Institut für Physik und Astronomie, Universität Potsdam
Wintersemester 2019/20

Quantenmechanik II
— Höhere Theoretische Physik —

Carsten Henkel

Baustelle 'Skriptum'

Überblick

Diese Notizen sind eine Sammlung von Bausteinen, die in den Jahren 2006 und 2007 entstanden sind, als die Vorlesung zum ersten und zweiten Mal gehalten wurde. Im WS 19/20 ist der Inhalt grundlegend umstrukturiert worden: Vielteilchenphysik, relativistische Quantenmechanik, Elementarteilchen. Nur der zweite Abschnitt (Relativität) ist für das WS 19/20 leicht überarbeitet worden (Details zu Diracs Lösung des Wasserstoff-Atoms). Material zu den anderen Abschnitten wurde auf BoxUP bereitgestellt.

Es gilt: diese Notizen haben keinen Anspruch auf Vollständigkeit, und Fehler können überall auftreten. Sie enthalten einige Abschnitte, die in der Vorlesung nicht durchgenommen wurden: zur Lektüre empfohlen, aber nicht Pflicht.

- Zusammenfassung der Grundlagen der Quantenmechanik
  → Kapitel 0
- Streutheorie
  → Kapitel 1: Wirkungsquerschnitt, Bornsche Näherung, S-Matrix, Partialwellenentwicklung,
  → Anhang 2 zeitabhängige Störungstheorie
- relativistische Quantenmechanik
  → Kapitel 3: Klein-Gordon-Gleichung, Dirac-Gleichung, Spinor-Transformation, 1-Teilchen-Zustände, diskrete Symmetrien
  → Kapitel 4: Symmetrien und Gruppen in der Quantenmechanik, Darstellungen der Drehgruppe, Addition von Drehimpulsen
- Einführung in die Quantenfeldtheorie
  → Kapitel 5: Phononen auf diskretem Gitter, kanonische Quantisierung, skalares, relativistisches Feld (Klein-Gordon-Feld), Zerfall von Teilchen, Quantisierung von Fermionen, Pauli-Prinzip und Vielteilchen-Theorie

Chapter 3

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.

3.1 Relativistic classical mechanics

Relativistic mechanics: the relevant equations of motion for a charged point particle can be derived from the following action

$$\begin{aligned} S &= -mc^2 \int \mathrm{d}\tau - e \int \mathrm{d}x^\mu A_\mu(x) \\ &= -mc^2 \int \mathrm{d}t \sqrt{1 - \dot{\mathbf{x}}^2/c^2} - e \int \mathrm{d}t \left(\phi(x) - \dot{\mathbf{x}} \cdot \mathbf{A}(x)\right) \end{aligned} \tag{3.1}$$

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

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = \frac{m\dot{\mathbf{x}}}{\sqrt{1 - \dot{\mathbf{x}}^2/c^2}} + e\mathbf{A} \tag{3.2}$$

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

Energy

$$E = \dot{\mathbf{x}}\mathbf{p} - L = \frac{mc^2}{\sqrt{1 - \dot{\mathbf{x}}^2/c^2}} + e\phi \tag{3.3}$$

combine into a ("covariant") 4-vector

$$p_\mu = (E/c, -\mathbf{p}) = m\frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} + eA_\mu \tag{3.4}$$

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

From these expressions, one can show that the Minkowski norm of the momentum satisfies the following 'dispersion relation' (a momentum-energy relation)

$$(p_\mu - eA_\mu)(p^\mu - eA^\mu) = \frac{(E-e\phi)^2}{c^2} - (\mathbf{p}-e\mathbf{A})^2 = (mc)^2 \tag{3.5}$$

Finally, the covariant form of the equation of motion of the charge

$$\frac{\mathrm{d}p_\mu}{\mathrm{d}\tau} = eF_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \tag{3.6}$$

with the so-called Faraday tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{3.7}$$

This is the 4-vector formulation of the Coulomb-Lorentz force.

3.2 Klein-Gordon equation

Cook book recipe

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

$$E = \frac{(\mathbf{p}-e\mathbf{A})^2}{2m} + e\phi \tag{3.8}$$

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

$$E \mapsto \mathrm{i}\hbar\partial_t, \qquad \mathbf{p} \mapsto -\mathrm{i}\hbar\nabla \tag{3.9}$$

and get the Schrödinger equation (in an electromagnetic field)

$$\mathrm{i}\hbar\partial_t\psi = -\frac{\hbar^2}{2m}[\nabla - \mathrm{i}(e/\hbar)\mathbf{A}]^2\psi + e\phi\psi \tag{3.10}$$

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 Eq.(3.5). The replacement rules (3.9) are equivalent to

$$p_\mu \mapsto \mathrm{i}\hbar\partial_\mu \tag{3.11}$$

Note that the signs come out correctly because the covariant derivative is $\partial_\mu = \partial/\partial x^\mu = ((1/c)\partial_t, \nabla)$. One gets

$$\begin{aligned}&[\mathrm{i}\hbar\partial_\mu - eA_\mu][\mathrm{i}\hbar\partial^\mu - eA^\mu]\psi \\ &= -\frac{\hbar^2}{c^2}(\partial_t + \mathrm{i}(e/\hbar)\phi)^2\psi + \hbar^2(\nabla - \mathrm{i}(e/\hbar)\mathbf{A})^2\psi = (mc)^2\psi\end{aligned} \tag{3.12}$$

or, in the free field case

$$\boxed{\left[\frac{1}{c^2}\partial_t^2 - \nabla^2 + \frac{m^2c^2}{\hbar^2}\right]\psi = 0} \tag{3.13}$$

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)

$$S = \int \mathrm{d}t\,\mathrm{d}^3x \left[\frac{\hbar^2}{c^2}\partial_t\psi^*\partial_t\psi - \hbar^2(\nabla\psi^*)\cdot(\nabla\psi) - (mc)^2\psi^*\psi\right] \tag{3.14}$$

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

$$\psi'(x') = \psi(x) = \psi(\Lambda^{-1}x'), \qquad \text{or } \psi' = \psi\circ\Lambda^{-1} \tag{3.15}$$

Compare to a vector field and coordinate rotation R:

$$E_i'(\mathbf{x}') = R_{ij}E_j(\mathbf{x}) = R_{ij}E_j(R^{-1}\mathbf{x}'), \qquad \text{or } \mathbf{E}' = R(\mathbf{E}\circ R^{-1}) \tag{3.16}$$

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 $\mathcal{L}$:

$$0 = \frac{\partial\mathcal{L}}{\partial\psi^*} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi^*)} \tag{3.17}$$

$$= -(mc)^2\psi - \hbar^2\partial_\mu\partial^\mu\psi \tag{3.18}$$

hence we recover again the Klein-Gordon equation.

Current conservation

Conserved current for Klein-Gordon field:

$$j^\mu = (nc, \mathbf{j}) = \frac{\mathrm{i}\hbar}{2m}\left(\psi^*\partial^\mu\psi - \partial^\mu\psi^*\psi\right) \tag{3.19}$$

easy to check that is conserved:

$$\partial_\mu j^\mu = \partial_t n + \nabla \cdot \mathbf{j} = 0 \tag{3.20}$$

from the Klein-Gordon equation (3.18) and its complex conjugate.

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

$$n = \frac{\mathrm{i}\hbar}{2mc^2}\left(\psi^*\partial_t\psi - \partial_t\psi^*\psi\right) = \frac{E}{mc^2}\psi^*\psi \gtrless 0 \tag{3.21}$$

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:

$$E = \pm\sqrt{\mathbf{p}^2c^2 + (mc^2)^2} \tag{3.22}$$

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)

$$\left(\partial_\mu\partial^\mu + m^2\right) G(x) = \delta(x) \tag{3.23}$$

Solution with frequency Fourier transform

$$G(x) = \frac{1}{4\pi|\mathbf{x}|}\int\frac{\mathrm{d}\omega}{2\pi}\exp[-\mathrm{i}\omega t + \mathrm{i}|\mathbf{x}|\sqrt{\omega^2 - m^2}] \tag{3.24}$$

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

$$G(\mathbf{x};\omega\to 0) = \frac{1}{4\pi|\mathbf{x}|}\mathrm{e}^{-m|\mathbf{x}|} \tag{3.25}$$

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.

Nonrelativistic limit: work directly on Klein-Gordon Eq.(3.13), factoring off $\mathrm{e}^{-\mathrm{i}mt}$ that corresponds to the rest mass.

Symmetries

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

$$\psi'(x') = \psi(x). \tag{3.26}$$

where $x' = \Lambda x$ are the 'new coordinates' of a space-time event after a Lorentz transformation Λ (a 4×4 matrix).

The Klein-Gordon action is invariant under complex conjugation:

$$\psi(x) \mapsto \psi^*(x). \tag{3.27}$$

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.

3.2.1 Gauge symmetry and minimal coupling

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

$$\psi'(x) = \mathrm{e}^{\mathrm{i}\varphi}\psi(x) \tag{3.28}$$

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

$$\psi'(x) = \mathrm{e}^{\mathrm{i}\varphi(x)}\psi(x) \tag{3.29}$$

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

$$\mathrm{i}\partial_\mu \psi' = \mathrm{e}^{\mathrm{i}\varphi(x)} \left(\mathrm{i}\partial_\mu - \partial_\mu \varphi\right) \psi \tag{3.30}$$

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 ($\hbar = 1$)

$$\begin{aligned} &\left(\mathrm{i}\partial_\mu - eA'_\mu\right) \psi' = \mathrm{e}^{\mathrm{i}\varphi(x)} \left(\mathrm{i}\partial_\mu - \partial_\mu\varphi - eA'_\mu\right) \psi \\ &= \mathrm{e}^{\mathrm{i}\varphi(x)} \left(\mathrm{i}\partial_\mu - eA_\mu\right) \psi \end{aligned} \tag{3.31}$$

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:

$$eA'_\mu = eA_\mu - \partial_\mu\varphi = e\left(A_\mu + \partial_\mu\chi\right) \tag{3.32}$$

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

This can be formalized by introducing the 'covariant derivative'

$$\mathrm{i}D_\mu := \mathrm{i}\partial_\mu - eA_\mu \tag{3.33}$$

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

$$S = \int \mathrm{d}t\, \mathrm{d}^3x \left[(D_\mu\psi)^*(D^\mu\psi) - (mc)^2\psi^*\psi\right] \tag{3.34}$$

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

$$\left\{ \begin{array}{rcl} \psi'(x) & = & \mathrm{e}^{-\mathrm{i}e\chi(x)/\hbar}\psi(x) \\ A'_\mu & = & A_\mu - \partial_\mu\chi \end{array} \right. \tag{3.35}$$

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,

$$(\mathrm{i}\partial_\mu - eA_\mu)\psi \mapsto (-\mathrm{i}\partial_\mu - eA_\mu)\psi^* = -(\mathrm{i}\partial_\mu + eA_\mu)\psi^* \tag{3.36}$$

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

3.3 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.

3.3.1 Heuristic *Ansatz*

Idea: the negative energies/frequencies in the Klein-Gordon equation arise because the equation contains second order derivatives in time. Find an equation with only first-order derivatives like the Schrödinger equation. *Ansatz* ($\hbar = c = 1$)

$$\mathrm{i}\partial_t \Psi = -\mathrm{i}\alpha^k \partial_k \Psi + \beta m \Psi \tag{3.37}$$

with numbers α^k and β to be fixed. We require that applying this differential operator twice on a plane wave with 4-momentum p^μ, one gets back the dispersion relation for a particle where m is the mass. Working out the square, this leads to

$$E^2 = \beta^2 m^2 + m \sum_k p^k \left(\alpha^k \beta + \beta \alpha^k\right) + \sum_{kl} p^k p^l \frac{1}{2} \left(\alpha^k \alpha^l + \alpha^l \alpha^k\right) \tag{3.38}$$

This reduced to $E^2 = m^2 + \mathrm{p}^2$ for all $\mathbf{p}$ only if

$$\beta^2 = 1 \tag{3.39}$$

$$\alpha^k \beta + \beta \alpha^k = 0 \tag{3.40}$$

$$\alpha^k \alpha^l + \alpha^l \alpha^k = 2\delta^{kl} \tag{3.41}$$

Hence, the α^k and β cannot be numbers. One has to work with non-commuting objects, and the simplest are matrices. They are called "Dirac matrices".

One can show that the minimum dimension one needs is four. The Dirac matrices are not unique, as a unitary transformation like $\beta \mapsto U\beta U^\dagger$ (use the same for the α^k's) leaves the relations (3.39–3.41) invariant. A particular representation that is often used is the "Dirac representation"

$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \alpha^k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \tag{3.42}$$

where the sub-blocks are 2×2 matrices, and the σ_k are the Pauli matrices. As a consequence, the Dirac wave function Ψ is also a four-component object. It is not a 4-vector like x^μ, however. Objects like this are called "spinors" or "spinor wavefunctions".

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

$$\frac{1}{c}\partial_t \mathbf{E} = \nabla \times c\mathbf{B} \tag{3.43}$$

$$\frac{1}{c}\partial_t c\mathbf{B} = -\nabla \times \mathbf{E} \tag{3.44}$$

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.

Covariant formulation

Covariance. One can sometimes read that the Dirac equation (3.37) is not invariant under parity or rotations since this changes the differential operator ∂_k, leading to a different form. This is not true, but the reason is subtle. In fact, we shall see that under this kind of transformations, the Dirac wave function Ψ must be changed as well. (It is not a 'scalar', nor is it a 4-vector.) The overall effect is similar to a change in the matrices β, α^k that exactly 'undoes' the change in the coordinate derivative. We shall study explicit examples when dealing with the rotation of the Dirac wave function and with Lorentz transformations.

A covariant formulation of the Dirac equation works with a 4-vector of matrices γ^μ, the so-called "gamma matrices" or Dirac matrices. One chooses $\gamma^0 = (\beta)^{-1} = \beta$ and gets

$$\left(\mathrm{i}\gamma^0\partial_t + \mathrm{i}\gamma^k\partial_k\right)\Psi = m\Psi \tag{3.45}$$

which is identical to Eq.(3.37) if we choose $\alpha^k = \gamma^0\gamma^k$. These matrices are called Dirac matrices and have in the standard representation the form

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \tag{3.46}$$

A Dirac equation that looks very covariant is then

$$\mathrm{i}\gamma^\mu\partial_\mu\Psi = m\Psi \tag{3.47}$$

The gamma matrices satisfy the following relations

$$\{\gamma^\mu,\, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \tag{3.48}$$

This is the mathematical structure of a so-called Clifford algebra.[1]

Note that the covariant vector $\partial_\mu = (\partial_t, \partial_k)$ is the derivative with respect to the contravariant coordinates x^μ: $\partial_\mu = \partial/\partial x^\mu$. This is so because for a scalar function φ, the differential

$$\mathrm{d}x^\mu\partial_\mu\varphi = \varphi(x^\mu + \mathrm{d}x^\mu) - \varphi(x^\mu) \tag{3.50}$$

is a scalar under Lorentz transformations. Observe that this is compatible with the energy-momentum operator in quantum mechanics:

$$E = \mathrm{i}\hbar\partial_t, \qquad p^k = -\mathrm{i}\hbar\frac{\partial}{\partial x^k}, \qquad p_\mu = (E, -p^k) = (\mathrm{i}\hbar\partial_t, \mathrm{i}\hbar\frac{\partial}{\partial x^k}) \tag{3.51}$$

[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

$$A_i \circ A_j + A_i \circ A_j = 2g_{ij}\,\mathbb{1} \tag{3.49}$$

with g_{ij} a diagonal matrix with entries $g_{11} = \ldots g_{pp} = 1$, $g_{p+1,p+1} = \ldots g_{nn} = -1$. The last q 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 in the algebra $\mathrm{Cl}(1,3)$ is the matrix multiplication.

Dirac Hamiltonian

We can write the Dirac equation in the Schrödinger-like form

$$\mathrm{i}\hbar\partial_t\Psi = H\Psi = \left(\gamma^0 mc^2 - \mathrm{i}\hbar c\gamma^0\gamma^i\partial_i\right)\Psi. \tag{3.52}$$

This defines the so-called Dirac Hamiltonian H, including factors $\hbar$ and c. 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. We have already seen the notation $\alpha^i = \gamma^0\gamma^i$ for the three hermitean matrices encountered here.

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

$$(\gamma^i)^\dagger = (\gamma^0\gamma^0\gamma^i)^\dagger = (\gamma^0\gamma^i)^\dagger(\gamma^0)^\dagger = (\gamma^0\gamma^i)\gamma^0 = -\gamma^i \tag{3.53}$$

because of the anti-commutation rule (3.48). Note that

$$(\alpha^i)^2 = (\gamma^0\gamma^i)^2 = \gamma^0\gamma^i\gamma^0\gamma^i = -\gamma^0\gamma^0\gamma^i\gamma^i = 1 \tag{3.54}$$

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

$$\mathrm{tr}\,(\gamma^i\gamma^0) \overset{(1)}{=} \mathrm{tr}\,(\gamma^0\gamma^i) \overset{(2)}{=} -\mathrm{tr}\,(\gamma^i\gamma^0) \tag{3.55}$$

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

$$\mathrm{tr}\,\gamma^0 = -\mathrm{tr}\,(\gamma^0\gamma^i\gamma^i) = -\mathrm{tr}\,(\gamma^i\gamma^0\gamma^i) = \mathrm{tr}\,(\gamma^0\gamma^i\gamma^i) = -\mathrm{tr}\,\gamma^0 \tag{3.56}$$

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

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}, \tag{3.57}$$

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

3.3.2 Plane-wave solutions

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

$$\mathrm{i}\hbar\partial_t\Psi = E\Psi_0\mathrm{e}^{-\mathrm{i}Et/\hbar} = mc^2\gamma^0\Psi_0\mathrm{e}^{-\mathrm{i}Et/\hbar} = mc^2\begin{pmatrix}\mathbb{1} & 0\\ 0 & -\mathbb{1}\end{pmatrix}\Psi_0\mathrm{e}^{-\mathrm{i}Et/\hbar} \tag{3.58}$$

We can thus distinguish four cases:

$$\text{'particles' with } E = mc^2\text{:} \qquad \Psi_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\text{'antiparticles' with } E = -mc^2\text{:} \qquad \Psi_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{3.59}$$

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 shall see later that the spin operator is given by

$$\mathbf{S} = \begin{pmatrix} \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix} \tag{3.60}$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices. The spinors (3.59) are eigenstates of the z-component of this operator, with eigenvalues $\pm 1/2$. Of course, any other two orthogonal two-component spinors would provide two independent solutions to the Dirac equation as well, they would be eigenstates to a different component of the spin operator, but still with eigenvalues $\pm 1/2$.

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 ($\Psi_0 = (\psi, \chi)^T$)

$$E\Psi_0 = m\gamma^0\Psi_0 + p^k\alpha^k\Psi_0, \qquad E\begin{pmatrix} \psi \\ \chi \end{pmatrix} = \begin{pmatrix} m & \mathbf{p}\cdot\boldsymbol{\sigma} \\ \mathbf{p}\cdot\boldsymbol{\sigma} & -m \end{pmatrix}\begin{pmatrix} \psi \\ \chi \end{pmatrix} \tag{3.61}$$

This gives, for example:

$$\chi = \frac{\mathbf{p}\cdot\boldsymbol{\sigma}}{E+m}\psi \tag{3.62}$$

and we can choose the 'upper' spinor ψ freely. We see that in the non-relativistic limit ($|E - m|, |\mathbf{p}| \ll m$), the lower spinor χ is much smaller in magnitude than the upper one. This is typical for the 'particle-like' solution. If it moves slowly enough, its 'admixture' of 'antiparticle-like' components (i.e., lower spinor) is small. In the opposite case (ultrarelativistic limit or nearly massless particle), both upper and lower spinors have entries that are comparable in magnitude. This is the case for neutrinos, for example.

The particle- and anti-particle like solutions with momentum $\mathbf{p}$ and energies $\pm E = (m^2 + \mathbf{p}^2)^{1/2}$ are finally given by

'particle', energy $E \geq m$:

$$\Psi(x) = N \begin{pmatrix} \psi_s \\ \dfrac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \psi_s \end{pmatrix} \exp(-\mathrm{i}Et + \mathrm{i}\mathbf{p} \cdot \mathbf{x}) \tag{3.63}$$

'antiparticle', energy $-E \leq -mc^2$:

$$\Psi(x) = N \begin{pmatrix} -\dfrac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \psi_s \\ \psi_s \end{pmatrix} \exp(\mathrm{i}Et + \mathrm{i}\mathbf{p} \cdot \mathbf{x}) \tag{3.64}$$

Here, ψ_s is a two-component spinor that specifies the spin state of the particle. The factors N are for normalization. The calculation for the antiparticle is similar to the one done here.

Another way is to to apply a Lorentz transformation to the solutions for a particle at rest. Starting from a 'particle-like' solution where $\chi_0 = 0$, the spinor transformation (3.122) we shall derive below gives

$$\begin{pmatrix} \psi \\ \chi \end{pmatrix} = \begin{pmatrix} \mathrm{ch}(\beta/2)\psi_s \\ \boldsymbol{\sigma} \cdot \mathbf{n}\, \mathrm{sh}(\beta/2)\psi_s \end{pmatrix} \tag{3.65}$$

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.(3.62) provided $\mathbf{n}$ is the unit vector along $\mathbf{p}$ and $\mathrm{sh}\,\beta = |\mathbf{p}|/m$. (Recall the difference between velocity and relativistic momentum, $p = \gamma m v$.)

3.3.3 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.3.2.1: 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 (3.52) in upper and lower two-component spinors, restore $\hbar$ and c

$$\begin{aligned} \mathrm{i}\hbar D_t \psi &= \quad mc^2 \psi - \mathrm{i}c\hbar \boldsymbol{\sigma} \cdot \mathbf{D} \chi \\ \mathrm{i}\hbar D_t \chi &= -mc^2 \chi - \mathrm{i}c\hbar \boldsymbol{\sigma} \cdot \mathbf{D} \psi \end{aligned} \tag{3.66}$$

We separate off a factor $\mathrm{e}^{-\mathrm{i}mc^2 t/\hbar}$ 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

$$\mathrm{i}\hbar D_t \tilde{\chi} = -2mc^2 \tilde{\chi} - \mathrm{i}c\hbar \boldsymbol{\sigma} \cdot \mathbf{D} \tilde{\psi} \tag{3.67}$$

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

$$\tilde{\chi} = -\frac{\mathrm{i}\hbar}{2mc}\boldsymbol{\sigma}\cdot\mathbf{D}\tilde{\psi} \tag{3.68}$$

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

For the first line of of (3.66), we thus get

$$\mathrm{i}\hbar D_t\tilde{\psi} = -\frac{\hbar^2}{2m}(\boldsymbol{\sigma}\cdot\mathbf{D})(\boldsymbol{\sigma}\cdot\mathbf{D})\tilde{\psi} \tag{3.69}$$

Now, the Pauli matrices have the property

$$\sigma_k\sigma_l = \delta_{kl} + \mathrm{i}\epsilon_{klm}\sigma_m \tag{3.70}$$

and therefore

$$\begin{aligned}\sigma_k D_k\sigma_l D_l &= \mathbf{D}\cdot\mathbf{D} + \frac{\mathrm{i}}{2}\sigma_m\epsilon_{klm}\left[D_k,\, D_l\right] \\ &= (\nabla - \mathrm{i}e\mathbf{A}/\hbar)^2 + \frac{e}{2\hbar}\sigma_m\epsilon_{klm}\left[\partial_k A_l - \partial_l A_k\right] \\ &= (\nabla - \mathrm{i}e\mathbf{A}/\hbar)^2 + \frac{e}{\hbar}\boldsymbol{\sigma}\cdot\mathbf{B}\end{aligned} \tag{3.71}$$

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

$$\mathrm{i}\hbar\partial_t\tilde{\psi} = -\frac{\hbar^2}{2m}(\nabla - \mathrm{i}e\mathbf{A}/\hbar)^2\tilde{\psi} - \frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}\tilde{\psi} + e\phi\tilde{\psi} \tag{3.72}$$

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:

$$H_{\mathrm{mag}} = -\frac{e}{2m}\mathbf{L}\cdot\mathbf{B} - \frac{e}{m}\mathbf{S}\cdot\mathbf{B} \tag{3.73}$$

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 (3.73) in terms of a magnetic moment, $H_{\mathrm{mag}} = -\boldsymbol{\mu}\cdot\mathbf{B}$:

$$\boldsymbol{\mu} = \mu_B\left(\mathbf{L} + g\mathbf{S}\right), \qquad g = 2, \qquad \mu_B = \frac{e}{2m}, \tag{3.74}$$

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.

3.3.4 Klein paradox

3.3.5 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, this can be done with the minimal coupling scheme (Secs. 3.2.1, 3.3.3). We take the vector potential in the rest frame of the nucleus:

$$A_\mu = (\phi(r), \mathbf{0}) \ , \qquad \phi(r) = -\frac{Ze}{4\pi\epsilon_0 r} \tag{3.75}$$

where Ze is the charge of the nucleus.

What are the quantum numbers that we may expect? They correspond to conservation laws in the potential (3.75): it is time-independent, hence energy E is conserved; it is invariant under parity, so eigenstates can be chosen to be even or odd; it is invariant under rotations around the nucleus, so angular momentum is conserved. The corresponding angular momentum operator is the "generator" of rotations: for the Dirac spinor, a rotation does affect both the spatial coordinates (via a rotation matrix R) and the spinor components (via the spinor transformation $S(R)$). As sketched in Sec. 3.6.1, Eq.(3.159), this leads to the following result for the angular momentum operator

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{3.76}$$

where $\mathbf{S}$ is the spin angular momentum of Eq.(3.60). From the rules for the "addition of angular momenta", we read off that the quantum number j corresponding to $\mathbf{J}^2$ can take the values $j = l \pm \frac{1}{2}$, knowing that the electron has spin $S = \frac{1}{2}$. The orbital quantum number l determines the parity of the state, we relate it in the following via the number $\varpi = \pm 1$ to the total angular momentum: $l = j + \varpi/2$. Examples for the smallest quantum numbers, in spectroscopic notation.

term	j	l	parity	ϖ	l'
$1\mathrm{s}_{1/2}$	$1/2$	0	even	-1	1
$2\mathrm{p}_{1/2}$	$1/2$	1	odd	$+1$	0
$2\mathrm{p}_{3/2}$	$3/2$	1	odd	-1	2

The quantum number l' applies to the lower spinor component.

It turns out that one can make the following *Ansatz* for the bi-spinor

$$\Psi(x) = \frac{\mathcal{N}}{r} \begin{pmatrix} F(r)\,\mathcal{Y}(\theta,\varphi) \\ \mathrm{i}G(r)\,\mathcal{Y}'(\theta,\varphi) \end{pmatrix} \tag{3.77}$$

where $\mathcal{Y} = \mathcal{Y}_{ljm}$ is the spinor that is generated by the addition of an angular momentum state l with a spin-$1/2$ spinor to give total quantum numbers j, m. The lower spinor $\mathcal{Y}' == \mathcal{Y}_{l'jm}$ is constructed in a similar way, but its parity is

different: it is based on adding an orbital angular momentum $l' = l - \varpi$ and a spin-$1/2$ to give also a j, m state. This behaviour appears because the lower components of the Dirac spinor represent the anti-particle solutions, and they have an 'intrinsic parity' opposite to the particle (see Table 3.1 on page 58).

The coupled angular momentum states are written in the following table first in ket form, specifying first the spin magnetic quantum number ($m_s = \pm 1/2$), then the orbital one ($m = -l, \ldots l$). Then, using the notation $|m_s\rangle = |+\frac{1}{2}\rangle = \binom{1}{0}$ and $|-\frac{1}{2}\rangle = \binom{0}{1}$, we give the two spinors. (No guarantee: check the rules given above.)

term	m	ket	$\mathcal{Y}$	$\mathcal{Y}'$
$1\mathrm{s}_{1/2}$	$+\frac{1}{2}$	$\lvert+\frac{1}{2},0\rangle$	$\begin{pmatrix} Y_{00} \\ 0 \end{pmatrix}$	$\begin{pmatrix} Y_{10} \\ -Y_{11} \end{pmatrix}$
$2\mathrm{p}_{1/2}$	$+\frac{1}{2}$	$a\lvert\frac{1}{2},0\rangle - b\lvert-\frac{1}{2},1\rangle$	$\begin{pmatrix} aY_{10} \\ -bY_{11} \end{pmatrix}$	$\begin{pmatrix} Y_{00} \\ 0 \end{pmatrix}$
$2\mathrm{p}_{3/2}$	$+\frac{3}{2}$	$\lvert\frac{1}{2},1\rangle$	$\begin{pmatrix} Y_{11} \\ 0 \end{pmatrix}$	$\begin{pmatrix} -a'Y_{21} \\ b'Y_{22} \end{pmatrix}$

With this Ansatz for the spinors, one derives from the Dirac equation

$$\begin{pmatrix} E + \dfrac{Z\alpha}{r} & \mathrm{i}\sigma \cdot \nabla \\ -\mathrm{i}\sigma \cdot \nabla & -E - \dfrac{Z\alpha}{r} \end{pmatrix} \Psi = m\Psi \tag{3.78}$$

where $\alpha = e^2/4\pi\varepsilon_0\hbar c \approx 1/137.04$ is the fine structure constant, the following coupled equations

$$\left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\varpi(j+\frac{1}{2})}{r}\right) F = \left(E + m + \frac{Z\alpha}{r}\right) G \tag{3.79}$$

$$\left(-\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\varpi(j+\frac{1}{2})}{r}\right) G = \left(E - m + \frac{Z\alpha}{r}\right) F \tag{3.80}$$

One tries, similar to the Coulomb problem, a polynomial *Ansatz* for bound states

$$F(r) = r^{\lambda}\mathrm{e}^{\,-\kappa r}\left(1 + a_1 r + \ldots\right), \qquad G(r) = r^{\lambda}\mathrm{e}^{\,-\kappa r}\left(1 + b_1 r + \ldots\right) \tag{3.81}$$

The decay constant κ can be found from the large-r behaviour of the system (3.80):

$$\kappa = \sqrt{m^2 - E^2} \tag{3.82}$$

Here we see that bound states should appear with energies $-m < E < m$ in the 'mass gap'. For the exponent λ in Eq.(3.81) that governs the small-r region, one finds

$$\lambda = \sqrt{(j+\tfrac{1}{2})^2 - (Z\alpha)^2} \tag{3.83}$$

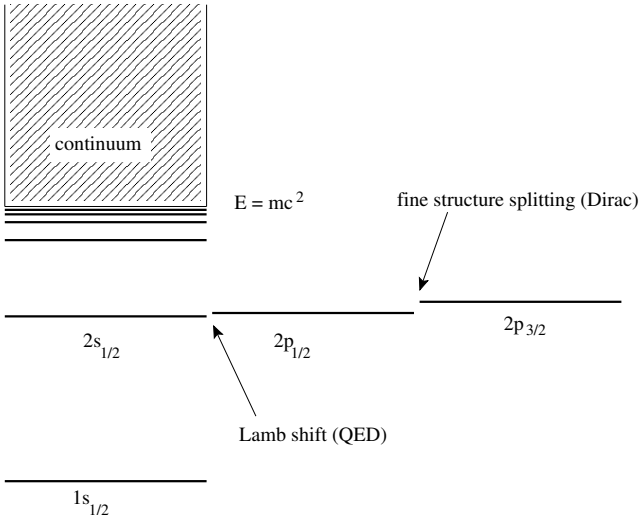

Figure 3.1: Energy levels of the hydrogen atom within the Dirac theory: the 'fine structure' between $2p_{1/2}$ and $2p_{3/2}$ appears. The Lamb shift (splitting between the levels $2s_{1/2}$ and $2p_{1/2}$) is not described by the Dirac equation, one needs QED for that.

which differs in practice very little from $j + \frac{1}{2}$.

Finally, one gets polynomials with a finite degree (the radial quantum number is $n' = n - j - \frac{1}{2}$) provided the energy takes the quantised values (see Fig.3.1)

$$E_{nj} = \frac{mc^2}{\left(1 + \dfrac{(Z\alpha)^2}{(n-\delta_j)^2}\right)^{1/2}}, \qquad \delta_j = j + \tfrac{1}{2} - \sqrt{(j+\tfrac{1}{2})^2 - (Z\alpha)^2} \tag{3.84}$$

The principal quantum number is $n = 1, 2, \ldots$, and the quantity δ_j is called a 'quantum defect': a deviation from integers in the denominator of this Balmer-like formula. Dirac's theory predicts successfully that the level $2p_{3/2}$ is splitting off the two others $2s_{1/2}$ and $2p_{1/2}$ (these two have the same values of n, j). Good point: this is the experimentally seen 'fine structure', and the non-relativistic calculation for the hydrogen atom could not explain this splitting. Since it is the quantum number j for the 'total angular momentum' that combines the orbital angular momentum and the electron spin, one is also talking about 'LS coupling'.

By expanding the root in Eq.(3.84) for small $Z\alpha$, one gets in the first order a correction to the electron's rest mass mc^2 which looks similar to the non-relativistic Hydrogen spectrum, but gets corrected by the quantum defect δ_j. In the next order, corrections of order $(Z\alpha)^2$ relative to the Balmer formula appear – these have also been observed in precision spectroscopy data and can be interpreted as the relativistic mass shift etc. Finally, these expansions fail for

highly charged nuclei like Lead or Uranium ($Z = 82$ or 92). In these systems, the bound states have extremely high binding energies, approaching the MeV range, and relativistic effects play an essential role.

At a closer look, the levels $2s_{1/2}$ and $2p_{1/2}$ 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.

3.4 Relativistic covariance and symmetry groups

3.4.1 Lagrangian

Lagrangian

$$S = \int \mathrm{d}t\, \mathrm{d}^3x \left[\frac{\mathrm{i}}{2} \overline{\Psi} \gamma^\mu \partial_\mu \Psi - \frac{\mathrm{i}}{2} \partial_\mu \overline{\Psi} \gamma^\mu \Psi - m \overline{\Psi} \Psi \right] \tag{3.85}$$

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

$$\boxed{\mathrm{i}\gamma^\mu \partial_\mu \Psi - m\Psi = 0} \tag{3.86}$$

Big advantage of this method: if we can construct the theory such that the action (3.85) is a covariant scalar with respect to Lorentz transformations, then the Dirac equation will be covariant as well. Note that for the moment, we are cheating since γ^μ only "looks like" a 4-vector, but is actually constant.

3.4.2 Probability current

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

$$-\mathrm{i}\partial_\mu \overline{\Psi} \gamma^\mu - m\overline{\Psi} = 0 \tag{3.87}$$

multiply from the right with Ψ, multiply (3.86) from the left with $\overline{\Psi}$ and take the difference:

$$\begin{aligned} 0 &= \overline{\Psi}\gamma^\mu \partial_\mu \Psi + (\partial_\mu \overline{\Psi})\gamma^\mu \Psi = \partial_\mu (\overline{\Psi}\gamma^\mu \Psi) \\ &\text{hence: } j^\mu = \overline{\Psi}\gamma^\mu \Psi \end{aligned} \tag{3.88}$$

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

$$n = j^0 = \overline{\Psi}\gamma^0 \Psi \geq 0 \tag{3.89}$$

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 (3.48), one possible solution is $\Gamma = \gamma^0$. Hence, we get

$$\overline{\Psi} = \Psi^\dagger \gamma^0, \qquad n = \Psi^\dagger \Psi \geq 0 \tag{3.90}$$

The spatial part is the probability current and given by

$$j^k = \overline{\Psi} \gamma^k \Psi = \Psi^\dagger \alpha^k \Psi \tag{3.91}$$

Note the analogy to the product between density $\Psi^\dagger \Psi$ and the velocity operator $\boldsymbol{\alpha}$ (see exercises).

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

$$(\gamma^\mu)^\dagger \gamma^0 = \gamma^0 \gamma^\mu \tag{3.92}$$

that follows from the fact that γ^k is anti-hermitean (3.53) and anti-commutes with the hermitean matrix γ^0.

The same relation pops up when one wants to show that the Euler-Lagrange equations for Ψ and $\overline{\Psi}$ are related to each other by complex conjugation. This is another way to get the solution $\Gamma = \gamma^0$ for the matrix in the adjoint spinor.

3.4.3 Spinor rotation

We now want to see how the Dirac spinor changes under a rotation of the coordinates. Our main goal is 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)

$$j_k = \Psi^\dagger \gamma^0 \gamma^k \Psi = \Psi^\dagger \alpha^k \Psi \tag{3.93}$$

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

$$\mathbf{x}' = R\mathbf{x}, \qquad j'_k(t, \mathbf{x}') = R_{kl} j_l(t, \mathbf{x}) \tag{3.94}$$

(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

$$j_k = \psi^\dagger \sigma_k \chi + \chi^\dagger \sigma_k \psi. \tag{3.95}$$

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

$$\psi^\dagger S^\dagger \sigma_k S \chi + \chi^\dagger S^\dagger \sigma_k S \psi \stackrel{!}{=} R_{kl} \left(\psi^\dagger \sigma_l \chi + \chi^\dagger \sigma_l \psi \right) \tag{3.96}$$

This is satisfied for all ψ, χ if we have

$$S^\dagger \sigma_k S = R_{kl} \sigma_l \tag{3.97}$$

This equation fixes $S(R)$ up to a global phase factor. You recognize here the transformation of the 2-component spinor discussed in Exercise 8.4. 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:

$$\Psi'(x') = \begin{pmatrix} \exp\left(\frac{\mathrm{i}}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}\right) & 0 \\ 0 & \exp\left(\frac{\mathrm{i}}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}\right) \end{pmatrix} \Psi(x) \tag{3.98}$$

From this equation, one can show that the Dirac spinor describes a particle with spin $1/2$. We shall see the precise proof of this statement in the chapter on symmetry groups. But let us first provide a proof of a special case of Eq.(3.98).

Infinitesimal rotation

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

$$R_{ij} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.99}$$

and the corresponding spinor transformation must satisfy:

$$S^\dagger \sigma_x S = \cos\theta\,\sigma_x + \sin\theta\,\sigma_y \tag{3.100}$$

and a similar equation including σ_y.

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

$$S^\dagger \sigma_x S = (\mathbb{1} - \mathrm{i}\theta J_z)\,\sigma_x\,(\mathbb{1} + \mathrm{i}\theta J_z) = \sigma_x + \theta\,\sigma_y \tag{3.101}$$

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 (3.101):

$$[J_z,\,\sigma_x] = \mathrm{i}\sigma_y. \tag{3.102}$$

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

$$J_z = \frac{1}{2}\sigma_z, \qquad S(\theta, z) = \mathbb{1} + \frac{\mathrm{i}}{2}\theta\sigma_z, \quad \theta \to 0. \tag{3.103}$$

By concatenating this transformation for many small angles, we get

$$S(\theta, z) = \lim_{N\to\infty}\left(\mathbb{1} + \tfrac{\mathrm{i}}{2}(\theta/N)\sigma_z\right)^N = \exp\left(\tfrac{\mathrm{i}}{2}\theta\sigma_z\right) \tag{3.104}$$

This coincides with the special case one obtains from (3.98) when the unit vector $\mathbf{n}$ is identified with the rotation axis.

[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.

If you do not feel comfortable with this calculation, you can introduce a one-parameter subgroup $S(\theta)$ of spinor transformations such that (3.100) holds. By differentiating this constraint with respect to θ, one gets a differential equation whose solution is the matrix exponential (3.104).

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:

$$S(\theta, z) = \cos(\theta/2) + \mathrm{i}\sigma_z \sin(\theta/2). \tag{3.105}$$

3.4.4 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.

Now, we have to implement the transformation for the 4-current. We require it to be a contravariant vector, hence

$$j'^\mu(x') \quad = \quad \Lambda^\mu{}_\nu j^\nu(x) \tag{3.106}$$

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

$$S^\dagger\gamma^0\gamma^\mu S = \Lambda^\mu{}_\nu\gamma^0\gamma^\nu. \tag{3.107}$$

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

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \mathrm{ch}\,\beta & \mathrm{sh}\,\beta & 0 & 0 \\ \mathrm{sh}\,\beta & \mathrm{ch}\,\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.108}$$

with $\mathrm{th}\,\beta = v/c$. Now, $\mu = 0$ and $\mu = x$ are the relevant equations in (3.107)

$$S^\dagger S \quad = \quad \mathrm{ch}\,\beta + \mathrm{sh}\,\beta\,\gamma^0\gamma^x = \mathrm{ch}\,\beta + \mathrm{sh}\,\beta\,\alpha^x \tag{3.109}$$

$$S^\dagger\gamma^0\gamma^x S = S^\dagger\alpha^x S \quad = \quad \mathrm{sh}\,\beta + \mathrm{ch}\,\beta\,\alpha^x \tag{3.110}$$

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

$$1 + \beta(K^\dagger + K) \quad = \quad 1 + \beta\alpha^x \tag{3.111}$$

$$\alpha^x + \beta(K^\dagger\alpha^x + \alpha^x K) \quad = \quad \alpha^x + \beta \tag{3.112}$$

These equations can be satisfied by taking $K = \frac{1}{2}\alpha^x$: K is hermitean, and from the fundamental properties (3.48), $K^\dagger\alpha^x + \alpha^x K = \mathbb{1}$. Exponentiating this infinitesimal spinor transformation, we get

$$S(\beta, x) = \exp(\tfrac{1}{2}\beta\alpha^x) = \mathrm{ch}(\beta/2) + \mathrm{sh}(\beta/2)\alpha^x. \tag{3.113}$$

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 $\mathbf{n}$.

3.4.5 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 $\overline{\Psi}\Psi$: we continue our conservative assumption that the spinor changes by a matrix S, but the Dirac matrices are the same $\gamma'^\mu = \gamma^\mu$. This term is thus a Lorentz scalar only when

$$S^\dagger \gamma^0 S = (\mathrm{ch}(\beta/2) + \mathrm{sh}(\beta/2)\alpha^x)\, \gamma^0\, (\mathrm{ch}(\beta/2) + \mathrm{sh}(\beta/2)\alpha^x) \tag{3.114}$$

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

$$S^\dagger \gamma^0 S = \left(\mathrm{ch}^2(\beta/2) - \mathrm{sh}^2(\beta/2)\right) \gamma^0 = \gamma^0 \tag{3.115}$$

and it follows that $\overline{\Psi}'(x')\Psi'(x') = \overline{\Psi}(x)\Psi(x)$: this transforms like a scalar field under Lorentz boosts. It is a simple exercise to show that this is a scalar under rotations as well (because the 2-spinor transformations are unitary).

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.

$$\begin{aligned} \overline{\Psi}'(x') \gamma^\mu \partial'_\mu \Psi'(x') &= \Psi^\dagger(x) S^\dagger \gamma^0 \gamma^\mu \frac{\partial}{\partial x'^\mu} S \Psi(x) \\ &= \Psi^\dagger(x) S^\dagger \gamma^0 \gamma^\mu S \frac{\partial}{\partial x'^\mu} \Psi(x) \\ &= \Psi^\dagger(x) \Lambda^\mu{}_\nu \gamma^0 \gamma^\nu \frac{\partial}{\partial x'^\mu} \Psi(x) \end{aligned} \tag{3.116}$$

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), and the transformation rules (3.107) 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

$$x^\nu = (\Lambda^{-1})^\nu{}_\mu x'^\mu. \tag{3.117}$$

Hence the chain rule for differentiation gives

$$\begin{aligned}\frac{\partial}{\partial x'^{\mu}}\Psi(x) &= \frac{\partial}{\partial x'^{\mu}}\Psi((\Lambda^{-1})^{\nu}{}_{\lambda}x'^{\lambda}) \\ &= (\Lambda^{-1})^{\nu}{}_{\lambda}\delta^{\lambda}{}_{\mu}\frac{\partial}{\partial x^{\nu}}\Psi(x) \\ &= (\Lambda^{-1})^{\lambda}{}_{\mu}\frac{\partial}{\partial x^{\lambda}}\Psi(x) \qquad (3.118)\end{aligned}$$

Taking the matrix elements of the Lorentz transformations out of the spinor product, we get from Eqs.(3.116, 3.118):

$$\begin{aligned}\overline{\Psi}'(x')\gamma^{\mu}\partial'_{\mu}\Psi'(x') &= (\Lambda^{-1})^{\lambda}{}_{\mu}\Lambda^{\mu}{}_{\nu}\Psi^{\dagger}(x)\gamma^{0}\gamma^{\nu}\frac{\partial}{\partial x^{\lambda}}\Psi(x) \\ &= \delta^{\lambda}{}_{\nu}\overline{\Psi}(x)\gamma^{\nu}\frac{\partial}{\partial x^{\lambda}}\Psi(x) \qquad (3.119)\end{aligned}$$

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

By the way, Eq.(3.118) is precisely the transformation for a covariant 4-vector. To see it, just compute

$$p_{\mu}x^{\mu} = p'_{\mu}x'^{\mu} = \tilde{\Lambda}_{\mu}{}^{\nu}p_{\nu}\Lambda^{\mu}{}_{\lambda}x^{\lambda} \quad \Rightarrow \quad \tilde{\Lambda}_{\mu}{}^{\nu}\Lambda^{\mu}{}_{\lambda} = \delta^{\nu}{}_{\lambda} \qquad (3.120)$$

Hence, the transformation matrix $\tilde{\Lambda}_{\mu}{}^{\nu}$ for the covariant vector p_{μ} is the transposed inverse of the Lorentz matrix $\Lambda^{\mu}{}_{\lambda}$ for the coordinates x^{μ}.

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

$$\Psi'(x') = \begin{pmatrix} \psi'(x') \\ \chi'(x') \end{pmatrix} = \begin{pmatrix} \exp\left(\frac{\mathrm{i}}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}\right) & 0 \\ 0 & \exp\left(\frac{\mathrm{i}}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}\right) \end{pmatrix}\begin{pmatrix} \psi(x) \\ \chi(x) \end{pmatrix} \qquad (3.121)$$

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

$$\Psi'(x') = \begin{pmatrix} \psi'(x') \\ \chi'(x') \end{pmatrix} = \begin{pmatrix} \mathrm{ch}(\beta/2) & \mathrm{sh}(\beta/2)\boldsymbol{\sigma}\cdot\mathbf{n} \\ \mathrm{sh}(\beta/2)\boldsymbol{\sigma}\cdot\mathbf{n} & \mathrm{ch}(\beta/2) \end{pmatrix}\begin{pmatrix} \psi(x) \\ \chi(x) \end{pmatrix} \qquad (3.122)$$

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.

3.5 Discrete symmetries

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

$$s^2 = c^2t^2 - \mathbf{x}^2 = c^2t^2 - (-\mathbf{x})^2 = c^2(-t)^2 - \mathbf{x}^2 \tag{3.123}$$

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.

3.5.1 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

$$\frac{\mathrm{d}}{\mathrm{d}\tau}p^k = eE^k\frac{\mathrm{d}t}{\mathrm{d}\tau} - eB^k{}_l\frac{\mathrm{d}x^l}{\mathrm{d}\tau} \tag{3.124}$$

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

$$P : \psi(t, \mathbf{x}) \mapsto \psi'(t', \mathbf{x}') = \psi(t', -\mathbf{x}') \tag{3.125}$$

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,

	t	$\mathbf{x}$	$\boldsymbol{\nabla}$	$\mathbf{p}$	$\mathbf{L}$	e	$\mathbf{E}$	$\mathbf{B}$	$\mathbf{A}$	ψ	Ψ
R		$R\mathbf{x}$	$(R^{-1})^T\boldsymbol{\nabla}$	$R\mathbf{p}$	$\det(R)R\mathbf{L}$		$R\mathbf{E}$	$\det(R)R\mathbf{B}$	$R\mathbf{A}$		$S(R)\Psi$
Λ	${\Lambda^0}_\mu x^\mu$	${\Lambda^k}_\mu x^\mu$	$(\Lambda^{-1})^T(\partial_t, \boldsymbol{\nabla})$	${\Lambda^k}_\mu p^\mu$	${\Lambda^\kappa}_\mu {\Lambda^\lambda}_\nu x^\mu p^\nu$		${\Lambda^\kappa}_\mu {\Lambda^\lambda}_\nu F^{\mu\nu}$		${\Lambda^k}_\mu A^\mu$	[–]	$S(\Lambda)\Psi$
P		$-\mathbf{x}$	$-\boldsymbol{\nabla}$	$-\mathbf{p}$			$-\mathbf{E}$		$-\mathbf{A}$		$\gamma^0\Psi$
T	$-t$			$-\mathbf{p}$	$-\mathbf{L}$			$-\mathbf{B}$	$-\mathbf{A}$	ψ^*	$-\mathrm{i}\Sigma_2\Psi^*$
C						$-e$				$[\psi^*]$	$\mathrm{i}\gamma^2\Psi^*$

Table 3.1: Transformation rules under (proper) rotations R, (proper) boosts Λ, parity P (mirror symmetry), time reversal T, and charge conjugation C. ψ: scalar Schrödinger wave function. Ψ: Dirac spinor wave function. No table entry: object is invariant.
Notes: the angular momentum $\mathbf{L}$ is a sub-matrix of an antisymmetric 4×4 tensor in relativity that transforms like a tensor. The Schrödinger equation is not Lorentz-invariant, of course. Charge conjugation does not really exist for the Schrödinger wave function, unless one also changes the mass $m \mapsto -m$.

while $\phi \mapsto \phi$ for the scalar potential.[4] 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'.

Again, the classical equation of motion (3.124) 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.(3.124) 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.

3.5.2 Overview

The following table gives an overview on the transformation properties of the different objects.

Die folgenden Abschnitte liefern mehr Details für diese Transformationen, speziell die Herleitung der Matrizen für den Dirac-Spinor. Im SS'07 waren diese Abschnitte nicht Teil der Vorlesung. Gehen Sie weiter zu Abschnitt 3.6.

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

3.5.3 Nonrelativistic quantum mechanics

Parity

The wave function transforms as given in Eq.(3.125) under the parity operator. With this, it is easy to check that the Schrödinger equation for a charged, nonrelativistic particle is covariant under parity:

$$\begin{aligned} \mathrm{i}\partial_{t'}\psi' &= \frac{\hbar^2}{2m}\left(-\mathrm{i}\nabla' - e'\mathbf{A}'\right)^2\psi' + e'\phi'\,\psi' \\ = \mathrm{i}\partial_t\psi(t,-\mathbf{x}') &= \frac{\hbar^2}{2m}\left(\mathrm{i}\nabla + e\mathbf{A}\right)^2\psi(t,-\mathbf{x}') + e\phi\,\psi(t,-\mathbf{x}') \end{aligned} \tag{3.126}$$

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 coordinates have to be transformed and therefore we deal with a different function $\phi'(\mathbf{x}') = \phi(-\mathbf{x}')$. This happens as required for a scalar field (3.125).

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:

$$T : \psi'(t', \mathbf{x}') = \psi^*(-t', \mathbf{x}') \tag{3.127}$$

To show that the Schrödinger equation is covariant, we write it down for the primed quantities and use the transformation laws (3.127) 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 $-m$.

3.5.4 Dirac equation

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

$$\gamma^0 \left(\mathrm{i}\,\partial_t - e\phi\right) \Psi + \boldsymbol{\gamma} \cdot \left(\mathrm{i}\nabla + e\mathbf{A}\right) \Psi = m\Psi \tag{3.128}$$

where $\boldsymbol{\gamma} = (\gamma^1, \gamma^2, \gamma^3)$ 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

$$S_P^{-1}\gamma^0 \left(\mathrm{i}\,\partial_t - e\phi\right) S_P\Psi + S_P^{-1}\boldsymbol{\gamma} \cdot \left(-\mathrm{i}\nabla - e\mathbf{A}\right) S_P\Psi = m\Psi \tag{3.129}$$

Hence, we need a matrix that satisfies

$$S_P^{-1}\gamma^0 S_P = \gamma^0, \qquad S_P^{-1}\gamma^k S_P = -\gamma^k \tag{3.130}$$

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

$$S_P = \mathrm{const}\,\gamma^0 \tag{3.131}$$

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

$$S_P^\dagger \gamma^0 S_P = \gamma^0 \tag{3.132}$$

which gives $|\mathrm{const}|^2 = 1$. We choose the simplest solution with $S_P = \gamma^0$ or

$$P : \Psi'(t', \mathbf{x}') = \gamma^0 \Psi(t', -\mathbf{x}') \tag{3.133}$$

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

$$S_T^{-1}\gamma^0 \left(-\mathrm{i}\,\partial_t - e\phi\right) S_T\Psi^* + S_T^{-1}\boldsymbol{\gamma} \cdot \left(\mathrm{i}\nabla - e\mathbf{A}\right) S_T\Psi^* = m\Psi^* \tag{3.134}$$

where the arguments $\Psi^*(-t', \mathbf{x}')$ are suppressed for brevity. taking the complex conjugate of this equation, we get the covariant derivatives right, but the Dirac matrices are getting conjugated:

$$S_T^{-1*}(\gamma^0)^* \left(\mathrm{i}\,\partial_t - e\phi\right) S_T^* \Psi + S_T^{-1*} \boldsymbol{\gamma}^{\cdot} \left(-\mathrm{i}\nabla - e\mathbf{A}\right) S_T^* \Psi = m\Psi \tag{3.135}$$

We thus require for the spinor transformation the properties

$$S_T^{-1*} \gamma^{0*} S_T^* = \gamma^0, \qquad S_T^{-1*} \gamma^{k*} S_T^* = -\gamma^k \tag{3.136}$$

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

$$\gamma^2 \gamma^{\mu *} \gamma^2 = \gamma^\mu \tag{3.137}$$

$$\gamma^0 \gamma^k \gamma^0 = -\gamma^k \tag{3.138}$$

$$\gamma_5 \gamma^\mu \gamma_5 = -\gamma^\mu \tag{3.139}$$

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

$$\gamma_5 = \mathrm{i}\,\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{3.140}$$

is a matrix that anti-commutes with the four Dirac matrices (see Eq.(3.139)).

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

$$S_T^* = -\mathrm{i}\,(\gamma^2)^*(\gamma^0)^*(\gamma_5)^* = \mathrm{i}\,\gamma^2\gamma^0\gamma_5 \tag{3.141}$$

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

$$\begin{aligned} (S_T)^{-1*}\gamma^{0*}S_T^* &= -\gamma_5\gamma^0\gamma^2\gamma^{0*}\gamma^2\gamma^0\gamma_5 \\ &\overset{(3.137)}{=} -\gamma_5\gamma^0\gamma^0\gamma^0\gamma_5 \\ &= -\gamma_5\gamma^0\gamma_5 \overset{(3.139)}{=} \gamma^0 \end{aligned} \tag{3.142}$$

And for the spatial components:

$$\begin{aligned} (S_T)^{-1*}\gamma^{k*}S_T^* &= -\gamma_5\gamma^0\gamma^2\gamma^{k*}\gamma^2\gamma^0\gamma_5 \\ &\overset{(3.137)}{=} -\gamma_5\gamma^0\gamma^k\gamma^0\gamma_5 \\ &\overset{(3.138)}{=} +\gamma_5\gamma^k\gamma_5 \\ &\overset{(3.139)}{=} -\gamma^k \end{aligned} \tag{3.143}$$

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

$$T : \Psi'(t', \mathbf{x}') = \mathrm{i}\,\gamma^2\gamma^0\gamma_5\Psi^*(-t', \mathbf{x}') = -\mathrm{i}\begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}\Psi^*(-t', \mathbf{x}') \tag{3.144}$$

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

$$C : \Psi'(t', \mathbf{x}') = \mathrm{i}\,\gamma^2\Psi^*(t', \mathbf{x}') = \mathrm{i}\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}\Psi^*(t', \mathbf{x}') \tag{3.145}$$

Indeed, the Dirac equation, after complex conjugation, yields

$$S_C^{-1*}\gamma^{\mu*}S_C^* = -\gamma^\mu \tag{3.146}$$

which can be satisfied with $S_C = \mathrm{i}\gamma^2$ because of Eq.(3.137) 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)

$$P : \Psi(x) = \begin{pmatrix} \psi_0 \\ \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\psi_0 \end{pmatrix}\mathrm{e}^{-\mathrm{i}Et+\mathrm{i}\mathbf{p}\cdot\mathbf{x}} \mapsto \begin{pmatrix} \psi_0 \\ \dfrac{-\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\psi_0 \end{pmatrix}\mathrm{e}^{-\mathrm{i}Et-\mathrm{i}\mathbf{p}\cdot\mathbf{x}} \tag{3.147}$$

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

$$T : \Psi(x) \mapsto -\mathrm{i}\begin{pmatrix} \sigma_2\psi_0^* \\ \dfrac{\sigma_2(\boldsymbol{\sigma}^*\cdot\mathbf{p})}{E+m}\psi_0^* \end{pmatrix}\mathrm{e}^{-\mathrm{i}Et-\mathrm{i}\mathbf{p}\cdot\mathbf{x}} \tag{3.148}$$

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 (3.147), 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:

$$C : \Psi(x) \mapsto \mathrm{i} \begin{pmatrix} \dfrac{\sigma_2(\boldsymbol{\sigma}^* \cdot \mathbf{p})}{E+m}\psi_0^* \\ -\sigma_2\psi_0^* \end{pmatrix} \mathrm{e}^{\mathrm{i}Et - \mathrm{i}\mathbf{p}\cdot\mathbf{x}} = -\mathrm{i} \begin{pmatrix} \dfrac{(\boldsymbol{\sigma} \cdot \mathbf{p})}{E+m}\sigma_2\psi_0^* \\ \sigma_2\psi_0^* \end{pmatrix} \mathrm{e}^{\mathrm{i}Et - \mathrm{i}\mathbf{p}\cdot\mathbf{x}} \tag{3.149}$$

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 (3.149) 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 is just swapping the 2-spinors ψ and χ.

3.6 Some notes on angular momentum

3.6.1 Angular momentum generates rotations

We have seen so far several objects that have commutation relations identical to the angular momentum operator $\mathbf{L}$,

$$[L_i,\, L_j] = \mathrm{i}\epsilon_{ijk}L_k. \tag{3.150}$$

We shall see why this is so. The basic reason is that the different operators (orbital angular momentum $\mathbf{L}$, spin $\mathbf{S}$ or $\frac{1}{2}\boldsymbol{\Sigma}$, total angular momentum $\mathbf{J} = \mathbf{L} + \frac{1}{2}\boldsymbol{\Sigma}$) are constructed to represent rotations.

Example: scalar field

Let us recall the transformation for a scalar field $\phi(\mathbf{x})$ under rotations, $\mathbf{x}' = R\mathbf{x}$:

$$\phi'(\mathbf{x}') = \phi(\mathbf{x}) \tag{3.151}$$

This means that at the points $\mathbf{x}$ and $\mathbf{x}'$ that correspond to each other under a rotation of the axes (they represent the same "*Punkt an sich*"), the fields ϕ and ϕ' have the same values. But they are not given by the same function, since Eq.(3.151) can be re-written as

$$\phi'(\mathbf{x}') = \phi(R^{-1}\mathbf{x}') \qquad \text{or} \qquad \phi' = \phi \circ R^{-1} \tag{3.152}$$

where $\circ$ denotes the concatenation (*Hintereinanderausführung*) of two functions.

Now consider an infinitesimal rotation. Checking carefully the sens of rotation, one has for a small rotation angle θ around an axis $\mathbf{n}$ (see Exercises):

$$\mathbf{x}' = R\mathbf{x} \approx \mathbf{x} - \theta\mathbf{n} \times \mathbf{x} \tag{3.153}$$

The inverse of this reads, to first order in θ:

$$\mathbf{x} = R^{-1}\mathbf{x}' \approx \mathbf{x}' + \theta\mathbf{n} \times \mathbf{x}' \tag{3.154}$$

Hence to compare the two functions ϕ' and ϕ, we have to compute

$$\phi'(\mathbf{x}') \approx \phi(\mathbf{x}' + \theta\mathbf{n} \times \mathbf{x}') \approx \phi(\mathbf{x}') + \theta(\mathbf{n} \times \mathbf{x}') \cdot \boldsymbol{\nabla}\phi(\mathbf{x}') \tag{3.155}$$

The differential operator here can be written in terms of the angular momentum $\mathbf{L}' = -\mathrm{i}\mathbf{x}' \times \mathbf{p}'$, and we get

$$\phi' \approx \phi + \mathrm{i}\theta\mathbf{n} \cdot \mathbf{L}'\phi \tag{3.156}$$

(We could have dropped the primes on the right-hand side everywhere by working with the name $\mathbf{x}$ for the argument of the function.) This equation expresses precisely what we mean by 'generate a rotation': for the scalar field, the change of the corresponding mathematical function under a small rotation of coordinates is requivalent to the application of the angular momentum operator.

This observation can be promoted to a general principle: for any object, the operator that occurs when the object transforms under a small rotation can be identified with the corresponding angular momentum operator.

Example: Dirac spinor field

For the Dirac spinor, we have found the transformation (written in analogy to Eq.(3.152)):

$$\Psi' = S(R)\Psi \circ R^{-1} \tag{3.157}$$

Now, we expand the right-hand side for a small rotation angle and get (recall that the matrix $\boldsymbol{\Sigma}$ has sub-blocks of Pauli matrices $\boldsymbol{\sigma}$ on the diagonal)

$$\begin{aligned} S(R)\Psi(R^{-1}\mathbf{x}) &\approx \left(\mathbb{1} + \frac{\mathrm{i}}{2}\theta\mathbf{n} \cdot \boldsymbol{\Sigma}\right)(1 + \mathrm{i}\theta\mathbf{n} \cdot \mathbf{L})\,\Psi(\mathbf{x}) \\ &\approx \left[\mathbb{1} + \mathrm{i}\theta\mathbf{n} \cdot \left(\mathbf{L} + \frac{1}{2}\boldsymbol{\Sigma}\right)\right]\Psi(\mathbf{x}) \end{aligned} \tag{3.158}$$

Here we see that the total angular momentum operator

$$\mathbf{J} = \mathbf{L} + \frac{1}{2}\boldsymbol{\Sigma} \tag{3.159}$$

contains an 'orbital part' ($\mathbf{L}$ acts like a derivative on the coordinate dependence) and an 'intrinsic part' ($\boldsymbol{\Sigma}$ acts as a matrix on the spinor components).

To be exact, one should write a unit matrix $\mathbb{1}$ in the spinor coordinates with the first term, but physicists usually forget about that notation. One important observation: the operators $\mathbf{L}$ and $\mathbf{\Sigma}$ commute since they act on different variables of the wave function. From, it follows that if $\mathbf{L}$ and $\frac{1}{2}\mathbf{\Sigma}$ fulfil the angular momentum commutation relations, then their sum does that as well.

You have checked in the exercises that the 'total angular momentum' $\mathbf{J}$ commutes with the (free) Dirac Hamiltonian, it is hence a conserved quantity. This is to be expected since the free Dirac equation is invariant under rotations. It is also true in a central potential, written in terms of $(\phi, \mathbf{A}) = (V(r), \mathbf{0})$ with $r = |\mathbf{x}|$, like in the case of the Hydrogen atom. The energy eigenvalues can indeed be found by looking for common eigenfunctions of the Hamiltonian and the angular momentum operators $\mathbf{J}^2$ and J_3. The corresponding eigenvalues are commonly denoted $j(j+1)$ and m_j, and different values of j are typically not degenerate in the energy spectrum (see Fig.3.1).

Commutators are invariant

The fact that the different angular momentum operators are the 'image' of mapping $S(R)$, resp., its linearization for infinitesimal rotations, has profound consequences. This mapping from the group of rotations into another group of matrices must be compatible with the multiplication of matrices. For example, when we take two rotations R_1 and R_2, we can apply succesively the transformation rules for a scalar field,

$$\phi' = \phi \circ R_1^{-1}; \qquad \phi'' = \phi' \circ R_2^{-1}; \tag{3.160}$$

or we can directly construct, by taking the product of rotations

$$\phi'' = \phi \circ (R_2 R_1)^{-1} \tag{3.161}$$

Now, by virtue of $(R_2R_1)^{-1} = R_1^{-1}R_2^{-1}$, one arrives in this way actually at the same function. A similar relation holds at the level of infinitesimal rotations. As a consequence:

> The commutation relations of any angular momentum operator are the same as those for infinitesimal rotations, and these are those of the 'angular momentum algebra'.

The constants ϵ_{ijk} that always appear when a commutator $[J_i, J_j]$ is formed, are called in mathematics the 'structure constants' of the 'algebra' that is spanned by the generators of the rotations.

3.6.2 Addition of angular momentum operators

This section is mainly intended as a reminder of the corresponding chapter in Quantum Mechanics I. We adopt here the Dirac bra-ket notation and write

$|m, m_s\rangle$ for the eigenvector of L_3 and S_3 with eigenvalues m and m_s. If you want to go back to the Dirac spinor, think of

$$\Psi(\mathbf{x}) = Y_{lm}(\theta, \varphi) \begin{pmatrix} \psi_s(r) \\ (\cdots) \end{pmatrix} \tag{3.162}$$

where the two-component spinor ψ_s can be chosen as eigenvector of the spin operator $\boldsymbol{\sigma}$, for example with eigenvalue $m_s = \pm 1/2$ for $\frac{1}{2}\sigma_3$. The angular dependence is carried by the spherical harmonic $Y_{lm}(\theta, \varphi)$, eigenfunction to the orbital angular momentum operators $\mathbf{L}^2$ and L_3. This spinor can, of course, depend on the radius. For example, one can try to find solutions to the free Dirac equation with fixed total angular momentum quantum numbers j and m_j. Then, as we shall see, one needs linear combinations of spinor-valued functions like in Eq.(3.162).

We now continue in a slightly more abstract way. The operators $\mathbf{L}$ and $\mathbf{S}$ can be any angular momentum operators. The operator $J_3 = L_3 + S_3$ commutes with L_3, S_3, and $\mathbf{L}^2$, $\mathbf{S}^2$. Hence the states m, m_s are already eigenstates of J_3 with eigenvalue $m+m_s$. It remains to construct eigenstates of $\mathbf{J}^2$. This operator does not commute with L_3 and S_3 (but with $\mathbf{L}^2$, $\mathbf{S}^2$).

The general result one gets is the following:

> *Addition of angular momenta.*
> The possible eigenvalues for $\mathbf{J}^2$ are $j(j+1)$ with $j = l + s, \ldots |l - s|$. For each value of j, one has the usual eigenvalues for $m_j = -j, \ldots j-1, j$. The eigenvectors $|j\, m_j\rangle$ are linear combinations of the $|m, m_s\rangle$ such that $m_j = m + m_s$.

Let us fill in some details. The state $|l, s\rangle$ corresponds to $m_j = l + s$, and this eigenvalue for J_3 can only occur when $j = l + s$. In fact, since one has

$$\mathbf{J}^2 = \mathbf{L}^2 + \mathbf{S}^2 + L_3 S_3 + \frac{1}{4}\left(L_+ S_- + L_- S_+\right) \tag{3.163}$$

with the ladder operators $L_\pm = L_1 \pm \mathrm{i} L_2$ and $S_\pm$, this state is also eigenstate of $\mathbf{J}^2$ with eigenvalue $j(j+1) = (l+s)(l+s+1)$. We can thus write

$$|l, s\rangle = |j = l + s, m_j = l + s\rangle \tag{3.164}$$

We have thus found a first state to act on with the ladder operator $J_- = L_- + S_-$. This gives the state

$$|j = l + s, m_j = l + s - 1\rangle \propto A|l - 1, s\rangle + B|l, s - 1\rangle \tag{3.165}$$

where the coefficients A, B are known from the action of L_- and S_-. By iteration, we thus form a set of basis states until $|j = l + s, m_j = -(l + s)\rangle$.

Another state, orthogonal to $|j = l + s, m_j = l + s - 1\rangle$ (and also eigenstate of $\mathbf{J}^2$) can be found by forming a suitable linear combination of the two

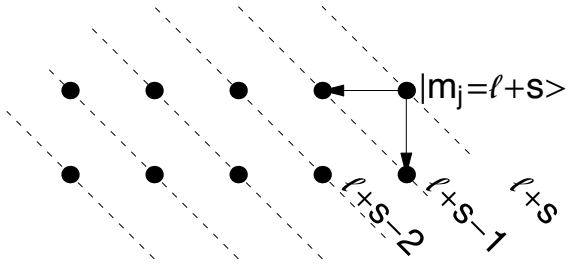

Figure 3.2: Addition of an angular momentum $l = 2$ and $s = 1/2$. The dots represent the basis vectors $|m, m_s\rangle$ in the so-called 'uncoupled basis'. The top left can be identified with an eigenvector $|l{+}s\rangle$ of J_3, with eigenvalue $m_j = l{+}s$. Acting on this vector with the ladder operator $J_- = L_- + S_-$, one gets the eigenvectors $|l + s - 1\rangle, \ldots | - (l + s)\rangle$, while $j = l + s$ is unchanged. A second, orthogonal vector with the same value for $m_j = l + s - 1$ can be found in the second diagonal and serves as a starting point for the ladder of eigenvectors with $j = l + s - 1$.

basis vectors in Eq.(3.165). This state forms the starting point for the subrepresentation with $j = l + s - 1$ (if this number is non-negative). This scheme goes on, with a new orthogonal vector found if the dimension of the space with $m + m_s = \text{const.}$ permits. It turns out that $j = |l - s|$ is the smallest value for j that can be found in this way.

A graphical representation of this procedure is given in Fig.3.2.

The expansion coefficients of the eigenstates of $\mathbf{J}^2$ and J_3 in the tensor product basis have a name on their own: the 'Clebsch-Gordan coefficients'. They are often written in the suggestive form

$$|j, m_j\rangle = \sum_{m,m_s} |m, m_s\rangle \, \langle l, m, s, m_s | j, m_j\rangle \tag{3.166}$$

where the quantum numbers l, s, j are written explicitly.

One can check that this construction is consistent from the viewpoint of dimensions: the total angular momentum $\mathbf{J}$ acts on a 'tensor product' Hilbert space with dimension $(2s + 1)(2l + 1)$ (for each Y_{lm}, there are $2s + 1$ spin components/possible eigenvectors). Now, in terms of the eigenvectors of $\mathbf{J}^2$ and J_3, the dimensions of the eigenspaces for the different values of j add up to the same numbers

$$(2s + 1)(2l + 1) = \sum_{j=|l-s|}^{l+s} (2j + 1) \tag{3.167}$$

as can be easily checked.

3.6.3 Generators of the Lorentz group

As a last illustration, let us have a look how the angular momentum operators are generalized when we allow for Lorentz transformations. Recall that a Lorentz transformation, $\mathrm{d}x'^{\mu} = \Lambda^{\mu}{}_{\nu}\mathrm{d}x^{\nu}$ leaves invariant the Minkowski distance (the travel time needed for light pulses)

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} \tag{3.168}$$

Hence, we find the following condition for the Lorentz matrices

$$\Lambda^{\mu}{}_{\nu}\Lambda^{\kappa}{}_{\lambda}g_{\mu\kappa} = g_{\nu\lambda} \tag{3.169}$$

Let us focus on an infinitesimal transformation,

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \theta\omega^{\mu}{}_{\nu} \tag{3.170}$$

where θ is a small parameter (rotation angle or relative velocity$/c$). Eq.(3.169), expanded to first order in θ, gives

$$\theta\left(\omega^{\mu}{}_{\nu}g_{\mu\lambda} + g_{\nu\kappa}\omega^{\kappa}{}_{\lambda}\right) = 0 \tag{3.171}$$

Now, the objects occurring here can be read as doubly covariant tensors with both indices at the bottom, $\omega_{\mu\nu} := g_{\mu\kappa}\omega^{\kappa}{}_{\nu}$, and indeed, we have

$$\theta\left(\omega_{\lambda\nu} + \omega_{\nu\lambda}\right) = 0 \tag{3.172}$$

hence these matrices are antisymmetric. As a test, you are invited to write down $\Lambda_{\mu\nu}$ for a Lorentz boost. It reduces indeed, for small v/c to the metric $g_{\mu\nu}$ and the first correction is antisymmetric. For rotation matrices, it is easy to show from $R^T R = \mathbb{1}$ that the generators J, $R = \mathbb{1} + \mathrm{i}\theta J$ must be antisymmetric as well.

There are six possibilities[5] to find antisymmetric matrices $\omega_{\mu\nu}$. We can easily identify them with three rotations and three boosts.
Rotation around axis $\mathbf{n}$: $\omega_{0k} = 0$ and $\omega_{kl} = n_m\epsilon_{kml}$.
Boost along axis $\mathbf{n}$: $\omega_{0k} = n_k = -\omega_{k0}$, $\omega_{kl} = 0$.

One can construct the action on scalar fields in this way: one finds three differential operators identical to the orbital angular momentum, and three others that mix time and space derivatives (for boosts).

For spinor fields, one also needs an 'intrinsic angular momentum' to transform the spinor components. Rotations correspond to

$$J = \frac{\mathrm{i}}{4}n_m\epsilon_{jmk}\left[\gamma^j, \gamma^k\right]$$

[5] One starts with $16 = 4 \times 4$, eliminates the diagonal elements, 12 and arrives at $12/2$ because of the antisymmetry.

the spin operator that you have already found, and

$$J = \frac{\mathrm{i}}{4} n_k \left[\gamma^0, \gamma^k\right]$$

This matrix is actually proportional to $n_k \alpha_k$, the generator of the spinor transformation we have found above.

Covariant expression for the spin

Finally, a covariant expression for the spin can be found from the following algebraic properties: the spin S is related to the eigenvalue of the following operator[6]

$$W_\mu W^\mu \Rightarrow -m^2 S(S+1) \tag{3.173}$$

that can be shown to commute with the generators of the Lorentz group. Actually, one works with the so-called Lorentz-Poincaré group that also contains spatial and temporal displacement. The generators for displacements are the derivatives ∂_μ or equivalently, the momentum operator p_μ. Its covariant square plays the role of eigenvalue for the mass

$$p_\mu p^\mu \Rightarrow m^2 \tag{3.174}$$

a condition that we know as the dispersion relation of the relativistic particle.

Finally, the operator W_μ in Eq.(3.173) is defined as a generalized rotation that leaves the momentum vector p_μ invariant (see the book by Itzykson & Zuber on quantum field theory)

$$W_\mu = -\frac{1}{2} \epsilon_{\mu\nu\kappa\lambda} S^{\nu\kappa} p^\lambda \tag{3.175}$$

where $S^{\nu\kappa}$ is the operator that generates the transformation of the field under Lorentz transformation (it is anti-symmetric in $\nu\kappa$) and p^λ the generator of translations. And $\epsilon_{\mu\nu\kappa\lambda}$ is the totally antisymmetric tensor with four indices and $\epsilon_{0123} = 1$.

[6]The notation $\Rightarrow$ indicates the possible eigenvalues for this operator.

Chapter 4

Extra: symmetries and groups

Dieses Kapitel enthält Information, die über den Stoff der Vorlesung im SS'07 hinausgeht.

4.1 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.

4.2 Continuous symmetry groups

4.2.1 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 $|\det U| = 1$.

4.2.2 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

$$L = \frac{\mathrm{d}R(0)}{\mathrm{d}\theta} \tag{4.1}$$

generator L

4.2.3 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)

$$R_1^{-1}R_2^{-1}R_1R_2 \approx \mathbb{1} + \theta_1\theta_2\,[L_1,\, L_2] \stackrel{!}{=} \mathbb{1} + \sum_k \alpha_k(\theta_1, \theta_2)L_k \tag{4.2}$$

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:

$$[L_i,\, L_j] = \sum_l f_{ijk}L_k, \qquad f_{ijk} = \frac{\partial^2\alpha_k}{\partial\theta_i\partial\theta_j} \tag{4.3}$$

Exercise: rotation group SO(3):

$$f_{ijk} = \epsilon_{ijk}, \qquad (L_i)_{kl} = -\epsilon_{ikl} \tag{4.4}$$

up to signs.

4.2.4 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:

$$S(R_1)S(R_2) = \mathrm{e}^{\mathrm{i}\varphi(R_1,R_2)}S(R_1R_2) \tag{4.5}$$

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 (4.5), $S(\mathbb{1})$ 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.(4.5) 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 $\mathrm{e}^{\mathrm{i}\varphi(R_1,R_2)} = 1$. This corresponds to a unitary (not projective) representation.

Example: scalar wave function

$$S(R)\psi := \psi \circ R^{-1}, \qquad \text{explicitly:} \quad [S(R)\psi](\mathbf{x}') := \psi(R^{-1}\mathbf{x}') \tag{4.6}$$

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:

$$\begin{aligned} \|\psi'\|^2 &= \int \mathrm{d}x' |\psi'(\mathbf{x}')|^2 \\ &= \int \mathrm{d}x' |\psi(R^{-1}\mathbf{x}')|^2 \qquad \text{substitute } \mathbf{y} = R^{-1}\mathbf{x}' \\ &= \int \mathrm{d}y |\det R| |\psi(\mathbf{y})|^2 \end{aligned} \tag{4.7}$$

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 $\mathbf{A}(\mathbf{x})$

$$S(R)\psi := R\mathbf{A} \circ R^{-1}, \qquad \text{explicitly:} \quad [S(R)\mathbf{A}](\mathbf{x}') := R\mathbf{A}(R^{-1}\mathbf{x}') \tag{4.8}$$

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

$$S(R)A = RAR^{-1} \tag{4.9}$$

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 $\Psi(\mathbf{x})$,

$$S(R)\Psi = S_R\Psi \circ R^{-1}, \qquad \text{explicitly:} \quad [S(R)\Psi](\mathbf{x}') := S_R\Psi(R^{-1}\mathbf{x}') \tag{4.10}$$

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

$$S(\theta) := S(R(\theta)) \tag{4.11}$$

for small θ, this gives $S \approx \mathbb{1} + \mathrm{i}\theta J$ 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

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

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

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

$$S(R_1^{-1}R_2^{-1}R_1R_2) = S(R_1^{-1})S(R_2^{-1})S(R_1)S(R_2) \tag{4.12}$$

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

$$\mathrm{i}\,[J_1,\, J_2] = f_{12k}J_k \tag{4.13}$$

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).

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

$$\Psi'(x') = S(R)\Psi(x) \tag{4.14}$$

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 (3.150):

$$[J_k,\, J_l] = \mathrm{i}\epsilon_{klm}J_m \tag{4.15}$$

This can be shown from the following requirement:[1]

$$S(R_1)S(R_2) = S(R_1R_2)\,\mathrm{e}^{\mathrm{i}\varphi(R_1R_2)} \tag{4.16}$$

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 4.)

Now, the commutation relations (4.15) 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 $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ 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 $-j, -j+1, \ldots, j$.

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).

4.3 (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.

$$\mathbb{1} \oplus R := \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R \end{pmatrix} \tag{4.17}$$

[1] There are cases where the phase $\varphi(R_1R_2)$ vanishes.

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,

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \tag{4.18}$$

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.

4.3.1 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 (4.13) 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

$$S'(R) = T^{-1} S(R) T \tag{4.19}$$

for all $R \in \mathrm{SO}(3)$.

As an alternative formulation, we can write

$$S'(R) T\psi = T S(R)\psi, \qquad \text{for all } R \text{ and all } \psi \tag{4.20}$$

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.(4.9) 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.(4.20), 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.

4.3.2 Application: addition of angular momenta

To illustrate the power of these concepts, let us consider the eigenvalue problem for the angular momentum in the Dirac equation: we have seen that the total angular momentum operator is given by

$$\mathbf{J} = \mathbf{L} + \mathbf{S} = \mathbf{L} + \frac{1}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix},$$

which is the sum of the 'orbital' angular momentum and the 'intrinsic' (or spin) angular momentum. Again, this operator generates the rotation of the spinor wave function. Recall that under a rotation, the spinor wave function is changed to

$$\Psi(\mathbf{r}) \mapsto S(R)\Psi(R^{-1}\mathbf{r}) \tag{4.21}$$

This can be written in the form

$$S(R)\Psi(R^{-1}\mathbf{r}) = \exp(\mathrm{i}\theta\mathbf{n}\cdot\mathbf{S})\exp(\mathrm{i}\theta\mathbf{n}\cdot\mathbf{L})\Psi(\mathbf{r}) \tag{4.22}$$

where the operator involving $\mathbf{S}$ acts only on the spinor components while $\mathbf{L}$ only acts on the spatial coordinates.

For an infinitesimal rotation, $\theta \to 0$, and we can check that the generator of the spatial rotation is the usual orbital angular momentum:

$$\begin{aligned} \Psi(R^{-1}\mathbf{r}) &\approx \Psi(\mathbf{r} + \mathbf{n}\times\mathbf{r}\theta + \mathcal{O}(\theta^2)) \\ &\approx \Psi(\mathbf{r}) + \theta(\mathbf{n}\times\mathbf{r})\cdot\nabla\Psi(\mathbf{r}) \\ &= \Psi(\mathbf{r}) + \mathrm{i}\theta\mathbf{n}\cdot(\mathbf{r}\times\mathbf{p})\Psi(\mathbf{r}) \end{aligned} \tag{4.23}$$

We have used the explicit form of an infinitesimal rotation around an axis $\mathbf{n}$, made a Taylor expansion of the wave function and re-wrote the ∇-operator in terms of the momentum operator.

Combining this with the infinitesimal rotation of the spinor components, we find that the total angular momentum generates the rotation of the spinor wave function:

$$\begin{aligned} S(R)\Psi(R^{-1}\mathbf{r}) &= (\mathbb{1} + \mathrm{i}\theta\mathbf{n}\cdot\mathbf{S})\,(1 + \mathrm{i}\theta\mathbf{n}\cdot\mathbf{L})\,\Psi(\mathbf{r}) \\ &\approx [\mathbb{1} + \mathrm{i}\theta\mathbf{n}\cdot(\mathbf{L}+\mathbf{S})]\,\Psi(\mathbf{r}) \end{aligned} \tag{4.24}$$

The conservation of $\mathbf{J}$ thus translates the fact that the Hamilton operator is covariant with respect to rotations.

Tensor product of group representations

We now observe that this kind of transformation corresponds to a tensor product of two group representations. This 'product' corresponds, on the level of matrix elements, to the product of matrix elements from the two representation matrices.

Let us first restrict to a subspace labelled with quantum numbers (l, m) for the spatial coordinate:

$$\Psi_\alpha(\mathbf{r}) = \sum_m c_{m\alpha}(r) Y_{lm}(\theta, \varphi) \tag{4.25}$$

where the Y_{lm} are the spherical harmonics (eigenfunctions of L_z and of $\mathbf{L}^2$). They have the property that under rotations, they are transformed into linear combinations of different m's, but the quantum number l is unchanged.[2]

The index α in Eq.(4.25) labels the spinor components. The full spinor rotation (4.22) now gives

$$\begin{aligned} \left[S(R)\Psi(R^{-1}\mathbf{r})\right]_\beta &= S_{\beta\alpha}(R) \exp(\mathrm{i}\theta \mathbf{n}\cdot\mathbf{L}) \Psi_\alpha(\mathbf{r}) \\ &= S_{\beta\alpha}(R) \exp(\mathrm{i}\theta \mathbf{n}\cdot\mathbf{L}) \sum_m c_{m\alpha}(r) Y_{lm} \\ &= S_{\beta\alpha}(R) \sum_m c_{m\alpha}(r) \sum_{m'} D_{m'm} Y_{lm'} \end{aligned} \tag{4.26}$$

we have written matrix components $D_{m'm}(R)$ for the image of the spherical harmonic Y_{lm}, expanded in the basis of the $Y_{lm'}$.

Eq.(4.26) is a linear combination of basis functions, as in Eq.(4.25). The coefficients of the transformed spinor wave function, $c'_{m'\beta}$ can be read off as

$$c'_{m'\beta} = S_{\beta\alpha}(R) D_{m'm}(R) c_{m\alpha} \tag{4.27}$$

This product of matrix elements can also be understood as the matrix element $U_{m'\beta,m\alpha}$, introducing a 'double index' $m\alpha$ to label the basis vectors (for each Y_{lm}, there are two possible spin eigenvectors, labelled by α).

We have already seen that the spinor transformation corresponds to a representation of the rotation group. For spinor wave functions, we now realize that this representation is of the form of a tensor product of spin and orbital representations.

The key question is now: *what are the irreducible blocks into which the $U_{m'\beta,m\alpha}$ matrices can be reduced?* We know from the theorems for group representations (Sec.4.3.1) that each block corresponds to a representation of the (j, m)-type. In addition, rotations in this block are generated by the total angular momentum $\mathbf{J}$, and therefore, the quantum number j corresponds to the possible values of $\mathbf{J}^2 = j(j+1)$. The irreducible blocks that we are looking for thus give us also the answer to the question: *given the spin and orbital angular momentum quantum numbers s, l, what are the possible eigenvalues j for the sum of these two angular momentum operators?*

The answer is the following:

[2] The subspace of wave functions spanned by the set of basis vectors Y_{lm} is called an 'invariant subspace' with respect to rotations.

In the tensor product representation generated by an orbital angular momentum with quantum number l and a spin angular momentum s, the possible values for the total angular momentum are $j = |l-s|, \ldots l+s$ and each value occurs exactly once.

One can check that this is consistent from the viewpoint of dimensions: the tensor product matrix $U_{m'\beta,m\alpha}$ acts on a 'tensor product' Hilbert space with dimension $(2s+1)(2l+1)$ (for each Y_{lm}, there are $2s+1$ spin components/possible eigenvectors). For the block-diagonal form, one can see that the dimensions of the invariant subspaces just add up to the total dimension so that and one has

$$(2s+1)(2l+1) = \sum_{j=|l-s|}^{l+s} (2j+1) \tag{4.28}$$

For the construction of the (j,m) subrepresentions in the tensor product space, details can be found in the textbooks. We have sketched the idea in Section 3.6.2, and continue with a group-theoretical argument below.

Group theory argument

Different approach, more elegant.

Definition *character* of a representation

$$\chi_U(R) = \operatorname{tr} U(R) \tag{4.29}$$

mapping from the group G into $\mathbb{C}$. Depends on the unitary matrices $U(R)$.

Key **Theorem**: the character of any representation of the rotation group depends only on the rotation angle θ and can be expanded in a unique way:

$$\chi_U(\theta) = \sum_j a_j \chi_j(\theta) \tag{4.30}$$

where χ_j is the character for a (j,m)-representation with coefficients

$$a_j = \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \sin^2(\theta/2)\chi_j(\theta)\chi_U(\theta) \tag{4.31}$$

The coefficients a_j are non-negative integers and 'count' how often the (j,m) representation occurs as sub-representation in the block-diagonal form of U.

Elements of a proof.

(1) The character only depends on the rotation angle.

For any rotation $R(\mathbf{n},\theta)$, one can find a rotation matrix that maps the rotation axis $\mathbf{n}$ onto the z-axis, hence

$$R(\mathbf{n},\theta) = T_{\mathbf{n}}^{-1} R(\mathbf{e}_z,\theta) T_{\mathbf{n}} \tag{4.32}$$

This is true because $T_{\mathbf{n}}$ just corresponds to a change of basis, and 'without loss of generality', one can choose another basis where the rotation axis coincides with the z-axis. Or: work it out with T some rotation matrix, starting from an explicit representation of $R(\mathbf{n}, \theta)$. This gives a rotation around the 'turned axis' $R(T\mathbf{n}, \theta)$. Hence with a suitable choice of T, the turned axis points along the z-axis.

Finally, from the properties of the trace,

$$\mathrm{tr}\left(T^{-1}R(\mathbf{e}_z,\theta)T\right) = \mathrm{tr}\left(TT^{-1}R(\mathbf{e}_z,\theta)\right) = \mathrm{tr}R(\mathbf{e}_z,\theta)$$

the character is unchanged by the basis change and does not depend on $\mathbf{n}$. And by the group homomorphism property, this translates to the unitary matrices $U(R)$ as well.

(2) The character for the (j, m)-representations are orthogonal and one can expand any character in this basis.
Consider the unitary matrix for a rotation around the z-axis:

$$D_{mm'}(\theta, \mathbf{e}_z) = \begin{pmatrix} \mathrm{e}^{\mathrm{i}j\theta} & 0 & & & \\ 0 & \ddots & & & \\ & & \mathrm{e}^{\mathrm{i}m\theta} & & \\ & & & \ddots & 0 \\ & & & 0 & \mathrm{e}^{-\mathrm{i}j\theta} \end{pmatrix} \tag{4.33}$$

This form directly follows from $D_{mm'}(\theta, \mathbf{e}_z) = \exp(\mathrm{i}\theta J_3)$ with the diagonal matrix J_3 in the $|j, m\rangle$ basis. Hence, taking the trace:

$$\chi_j(\theta) = \sum_{m=-j}^{j} \mathrm{e}^{\mathrm{i}m\theta} = \frac{\sin[(j+\frac{1}{2})\theta]}{\sin(\theta/2)} \tag{4.34}$$

after summation of the geometric series.

(3) Now, the functions $\sin[(j + \frac{1}{2})\theta]$ are a Fourier basis for odd functions of θ that are periodic with period 4π. Hence, with the weight $\sin^2(\theta/2)$ given in Eq.(4.31), the characters $\chi_j(\theta)$ are orthogonal. Conversely, Eqs.(4.30, 4.31) give the Fourier expansion of a character $\chi_U(\theta)$.

That any character is even in θ, follows from the fact that the rotation $R(\mathbf{n}, -\theta)$ is the same as $R(-\mathbf{n}, \theta)$, and applying Eq.(4.32). That it must be periodic with period 4π can be seen for the (j, m)-characters (4.34). A period of 2π is obvious for integer j because here the wavefunctions must be periodic in the azimuthal angle φ. A rotation of 4π is needed to remove a factor -1 for a half-integer angular momentum: this is characteristic for spin-1/2 particles, for example.

(4) The character for a representation does not change when the representation is brought into block-diagonal form (this is an equivalent representation, $T^{-1}UT$ that has the same trace). In addition, the character/trace of a block-diagonal matrix is the sum of the characters/traces of the sub-blocks. Hence

from our basic theorem, in the expansion (4.30) only χ_j's can appear, and the coefficients a_j count how often. These coefficients also must be integers because for $\theta = 0$, $\chi_U(0)$ is the dimension of the representation that is the sum of the dimensions of its sub-representations.

Application: angular momentum addition. Let us apply this theorem to the tensor product representation generated by $\mathbf{L}$ and $\mathbf{S}$ with quantum numbers l and s. The unitary matrices of the tensor product have the following trace

$$\begin{aligned} \mathrm{tr}\,(U_{m\alpha,m'\beta}) &= \sum_{m\alpha} U_{m\alpha,m\alpha} \\ &= \sum_{m\alpha} S_{\alpha\alpha} D_{mm} \\ &= \sum_{\alpha} S_{\alpha\alpha} \sum_{m} D_{mm} \\ &= \mathrm{tr}(S)\mathrm{tr}(D) \end{aligned} \tag{4.35}$$

hence the character $\chi_U(\theta)$ is here the product of $\chi_l(\theta)$ and $\chi_s(\theta)$.

The expansion coefficients are thus given by the integral (4.31)

$$\begin{aligned} a_j &= \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \chi_j(\theta) \sin[(l+\tfrac{1}{2})\theta] \sin[(s+\tfrac{1}{2})\theta] \\ &= \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \frac{1}{2} \left[\cos[(l+s+1)\theta] - \cos[(l-s)\theta]\right] \sum_{m=-j}^{j} \mathrm{e}^{\mathrm{i}m\theta} \end{aligned} \tag{4.36}$$

Now these kind of integrals give a nonzero result only of the 'frequency' m of the exponentials match the frequencies $\pm(l-s)$ and $\pm(l+s+1)$ of the cosine functions. Hence, the integral is zero if $j < |l-s|$. It is also zero if $j \geq l+s+1$ because in that case two terms in the sum over m are nonzero, but they occur with opposite sign (note the minus sign between the cosines).

Hence, one finds the allowed values for the total angular momentum quantum number j:

$$a_j = 1 \quad \text{for} j = |l-s|, \ldots l+s \tag{4.37}$$

and in agreement with the result announced above, each value occurs exactly once in the reduction of the representation.

4.4 Miscellaneous

Remark. For the rotation group, there is a unitary representation in SU(2) with the property

$$(R^T \mathbf{x}) \cdot \boldsymbol{\sigma} = \mathbf{x} \cdot S^\dagger(R) \boldsymbol{\sigma} S(R) \tag{4.38}$$

for all 3D vectors $\mathbf{x}$ and the 'vector of Pauli matrices' $\boldsymbol{\sigma}$. We have seen in the lecture that for a rotation angle θ and an axis $\mathbf{n}$,

$$S(R) = \exp[\mathrm{i}(\boldsymbol{\sigma} \cdot \mathbf{n})\theta/2] \tag{4.39}$$

so that the Pauli matrices $\sigma/2$ are the generators of this representation.[3] In a previous exercise (spin coupling to a magnetic field), you have in fact checked this relation, both globally and infinitesimally.

It is quite involved to show that the so-defined $S(R)$ is actually a group homomorphism because one has to identify how rotations defined by $\mathbf{n}, \theta$ multiply. Conversely, Eq.(4.38) can be used to define a matrix R for a given SU(2) matrix S. Note that S and $-S$ are mapped onto the same rotation matrix. This sign ambiguity is typical for the two-component spinors for a spin $1/2$ particle. In the exercises, you show that the so-defined R is an orthogonal matrix (keeps $|\mathbf{x}|$ invariant) and that it preserves the handedness (its determinant is $+1$). In addition, the composition is compatible with the multiplication (composition) in the rotation group.

[3]The construction (4.38) only works because the 'spatial vector' $\mathbf{x}$ has three (real) components and there are three generators for SU(2) (and for SO(3)). For the four-dimensional Minkowski space-time, something different had to be done because there are six generators for the Lorentz group (rotations and boosts).