = \frac{1}{4\pi|\mathbf{x}|} \mathrm{e}^{-\mathrm{i}(m+E)t + \mathrm{i}|\mathbf{x}|\sqrt{2mE}}. \tag{2.26}$$

The oscillation $\mathrm{e}^{-\mathrm{i}mt}$ corresponds to the rest mass, it is factored out in the non-relativistic limit.

Symmetries

The Klein-Gordon field describes a 'scalar' field: under a Lorentz transformation or a rotation, the amplitude changes to

$$\psi'(x') = \psi(x). \tag{2.27}$$

The Klein-Gordon action is invariant under complex conjugation:

$$\psi(x) \mapsto \psi^*(x). \tag{2.28}$$

This is related to the operation called 'charge conjugation'. Two cases can occur: either the field ψ is real, then it is itself invariant under conjugation: the field is electrically neutral. If the field is explicitly complex, then ψ and ψ^* differ and describe fields (particles) with an opposite electric charge. The field ψ^* satisfies the same equation (if $A_\mu = 0$), in particular, it has the same mass.

2.2.1 Gauge symmetry

For the complex Klein-Gordon field, the action is also invariant under a 'global phase factor'

$$\psi'(x) = \mathrm{e}^{\mathrm{i}\varphi}\psi(x) \tag{2.29}$$

It is a very deep principle of field theory that this symmetry can be formulated 'locally' and that it provides a natural way to couple to the electromagnetic field. This principle opened the way to what it now called the 'standard model': the weak and strong interactions are introduced by postulating similar local symmetry transformations.

A local phase transformation is given by

$$\psi'(x) = \mathrm{e}^{\mathrm{i}\varphi(x)}\psi(x) \tag{2.30}$$

Since now $\varphi(x)$ is position-dependent, the momentum operator acts on ψ' like

$$\mathrm{i}\partial_\mu \psi' = \mathrm{e}^{\mathrm{i}\varphi(x)} \left(\mathrm{i}\partial_\mu - \partial_\mu \varphi\right) \psi \tag{2.31}$$

where an extra term appears. This looks a little bit like the vector potential that appears in the minimal coupling scheme. Let us consider the combination ($\hbar = 1$)

$$\begin{aligned} \left(\mathrm{i}\partial_\mu - eA'_\mu\right) \psi' &= \mathrm{e}^{\mathrm{i}\varphi(x)} \left(\mathrm{i}\partial_\mu - \partial_\mu\varphi - eA'_\mu\right) \psi \\ &= \mathrm{e}^{\mathrm{i}\varphi(x)} \left(\mathrm{i}\partial_\mu - eA_\mu\right) \psi \end{aligned} \tag{2.32}$$

In the last line, we have forced this to 'look like' another vector potential A_μ by the identification $\partial_\mu\varphi + eA'_\mu = eA_\mu$ which looks like a gauge transformation:

$$eA'_\mu = eA_\mu - \partial_\mu\varphi = e\left(A_\mu + \partial_\mu\chi\right) \tag{2.33}$$

if we take $\varphi = -e\chi$ where χ is the 'gauge function'.

This can be formalized by introducing the 'covariant derivative'

$$\mathrm{i}D_\mu := \mathrm{i}\partial_\mu - eA_\mu \tag{2.34}$$

and we find that the following Klein-Gordon action is covariant

$$S = \int \mathrm{d}t\, \mathrm{d}^3x \left[(D_\mu \psi)^* (D^\mu \psi) - (mc)^2 \psi^* \psi \right] \tag{2.35}$$

provided that one makes *simultaneously* a local phase transformation and a gauge transformation:

$$\begin{cases} \psi'(x) & = & \mathrm{e}^{-\mathrm{i}e\chi(x)/\hbar} \psi(x) \\ A'_\mu & = & A_\mu - \partial_\mu \chi \end{cases} \tag{2.36}$$

The scalar quantity e in the phase factor can now be identified with the electric charge of the field.

We now see also that the complex conjugate field ψ satisfies a field equation with the opposite sign of e: indeed, under complex conjugation,

$$(\mathrm{i}\partial_\mu - eA_\mu)\psi \mapsto (-\mathrm{i}\partial_\mu - eA_\mu)\psi^* = -(\mathrm{i}\partial_\mu + eA_\mu)\psi^* \tag{2.37}$$

the global sign disappears from both the action and the wave equation since these are quadratic in D_μ.

2.3 Dirac equation

Motivation: find the 'square root' of the wave operator $\partial_\mu \partial^\mu$ to get something linear in ∂_t. This should help avoiding solutions with both signs in energy.

2.3.1 Lagrangian

Lagrangian

$$S = \int \mathrm{d}t\, \mathrm{d}^3x \left[\frac{\mathrm{i}}{2} \overline{\Psi} \gamma^\mu \partial_\mu \Psi - \frac{\mathrm{i}}{2} \partial_\mu \overline{\Psi} \gamma^\mu \Psi - m \overline{\Psi} \Psi \right] \tag{2.38}$$

with a field Ψ. The object $\overline{\Psi}$ (the 'adjoint' field) is something like a hermitian conjugate of Ψ. This action is linear in ∂_μ and 'symmetric' in $t, \mathbf{x}$, although the quantities γ^μ remain mysterious. Shall see: 4×4 matrices.

Euler Lagrange equation gives the Dirac equation

$$\boxed{\mathrm{i}\gamma^\mu \partial_\mu \Psi - m\Psi = 0} \tag{2.39}$$

2.3.2 Dispersion relation

Require that dispersion relation for particle with mass m is satisfied. Make *Ansatz* $\Psi(x) = \Psi_0 \exp(-\mathrm{i} p_\mu x^\mu)$ and find

$$\gamma^\mu p_\mu = m \tag{2.40}$$

square this and get

$$\gamma^\mu \gamma^\nu p_\mu p_\nu = \frac{1}{2} \left\{ \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu \right\} p_\mu p_\nu \stackrel{!}{=} g^{\mu\nu} p_\mu p_\nu = (mc)^2 \tag{2.41}$$

Analogy in electrodynamics: the components of the 4-vector potential all satisfy the (scalar!) wave equation with the D'Alembert operator.

Constraint for γ-operators:

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \tag{2.42}$$

Fundamental relation, cannot be satisfied with c-numbers, need matrices. Mathematics: structure of a Clifford algebra.[1] Simple consequences:

$$(\gamma^0)^2 = 1 \tag{2.44}$$
$$(\gamma^i)^2 = -1, \qquad \text{for } i = 1, 2, 3 \tag{2.45}$$
$$\gamma^0\gamma^i = -\gamma^i\gamma^0 \tag{2.46}$$

Last equation: the γ's cannot be c-numbers. One needs matrices.

2.3.3 Probability current

Next: conserved current. Need Dirac equation for adjoint spinor, from Euler-Lagrange equation

$$-\mathrm{i}\partial_\mu\overline{\Psi}\gamma^\mu - m\overline{\Psi} = 0 \tag{2.47}$$

multiply from the right with Ψ, multiply (2.39) from the left with $\overline{\Psi}$ and take the difference:

$$0 = \overline{\Psi}\gamma^\mu\partial_\mu\Psi + (\partial_\mu\overline{\Psi})\gamma^\mu\Psi = \partial_\mu(\overline{\Psi}\gamma^\mu\Psi)$$
$$\text{hence: } j^\mu = \overline{\Psi}\gamma^\mu\Psi \tag{2.48}$$

The conserved current is bilinear in Ψ and $\overline{\Psi}$, with no derivative. We know require that the probability density be positive:

$$n = j^0 = \overline{\Psi}\gamma^0\Psi \geq 0 \tag{2.49}$$

Let us define the 'adjoint' field by $\overline{\Psi} = \Psi^\dagger\Gamma$ with some factor (matrix) Γ to be specified. The density n is positive if $\Gamma\gamma^0$ is positive, for example if $\Gamma\gamma^0 = 1$. From the anti-commutation relation (2.44), one possible solution is $\Gamma = \gamma^0$. Hence, we get

$$\overline{\Psi} = \Psi^\dagger\gamma^0, \qquad n = \Psi^\dagger\Psi \geq 0 \tag{2.50}$$

We check below that the spatial part, $j^i = \overline{\Psi}\gamma^i\Psi$ is a vector.

[1] **Clifford algebra.** Definition: an algebra is a vector space where a 'product' $A \circ B$ is defined. Let $\mathbb{1}, A_1, \ldots A_n$ be the 'generators' of this vector space, i.e., any element of the algebra can be written as a linear combination of products of the A_i. The algebra is called a Clifford algebra $\mathrm{Cl}(p, q)$, $n = p + q$ if

$$A_i \circ A_j + A_i \circ A_j = 2g_{ij}\mathbb{1} \tag{2.43}$$

with g_{ij} a diagonal matrix with entries $g_{11} = \ldots g_{pp} = -1$, $g_{p+1,p+1} = \ldots g_{nn} = 1$. The first p generators are thus 'square roots of -1', similar to the complex unit i. The Dirac matrices are in the vector space of complex 4×4 matrices, $p = 1, q = 3$, the product of the algebra is the matrix multiplication.

2.3.4 Dirac Hamiltonian

We can write the Dirac equation in the form

$$\mathrm{i}\partial_t\Psi = H\Psi = \left(\gamma^0 m - \mathrm{i}\gamma^0\gamma^i\partial_i\right)\Psi. \tag{2.51}$$

We would like this Hamiltonian to be hermitean, to preserve the probability density $\Psi^\dagger\Psi$. Hence: the matrices γ^0 and $\gamma^0\gamma^i$ are hermitean. One sometimes uses the notation $\alpha^i = \gamma^0\gamma^i$ for the three hermitean matrices encountered here.

We can check that γ^i is anti-hermitean:

$$(\gamma^i)^\dagger = (\gamma^0\gamma^0\gamma^i)^\dagger = (\gamma^0\gamma^i)^\dagger(\gamma^0)^\dagger = (\gamma^0\gamma^i)\gamma^0 = -\gamma^i \tag{2.52}$$

because of the anti-commutation rule (2.46). Note that

$$(\alpha^i)^2 = (\gamma^0\gamma^i)^2 = \gamma^0\gamma^i\gamma^0\gamma^i = -\gamma^0\gamma^0\gamma^i\gamma^i = 1 \tag{2.53}$$

The γ matrices are traceless. We show this first for the α^i:

$$\mathrm{tr}\,(\gamma^i\gamma^0) \overset{(1)}{=} \mathrm{tr}\,(\gamma^0\gamma^i) \overset{(2)}{=} -\mathrm{tr}\,(\gamma^i\gamma^0) \tag{2.54}$$

Step (1): cyclic permutation under the trace. Step (2): anti-commutation rule.

$$\mathrm{tr}\,\gamma^0 = -\mathrm{tr}\,(\gamma^0\gamma^i\gamma^i) = -\mathrm{tr}\,(\gamma^i\gamma^0\gamma^i) = \mathrm{tr}\,(\gamma^0\gamma^i\gamma^i) = -\mathrm{tr}\,\gamma^0 \tag{2.55}$$

and similarly for γ^i.

If we diagonalize the hermitean matrices γ^0 and the $\gamma^0\gamma^i$, the trace is the sum of the eigenvalues. In addition, since $(\gamma^0)^2 = 1 = (\gamma^0\gamma^i)^2$, the eigenvalues are ± 1. So we need at least matrices of dimension 2×2 to have eigenvalues in pairs with opposite sign.

The 2×2 case is too restrictive to construct the matrices we need: we do not have four linearly independent matrices that are all different from the unit matrix. The best one could do is to use Pauli matrices that have the property

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}, \tag{2.56}$$

are hermitean with trace zero. But the only candidate for γ^0 would be the unit matrix that does not anti-commute.

To implement the anti-commutation rules (2.42), the minimum dimension is four. A particular representation of the Dirac matrices is given by

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{2.57}$$

$$\gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{2.58}$$

where the σ_i are the 2×2 Pauli matrices. The matrices α^i are given by

$$\alpha^i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{2.59}$$

they are hermitean because the σ_i's are hermitean.

2.3.5 Vectors, spins, rotations

Spinor rotation

We now want to ensure that the probability current transforms like a vector. The current is given by (we write the spatial indices as lower or upper indices, there is no sign change implied here as with greek indices)

$$j_k = \Psi^\dagger \gamma^0 \gamma^k \Psi = \Psi^\dagger \alpha^k \Psi \tag{2.60}$$

This is a vector if for a coordinate rotation (with rotation matrix R with entries R_{kl}) we have the following transformation

$$\mathbf{x}' = R\mathbf{x}, \qquad j'_k(t, \mathbf{x}') = R_{kl} j_l(t, \mathbf{x}) \tag{2.61}$$

(the 'new components' j'_k are given by the same formula as the 'new coordinates' $R\mathbf{x}$). To implement this formula, let us first write the Dirac spinor in the form $\Psi = (\psi, \chi)^T$ so that

$$j_k = \psi^\dagger \sigma_k \chi + \chi^\dagger \sigma_k \psi. \tag{2.62}$$

We have to assume that these spinors change unter a rotation, $\psi'(x') = S(R)\psi(x)$, $\chi'(x') = S(R)\chi(x)$ where S is a complex matrix. We then get

$$\psi^\dagger S^\dagger \sigma_k S \chi + \chi^\dagger S^\dagger \sigma_k S \psi \stackrel{!}{=} R_{kl} \left(\psi^\dagger \sigma_l \chi + \chi^\dagger \sigma_l \psi \right) \tag{2.63}$$

This is satisfied for all ψ, χ if we have

$$S^\dagger \sigma_k S = R_{kl} \sigma_l \tag{2.64}$$

This equation fixes $S(R)$ up to a global phase factor. You recognize here the transformation of the 2-component spinor discussed in Exercise 6.3. The general result worked out there gives the transformation $\exp\left(\frac{\mathrm{i}}{2}\theta \boldsymbol{\sigma} \cdot \mathbf{n}\right)$ for a rotation about an angle θ around the axis $\mathbf{n}$ (this is a unit vector). The Dirac 4-spinor therefore transforms in the following way under a rotation:

$$\Psi'(x') = \begin{pmatrix} \exp\left(\frac{\mathrm{i}}{2}\theta \boldsymbol{\sigma} \cdot \mathbf{n}\right) & 0 \\ 0 & \exp\left(\frac{\mathrm{i}}{2}\theta \boldsymbol{\sigma} \cdot \mathbf{n}\right) \end{pmatrix} \Psi(x) \tag{2.65}$$

From this equation, one can show that the Dirac spinor describes a particle with spin $1/2$. We need some technicalities to make this statement. But let us first provide a proof of a special case of Eq.(2.65).

Infinitesimal rotation

We focus on a rotation around the z-axis. The corresponding rotation matrix is given by

$$R_{ij} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.66}$$

and the corresponding spinor transformation must satisfy:

$$S^\dagger \sigma_x S = \cos\theta\, \sigma_x + \sin\theta\, \sigma_y \tag{2.67}$$

and a similar equation including σ_y.

If the rotation angle θ is small, we have to first order in θ:

$$S^\dagger \sigma_x S = (\mathbb{1} - \mathrm{i}\theta J_z)\, \sigma_x \,(\mathbb{1} + \mathrm{i}\theta J_z) = \sigma_x + \theta\, \sigma_y \tag{2.68}$$

we have written S as the unit matrix[2] plus a correction to first order in θ. We try to work with a hermitean matrix J_z. The matrix J_z is called the 'generator' of the spinor transformation. We get from (2.68):

$$[J_z,\, \sigma_x] = \mathrm{i}\sigma_y. \tag{2.69}$$

If we exclude a c-number valued contribution[3], this relation is satisfied if we take

$$J_z = \frac{1}{2}\sigma_z, \qquad S(\theta, z) = \mathbb{1} + \frac{\mathrm{i}}{2}\theta\sigma_z, \quad \theta \to 0. \tag{2.70}$$

By concatenating this transformation for many small angles, we get

$$S(\theta, z) = \lim_{N\to\infty} \left(\mathbb{1} + \tfrac{\mathrm{i}}{2}(\theta/N)\sigma_z\right)^N = \exp\left(\tfrac{\mathrm{i}}{2}\theta\sigma_z\right) \tag{2.71}$$

This coincides with the special case one obtains from (2.65) when the unit vector $\mathbf{n}$ is identified with the rotation axis.

If you do not feel comfortable with this calculation, you can introduce a one-parameter subgroup $S(\theta)$ of spinor transformations such that (2.95) holds. By differentiating this constraint with respect to θ, one gets a differential equation whose solution is the matrix exponential (2.71).

Since $\sigma_z^2 = \mathbb{1}$, the matrix exponential can be expanded and re-summed easily. This gives the matrix elements of the spinor transformation:

$$S(\theta, z) = \cos(\theta/2) + \mathrm{i}\sigma_z \sin(\theta/2). \tag{2.72}$$

The spin of a field or particle

The spin of a particle is defined in terms of the 'transformation properties' of its 'wave function' under rotations. For example, the Klein-Gordon wave function does not change under a Lorentz transformation, nor under a rotation (a subgroup of the Lorentz group), $\psi'(x') = \psi(x)$. This corresponds to a 'scalar' or 'spin 0' particle.

For an arbitrary field Ψ, considerations similar to those we followed for the Dirac field, lead to a transformation of the form

$$\Psi'(x') = S(R)\Psi(x) \tag{2.73}$$

where $S(R)$ is a possibly complex matrix that corresponds to the rotation matrix. Now, the generators of this transformation, J_x, J_y, J_z (for infinitesimal rotations around the x, y, z axis) behave like an angular momentum operators:

$$[J_k,\, J_l] = \mathrm{i}\epsilon_{klm} J_m \tag{2.74}$$

This can be shown from the following requirement:[4]

$$S(R_1)S(R_2) = S(R_1R_2)\, \mathrm{e}^{\mathrm{i}\varphi(R_1R_2)} \tag{2.75}$$

[2] Note that this corresponds to a specific choice for the global phase of S.
[3] This would again correspond to a global phase in S.
[4] There are cases where the phase $\varphi(R_1R_2)$ vanishes.

that translates the fact that rotations (and Lorentz transformations) form a group and can be concatenated. The corresponding field transformations should be compatible with this group structure. The mapping $S : R \mapsto S(R)$ is a 'group homomorphism' and the transformation matrices S provide a 'representation' of the rotation group. (More details in Chapter 3.)

Now, the commutation relations (2.74) provide constraints on the possible matrices that can be used for the generators J_k. In the quantum mechanics lecture I, you have seen that $J^2 := J_x^2 + J_y^2 + J_z^2$ commutes with all of the J_k's and that its eigenvalues are of the form $j(j + 1)$ (when $\hbar = 1$). For a given particle, there is a specific value for j that occurs here. The possible values are $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ and the dimension of the matrix/the generators is $(2j + 1) \times (2j + 1)$. The corresponding 'wave functions' on which the spinor transformations act, also have $(2j+1)$ entries. The corresponding basis vectors can be identified with eigenvectors of the spin operator J_z with eigenvalues $-j, -j + 1, \dots, j$.

In the case $j = 0$, the S reduce to c-numbers and excluding global phases, the field does not change under a rotation (scalar field). The case $j = 1/2$ (spin 1/2) corresponds to spinors with two components, 'spin up' and 'spin down'. The transformation matrices are generated by the Pauli matrices. The case $j = 1$ is related to the 'trivial choice' $S = R$ with 3×3 matrices. These wave functions are the familiar vector fields you know from electrodynamics.

The fact that the Dirac spinor has four and not two components arises because we also require Lorentz invariance. The group of Lorentz transformations has a structure different from the rotation group (it contains twice as much subgroups or generators as the rotation group). We shall see this in a second when we construct the transformation under a 'boost' (special Lorentz transformation).

Let us also mention that the 'photon wavefunction' can be identified with the vector fields $\mathbf{E}$ and $c\mathbf{B}$. In fact, the Maxwell equations

$$\begin{aligned} \frac{1}{c}\partial_t \mathbf{E} &= \nabla \times c\mathbf{B} & (2.76)\\ \frac{1}{c}\partial_t c\mathbf{B} &= -\nabla \times \mathbf{E} & (2.77) \end{aligned}$$

provide another way to 'take the square root' of the D'Alembert operator. We work on the subspace of functions with $\nabla \cdot \mathbf{E} = \nabla \cdot \mathbf{B} = 0$. This constraint holds at all times if it holds at one time.

2.3.6 Boosting spinors

Now, how can we make the Dirac spinor behave in a relativistic way? Our basic requirement will be that the probability current j^μ be a 4-vector. For a Lorentz transformation Λ with matrix elements $\Lambda^\mu{}_\nu$, this means that

$$\begin{aligned} x'^\mu &= \Lambda^\mu{}_\nu x^\nu \qquad \text{or} \quad x' = \Lambda x, & (2.78)\\ j'^\mu(x') &= \Lambda^\mu{}_\nu j^\nu(x) & (2.79) \end{aligned}$$

Again the usual interpretation: at the 'new coordinate' x', the 'new components' j' are related to the 'old' ones in the same way as the coordinates.

Let us write $S = S(\Lambda)$ for the spinor transformation. We are not sure yet that it is unitary, but it should be invertible. The probability current $j^\mu = \overline{\Psi}\gamma^\mu\Psi = \Psi^\dagger\gamma^0\gamma^\mu\Psi$ is a 4-vector if we have

$$S^\dagger\gamma^0\gamma^\mu S = \Lambda^\mu{}_\nu\gamma^0\gamma^\nu. \tag{2.80}$$

To solve this equation, let us focus on a Lorentz boost along the x-axis ($\mathrm{sh} = \sinh$, $\mathrm{ch} = \cosh$, $\mathrm{th} = \tanh$)

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \mathrm{ch}\,\beta & \mathrm{sh}\,\beta & 0 & 0 \\ \mathrm{sh}\,\beta & \mathrm{ch}\,\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.81}$$

with $\mathrm{th}\,\beta = v/c$. Now, $\mu = 0$ and $\mu = x$ are the relevant equations in (2.80)

$$S^\dagger S = \mathrm{ch}\,\beta + \mathrm{sh}\,\beta\,\gamma^0\gamma^x = \mathrm{ch}\,\beta + \mathrm{sh}\,\beta\,\alpha^x \tag{2.82}$$

$$S^\dagger\gamma^0\gamma^x S = S^\dagger\alpha^x S = \mathrm{sh}\,\beta + \mathrm{ch}\,\beta\,\alpha^x \tag{2.83}$$

The first equation shows that S is *not* a unitary matrix. We make nevertheless the ansatz $S = \mathbb{1} + \beta K$ for an infinitesimal Lorentz transformation. This gives

$$1 + \beta(K^\dagger + K) = 1 + \beta\alpha^x \tag{2.84}$$

$$\alpha^x + \beta(K^\dagger\alpha^x + \alpha^x K) = \alpha^x + \beta \tag{2.85}$$

These equations can be satisfied by taking $K = \frac{1}{2}\alpha^x$: K is hermitean, and from the fundamental properties (2.42), $K^\dagger\alpha^x + \alpha^x K = \mathbb{1}$. Exponentiating this infinitesimal spinor transformation, we get

$$S(\beta, x) = \exp(\tfrac{1}{2}\beta\alpha^x) = \mathrm{ch}(\beta/2) + \mathrm{sh}(\beta/2)\alpha^x. \tag{2.86}$$

For a boost along an axis $\mathbf{n}$, one has to replace $\alpha^x \to n^k\alpha^k$. This comes from the fact that $(n^k\alpha^k)^2 = \mathbb{1}$ for all unit vectors $\mathbf{n}$.

The action is scalar

To complete the construction of the Lorentz-covariant Dirac equation, let us show that the Lagrangian for the Dirac equation is a Lorentz scalar. Then the Dirac equation (via the Euler-Lagrange equations) will be covariant also.

In the Lagrangian

$$S = \int \mathrm{d}t\,\mathrm{d}^3x \left[\frac{\mathrm{i}}{2}\overline{\Psi}\gamma^\mu\partial_\mu\Psi - \frac{\mathrm{i}}{2}\partial_\mu\overline{\Psi}\gamma^\mu\Psi - m\overline{\Psi}\Psi\right]$$

First the term $\overline{\Psi}\Psi$: to work this out, we start with the conservative assumption that the Dirac matrices are the same $\gamma'^\mu = \gamma^\mu$: this turns out to work. We then need

$$S(\beta, x)^\dagger\gamma^0 S(\beta, x) = (\mathrm{ch}(\beta/2) + \mathrm{sh}(\beta/2)\alpha^x)\,\gamma^0\,(\mathrm{ch}(\beta/2) + \mathrm{sh}(\beta/2)\alpha^x) \tag{2.87}$$

One term gives $\mathrm{ch}^2(\beta/2)\gamma^0$. The mixed term vanishes because α^x and γ^0 anticommute. The last term gives $\alpha^x\gamma^0\alpha^x = -\gamma^0(\alpha^x)^2 = -\gamma^0$. Hence

$$S(\beta, x)^\dagger\gamma^0 S(\beta, x) = \left(\overset{2}{\mathrm{ch}}(\beta/2) - \overset{2}{\mathrm{sh}}(\beta/2)\right)\gamma^0 = \gamma^0 \tag{2.88}$$

and it follows that $\overline{\Psi}'(x')\Psi'(x') = \overline{\Psi}(x)\Psi(x)$: this transforms like a scalar field under Lorentz boosts. In the exercises, you show that this is a scalar under rotations as well (because the 2-spinor transformations are unitary). An alternative formulation of this proof is given in Eq.(2.97).

So let us consider the term with the derivatives. One of them will be sufficient because one can show that the other one is the complex conjugate.

$$\begin{aligned}
\overline{\Psi}'(x')\gamma^\mu\partial'_\mu\Psi'(x') &= \Psi^\dagger(x)S^\dagger\gamma^0\gamma^\mu\frac{\partial}{\partial x'^\mu}S\Psi(x) \\
&= \Psi^\dagger(x)S^\dagger\gamma^0\gamma^\mu S\frac{\partial}{\partial x'^\mu}\Psi(x) \\
&= \Psi^\dagger(x)\Lambda^\mu{}_\nu\gamma^0\gamma^\nu\frac{\partial}{\partial x'^\mu}\Psi(x)
\end{aligned} \tag{2.89}$$

we have used that the transformation matrix S does not depend on the coordinates (going beyond that would lead us from special to general relativity theory), and the transformation rules (2.80) for the γ matrices.

Note that these rules hold in this form for both rotations and boosts. For rotations, Λ is a matrix that acts trivially on the time coordinate, with the rotation matrix being a 3×3 block for the spatial coordinates.

Now, how does the derivative ∂'_μ transform? We have to re-express the argument $x = \Lambda^{-1}x'$, or in components

$$x^\nu = (\Lambda^{-1})^\nu{}_\mu x'^\mu. \tag{2.90}$$

Hence the chain rule for differentiation gives

$$\begin{aligned}
\frac{\partial}{\partial x'^\mu}\Psi(x) &= \frac{\partial}{\partial x'^\mu}\Psi((\Lambda^{-1})^\nu{}_\lambda x'^\lambda) \\
&= (\Lambda^{-1})^\nu{}_\lambda\delta^\lambda{}_\mu\frac{\partial}{\partial x^\nu}\Psi(x) \\
&= (\Lambda^{-1})^\lambda{}_\mu\frac{\partial}{\partial x^\lambda}\Psi(x)
\end{aligned} \tag{2.91}$$

Taking the matrix elements of the Lorentz transformations out of the spinor product, we get from Eqs.(2.89, 2.91):

$$\begin{aligned}
\overline{\Psi}'(x')\gamma^\mu\partial'_\mu\Psi'(x') &= (\Lambda^{-1})^\lambda{}_\mu\Lambda^\mu{}_\nu\Psi^\dagger(x)\gamma^0\gamma^\nu\frac{\partial}{\partial x^\lambda}\Psi(x) \\
&= \delta^\lambda{}_\nu\overline{\Psi}(x)\gamma^\nu\frac{\partial}{\partial x^\lambda}\Psi(x)
\end{aligned} \tag{2.92}$$

The second line follows because we deal with the Lorentz matrix and its inverse. Hence indeed, this term of the action is covariant.

To show that the two terms in the action with derivates are the complex conjugates of each other, you need the property

$$(\gamma^\mu)^\dagger\gamma^0 = \gamma^0\gamma^\mu \tag{2.93}$$

that follows from the fact that γ^k is anti-hermitean (2.52) and anti-commutes with the hermitean matrix γ^0.

Formulas for spinor transformations

To summarize the spinor transformations we found, consider first a coordinate rotation around the axis $\mathbf{n}$ with angle θ. One then has

$$\Psi'(x') = \begin{pmatrix} \psi'(x') \\ \chi'(x') \end{pmatrix} = \begin{pmatrix} \exp\left(\frac{\mathrm{i}}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}\right) & 0 \\ 0 & \exp\left(\frac{\mathrm{i}}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}\right) \end{pmatrix} \begin{pmatrix} \psi(x) \\ \chi(x) \end{pmatrix} \tag{2.94}$$

For a Lorentz boost along the direction $\mathbf{n}$ with a relative velocity $v = c\,\mathrm{th}\,\beta$:

$$\Psi'(x') = \begin{pmatrix} \psi'(x') \\ \chi'(x') \end{pmatrix} = \begin{pmatrix} \mathrm{ch}(\beta/2) & \mathrm{sh}(\beta/2)\boldsymbol{\sigma}\cdot\mathbf{n} \\ \mathrm{sh}(\beta/2)\boldsymbol{\sigma}\cdot\mathbf{n} & \mathrm{ch}(\beta/2) \end{pmatrix} \begin{pmatrix} \psi(x) \\ \chi(x) \end{pmatrix} \tag{2.95}$$

Finally, let us check that the quantity $\overline{\Psi}\Psi$ really transforms like a Lorentz scalar:

$$\Psi^\dagger S^\dagger \gamma^0 S \Psi \stackrel{?}{=} \Psi^\dagger \gamma^0 \Psi \tag{2.96}$$

This is satisfied if we have $S^\dagger\gamma^0 S = \gamma^0$, or in other words:

$$\gamma^0 S^\dagger \gamma^0 = S^{-1} \tag{2.97}$$

Now, insert the explicit result (2.86) and use that the α^k are hermitean:

$$\begin{aligned} \gamma^0 S^\dagger \gamma^0 &= \gamma^0 \left(\mathrm{ch}(\beta/2) + \mathrm{sh}(\beta/2) n^k \alpha^k\right)\gamma^0 \\ &= \mathrm{ch}(\beta/2) - \mathrm{sh}(\beta/2) n^k \alpha^k \\ &= \exp(-\tfrac{1}{2}\beta n^k \alpha^k) = S^{-1} \end{aligned} \tag{2.98}$$

We have used the fact that γ^0 and the α^k anti-commute.

There are other quantities that can be constructed from the Dirac matrices: four-tensors with two and three indices. The quantity $\gamma^0\gamma^x\gamma^y\gamma^z$ transforms like a 'pseudoscalar' (it changes sign under parity and/or time reversal) and is related to the 'handedness' (chirality) of the Dirac spinors. This is an important concept in the standard model, but beyond the scope of this lecture.

2.3.7 Solutions to the Dirac equation

Plane waves

Let us consider first a particle at rest. This corresponds to a plane wave with $p_\mu = (E, \mathbf{0})$. The wave function is thus only time-dependent, $\Psi(t, \mathbf{x}) = \Psi_0\,\mathrm{e}^{-\mathrm{i}Et/\hbar}$, and the Dirac Hamiltonian gives

$$\mathrm{i}\hbar\partial_t\Psi = E\Psi_0\mathrm{e}^{-\mathrm{i}Et/\hbar} = mc^2\gamma^0\Psi_0\mathrm{e}^{-\mathrm{i}Et/\hbar} = mc^2 \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \Psi_0\mathrm{e}^{-\mathrm{i}Et/\hbar} \tag{2.99}$$

We can thus distinguish four cases:

$$\begin{aligned} &\text{'particles' with } E = mc^2: && \Psi_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \Psi_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\ &\text{'antiparticles' with } E = -mc^2: && \Psi_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{or} \quad \Psi_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned} \tag{2.100}$$

So the Dirac equation also suffers from 'solutions with negative energy', as the Klein-Gordon equation did! One advantage is that with Pauli's exclusion principle, one can construct an interpretation in terms of antiparticles as 'holes' in the filled 'Fermi sea' of negative energy solutions.

The two independent, particle-like solutions are related to the spin. This is a big plus in favour of the Dirac equation (outweighing the problems with the negative energy solutions). We have already found that the spin operator (the generator of the rotation (2.65)) is given by

$$\boldsymbol{\Sigma} = \begin{pmatrix} \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix} \tag{2.101}$$

and the spinors (2.100) are eigenstates of the z-component of this operator, with eigenvalues $\pm 1/2$. Of course, any other orthogonal orientations of the spin would provide two independent solutions to the Dirac equation as well.

How can we construct the wave function for a plane wave with momentum $\mathbf{p} \neq 0$? The simple way is to solve the Dirac equation ($\Psi_0 = (\psi, \chi)^T$)

$$E\Psi_0 = m\gamma^0\Psi_0 + p^k\alpha^k\Psi_0, \qquad E\begin{pmatrix} \psi \\ \chi \end{pmatrix} = \begin{pmatrix} m & \mathbf{p}\cdot\boldsymbol{\sigma} \\ \mathbf{p}\cdot\boldsymbol{\sigma} & -m \end{pmatrix}\begin{pmatrix} \psi \\ \chi \end{pmatrix} \tag{2.102}$$

This gives, for example:

$$\chi = \frac{\mathbf{p}\cdot\boldsymbol{\sigma}}{E+m}\psi \tag{2.103}$$

and we can choose the 'upper' spinor ψ freely.

Another way is to apply a boost to the solutions for a particle at rest. Starting from a 'particle-like' solution where $\chi_0 = 0$, the spinor transformation (2.95) gives

$$\begin{pmatrix} \psi \\ \chi \end{pmatrix} = \begin{pmatrix} \mathrm{ch}(\beta/2)\psi_0 \\ \boldsymbol{\sigma}\cdot\mathbf{n}\,\mathrm{sh}(\beta/2)\psi_0 \end{pmatrix} \tag{2.104}$$

It is straightforward to check that the ratio between χ and ψ, $\boldsymbol{\sigma}\cdot\mathbf{n}\,\mathrm{th}(\beta/2)$, is the same as in Eq.(2.103) provided $\mathbf{n}$ is the unit vector along $\mathbf{p}$ and $\mathrm{sh}\,\beta = |\mathbf{p}|/m$. (Recall the difference between velocity and momentum, $p = \gamma m v$.)

Dirac theory for hydrogen

The Dirac equation can also be solved for the hydrogen atom. We need, of course, to couple the electron to the electromagnetic field, as discussed in the next section. Then, one can derive in spherical coordinates a radial equation, make a polynomial ansatz and so on to get the quantized energy levels. Here, the Dirac equation successfully describes that the l levels are not degenerate. The 'fine structure' is getting apparent with 2s and 2p changing into $2\mathrm{s}_{1/2}$, $2\mathrm{p}_{1/2}$, and $2\mathrm{p}_{3/2}$. The index $j = 1/2, 3/2$ is the 'total angular momentum', a quantum number that combines the orbital angular momentum and the electron spin. For this reason, one is also talking about 'LS coupling'.

If you are interested, you can find these calculations in the books. The energy eigenvalues are given by (see Fig.2.1)

$$E_{nj} = \frac{mc^2}{\left(1 + \dfrac{\alpha^2}{(n-\delta_j)^2}\right)^{1/2}}, \qquad \delta_j = j + \tfrac{1}{2} - \sqrt{(j+\tfrac{1}{2})^2 - \alpha^2} \tag{2.105}$$

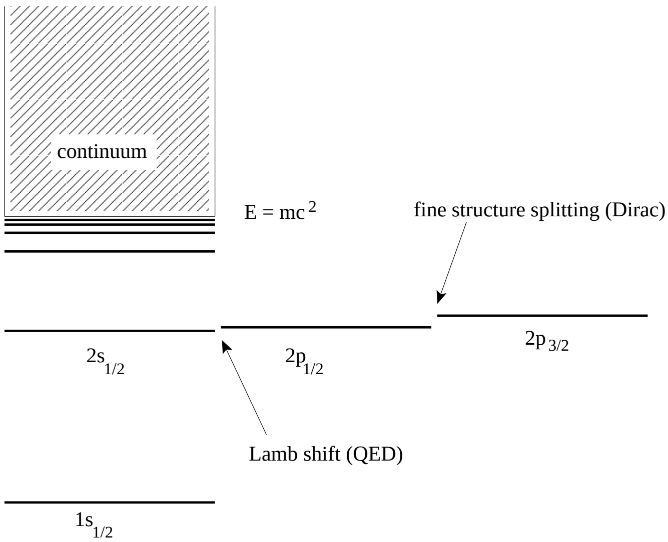

Figure 2.1: Energy levels of the hydrogen atom within the Dirac theory: the 'fine structure' between $2p_{1/2}$ and $2p_{3/2}$ appears. The Lamb shift (splitting between the levels $2s_{1/2}$ and $2p_{1/2}$) is not described by the Dirac equation, one needs QED for that.

where $\alpha = e^2/4\pi\varepsilon_0\hbar c \approx 1/137.04$ is the fine structure constant. The quantum numbers are $n = 1, 2, \ldots$, $j = \frac{1}{2}, \frac{3}{2}, \ldots n - \frac{1}{2}$. Dirac's theory predicts that the level $2\mathrm{p}_{3/2}$ is splitting off the two others $2\mathrm{s}_{1/2}$ and $2\mathrm{p}_{1/2}$ (these two have the same values of n, j). Good point: this is experimentally seen, and the non-relativistic calculation for the hydrogen atom could not explain this 'fine structure splitting'.

At a closer look, the levels $2\mathrm{s}_{1/2}$ and $2\mathrm{p}_{1/2}$ are also slightly split: this is called the 'Lamb shift' and is due to quantum electrodynamics (QED). The electron is surrounded by a 'cloud of virtual photons' that changes its mass, charge and so on.

2.3.8 Minimal coupling and non-relativistic limit

Let us discuss another successful prediction of the Dirac equation: the 'anomalous magnetic moment' of the electron.

Minimal coupling or local U(1) invariance, see Sec.2.2.1: change $\mathrm{i}\partial_\mu$ into $\mathrm{i}D_\mu = \mathrm{i}\partial_\mu - eA_\mu/\hbar$. 4-vector potential $A_\mu = (\phi/c, -\mathbf{A})$, covariant derivative $D_\mu = (D_t, \mathbf{D})$.

Split Dirac equation (2.51) in upper and lower two-component spinors, restore $\hbar$ and c

$$\begin{aligned} \mathrm{i}\hbar D_t\psi &= mc^2\psi - \mathrm{i}c\hbar\boldsymbol{\sigma}\cdot\mathbf{D}\chi \\ \mathrm{i}\hbar D_t\chi &= -mc^2\chi - \mathrm{i}c\hbar\boldsymbol{\sigma}\cdot\mathbf{D}\psi \end{aligned} \tag{2.106}$$

We separate off a factor $\mathrm{e}^{-\mathrm{i}mc^2t/\hbar}$ from *both* ψ and χ and assume that the remaining functions $\tilde{\psi}$ and $\tilde{\chi}$ are 'slowly varying' on the time scale set by $\hbar/mc^2$.

We also require that the vector potential (more precisely $eA_\mu c$) is small compared to the rest energy. This gives for the second equation

$$\mathrm{i}\hbar D_t \tilde{\chi} = -2mc^2 \tilde{\chi} - \mathrm{i}c\hbar \boldsymbol{\sigma} \cdot \mathbf{D} \tilde{\psi} \tag{2.107}$$

where the left hand side is by assumption negligible compared to the first term on the right hand side. Therefore

$$\tilde{\chi} = -\frac{\mathrm{i}\hbar}{2mc} \boldsymbol{\sigma} \cdot \mathbf{D} \tilde{\psi} \tag{2.108}$$

This gives the small admixture of 'antiparticle' to the particle wave function.

For the first line of of (2.106), we thus get

$$\mathrm{i}\hbar D_t \tilde{\psi} = -\frac{\hbar^2}{2m} (\boldsymbol{\sigma} \cdot \mathbf{D})(\boldsymbol{\sigma} \cdot \mathbf{D}) \tilde{\psi} \tag{2.109}$$

Now, the Pauli matrices have the property

$$\sigma_k \sigma_l = \delta_{kl} + \mathrm{i}\epsilon_{klm} \sigma_m \tag{2.110}$$

and therefore

$$\begin{aligned} \sigma_k D_k \sigma_l D_l &= \mathbf{D} \cdot \mathbf{D} + \frac{\mathrm{i}}{2} \sigma_m \epsilon_{klm} \left[D_k,\, D_l \right] \\ &= (\nabla - \mathrm{i}e\mathbf{A}/\hbar)^2 + \frac{e}{2\hbar} \sigma_m \epsilon_{klm} \left[\partial_k A_l - \partial_l A_k \right] \\ &= (\nabla - \mathrm{i}e\mathbf{A}/\hbar)^2 + \frac{e}{\hbar} \boldsymbol{\sigma} \cdot \mathbf{B} \end{aligned} \tag{2.111}$$

Putting everything together, we have found the so-called Pauli equation

$$\mathrm{i}\hbar \partial_t \tilde{\psi} = -\frac{\hbar^2}{2m} (\nabla - \mathrm{i}e\mathbf{A}/\hbar)^2 \tilde{\psi} - \frac{e\hbar}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} \tilde{\psi} + e\phi \tilde{\psi} \tag{2.112}$$

The key result is how the spin couples to the magnetic field. Let us write $\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma}$ for the spin operator. Collect the spin-magnetic field coupling with the one involving the orbital angular momentum $\mathbf{L}$ (see exercise), we have:

$$H_{\text{mag}} = -\frac{e}{2m} \mathbf{L} \cdot \mathbf{B} - \frac{e}{m} \mathbf{S} \cdot \mathbf{B} \tag{2.113}$$

where the spin $\mathbf{S}$ appears with a prefactor e/m that is two times larger than for $\mathbf{L}$. If one writes the magnetic interaction (2.113) in terms of a magnetic moment, $H_{\text{mag}} = -\boldsymbol{\mu} \cdot \mathbf{B}$:

$$\boldsymbol{\mu} = \mu_B \left(\mathbf{L} + g\mathbf{S} \right), \qquad g = 2, \qquad \mu_B = \frac{e}{2m}, \tag{2.114}$$

the factor $g = 2$ in front of the spin is called the 'gyromagnetic ratio' of the electron spin. Here, μ_B is called the 'Bohr magneton'. In the early days of quantum theory, only the orbital angular momentum was known, and it was regarded as 'anomalous' that some atomic energy levels should show shifts in a magnetic field whose splittings are twice as large, namely $g\mu_B \hbar |\mathbf{B}|$. In fact, one was observing the Zeeman effect for a state with, for example, $l = 0$ and $S = 1/2$.

Again, a closer look reveals that $g \approx 2 + \alpha/\pi + \mathcal{O}(\alpha^2) + \ldots$. The correction to first order in the fine structure constant α is due to quantum electrodynamics and 'virtual photons' surrounding the electron.

2.4 Discrete symmetries

The Minkowski distance is also invariant under sign changes of space or time:

$$s^2 = c^2t^2 - \mathbf{x}^2 = c^2t^2 - (-\mathbf{x})^2 = c^2(-t)^2 - \mathbf{x}^2 \tag{2.115}$$

These transformations are *not* included in the rotation and Lorentz matrices that we considered so far. Indeed, the space inversion $\mathbf{x} \mapsto -\mathbf{x}$ has a determinant -1, hence it is not a proper rotation matrix. And for $t \mapsto -t$, the 00 element of the Lorentz matrix $\Lambda^{\mu}{}_{\nu}$ is negative which did not occur in the Lorentz boosts we considered so far. The full group of transformations that leaves the Minkowski distance s^2 invariant thus contains four times as much elements: for each Λ, one has in addition $P\Lambda$, $T\Lambda$, and $PT\Lambda$ where P is the 'parity operator' (space inversion) and T denotes 'time reversal'.

We discuss now these operations separately and determine how they can be implemented for the Schrödinger and Dirac equations. Note that both operations $K = T, P$ satisfy $K^2 = 1$ on the level of coordinate transformation. This will no longer be true in the quantum mechanics setting.

2.4.1 Coordinates and vectors

Parity

The parity operator P takes the mirror image of the spatial coordinates. The following vectors ('polar vectors') change sign under P: $\mathbf{x}$, $\mathbf{p}$, ∇, $\mathbf{A}$, $\mathbf{E}$. The magnetic field $\mathbf{B}$ and the angular momentum $\mathbf{L}$ do *not* change sign (pseudovectors or axial vectors).

As a consequence, the classical equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}\tau}p^k = eE^k\frac{\mathrm{d}t}{\mathrm{d}\tau} - eB^k{}_l\frac{\mathrm{d}x^l}{\mathrm{d}\tau} \tag{2.116}$$

is covariant under P. (τ is the proper time, and the matrix $B^k{}_l$ acts like a vector product: $B_{kl}u^l = (\mathbf{B} \times \mathbf{u})^l$. It is invariant under parity.)

A scalar wave function transforms under parity according to

$$P : \psi(t, \mathbf{x}) \mapsto \psi'(t', \mathbf{x}') = \psi(t', -\mathbf{x}') \tag{2.117}$$

which is the usual rule for a scalar function under a coordinate transformation.

Time reversal

Under time reversal, $t \mapsto t' = -t$. This flips in particular the signs of all momenta (the coordinates being unchanged): $\mathbf{x} \mapsto \mathbf{x}$, $\mathbf{p} \mapsto -\mathbf{p}$. Since charges at rest (in motion) create electric (magnetic) fields, respectively, we have $\mathbf{E} \mapsto \mathbf{E}$ and $\mathbf{B} \mapsto -\mathbf{B}$. This is consistent with $\mathbf{A} \mapsto -\mathbf{A}$ for the vector potential, while $\phi \mapsto \phi$ for the scalar potential.[5] Hence, also the kinematic momentum $m\mathbf{v} = \mathbf{p} - e\mathbf{A}$ changes sign under time reversal. This corresponds to the intuitive picture of 'running the film backwards'.

[5]Recall that $\mathbf{E} = -\partial_t\mathbf{A} - \nabla\phi$.

Again, the classical equation of motion (2.116) is covariant under time reversal: the coordinate time is flipped, $\mathrm{d}t' = -\mathrm{d}t$, but the proper time $\mathrm{d}\tau$ is invariant.

We shall see that under time reversal,

$$T : \psi(t, \mathbf{x}) \mapsto \psi'(t', \mathbf{x}') = \psi^*(-t', \mathbf{x}')$$

where ψ^* is the complex conjugate wave function.

Charge conjugation

This operation flips the sign of the charge $e \mapsto e' = -e$. Eq.(2.116) is covariant only if we also take $m' = -m$, but this does not make much sense in the classical context. In the quantum context, with 'antiparticles', this is easier to understand, but one has to go to the relativistic level. We shall see that it is actually not necessary to change the sign of m.

2.4.2 Nonrelativistic quantum mechanics

Parity

The wave function transforms as given in Eq.(2.117) under the parity operator. With this, it is easy to check that the Schrödinger equation for a charged, nonrelativistic particle is covariant under parity:

$$\begin{aligned}
\mathrm{i}\partial_{t'}\psi' &= \frac{\hbar^2}{2m}\left(-\mathrm{i}\nabla' - e'\mathbf{A}'\right)^2 \psi' + e'\phi'\,\psi' \\
= \mathrm{i}\partial_t \psi(t, -\mathbf{x}') &= \frac{\hbar^2}{2m}\left(\mathrm{i}\nabla + e\mathbf{A}\right)^2 \psi(t, -\mathbf{x}') + e\phi\,\psi(t, -\mathbf{x}')
\end{aligned} \qquad (2.118)$$

and this is identical to the original Schrödinger equation, evaluated at $t, \mathbf{x}$.

Note that letting the scalar potential ϕ be invariant under parity is consistent with the rule for the electric field vector. A subtlety: of course, the function ϕ has to be evaluated at the transformed coordinates, as required for a scalar field (2.117).

The familiar picture of the 'even potential well' with even and odd eigenfunctions only applies if under this transformation, the function is actually the same, $\phi'(\mathbf{x}') = \phi(\mathbf{x}')$. This illustrates the subtle difference between an equation being covariant (of the same form) or invariant (does not change in numerical value).

Time reversal

Time reversal is already more tricky here. Definitely, the derivative $\partial_t \mapsto -\partial_t$. In addition, the vector potential $\mathbf{A} \mapsto -\mathbf{A}$ as mentioned above. To compensate for both changes a single trick is sufficient: take the complex conjugate of the Schrödinger equation. The transformation under time reversal is therefore, as announced before:

$$T : \psi'(t', \mathbf{x}') = \psi^*(-t', \mathbf{x}') \qquad (2.119)$$

To show that the Schrödinger equation is covariant, we write it down for the primed quantities and use the transformation laws (2.119) and the rules for the electromagnetic potentials. Taking the complex conjugate of the resulting equation leads to the Schrödinger equation for the unprimed quantities.

Charge conjugation

The prescription here would be $e \mapsto e' = -e$ and $\psi(t, \mathbf{x}) \mapsto \psi^*(t, \mathbf{x})$. Note that nothing happens with the coordinates. This operation, however, does *not* work at the level of non-relativistic quantum mechanics: unless one also flips m into $-m$.

2.4.3 Dirac equation

The Dirac equation in an electromagnetic potential can be written in the form

$$\gamma^0 \left(\mathrm{i}\,\partial_t - e\phi\right) \Psi + \boldsymbol{\gamma} \cdot \left(\mathrm{i}\nabla + e\mathbf{A}\right) \Psi = m\Psi \tag{2.120}$$

where $\boldsymbol{\gamma} = (\gamma^1, \gamma^2, \gamma^3)$ is formally a 3-vector and contains $\pm\boldsymbol{\sigma}$ as off-diagonal elements.

Parity

The spinor transforms like

$$P : \Psi'(t', \mathbf{x}') = S_P \Psi(t', -\mathbf{x}')$$

where the 4×4 matrix S_P remains to be found. We require that the Dirac equation be covariant under the changes $\nabla \mapsto -\nabla$ and $\mathbf{A} \mapsto -\mathbf{A}$. Writing the Dirac equation for Ψ' and using the spinor transformation, one gets

$$S_P^{-1} \gamma^0 \left(\mathrm{i}\,\partial_t - e\phi\right) S_P \Psi + S_P^{-1} \boldsymbol{\gamma} \cdot \left(-\mathrm{i}\nabla - e\mathbf{A}\right) S_P \Psi = m\Psi \tag{2.121}$$

Hence, we need a matrix that satisfies

$$S_P^{-1} \gamma^0 S_P = \gamma^0, \qquad S_P^{-1} \gamma^k S_P = -\gamma^k \tag{2.122}$$

This matrix is easy to find: due to the basic anti-commutation relations between the Dirac matrices, one has

$$S_P = \mathrm{const}\gamma^0 \tag{2.123}$$

where the constant is not yet fixed. We fix it by the requirement that the Lorentz scalar $\overline{\Psi}\Psi$ be invariant under parity. This amounts to

$$S_P^\dagger \gamma^0 S_P = \gamma^0 \tag{2.124}$$

which gives $|\mathrm{const}|^2 = 1$, hence a freedom up to a phase factor. We choose the simplest solution with $S_P = \gamma^0$ or

$$P : \Psi'(t', \mathbf{x}') = \gamma^0 \Psi(t', -\mathbf{x}') \tag{2.125}$$

In this case, $(S_P)^2 = \mathbb{1}$. (Note that a global phase factor could have occurred here.)

Time reversal

As in the Schrödinger case, time reversal involves the complex conjugate of the (spinor) wave function, in addition to the sign flip in the time coordinate:

$$T : \Psi'(t', \mathbf{x}') = S_T \Psi^*(-t', \mathbf{x}')$$

A calculation similar to the previous one, using the transformation of the electromagnetic potential under time reversal, leads to

$$\begin{aligned} & S_T^{-1}\gamma^0 \left(-\mathrm{i}\,\partial_t - e\phi\right) S_T \Psi^*(-t', \mathbf{x}') \\ & + S_T^{-1}\boldsymbol{\gamma}\cdot\left(\mathrm{i}\nabla - e\mathbf{A}\right) S_T \Psi^*(-t', \mathbf{x}') = m\Psi^*(-t', \mathbf{x}') \end{aligned} \tag{2.126}$$

taking the complex conjugate of this equation, we get the covariant derivatives right, but the Dirac matrices are getting conjugated:

$$\begin{aligned} & S_T^{-1*}(\gamma^0)^* \left(\mathrm{i}\,\partial_t - e\phi\right) S_T^* \Psi(-t', \mathbf{x}') \\ & + S_T^{-1*}\boldsymbol{\gamma}^{\cdot} \left(-\mathrm{i}\nabla - e\mathbf{A}\right) S_T^* \Psi(-t', \mathbf{x}') = m\Psi(-t', \mathbf{x}') \end{aligned} \tag{2.127}$$

We thus require for the spinor transformation the properties

$$S_T^{-1*}\gamma^{0*}S_T^* = \gamma^0, \qquad S_T^{-1*}\gamma^{k*}S_T^* = -\gamma^k \tag{2.128}$$

This is already more difficult to solve. To proceed, we need the following identities:

$$\gamma^2\gamma^{\mu*}\gamma^2 = \gamma^\mu \tag{2.129}$$

$$\gamma^0\gamma^k\gamma^0 = -\gamma^k \tag{2.130}$$

$$\gamma_5\gamma^\mu\gamma_5 = -\gamma^\mu \tag{2.131}$$

that you are invited to prove in the exercises. Here, γ^2 is the (only) Dirac matrix with imaginary matrix elements (in the usual representation), and

$$\gamma_5 = \mathrm{i}\,\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{2.132}$$

is a matrix that anti-commutes with the four Dirac matrices (see Eq.(2.131)).

Let us try the strange combination $S_T = \mathrm{i}\,\gamma^2\gamma^0\gamma_5$. One has

$$S_T^* = -\mathrm{i}\,(\gamma^2)^*(\gamma^0)^*(\gamma_5)^* = \mathrm{i}\,\gamma^2\gamma^0\gamma_5 \tag{2.133}$$

since γ^0 and γ_5 are real and γ^2 is purely imaginary. In addition, $(S_T)^{-1*} = \mathrm{i}\,\gamma_5\gamma^0\gamma^2$ because $(\gamma^2)^2 = -\mathbb{1}$, and $(\gamma^0)^2 = (\gamma_5)^2 = +\mathbb{1}$. We thus have to compute

$$\begin{aligned} (S_T)^{-1*}\gamma^{0*}S_T^* &= -\gamma_5\gamma^0\gamma^2\gamma^0\gamma^2\gamma^0\gamma_5 \\ &\overset{(2.129)}{=} -\gamma_5\gamma^0\gamma^0\gamma^0\gamma_5 \\ &= -\gamma_5\gamma^0\gamma_5 \overset{(2.131)}{=} \gamma^0 \end{aligned} \tag{2.134}$$

And for the spatial components:

$$\begin{aligned} (S_T)^{-1*}\gamma^{k*}S_T^* &= -\gamma_5\gamma^0\gamma^2\gamma^{k*}\gamma^2\gamma^0\gamma_5 \\ &\overset{(2.129)}{=} -\gamma_5\gamma^0\gamma^k\gamma^0\gamma_5 \\ &\overset{(2.130)}{=} +\gamma_5\gamma^k\gamma_5 \\ &\overset{(2.131)}{=} -\gamma^k \end{aligned} \tag{2.135}$$

Note that one really needs the three matrices here: γ^2 'undoes' the complex conjugation, γ^0 flips the sign of the spatial components only, and γ_5 arranges for a global sign change.

So in short, time reversal on the level of the Dirac equation amounts to

$$T : \Psi'(t', \mathbf{x}') = \mathrm{i}\,\gamma^2\gamma^0\gamma_5\Psi^*(-t', \mathbf{x}') = -\mathrm{i}\begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}\Psi^*(-t', \mathbf{x}') \tag{2.136}$$

Note that applying this transformation twice gives a global minus sign.

Charge conjugation

The final transformation is charge conjugation. We require that the Dirac equation be covariant with respect to a flipped charge $e \mapsto e' = -e$. To get the covariant derivatives $\mathrm{i}\partial_\mu - e'A_\mu$ right, we need a complex conjugation. The remaining transformation can be found as

$$C : \Psi'(t', \mathbf{x}') = \mathrm{i}\,\gamma^2\Psi^*(t', \mathbf{x}') = \mathrm{i}\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}\Psi^*(t', \mathbf{x}') \tag{2.137}$$

Indeed, the Dirac equation, after complex conjugation, yields

$$S_C^{-1*}\gamma^{\mu*}S_C^* = -\gamma^\mu \tag{2.138}$$

which can indeed be satisfied with $S_C = \mathrm{i}\gamma^2$ because of Eq.(2.129) and $S_C^* = (S_C)^{-1*} = \mathrm{i}\gamma^2$.

Note that under charge conjugation, upper and lower spinors are getting exchanged (and σ_2 is applied). This transforms 'particle-like' solutions into 'antiparticle-like' solutions.

Dirac spinor transformation

Let us consider as an example a 'particle-like' plane wave solution and apply the discrete transformations found so far.

Parity changes the sign of the spatial momentum (change sign of $\mathbf{x}$ and apply γ^0)

$$P : \Psi(x) = \begin{pmatrix} \psi_0 \\ \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\psi_0 \end{pmatrix} \mathrm{e}^{-\mathrm{i}Et+\mathrm{i}\mathbf{p}\cdot\mathbf{x}} \mapsto \begin{pmatrix} \psi_0 \\ \dfrac{-\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\psi_0 \end{pmatrix} \mathrm{e}^{-\mathrm{i}Et-\mathrm{i}\mathbf{p}\cdot\mathbf{x}} \tag{2.139}$$

Time reversal (flip sign of t, take complex conjugate and apply $-\mathrm{i}\,\sigma_2$ to both spinor components)

$$T : \Psi(x) \mapsto -\mathrm{i}\begin{pmatrix} \sigma_2\psi_0^* \\ \dfrac{\sigma_2(\boldsymbol{\sigma}^*\cdot\mathbf{p})}{E+m}\psi_0^* \end{pmatrix} \mathrm{e}^{-\mathrm{i}Et-\mathrm{i}\mathbf{p}\cdot\mathbf{x}} \tag{2.140}$$

From the Pauli matrices, one has (exercise!) $\sigma_2(\boldsymbol{\sigma}^*\cdot\mathbf{p}) = -(\boldsymbol{\sigma}\cdot\mathbf{p})\sigma_2$ so that this spinor is nearly identical to the spatial mirror image (2.139), up to a global phase and the change $\psi_0 \mapsto \sigma_2\psi_0^*$. One can show (exercise!) that this flips the sign of the spin quantum number. (This is compatible with the mapping $\mathbf{L} \mapsto -\mathbf{L}$ under time reversal.)

Charge conjugation:

$$C : \Psi(x) \mapsto \mathrm{i} \begin{pmatrix} \frac{\sigma_2(\boldsymbol{\sigma}^* \cdot \mathbf{p})}{E+m} \psi_0^* \\ -\sigma_2 \psi_0^* \end{pmatrix} \mathrm{e}^{\mathrm{i}Et - \mathrm{i}\mathbf{p}\cdot\mathbf{x}} = -\mathrm{i} \begin{pmatrix} \frac{(\boldsymbol{\sigma} \cdot \mathbf{p})}{E+m} \sigma_2 \psi_0^* \\ \sigma_2 \psi_0^* \end{pmatrix} \mathrm{e}^{\mathrm{i}Et - \mathrm{i}\mathbf{p}\cdot\mathbf{x}} \tag{2.141}$$

This is an antiparticle-like solution with flipped spin. Indeed, if we apply a boost to $(0, \chi_0)^T \mathrm{e}^{+\mathrm{i}mt}$, we get

$$\Psi(x) = \begin{pmatrix} (\boldsymbol{\sigma} \cdot \mathbf{n}) \operatorname{sh}(\beta/2) \chi_0 \\ \operatorname{ch}(\beta/2) \chi_0 \end{pmatrix} \mathrm{e}^{\mathrm{i}m \operatorname{ch}(\beta) t - \mathrm{i}m\mathbf{n} \operatorname{sh}(\beta) \cdot \mathbf{x}}$$

which is the same as (2.141) up to a normalization, provided $E = m \operatorname{ch} \beta$ and $\mathbf{p} = m\mathbf{n} \operatorname{sh} \beta$. Note also that these are negative energy solutions at the energy $-E$.

Exercise: the complete transformation CPT gives just swap of the 2-spinors ψ and χ.

Chapter 3

Extra: symmetries and groups

3.1 Definition

definition of symmetry operation: probabilities are invariant.

check that this is true for charge conjugated spinor: maps onto complex conjugate probability amplitudes.

Wigner's theorem: either linear unitary or antilinear unitary, determined up to a phase factor.

'anti-linear operation'

$$K\left(\alpha|0\rangle+\beta|1\rangle\right)=\alpha^{*}K|0\rangle+\beta^{*}K|1\rangle$$

examples: time reversal, charge conjugation.

Symmetry group: set of symmetry operations. Is closed under composition/concatenation (*hintereinander ausführen*).

Definition of a group.

Discrete groups Examples: P, T, C because C C is $\mathbb{1}$ up to phase. Together, they span a group with 8 elements.

3.2 Continuous symmetry groups

3.2.1 Examples

rotations SO(d), geometrical definition (scalar product invariant). $R^{T}R=\mathbb{1}$ and $\det(R)=+1$.

Dimension of this group: $\frac{1}{2}d(d-1)$. $R^{T}R=\mathbb{1}$ is a symmetric matrix equation.

complex 'rotations': SU(n), complex scalar product invariant. $U^{\dagger}U=\mathbb{1}$ and $\det U=+1$.

Dimension: $n^{2}-1$. $U^{\dagger}U=\mathbb{1}$ is a hermitean matrix equation, and fixes only $|\det U|=1$.

3.2.2 Generators

one-parameter subgroup. convenient parametrization $R(\theta_1)R(\theta_2) = R(\theta_1+\theta_2)$ with modulo addition if required.

generator of this subgroup, infinitesimal operation. $R(\theta) = \exp(\theta L)$ with

$$L = \frac{\mathrm{d}R(0)}{\mathrm{d}\theta} \tag{3.1}$$

generator L

3.2.3 Lie algebra

commutator and structure constants: $R(\theta_1, \ldots)$, generators L_i.

Multiplication of group elements: addition in Lie algebra.

Only structure where multiplication of group elements occurs (vanishing linear order)

$$R_1^{-1}R_2^{-1}R_1R_2 \approx \mathbb{1} + \theta_1\theta_2\,[L_1,\,L_2] \stackrel{!}{=} \mathbb{1} + \sum_k \alpha_k(\theta_1,\theta_2)L_k \tag{3.2}$$

use that all generators span the (tangent space at the unit element of the) group. Coefficients α_k can be zero, then L_1 and L_2 commute (the corresponding subgroups also). In general case:

$$[L_i,\,L_j] = \sum_l f_{ijk}L_k, \qquad f_{ijk} = \frac{\partial^2\alpha_k}{\partial\theta_i\partial\theta_j} \tag{3.3}$$

Exercise: rotation group SO(3):

$$f_{ijk} = \epsilon_{ijk}, \qquad (L_i)_{kl} = -\epsilon_{ikl} \tag{3.4}$$

up to signs.

3.2.4 Group representation

Rotation group in quantum mechanics: group representation. For each rotation R, there exists a unitary (or anti-unitary) symmetry operator $S(R)$. The $S(R)$ form a group. Convenient to choose the mapping S as compatible with group products:

$$S(R_1)S(R_2) = \mathrm{e}^{\mathrm{i}\varphi(R_1,R_2)}S(R_1R_2) \tag{3.5}$$

phase factor $\mathrm{e}^{\mathrm{i}\varphi(R_1,R_2)}$ as allowed by Wigner's theorem. Mathematical structure: homomorphism between groups, (projective) group representation. Always comes with a vector space $\mathcal{H}$ containing the ψ on which the $S(R)$ act.

Continuous group: from (3.5), $S(\mathbb{1})$ acts up to a phase factor like the identity on $S(R_2)$. Remove this phase by redefinition and get $S(\mathbb{1}) = \mathbb{1}$. (Note: unit matrices on different spaces.) In particular, this is a linear and unitary operator. By continuity, $S(R)$ is unitary. (Anti-linear operators can only occur in a finite/discrete group.)

Definition if $S(R)$ unitary: Eq.(3.5) defines a (projective) unitary representation of the rotation group. In many cases, one can choose unitary matrices in the group SU(n), and then the phase factor $e^{i\varphi(R_1,R_2)} = 1$. This corresponds to a unitary (not projective) representation.

Example: scalar wave function

$$S(R)\psi := \psi \circ R^{-1}, \qquad \text{explicitly:} \quad [S(R)\psi](\mathbf{x}') := \psi(R^{-1}\mathbf{x}') \tag{3.6}$$

identical to the transformation of a 'scalar field'. Before, we had the notation $\mathbf{x}' = R\mathbf{x}$ and $\psi'(\mathbf{x}') = \psi(\mathbf{x}) = \psi(R^{-1}\mathbf{x}')$. For complex-valued ψ, this is a unitary representation. Check that the norm of ψ is unchanged: interpret in terms of the L^2-norm:

$$\begin{aligned} \|\psi'\|^2 &= \int \mathrm{d}x' |\psi'(\mathbf{x}')|^2 \\ &= \int \mathrm{d}x' |\psi(R^{-1}\mathbf{x}')|^2 \qquad \text{substitute } \mathbf{y} = R^{-1}\mathbf{x}' \\ &= \int \mathrm{d}y |\det R| |\psi(\mathbf{y})|^2 \end{aligned} \tag{3.7}$$

Now, $|\det R| = 1$ for a rotation matrix, hence this is equal to $\|\psi\|^2$.

Example: trivial representation $S(R) = 1$ for all R.

Example: vector field $\mathbf{A}(\mathbf{x})$

$$S(R)\psi := R\mathbf{A} \circ R^{-1}, \qquad \text{explicitly:} \quad [S(R)\mathbf{A}](\mathbf{x}') := R\mathbf{A}(R^{-1}\mathbf{x}') \tag{3.8}$$

Example: on the space of 3×3 matrices A,

$$S(R)A = RAR^{-1} \tag{3.9}$$

equivalent to the transformation for a linear vector field $\mathbf{A}(\mathbf{x}) = A\mathbf{x}$ or more generally for a 'tensor of rank two': $S(R)A$ gives the components of the tensor in a rotated basis.

All these: 'real representations'.

Example: spinor wave function $\Psi(\mathbf{x})$,

$$S(R)\Psi = S_R\Psi \circ R^{-1}, \qquad \text{explicitly:} \quad [S(R)\Psi](\mathbf{x}') := S_R\Psi(R^{-1}\mathbf{x}') \tag{3.10}$$

with S_R the spinor transformation matrix; for the Dirac equation this is a unitary one.

Generators of group representation: take the image of a one-parameter subgroup

$$S(\theta) := S(R(\theta)) \tag{3.11}$$

for small θ, this gives $S \approx \mathbb{1} + i\theta J$ with a generator J in the Lie algebra of the representation. Convention with i: J hermitean iff S unitary. Hence as in the SO(d) case, only the derivative near the unit element is needed

$$iJ = \left.\frac{\mathrm{d}S}{\mathrm{d}\theta}\right|_{\theta=0}$$

to generate the subgroup: $S(\theta) = \exp(i\theta J)$.

The same commutation relations apply. As for the Lie algebra of the group itself, work out

$$S(R_1^{-1}R_2^{-1}R_1R_2) = S(R_1^{-1})S(R_2^{-1})S(R_1)S(R_2) \tag{3.12}$$

Left hand side gives image of one-parameter subgroup generated by $f_{12k}L_k$ with small parameter $\theta_1\theta_2$, hence $\mathrm{i}\theta_1\theta_2 f_{12k}J_k$. Right hand side gives commutator $\theta_1\theta_2\,[\mathrm{i}J_1,\ \mathrm{i}J_2]$, hence

$$\mathrm{i}\,[J_1,\ J_2] = f_{12k}J_k \tag{3.13}$$

hence the same structure constants as the original group. Brings restrictions for the possible 'representation spaces' and possible unitary matrices that are allowed in $S(R)$.

Remark: commutator $\mathrm{i}[J_k,\ J_l]$ is natural product in Lie algebra of SU(n): it is hermitean and of trace zero (in the finite-dimensional case).

3.3 (Ir)reducible representations of the rotation group

Idea: break up 'large' matrices $S(R)$ into smaller ones.

Example: Dirac spinor transformation under rotation.

Example: the rotation subgroup of the Lorentz group $S(R) = \mathbb{1} \oplus R$, acting on Minkowski four-vectors.

$$\mathbb{1} \oplus R := \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R \end{pmatrix} \tag{3.14}$$

Expand scalar wave function in eigenfunctions of the spherical harmonics $Y_{lm}(\theta, \varphi)$ with orbital angular momentum quantum numbers l, m. The quantum number l of the basis functions does not change by rotation. The action of the rotation group thus breaks up in separate subspaces of dimension $2l + 1$.

All these examples: group is represented by block-diagonal unitary matrices in suitable basis. Block-diagonal matrices are examples of a "direct sum" $A \oplus B$ of two operators. One constructs the underlying vector space by joining two basis sets for $\mathcal{H}_A$ and $\mathcal{H}_B$. In the matrix for the direct sum,

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \tag{3.15}$$

On the "first half" (upper components of the vector), the operator A acts, on the second half, the operator B acts. If we use the suggestive notation $a + b$ for for a vector in $\mathcal{H}_{A\oplus B}$, we have

$$(A \oplus B)(a + b) = Aa + Bb$$

which makes perfectly transparent that we simply add the actions of A and B.

3.3.1 All representations

The concept of 'breaking up' a representation is very useful because one can work in a small subspace of the original Hilbert space. This has been used, for example, for a rotationally symmetric Hamilton operator where the eigenfunctions could be found in the smaller subspace of a fixed quantum number for the angular momentum operator $\mathbf{L}^2$.

We recall from the QM lecture the following fundamental property that derives from the commutation relations (3.13) for the unitary matrices that represent the rotation group:

> For each $n = 2j+1 \geq 0$, there exists a vector space of dimension n spanned by the vectors $|m\rangle$, $m = -j, ..., j-1, j$ that transform into each other by action of the generators J_x, J_y, J_z. The generators have a simple band-diagonal form: J_z is diagonal, and $J_\pm = J_x \pm \mathrm{i} J_y$ have entries on the diagonal just above or below the central diagonal. The rotation group is thus represented by $n \times n$ dimensional unitary matrices that are generated by the subgroups $\exp[\mathrm{i}\theta \mathbf{n} \cdot \mathbf{J}]$. We shall call this representation in the following the "(j, m)-representation".

One can show the following two theorems:

Theorem. Any irreducible, finite-dimensional representations of the rotation group SO(3) is equivalent to a (j, m)-representation for some value of $j = 0, \frac{1}{2}, 1, ...$.

Theorem. Any unitary, finite-dimensional representation of the rotation group can be decomposed into a direct sum of some of the (j, m)-representations.

We are not going to demonstrate these theorems in detail. Even in some of the advanced physics books, there are some points that are left open, and there is not the time here to fill all the gaps.

Equivalence of representations

The word 'equivalent' in the previous theorem has the following meaning: two representations $S(R)$ and $S'(R)$ are called equivalent if a non-singular matrix T exists with the property that

$$S'(R) = T^{-1} S(R) T \tag{3.16}$$

for all $R \in$ SO(3).

As an alternative formulation, we can write

$$S'(R) T \psi = T S(R) \psi, \qquad \text{for all } R \text{ and all } \psi \tag{3.17}$$

Now, this looks just like a change of basis for the Hilbert space: instead of the components ψ, one works with $T\psi$. Compare to Eq.(3.9) for rotation acting on a matrix. So, the equivalent representation S' is 'essentially the same' representation as S.

A representation whose matrices can be brought in block-diagonal form (for all group elements!) by a change of basis is called **reducible**. It is then possible to restrict the underlying vector space to the upper basis vectors: on this subspace, the upper block of the representation matrices gives a well-defined representation with a smaller dimension. Such a sub-space is called an **invariant subspace** $\mathcal{K}$: if $\psi \in \mathcal{K}$, then $S(R)\psi \in \mathcal{K}$ for all group elements $S(R)$ applied to it.

This leads us the concept of an *irreducible representation.* This is a representation where all invariant subspaces are 'trivial' (they are either the full space $\mathcal{H}$ or the space $\{0\}$ spanned by the null vector 0.

The second of the previous theorems shows that any unitary representation is either irreducible or completely reducible.

Two nice observations that help in proving these results are the following:

1. if T is a linear operator that satisfies Eq.(3.17), then its kernel $\ker T = \{\psi \in \mathcal{H} | T\psi = 0\}$ and its image $T\mathcal{H}$ are invariant subspaces for the representations S and S', respectively.

2. if $\mathcal{K}$ is an invariant subspace of a unitary representation, then its orthogonal complement $\mathcal{K}^{\perp} = \{\psi \in \mathcal{H} | \langle\psi|\chi\rangle = 0 \text{ for all } \chi \in \mathcal{K}\}$ is invariant as well.

3.3.2 Application: addition of angular momenta