Relativistic Waves and Quantum Fields

Andreas Brandhuber

Department of Physics
Queen Mary, University of London
Mile End Road, London, E1 4NS
United Kingdom
email: a.brandhuber@qmul.ac.uk

Abstract

Lecture notes for the course Relativistic Waves and Quantum Fields MSci 4242/QMUL PHY-415. (These are preliminary notes and may contain typos, so use with care!)

Contents

Chapter 0

Why Quantum Field Theory?

The fundamental challenge of every physical theory is to describe Nature and its phenomena in as much detail as possible. Furthermore, this should be possible for the microcosm and the macrocosm i.e. we are looking for theories that are valid at short distance scales and long distance scales.

Classical physics is able to give a satisfactory description of many physical phenomena, and its validity is not limited to the macroscopic regime alone. Certain aspects of the microcosm can be treated with classical physics, e.g. certain aspects of kinetic gas theory. But in general classical physics gives only an approximation that breaks down when we reach a certain energy or distance scale.

The deficiencies of the classical viewpoint are most pronounced in the physics of the microcosm and in the subnuclear regime. Many phenomena of molecular physics, atomic physics, nuclear physics and elementary particle physics simply cannot be explained using ideas from classical physics. Also the inclusion of Quantum Mechanics (QM) alone is often insufficient, to explain physical problems of the microcosm. As was realised in the last century, only the inclusion of Einstein's Theory of Special Relativity (SR) leads to the desired success.

Today Quantum Field Theory (QFT) provides *the* unified framework to describe all particles and its interactions (forces) that we observe in Nature, including electromagnetism, the weak and the strong nuclear force, and probably gravity. QFT was born out of the attempt to combine the rules of Quantum Mechanics with the principles of Special Relativity and field theory. The prime example for the success of QFT is Quantum Electrodynamics (QED), which with its extremely accurate predictions of physical quantities, like the anomalous magnetic moment of the electron, is also called "the jewel of physics".

Chapter 1

Quantum Mechanics and Special Relativity

In this chapter we want to review and introduce basic concepts of Quantum Mechanics and Special Relativity. At this stage we treat these two topics separately before we attempt to unify them in Chapter 2 trying to preserve most features of usual non-relativistic QM. This will lead only to partial success since a couple of conceptual problems will arise. As was understood in the last century, the main reason for these problems is that the inclusion of SR implies the existence of anti-particles and leads to a multi-particle theory in which probability is not preserved, because particles can be destroyed and created (death and birth of particles). These issues can be addressed properly if we adopt the formalism of Quantum Field Theory, which will be developed from Chapter 3 on.

1.1 Quantum Mechanics

1.1.1 Principles of Non-Relativistic Quantum Mechanics

In this section I will remind you of the most important axioms of Quantum Mechanics (QM). I expect that you are familiar with most of these facts, but it is good to repeat them to see which ones we will be forced to give up once we include SR.

Axioms of QM:

- The **state of the system** is represented by a **wave function** $|\Psi\rangle$ (also called a vector in Hilbert space). Note that $\Psi = \Psi(q_i, s_i, t)$ is a function of the coordinates of individual particles q_i, some internal degrees of freedom like spin s_i and, if we are in the Schödinger picture also on time t (if you are not familiar with this concept either wait until I will review it later in the course or check your favourite book on QM). Most importantly Ψ itself has no direct physical meaning whatsover; only expression like the probability density $\rho = |\Psi|^2 \geq 0$.

- **Physical observables** are in one-to-one correspondence with **linear, Hermitian operators**. Hermiticity $\widehat{A} = \widehat{A}^\dagger$ implies that all **eigenvalues** of the operators are real. Example: momentum $\widehat{p}_x = -i\hbar\frac{\partial}{\partial x}$.

- The system can be in an **eigenstate** ψ_n which obeys $\widehat{A}\psi_n = \lambda_n\psi_n$ with real eigenvalue λ_n. A general state Ψ can be written as a linear combination of a **complete set of eigenfunctions** of a complete set of commuting operators $[\widehat{A}_i, \widehat{A}_j] = 0$ (thanks to linearity of the QM operators). So $\Psi = \sum_n a_n\psi_n$ where the a_n are complex coefficients. If the eigenfunctions are properly *orthonormalised* (= orthogonal and normalised to one) i.e. $\int d^3x\psi_n^*\psi_m = \delta_{nm}$ then $|a_n|^2$ is the probability to find the system in the state ψ_n.

- The result of a QM measurement of $\widehat{A}$ is one of its eigenvalues λ_n (with probability $|a_n|^2$). The QM **expectation value** — the average over many measurements — is

$$\langle A\rangle = \langle\Psi|\widehat{A}|\Psi\rangle = \int d^3x\Psi^*\widehat{A}\Psi = \sum_n |a_n|^2\lambda_n\,. \tag{1.1.1}$$

- The **time evolution** of the system is desribed by the Schrödinger equation $i\hbar\frac{\partial\Psi}{\partial t} = \widehat{H}\Psi$. The QM Hamiltionian $\widehat{H}$ is obtained from its classical counterpart H by the replacements $\vec{p} \to -i\hbar\vec{\nabla}$ and $E \to +i\hbar\frac{\partial}{\partial t}$. For example: $H = \frac{\vec{p}^2}{2m} + V(\vec{x}) \to H = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x})$. Since $\widehat{H}$ is linear we can apply the superposition principle to wavefunctions. Furthermore, Hermiticiy $\widehat{H} = \widehat{H}^\dagger$ implies the conservation of probability. <u>Proof:</u>

$$\begin{aligned}\frac{d}{dt}\langle\Psi|\Psi\rangle &= \frac{d}{dt}\int\Psi^*\Psi = \int(\frac{\partial\Psi^*}{\partial t}\Psi + \Psi^*\frac{\partial\Psi}{\partial t}\Psi)\\ &= \int((-\frac{i}{\hbar}\widehat{H}\Psi)^*\Psi + \Psi^*(-\frac{i}{\hbar}\widehat{H}\Psi))\\ &= \frac{i}{\hbar}\int((\widehat{H}\Psi)^*\Psi - \Psi^*\widehat{H}\Psi) = \frac{i}{\hbar}(\langle\widehat{H}\Psi|\Psi\rangle - \langle\Psi|\widehat{H}\Psi\rangle) = 0\,.\end{aligned} \tag{1.1.2}$$

1.1.2 Symmetries and Angular Momentum

Symmetries and Conservation Laws

In Quantum Mechanics (QM), for an observable A without explicit time dependence[1] we have

$$i\hbar\frac{d\langle A\rangle}{dt} = \langle\Psi|[\widehat{A},\widehat{H}]|\Psi\rangle\,, \tag{1.1.3}$$

for any state $|\Psi\rangle$ and where $\langle A\rangle = \langle\Psi|\widehat{A}|\Psi\rangle$, is the QM expectation value of A.
Proof:

$$\begin{aligned}
i\hbar\frac{d\langle A\rangle}{dt} &= i\hbar\frac{d}{dt}\int \Psi^*\widehat{A}\Psi \\
&= i\hbar\int \frac{\partial\Psi^*}{\partial t}\widehat{A}\Psi + \Psi^*\widehat{A}\frac{\partial\Psi}{\partial t} \\
&= \int -(i\hbar\frac{\partial\Psi}{\partial t})^*\widehat{A}\Psi + \Psi^*\widehat{A}(i\hbar\frac{\partial\Psi}{\partial t}) \\
&= \int -(\widehat{H}\Psi)^*\widehat{A}\Psi + \Psi^*\widehat{A}\widehat{H}\Psi \\
&= \int -\Psi^*\widehat{H}\widehat{A}\Psi + \Psi^*\widehat{A}\widehat{H}\Psi \\
&= \langle\Psi|\left[\widehat{A},\widehat{H}\right]|\Psi\rangle\,.
\end{aligned} \tag{1.1.4}$$

Therefore, if

$$[\widehat{A},\widehat{H}] = 0\,, \tag{1.1.5}$$

then

$$\frac{d\langle A\rangle}{dt} = 0\,, \tag{1.1.6}$$

and we say that A is a conserved quantitity or constant of motion.

A **symmetry** is a transformation on the coordinates of a system which leaves the Hamiltonian H invariant. We shall see that **conservation laws** are the consequence of **symmetries of a system**. Symmetries are very powerful since they can be used to derive results for a system even when we do not know the details of the dynamics involved.

[1] i.e. the QM operator $\widehat{A}$ corresponding to the observable A obeys $\partial\widehat{A}/\partial t = 0$.

Translational Invariance

First consider a single particle. If $\vec{x}$ is the position vector, then a translation is the operation

$$\vec{x} \to \vec{x}' = \vec{x} + \vec{a}\,. \tag{1.1.7}$$

If $\widehat{H}$ is invariant then

$$\widehat{H}(\vec{x}') = \widehat{H}(\vec{x} + \vec{a}) = \widehat{H}(\vec{x})\,. \tag{1.1.8}$$

For an infinitesimal displacement we can make a Taylor expansion[2]

$$\widehat{H}(\vec{x} + \vec{a}) \cong \widehat{H}(\vec{x}) + \vec{a} \cdot \vec{\nabla} \widehat{H}(\vec{x})\,, \tag{1.1.9}$$

ignoring higher powers of $\vec{a}$. Thus, if $\widehat{H}$ is invariant,

$$0 = \widehat{H}(\vec{x} + \vec{a}) - \widehat{H}(\vec{x}) = \vec{a} \cdot \vec{\nabla} \widehat{H}(\vec{x})\,. \tag{1.1.10}$$

In general for the momentum operator $\widehat{\vec{P}}$ and any other operator $\widehat{O}(\vec{x})$ we have

$$\begin{aligned} [\widehat{\vec{P}}, \widehat{O}]\Psi &= [-i\hbar\vec{\nabla}, \widehat{O}]\Psi \\ &= -i\hbar\vec{\nabla}(\widehat{O}\Psi) - \widehat{O}(-i\hbar\Psi) \\ &= -i\hbar(\vec{\nabla}\widehat{O})\Psi\,, \end{aligned} \tag{1.1.11}$$

where we have suppressed the explicit $\vec{x}$ and t dependence. Since eqn. (1.1.11) is true for arbitrary wavefunctions Ψ

$$[\widehat{\vec{P}}, \widehat{O}(\vec{x})] = -i\hbar\vec{\nabla}\widehat{O}(\vec{x})\,. \tag{1.1.12}$$

In particular for $\widehat{O} = \widehat{H}$

$$[\widehat{\vec{P}}, \widehat{H}] = -i\hbar\vec{\nabla}\widehat{H}\,. \tag{1.1.13}$$

Now,

$$0 = -i\hbar\vec{a} \cdot \vec{\nabla}\widehat{H} = \vec{a} \cdot [\widehat{\vec{P}}, \widehat{H}] \tag{1.1.14}$$

and since this is true for an arbitrary displacement vector $\vec{a}$ we find

$$[\widehat{\vec{P}}, \widehat{H}] = 0\,. \tag{1.1.15}$$

We conclude that momentum is a conserved quantity $\frac{d\langle\vec{P}\rangle}{dt} = 0$ if the $\widehat{H}$ is translationally invariant.

Take e.g.

$$\widehat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x}) \tag{1.1.16}$$

[2] The symbol $\cong$ indicates that we only expand to first order in $\vec{a}$ and suppress all higher order terms.

For the translation, $\vec{x}' = \vec{x} + \vec{a}$, in particular $x' = x + a$

$$\Rightarrow \frac{\partial}{\partial x} = \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} = \frac{\partial}{\partial x'} \tag{1.1.17}$$

and similarly for y and z. Thus $\vec{\nabla}$ and $\vec{\nabla}^2$ are invariant under translations and

$$\begin{aligned} \widehat{H}(\vec{x}') &= \widehat{H}(\vec{x}+\vec{a}) = -\frac{\hbar^2}{2m}\vec{\nabla}'^2 + V(\vec{x}') \\ &= -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x}+\vec{a})\,. \end{aligned} \tag{1.1.18}$$

Hence, for a translationally invariant Hamiltonian we must require

$$V(\vec{x}+\vec{a}) = V(\vec{x})\,, \tag{1.1.19}$$

which is only true for a (trivial) constant potential, i.e. for a free particle. Thus, the momentum of a free particle is conserved in QM in the sense

$$\frac{d\langle\vec{P}\rangle}{dt} = 0\,. \tag{1.1.20}$$

Consider now a two particle system (easily generalised to N particles). If the two particles have position vectors $\vec{x}_1$ and $\vec{x}_2$, the invariance condition for the translation of the system through $\vec{a}$ reads

$$\widehat{H}(\vec{x}_1,\vec{x}_2) = \widehat{H}(\vec{x}_1+\vec{a},\vec{x}_2+\vec{a})\,. \tag{1.1.21}$$

Then for an infinitesimal translation

$$\widehat{H}(\vec{x}_1+\vec{a},\vec{x}_2+\vec{a}) \cong \widehat{H}(\vec{x}_1,\vec{x}_2) + \vec{a}\cdot\vec{\nabla}_1\widehat{H}(\vec{x}_1,\vec{x}_2) + \vec{a}\cdot\vec{\nabla}_2\widehat{H}(\vec{x}_1,\vec{x}_2)\,, \tag{1.1.22}$$

we find that translational invariance implies

$$0 = \vec{a}\cdot(\vec{\nabla}_1+\vec{\nabla}_2)\widehat{H}(\vec{x}_1,\vec{x}_2)\,, \tag{1.1.23}$$

and the total momentum operator is

$$\widehat{\vec{P}} = \widehat{\vec{P}}_1 + \widehat{\vec{P}}_2\,, \tag{1.1.24}$$

where

$$\widehat{\vec{P}}_1 = -i\hbar\vec{\nabla}_1 \;\text{ and }\; \widehat{\vec{P}}_2 = -i\hbar\vec{\nabla}_2\,. \tag{1.1.25}$$

Identical to the one particle case, for any operator $\widehat{O}(\vec{x}_1,\vec{x}_2)$,

$$[\widehat{\vec{P}}_1,\widehat{O}(\vec{x}_1,\vec{x}_2)] = -i\hbar\vec{\nabla}_1\widehat{O}(\vec{x}_1,\vec{x}_2) \;\text{ and }\; [\widehat{\vec{P}}_2,\widehat{O}(\vec{x}_1,\vec{x}_2)] = -i\hbar\vec{\nabla}_2\widehat{O}(\vec{x}_1,\vec{x}_2)\,, \tag{1.1.26}$$

and so

$$[\widehat{\vec{P}}, \widehat{O}(\vec{x}_1, \vec{x}_2)] = -i\hbar(\vec{\nabla}_1 + \vec{\nabla}_2)\widehat{O}(\vec{x}_1, \vec{x}_2)\,. \tag{1.1.27}$$

This is true in particular for $\widehat{O} = \widehat{H}$. Thus,

$$0 = -i\hbar\vec{a}\cdot(\vec{\nabla}_1 + \vec{\nabla}_2)\widehat{H} = \vec{a}\cdot[\widehat{\vec{P}}, \widehat{H}]\,. \tag{1.1.28}$$

Since this must be true for arbitrary translation vector $\vec{a}$, we have

$$[\widehat{\vec{P}}, \widehat{H}] = 0\,, \tag{1.1.29}$$

and total momentum is conserved in the sense of QM i.e. $\frac{d\langle\vec{P}\rangle}{dt} = 0$.

Rotational Invariance

Just as translational invariande is associated with conservation of momentum, it turns out that rotational invariance is associated with conservation of angular momentum (AM).

Take spherical polar coordinates and take the axis of rotation to be the z-axis. Specify the position vector $\vec{x}$ in spherical polar coordinates (r, θ, ϕ) of the point. Then the symmetry operation $\vec{x} \to \vec{x}'$ corresponding to a rotation by an angle α about the $z-$axis is

$$(r, \theta, \phi) \longrightarrow (r', \theta', \phi') = (r, \theta, \phi + \alpha)\,. \tag{1.1.30}$$

For the Hamiltonian to be invariant under rotations about the z-axis

$$\begin{aligned} & \widehat{H}(\vec{x}') = \widehat{H}(\vec{x}) \\ \Longleftrightarrow\quad & \widehat{H}(r', \theta', \phi') = \widehat{H}(r, \theta, \phi) \\ \Longleftrightarrow\quad & \widehat{H}(r, \theta, \phi + \alpha) = \widehat{H}(r, \theta, \phi)\,. \end{aligned} \tag{1.1.31}$$

For an infinitesimal rotation

$$\widehat{H}(r, \theta, \phi + \alpha) \cong \widehat{H}(r, \theta, \phi) + \alpha\frac{\partial}{\partial\phi}\widehat{H}(r, \theta, \phi)\,, \tag{1.1.32}$$

and, hence, for invariance of $\widehat{H}$

$$\begin{aligned} 0 &= \widehat{H}(r, \theta, \phi + \alpha) - \widehat{H}(r, \theta, \phi) = \alpha\frac{\partial}{\partial\phi}\widehat{H}(r, \theta, \phi) \\ \longrightarrow\quad & \frac{\partial}{\partial\phi}\widehat{H}(r, \theta, \phi) = 0\,. \end{aligned} \tag{1.1.33}$$

The $z-$component of the orbital AM operator written in spherical coordinates is[3]

$$\widehat{L}_z = -i\hbar\frac{\partial}{\partial\phi} \tag{1.1.34}$$

[3] In cartesian coordinates $\vec{x} = (x, y, z)$: $\widehat{L}_z = (\widehat{\vec{x}} \times \widehat{\vec{P}})_z = -i\hbar(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})$. The other components, $\widehat{L}_x$ and $\widehat{L}_y$, can be obtained by cyclic permutation of (x, y, z).

In general, for any other operator $\widehat{O}$,

$$[\widehat{L}_z, \widehat{O}]\Psi = -i\hbar[\frac{\partial}{\partial\phi}, \widehat{O}] = -i\hbar(\frac{\partial}{\partial\phi}(\widehat{O}\Psi) - \widehat{O}\frac{\partial}{\partial\phi}\Psi) = -i\hbar\frac{\partial\widehat{O}}{\partial\phi}\Psi\,. \tag{1.1.35}$$

This is true for arbitrary wavefuntions Ψ, thus

$$[\widehat{L}_z, \widehat{O}] = -i\hbar\frac{\partial\widehat{O}}{\partial\phi}\,, \tag{1.1.36}$$

and in particular for $\widehat{O} = \widehat{H}$,

$$[\widehat{L}_z, \widehat{H}] = -i\hbar\frac{\partial\widehat{H}}{\partial\phi}\,. \tag{1.1.37}$$

If the Hamiltonian is invariant under rotation about the z-axis, we now conclude that

$$[\widehat{L}_z, \widehat{H}] = 0\,. \tag{1.1.38}$$

We can define angles ϕ_x and ϕ_y analogous to $\phi_z \equiv \phi$ for rotations about the x and y-axis. If $\widehat{H}$ is also invariant under rotations about the x and y-axis we will conclude that

$$[\widehat{L}_x, \widehat{H}] = [\widehat{L}_y, \widehat{H}] = [\widehat{L}_z, \widehat{H}] = 0 \;\text{ i.e. }\; [\widehat{\vec{L}}, \widehat{H}] = 0\,. \tag{1.1.39}$$

Thus AM is a constant of motion

$$\frac{d\langle\vec{L}\rangle}{dt} = 0\,. \tag{1.1.40}$$

Any rotation can be built out of successive rotations about the x, y and z-axis.

Whenever $\widehat{H}$ is invariant under arbitrary rotations the AM $\vec{L}$ is a conserved quantity. We must construct such an $\widehat{H}$ out of scalars i.e. invariants under rotations. The simplest examples of scalars are the magnitude (length) of a vector or the scalar product of two vectors. For example, consider the Hamiltonian for a particle moving in a central potential such as the Coulomb potential

$$\widehat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(|\vec{x}|)\,. \tag{1.1.41}$$

Because $\vec{\nabla}^2 = \vec{\nabla}\cdot\vec{\nabla}$ and $|\vec{x}|$ are scalars, so is the Hamiltonian and orbital AM is conserved.

This discussion generalizes immediately to two (or more) particles:

For 2 particles, the Hamiltonian is a function of two sets of spherical coordinates (r_1, θ_1, ϕ_1) and (r_2, θ_2, ϕ_2) so that

$$\widehat{H} = \widehat{H}(r_1, \theta_1, \phi_1; r_2, \theta_2, \phi_2)\,. \tag{1.1.42}$$

The invariance condition for rotation of the system by an angle α about the z-axis is

$$\widehat{H}(r_1,\theta_1,\phi_1;r_2,\theta_2,\phi_2)=\widehat{H}(r_1,\theta_1,\phi_1+\alpha;r_2,\theta_2,\phi_2+\alpha)\,, \tag{1.1.43}$$

and for an infinitesimal rotation

$$\widehat{H}(r_1,\theta_1,\phi_1+\alpha;r_2,\theta_2,\phi_2+\alpha)\cong\widehat{H}(r_1,\theta_1,\phi_1;r_2,\theta_2,\phi_2)+\alpha\partial_{\phi_1}\widehat{H}+\alpha\partial_{\phi_2}\widehat{H}\,, \tag{1.1.44}$$

where we defined $\partial_\phi\equiv\frac{\partial}{\partial\phi}$. Invariance of $\widehat{H}$ gives

$$\begin{aligned} 0 &= \widehat{H}(r_1,\theta_1,\phi_1+\alpha;r_2,\theta_2,\phi_2+\alpha)-\widehat{H}(r_1,\theta_1,\phi_1;r_2,\theta_2,\phi_2)\\ &= \alpha(\partial_{\phi_1}+\partial_{\phi_2})\widehat{H}\\ &\longrightarrow (\partial_{\phi_1}+\partial_{\phi_2})\widehat{H}=0\,. \end{aligned} \tag{1.1.45}$$

The z-components of the orbital AM operator for the two particles are

$$\widehat{L}_{1z}=-i\hbar\partial_{\phi_1}\ \ ,\ \ \widehat{L}_{2z}=-i\hbar\partial_{\phi_2}\,. \tag{1.1.46}$$

The z-component of the total orbital AM is:

$$\widehat{L}_z=\widehat{L}_{1z}+\widehat{L}_{2z}\,. \tag{1.1.47}$$

Much as before, for any operator O,

$$[\widehat{L}_z,\widehat{O}]=-i\hbar(\partial_{\phi_1}+\partial_{\phi_2})\widehat{O}\,, \tag{1.1.48}$$

and in particular

$$[\widehat{L}_z,\widehat{H}]=-i\hbar(\partial_{\phi_1}+\partial_{\phi_2})\widehat{H}\,. \tag{1.1.49}$$

If the Hamiltonian is invariant under rotations about the z-axis, we now conclude that

$$[\widehat{L}_z,\widehat{H}]=0\,. \tag{1.1.50}$$

By also considering rotations about the x and y-axis we conclude that if $\widehat{H}$ is invariant under arbitrary rotations then

$$[\widehat{\vec{L}},\widehat{H}]=0\,, \tag{1.1.51}$$

so total AM is a constant of motion

$$\frac{d\langle\vec{L}\rangle}{dt}=0\,. \tag{1.1.52}$$

As usual the orbital AM operators obey the ($SO(3)$ or $SU(2)$) algebra

$$[\widehat{L}_x,\widehat{L}_y]=i\hbar\widehat{L}_z\ \ ,\ \ [\widehat{L}_x,\widehat{L}_y]=i\hbar\widehat{L}_z\ \ ,\ \ [\widehat{L}_x,\widehat{L}_y]=i\hbar\widehat{L}_z\,. \tag{1.1.53}$$

Consequently, it is only possible to know simultaneously the values of one component $\vec{L}$ and $\vec{L}^2$, for example, L_z and $\vec{L}^2$. Therefore, we may take the conserved quantities to be L_z and $\vec{L}^2$.

An atom with atomic number Z with Coulomb forces between the nucleus and the electrons and between the electrons is an example of a system with the necessary rotational invariance for conservation of AM. In this case:

$$\widehat{H} = -\frac{\hbar^2}{2m}\sum_{i=1}^{Z}\vec{\nabla}_i^2 - \sum_{i=1}^{Z}\frac{Ze^2}{4\pi\epsilon_0|\vec{x}_i|} + \sum_{i,j=1,i<j}^{Z}\frac{e^2}{4\pi\epsilon_0|\vec{x}_i - \vec{x}_j|}\,, \tag{1.1.54}$$

where $\vec{\nabla}_i$ acts on the coordinates of the i-th electron, and $\vec{x}_i$ is the position vector of the i-th electron relative to the nucleus. As before, $\vec{\nabla}_i^2$ and $|\vec{x}_i|$ are scalars (rotationally invariant) and so is $|\vec{x}_i - \vec{x}_j|$, so that $\widehat{H}$ is rotationally invariant.

Remark: Generally for a system with spin it is the total AM

$$\widehat{\vec{J}} = \widehat{\vec{L}} + \widehat{\vec{S}} \tag{1.1.55}$$

that commutes with the Hamiltonian

$$[\widehat{\vec{J}}, \widehat{H}] = 0 \tag{1.1.56}$$

if $\widehat{H}$ is rotationally invariant, and then $\langle J_z\rangle$ and $\langle \vec{J}^2\rangle$ are constants of motion.

For the atomic (non-relativistic) Hamiltonian above, $\langle L_z\rangle$ and $\langle \vec{L}^2\rangle$ are also constants of motion, because the spin does not appear explicitly in the Hamiltion. This is a manifestation of the fact that spin is an effect of Special Relativity as we will see later in the course. However, if we include the Spin-Orbit interaction due to Relativistic effects

$$\widehat{H}_{Spin-Orbit} = \frac{1}{2m_e^2c^2}\frac{1}{r}\frac{dV}{dr}\widehat{\vec{L}}\cdot\widehat{\vec{S}}\,, \tag{1.1.57}$$

for an electron moving in a central potential $V(r)$, then $\widehat{\vec{L}}$ does not commute with the complete Hamiltonian. Then, only $\langle J_z\rangle$ and $\langle \vec{J}^2\rangle$ are conserved quantities.

1.2 Special Relativity

Topics we are covering in this section include: Lorentz transformations, 4-vectors, covariance, contravariance, forminvariance, Lorentz and Poincare group, Maxwell equations as an example of a relativistic wave equations.

Einstein's Theory of Special Relativity (1905) rests on two postulates:

1. The speed of light c in vacuum is absolut. It is the same in all inertial frames (these are reference frames moving at constant velocity with respect to each other).
2. Principle of Relativity. The laws of physics are the same in all inertial frames. Or in other words there is no preferred reference frame and, in particular, there is no absolute time.

Two observers in two different inertial frames K and K' using coordinate systems (t, x, y, z) and (t', x', y', z') should agree on the predicted results of all experiments. In other words the form of all dynamical equations should be invariant (Forminvariance).

An important example are the Maxwell equations written for $\vec{E}$ and $\vec{B}$. The fields *do* transform under Lorentz transformations, but the Maxwell equations take the same form when written in terms of the new (t', x', y', z') coordinates and the Lorentz transformed fields $\vec{E}'$ and $\vec{B}'$. Actually, if Maxwell had looked more carefully he should have discovered special relativity, since Maxwell equations are not forminvariant under Galilei transformations but only under Lorentz transformations!

1.2.1 Lorentz Transformations

You are all familiar with Lorentz transformations (LT), say a boost in the x direction:

$$\begin{aligned} ct' &= \frac{ct - vx/c}{\sqrt{1 - v^2/c^2}} = \gamma(ct - \beta x) \\ x' &= \frac{x - vt}{\sqrt{1 - v^2/c^2}} = \gamma(-\beta ct + x) \\ y' &= y \\ z' &= z \end{aligned} \tag{1.2.1}$$

with $\beta = v/c$ and $\gamma = 1/\sqrt{1 - v^2/c^2}$. This LT can be neatly rewritten in matrix form as

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} . \tag{1.2.2}$$

Note that general LTs are linear transformations of the coordinates. If we were to consider general relativity we would have to relax this condition and consider general non-linear transformations.

An important consequence of Einstein's postulates is that the quantity

$$(ct)^2 - x^2 - y^2 - z^2 \tag{1.2.3}$$

is Lorentz invariant. Try this with the LT given above

$$\begin{aligned}
(ct')^2-(x')^2-(y')^2-(z')^2 &= \gamma^2(c^2t^2-2\beta ctx+\beta^2x^2)+\gamma^2(\beta^2c^2t^2+x^2-2c\beta tx)-y^2-z^2 \\
&= \gamma^2(c^2(1-\beta^2)t^2-(1-\beta^2)x^2)-y^2-z^2 \\
&= c^2t^2-x^2-y^2-z^2 \ , \qquad (1.2.4)
\end{aligned}$$

where we have used $\gamma^2(1-\beta^2)=1$.

This explains why the speed of light is constant in two frames K and K'. Assume that a spherical light wave starts at $t=0$ expanding from the origin $x=y=z=0$. Then for $t>0$ the equation $c^2t^2-x^2-y^2-z^2$ describes a spherical wave front with radius $r=ct$ i.e. it is expanding with the speed of light. Obviously, using the identity derived above this holds true also in the frame K' and actually for any other frame, hence the speed of light is consant.

It is very useful to think of Lorentz Transformations (LTs) as generalised rotations. In order to see this remember $\gamma^2-\gamma^2\beta^2=1=\gamma^2-(\beta\gamma)^2$ and compare this with the identity $\cosh^2\omega-\sinh^2\omega=1$. Hence, we can set

$$\gamma\equiv\cosh\omega \ , \ \beta\gamma\equiv\sinh\omega\,, \qquad (1.2.5)$$

and the boost in the x direction can be written as

$$\begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (1.2.6)$$

where similar expressions exist for boosts in the y and z directions. Compare this now with a spatial rotation around, say, the z axis by an angle θ

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (1.2.7)$$

where $\cos^2\theta+\sin^2\theta=1$. Note that under a spatial rotation also $x^2+y^2+z^2$ is invariant. Also there are similar transformation matrices for rotations around the x and y axis. This implies that the *Lorentz group* is six dimensional: there are 3 proper LTs (boost) and 3 spatial rotations. Why is this a group? Because any combination of two LTs which corresponds to matrix multiplication of the transformation matrices produces another LT.

1.2.2 4-Vector Notation and Tensors

Recall that $c^2t^2-x^2-y^2-z^2$ is rather similar to the dot-product of 3-vectors $\vec{x}\cdot\vec{x}=x^2+y^2+z^2$ (which is invariant under spatial rotations) except for the *funny* minus signs.

In order to deal with this one introduces a metric tensor η, which is a 4-by-4 matrix, and two types of 4-vectors, covariant ones and contravariant ones.

We can write the relativistic inner product $c^2t^2 - x^2 - y^2 - x^2$ as

$$(ct, x, y, z) \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}_{\equiv \eta} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.2.8}$$

Another way to deal with this is to introduce two types of vectors

$$\begin{aligned} x^\mu &\equiv (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \vec{x}) && \text{contravariant} \\ x_\mu &\equiv (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) = (ct, -\vec{x}) && \text{covariant} \end{aligned} \tag{1.2.9}$$

so that

$$\sum_{\mu=0}^{3} x^\mu x_\mu = c^2t^2 - x^2 - y^2 - x^2 \equiv x \cdot x\,. \tag{1.2.10}$$

Also if we introduce the Minkowski metric $\eta^{\mu\nu}$ then

$$\sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \eta^{\mu\nu} x_\nu x_\mu = c^2t^2 - x^2 - y^2 - x^2 \equiv x \cdot x\,, \tag{1.2.11}$$

because

$$\sum_{\nu=0}^{3} \eta^{\mu\nu} x_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = x^\mu\,. \tag{1.2.12}$$

The metric tensor $\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ and the inverse metric $\eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ are used to *raise* and *lower* indices:

$$\begin{aligned} \sum_{\nu=0}^{3} \eta^{\mu\nu} x_\nu &= x^\mu\,, \\ \sum_{\nu=0}^{3} \eta_{\mu\nu} x^\nu &= x_\mu\,. \end{aligned} \tag{1.2.13}$$

Note also that the metric is symmetric $\eta_{\mu\nu} = \eta_{\nu\mu}$ and

$$\sum_{\alpha=0}^{3} \eta_{\mu\alpha}\eta^{\alpha\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv \delta_\mu{}^\nu\,. \tag{1.2.14}$$

Finally, one introduces the Einstein summation convention in order to avoid writing the boring sums over and over again: repeated upper and lower index pairs are automatically summed over (unless stated otherwise). Note that this only makes sense if one index is upper and the other is lower, otherwise one gets expressions that do not transform properly under LTs.

Examples

$$c^2t^2 - x^2 - y^2 - z^2 = x^\mu x_\mu = x_\mu x^\mu = \eta^{\mu\nu} x_\mu x_\nu \tag{1.2.15}$$

$$x^\mu = \eta^{\mu\nu} x_\nu \tag{1.2.16}$$

$$x_\mu = \eta_{\mu\nu} x^\nu \tag{1.2.17}$$

1.2.3 Lorentz transformations, again

Now LTs can be written very compactly as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \,, \tag{1.2.18}$$

where Λ can be any LT matrix.

This was for a contravariant vector, what about a covariant one?

$$x'_\alpha = \eta_{\alpha\mu} x'\mu = \eta_{\alpha\mu} \Lambda^\mu{}_\rho x^\rho = \eta_{\alpha\mu} \Lambda^\mu{}_\rho \eta^{\rho\beta} x_\beta \equiv \Lambda_\alpha{}^\beta x_\beta \tag{1.2.19}$$

Therefore,

$$x'_\alpha = \Lambda_\alpha{}^\beta x_\beta \;\text{ with }\; \Lambda_\alpha{}^\beta = \eta_{\alpha\mu} \Lambda^\mu{}_\nu \eta^{\nu\beta} \,. \tag{1.2.20}$$

Now the Lorentz transformations are those Λ's that leave $x^\mu x_\mu$ invariant. Note that a priori Λ has 16 components which should be reduced to 6. Let's see how this works. We require that the norm of a 4-vector x^μ is Lorentz invariant. So we have

$$\eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta x^\alpha x^\beta = \eta_{\alpha\beta} x^\alpha x^\beta \,, \tag{1.2.21}$$

which implies

$$\eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = \eta_{\alpha\beta} \,. \tag{1.2.22}$$

This is a matrix equation for Λ with 16 components. But only 10 components are independent because both sides of the equation are symmetric under exchange of α and β. Hence the number of independent components of Λ is reduced to $16 - 10 = 6$. This is just the right number to incorporate 3 boosts and 3 spatial rotations!

General tensors with n contravariant indices and m covariant indices transform as follows

$$T'^{\mu_1 \ldots \mu_n}{}_{\nu_1 \ldots \nu_m} = \Lambda^{\mu_1}{}_{\alpha_1} \ldots \Lambda^{\mu_n}{}_{\alpha_n} \Lambda_{\nu_1}{}^{\beta_1} \ldots \Lambda_{\nu_m}{}^{\beta_m} T^{\alpha_1 \ldots \alpha_n}{}_{\beta_1 \ldots \beta_m} \,. \tag{1.2.23}$$

This is the general transformation law for tensors that transform covariantly, however we usually do not account tensors with more than two indices (unless we consider General Relativity).

It is also important to know how differentials and partial derivatives transform under LTs.

From

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{1.2.24}$$

we learn that

$$dx'^\mu = \Lambda^\mu{}_\nu dx^\nu\,, \tag{1.2.25}$$

so dx'^μ transforms like a contra-variant vector. It is also known that for general coordinate transformations we have

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu\,, \tag{1.2.26}$$

where for the special case of LTs $\frac{\partial x'^\mu}{\partial x^\nu} = \Lambda^\mu{}_\nu$.

Furthermore, application of the chain rule yields

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = \left(\frac{\partial x'^\mu}{\partial x^\nu}\right)^{-1}\frac{\partial}{\partial x^\nu}\,. \tag{1.2.27}$$

Hence, partial derivatives $\frac{\partial}{\partial x^\mu}$ transform inversely to dx^μ i.e. they transform like a co-variant vector (index downstairs). It is common to introduce the co-variant 4-vector

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \vec{\nabla}\right) \tag{1.2.28}$$

and the contra-variant vector

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial t}, -\vec{\nabla}\right)\,. \tag{1.2.29}$$

Important Lorentz invariant contractions:

- The d'Alembertian operator or "box"

$$\partial_\mu \partial^\mu \equiv \partial \cdot \partial = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 = \Box \tag{1.2.30}$$

  is a relativistic wave operator.

- The energy momentum 4-vector is $p^\mu = (E/c, \vec{p})$ from which we can form the invariant $p^\mu p_\mu = p \cdot p = p^2 = \frac{E^2}{c^2} - \vec{p}^2 = m^2c^2$. Another important invariant is $p \cdot x = p^\mu x_\mu = Et - \vec{p}\cdot\vec{x}$. Note that $\Psi = \exp(\pm ip \cdot x)$ is a plane wave solution to the wave equation $\Box\Psi = 0$ if $p^2 = 0$.

- Note that any dot product of two 4-vectors is a Lorentz invariant $a \cdot b = a^\mu b_\mu = \eta^{\mu\nu} a_\mu b_\nu$.

1.2.4 Lorentz Group

We will set $\hbar = c = 1$ from now on.

Let us think a moment about infinitesimal LTs $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu$, where $\omega^\mu{}_\nu$ is small and obeys $\omega^{\mu\nu} = -\omega^{\nu\mu}$.

For a boost in the x-direction

$$\begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \sim \mathbb{I}_4 - i\omega K_x\,, \text{ with } K_x = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{1.2.31}$$

where we call K_x the generator of the boost in the x-direction. (Check for yourself that $\exp(-i\omega K_x)$ gives back the finite Lorentz transformation.) Similarly we find the generators for boosts in the y and z direction.

$$K_y = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; K_z = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}. \tag{1.2.32}$$

For a rotation around the x-axis

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\phi & -\sin\phi \\ 0 & 0 & \sin\phi & \cos\phi \end{pmatrix} \sim \mathbb{I}_4 - i\phi J_x\,, \text{ with } J_x = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \tag{1.2.33}$$

where J_x is the generator for rotations around the x axis. Similarly we find the generators for rotations around the y and z axis.

$$J_y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \; J_z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.2.34}$$

The boost and rotation generators obey the following set of commutation relations.

$$\begin{aligned} [J_x, J_y] &= iJ_z \text{ ,+cyclic permutations} \\ [K_x, K_y] &= -iJ_z \text{ ,+cyclic permutations} \\ [J_x, K_x] &= [J_y, K_y] = [J_z, K_z] = 0 \\ [J_x, K_y] &= iK_z \text{ ,+cyclic permutations} \end{aligned} \tag{1.2.35}$$

Note that the first line is just the standard $SU(2)$ algebra for angular momentum operators.

This commutator algebra reveals a more interesting structure if we introduce the linear combinations

$$\begin{aligned}\vec{N} &= \frac{1}{2}\left(\vec{J}+i\vec{K}\right) \\ \vec{M} &= \frac{1}{2}\left(\vec{J}-i\vec{K}\right) .\end{aligned} \tag{1.2.36}$$

Note that these are two 3-vectors where each component of the 3-vectors is a 4-by-4 matrix. These two sets of matrix generators obey the following commutator algebra

$$\begin{aligned}[N_x, M_x] &= [N_x, M_y] = [N_x, M_z] = \ldots = 0 \\ [N_x, N_y] &= iN_z \text{ , +cyclic permutations} \\ [M_x, M_y] &= iM_z \text{ , +cyclic permutations} ,\end{aligned} \tag{1.2.37}$$

which are two commuting sets of $SU(2)$ algebras.

Until now we have considered a particular representation of the operators $K_{x,y,z}$ and $J_{x,y,z}$ in terms of 4-by-4 matrices which are the generators of boosts and rotations acting on 4-vectors (and tensors by multiple action of Λs). But this algebra could be represented by differential operators (as we are used to from QM) or different sets of 4-by-4 matrices (as we will find later for the spinors which are spin 1/2 particles). All these have one thing in common: they obey the same algebra. So let us now consider the operators abstractly and assume that the only thing we know about them is that they obey the Lorentz algebra (1.2.37).

Then we know on general grounds that $\vec{N}^2$ and $\vec{M}^2$ have eigenvalues $j_N(j_N+1)$ and $j_M(j_M+1)$ respectively, where j_N, j_M can be integer or half-integer. This statement is true for any set of matrices or operators that obey the same algebra as above and is not restricted to our particular choice of generators for boost and rotations. Representations (particles) are labelled by (j_N, j_M), where the spin of the particle is $j_N + j_M$.

Important examples include:

- $(0,0)$: spin zero, scalar particle
- $(\frac{1}{2},0)$: spin $\frac{1}{2}$, left handed fermion
- $(0,\frac{1}{2})$: spin $\frac{1}{2}$, right handed fermion
- $(\frac{1}{2},\frac{1}{2})$: spin 1, transforms like a 4-vector, e.g. photon, gluon, W and Z bosons

1.2.5 Poincaré Group

This is a bonus section and will NOT be part of the exam. A nice exposition of the material in this and the previous section can be found in the book by Ryder [4].

So far we have only considered rotations and *generalised* rotations i.e. boosts, but you might have wondered what has happened to translations in space and time that we discussed in the context of non-relativistic QM. We will remedy this situation now and introduce the Poincaré Group which is the natural extension of the Lorentz group including translations in space-time:

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu} \tag{1.2.38}$$

The first part of this transformation are just the LTs, which are generated by $J_{x,y,z}$ and $K_{x,y,z}$ and the second, inhomogenous, part are the translations generated by the 4-momentum p^{μ}.

In QM we are used to expressions for $\widehat{J}_{x,y,z}$ or $\widehat{\vec{p}}$ in terms of differential operators. For example $\vec{p} \to \widehat{\vec{p}} = -i\vec{\nabla}$ and $E \to i\partial/\partial t$ which can be combined into the 4-vector

$$p^{\mu} \to \widehat{p}^{\mu} = i\nabla^{\mu} \, . \tag{1.2.39}$$

The QM angular momentum operators are $\widehat{J}_x = -i\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$, $\widehat{J}_y = -i\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$, $\widehat{J}_z = -i\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$. Similar operators exist for the boosts (however they are not hermitian): $\widehat{K}_x = i\left(t\frac{\partial}{\partial x} + x\frac{\partial}{\partial t}\right)$, $\widehat{K}_y = i\left(t\frac{\partial}{\partial y} + y\frac{\partial}{\partial t}\right)$, $\widehat{K}_z = i\left(t\frac{\partial}{\partial z} + z\frac{\partial}{\partial t}\right)$.

Notably, this set of operators obeys the same algebra (1.2.35) that we found in the last section for the generators $J_{x,y,z}$ and $K_{x,y,z}$ in matrix form.

These differential operators are defined through their action on functions, while the matrices act directly on the coordinates which appear as arguments of spacetime dependent functions. So these two viewpoints are naturally related, e.g. consider an infinitesimal boost in the x-direction. Recalling

$$\begin{aligned} \Lambda^{\mu}{}_{\nu} &= \delta^{\mu}{}_{\nu} + i\omega_x (K_x)^{\mu}{}_{\nu} \\ x'^{\mu} &= x^{\mu} + i\omega_x (K_x)^{\mu}{}_{\nu} x^{\nu} \, , \end{aligned} \tag{1.2.40}$$

it is easy to see that (infinitesimally)

$$i\omega_x \widehat{K}_x f(x) = f(x') - f(x) \, , \tag{1.2.41}$$

where the argument x of $f(x)$ stands for a point in spacetime i.e. a 4-vector.

Another common form of the algebra of the $K's$ and $J's$ is written in terms of the anti-symmetric tensor $\widehat{M}_{\mu\nu}$ which is defined as follows:

$$\begin{aligned}\widehat{M}_{ij} &= -\widehat{M}_{ji} = \epsilon_{ijk}\widehat{J}_k \\ \widehat{M}_{i0} &= -\widehat{M}_{0i} = -\widehat{K}_i \end{aligned} \tag{1.2.42}$$

where $i, j, k = 1 \ldots 3$. Now the Lorentz algebra becomes

$$\left[\widehat{M}_{\mu\nu}, \widehat{M}_{\rho\sigma}\right] = i\left(\eta_{\nu\rho}\widehat{M}_{\mu\sigma} - \eta_{\mu\rho}\widehat{M}_{\nu\sigma} + \eta_{\mu\sigma}\widehat{M}_{\nu\rho} - \eta_{\nu\sigma}\widehat{M}_{\mu\rho}\right) . \tag{1.2.43}$$

Including translations generated by $\widehat{p}_\mu = i\partial_\mu$ one finds the additional commutation relations

$$[\widehat{p}_\mu, \widehat{p}_\nu] = 0 , \tag{1.2.44}$$

$$\left[\widehat{p}_\mu, \widehat{M}_{\rho\sigma}\right] = i\left(\eta_{\mu\rho}\widehat{p}_\sigma - \eta_{\mu\sigma}\widehat{p}_\rho\right) . \tag{1.2.45}$$

This is the Poincaré algebra (the Poincaré group is obtained by exponentiating the generators). A mathematical problem with the Lorentz group is that it does not have finite dimensional unitary representations; the representations we listed in the last section are finite dimensional but non-unitary! The problem comes from the boosts, whose generators are not Hermitian and hence their exponentiation gives non-unitary matrices or operators. But in QM we seek unitary finite dimensional representations and here the Poincaré group comes to the rescue, since it does have finite dimensional unitary representations corresponding to particles with various spins.

This result allows us to answer with mathematical rigour the following fundamental question: What is a particle?
The mathematicians answer: An irreducible representation of the Poincaré group.

In order to classify representations we have to look for operators that commute with all operators of the Poincaré group. (This is the analogue to the operator $\widehat{\vec{J}}^2$ for the angular momentum operators.) It turns out that there exist exactly two such operators, which are also called Casimir operators of the Poincaré algebra:

$$C_1 = \widehat{p}^\mu \widehat{p}_\mu , \tag{1.2.46}$$

which for a momentum eigenstate becomes $p^2 = E^2 - \vec{p}^2 = m^2$, so this is just the mass squared. Now interestingly $\widehat{\vec{J}}^2$ is not a Casimir since it does not commute with boosts $\widehat{\vec{K}}$. To constructe the second Casimir one introduces the Pauli-Lubanski vector $\widehat{W}_\mu$

$$\begin{aligned}\widehat{W}_\mu &= -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\widehat{M}^{\nu\rho}\widehat{P}^\sigma \\ \widehat{W}_\mu \widehat{P}^\mu &= 0 , \end{aligned} \tag{1.2.47}$$

so that

$$C_2 = \widehat{W}_\mu \widehat{W}^\mu . \tag{1.2.48}$$

As an example consider a massive, spin s particle in its rest frame where $p^\mu = (m, \vec{0})$. Now in the rest frame the orbital angular momentum $\vec{L} = 0$ so that the total angular momentum $\vec{J} = \vec{L} + \vec{S} = \vec{S}$. For W_μ we find

$$\begin{aligned} W_0 &= -\frac{1}{2}\epsilon_{0ijk}J^{ij}P^k = 0 \\ W_i &= -\frac{1}{2}\epsilon_{i\mu\nu\rho}J^{\mu\nu}P^\rho = -\frac{m}{2}\epsilon_{i\mu\nu 0}J^{\mu\nu} = -\frac{m}{2}\epsilon_{ijk}J^{jk} = -mJ_i \end{aligned} \tag{1.2.49}$$

and, hence,

$$C_2 = W_\mu W^\mu = (W_0)^2 - \vec{W}^2 = -m^2\vec{J}^2 = -m^2\vec{S}^2 = -m^2 s(s+1)\,. \tag{1.2.50}$$

So we find that a relativistic particle is characterised by its mass m and its spin s.

1.2.6 The Maxwell Equation — First Glimpse at a Relativistic Wave Equation

Remember Maxwell's equations (where again $\hbar = c = 1$)

$$\begin{aligned} \vec{\nabla}\cdot\vec{E} &= \rho \\ \vec{\nabla}\cdot\vec{B} &= 0 \\ \vec{\nabla}\times\vec{E} &= -\frac{\partial\vec{B}}{\partial t} \\ \vec{\nabla}\times\vec{B} &= \frac{\partial\vec{E}}{\partial t} + \vec{j} \end{aligned} \tag{1.2.51}$$

This does not look manisfestly Lorentz covariant since it is written in terms of 3-vectors $\vec{E}$, $\vec{B}$ and $\vec{j}$.

Let us introduce the 4-vector potential $A^\mu = (\phi, \vec{A})$ in terms of which $\vec{B} = \vec{\nabla}\times\vec{A}$ and $\vec{E} = -\frac{\partial\vec{A}}{\partial t} - \vec{\nabla}\phi$. This implies the 2nd and 3rd Maxwell equation since

$$\begin{gathered} \vec{\nabla}\cdot\vec{B} = \vec{\nabla}\cdot(\vec{\nabla}\times\vec{A}) = 0\,, \\ \vec{\nabla}\times(-\frac{\partial\vec{A}}{\partial t} - \vec{\nabla}\phi) = -\vec{\nabla}\times\frac{\partial\vec{A}}{\partial t} = -\frac{\partial\vec{B}}{\partial t}\,. \end{gathered} \tag{1.2.52}$$

Now we define the anti-symmetric electro-magnetic fieldstrength tensor $F^{\mu\nu}$ (remember $\partial^\mu = (\frac{\partial}{\partial t}, -\vec{\nabla})$):

$$\begin{aligned} F^{\mu\nu} &= -F^{\nu\mu} = \partial^\mu A^\nu - \partial^\nu A^\mu \\ F^{0i} &= \partial^0 A^i - \partial^i A^0 = (\frac{\partial\vec{A}}{\partial t} + \vec{\nabla}\phi)_i = -E_i \\ F^{ij} &= \partial^i A^j - \partial^j A^i = -\epsilon^{ijk}B_k\,. \end{aligned} \tag{1.2.53}$$

So

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix} \tag{1.2.54}$$

Under Lorentz transformations

$$F^{\mu\nu} \to F'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta} \tag{1.2.55}$$

which can also be written as a matrix equation $F \to F' = \Lambda \cdot F \cdot \Lambda^T$. The second and the third Maxwell equation are encoded in the so-called Bianchi identity

$$\partial^\mu F^{\nu\rho} + \partial^\nu F^{\rho\mu} + \partial^\rho F^{\mu\nu} = 0, \tag{1.2.56}$$

while the first and the fourth Maxwell equation are encoded in the equation

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{1.2.57}$$

with $j^\nu = (\rho, \vec{j})$. This is an example of an equation that is forminvariant under Lorentz transformations, i.e. it transforms covariantly under LTs.

An important feature of the electromagnetic fieldstrength tensor $F^{\mu\nu}$ is that it is invariant under *gauge transformations*. Gauge transformations are redefinition of A^μ by a total derivative

$$A^\mu \to A^\mu + \partial^\mu \chi(x)\,. \tag{1.2.58}$$

Proof: under such a gauge transformation

$$\begin{aligned} F^{\mu\nu} \quad \to \quad & \partial^\mu(A^\nu + \partial^\nu\chi(x)) - \partial^\nu(A^\mu + \partial^\mu\chi(x)) = \\ & F^{\mu\nu} + (\partial^\mu\partial^\nu - \partial^\nu\partial^\mu)\chi(x) = F^{\mu\nu}\,. \end{aligned} \tag{1.2.59}$$

This gauge invariance implies that out of the four degrees of freedom (components) of A^μ one is redundant and we can impose a convenient gauge condition to eliminate this redundancy. A popular choice is the so-called Lorentz gauge $\partial \cdot A = \partial_\mu A^\mu = 0$ because it is a Lorentz invariant condition.

In this gauge

$$\partial_\mu F^{\mu\nu} = \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = j^\nu \tag{1.2.60}$$

becomes

$$\Box A^\nu = j^\nu\,. \tag{1.2.61}$$

Furthermore, for $j^\nu = 0$ (in vacuo) we obtain

$$\Box A^\nu\,, \tag{1.2.62}$$

which is a relativistic wave equation for massless spin 1 particles called photons. The equation implies in particular that the waves/photons are travelling at the speed of light as any massless particle does (independent of its spin).

Chapter 2

Relativistic Quantum Mechanics

2.1 The Klein-Gordon Equation

In non-relativistic QM, the free Hamiltonian $H = E = \frac{\vec{p}^2}{2m}$ is quantised by the substitution

$$H \to i\hbar\frac{\partial}{\partial t} \ , \ \ \vec{p} \to -i\hbar\vec{\nabla} \tag{2.1.1}$$

to give the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\vec{\nabla}^2\Psi \, . \tag{2.1.2}$$

A relativistic free particle has Hamiltonian

$$H = E = \sqrt{\vec{p}^2c^2 + m^2c^4} \tag{2.1.3}$$

and hence the same, naive substitution gives

$$i\hbar\frac{\partial\Psi}{\partial t} = \sqrt{m^2c^4 - \hbar^2c^2\vec{\nabla}^2}\Psi \, . \tag{2.1.4}$$

But what to do about the square root of the operator? One interpretation is to make a series expansion, but then we get a Hamiltonian with derivatives of arbitrarily high order.

A more sensible route is to start from $H^2 = \vec{p}^2c^2 + m^2c^4$ to get

$$\begin{aligned}
-\hbar^2\frac{\partial^2}{\partial t^2}\Psi &= \left(-\hbar^2c^2\vec{\nabla}^2 + m^2c^4\right)\Psi \\
&\Rightarrow \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2\right)\Psi + \left(\frac{mc}{\hbar}\right)^2\Psi = 0 \\
&\Rightarrow \left(\Box + \left(\frac{mc}{\hbar}\right)^2\right)\Psi = 0
\end{aligned} \tag{2.1.5}$$

By analogy with the Schrödinger equation it is possible to derive a continuity equation for the Klein-Gordon (KG) equation

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \tag{2.1.6}$$

where ρ is the probability density and $\vec{J}$ is the probability current.[1]

Derivation: Subtracting the following two equations

$$\begin{aligned} \Psi^* \left(\Box + \left(\frac{mc}{\hbar} \right)^2 \right) \Psi &= 0 \\ \Psi \left(\Box + \left(\frac{mc}{\hbar} \right)^2 \right) \Psi^* &= 0 \end{aligned} \tag{2.1.7}$$

gives

$$\Psi^* \partial_\mu \partial^\mu \Psi - \Psi \partial_\mu \partial^\mu \Psi^* = 0 \, . \tag{2.1.8}$$

Furthermore, this implies

$$\begin{aligned} \partial_\mu \left(\Psi^* \partial^\mu \Psi - \Psi \partial^\mu \Psi^* \right) &= 0 \\ \Rightarrow \frac{\partial}{\partial t} \left(\frac{1}{c^2} \left(\Psi^* \partial_t \Psi - \Psi \partial_t \Psi^* \right) \right) - \vec{\nabla} \cdot \left(\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^* \right) &= 0 \, . \end{aligned} \tag{2.1.9}$$

After multiplying through by ic^2, in order to make ρ real, we can write this as desired as a continuity equation with

$$\rho = i \left(\Psi^* \partial_t \Psi - \Psi \partial_t \Psi^* \right) \;\text{ and }\; \vec{J} = -ic^2 \left(\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^* \right) \, . \tag{2.1.10}$$

Now because of the negative sign between the two terms in ρ, the probability density can both take positive and negative values (in contrast to non-relativistic QM where $\rho = |\Psi|^2$ is positive definite)! This is an absurd and nonsensical result for a probability density!!

Junk the KG equation for the moment and try harder. Schrödinger was the first to write down this relativistic wave equation, but discarded it for a different reason; the spectrum is not bounded from below.

2.2 The Dirac Equation

Let us go back to our starting point

$$\widehat{H} \Psi = i\hbar \frac{\partial \Psi}{\partial t} \tag{2.2.1}$$

[1] Integrating the probability density over a volume V bounded by the surface S we find $\int_V \frac{\partial \rho}{\partial t} d^3x = \frac{d}{dt} \int_V \rho d^3x = -\int_V \vec{\nabla} \cdot \vec{j} d^3x = -\int_S \vec{j} \cdot d\vec{S}$. This implies that probability cannot be created or destroyed; it can only flow from one point to another.

and try to give a meaning to the square root in

$$\widehat{H} = \sqrt{m^2c^4 + \widehat{\vec{p}}c^2} = \sqrt{m^2c^4 - \hbar^2c^2\vec{\nabla}^2}\,. \tag{2.2.2}$$

Since the equation is *linear* in $\frac{\partial}{\partial t}$, Lorentz covariance suggests it should be linear also in the $\frac{\partial}{\partial x^i}$, $i = 1, 2, 3$.

So we write

$$\begin{aligned} \widehat{H} &= c\vec{\alpha}\cdot\widehat{\vec{p}} + \beta mc^2 \\ &= -i\hbar c\vec{\alpha}\cdot\vec{\nabla} + \beta mc^2\,, \end{aligned} \tag{2.2.3}$$

where α^i and β are coefficients to be determined. More explicitly this equation can be written

$$\widehat{H} = -i\hbar c\left(\alpha^1\frac{\partial}{\partial x^1} + \alpha^2\frac{\partial}{\partial x^2} + \alpha^3\frac{\partial}{\partial x^3}\right) + \beta mc^2\,. \tag{2.2.4}$$

Now we determine the coefficients α^i and β by requiring that this linear operator "squares" to the KG operator

$$\widehat{H}^2 = -\hbar^2c^2\vec{\nabla}^2 + m^2c^4\,. \tag{2.2.5}$$

We find

$$\begin{aligned} \widehat{H}^2 \;=\; & -\hbar^2c^2\left((\alpha^1)^2\frac{\partial^2}{\partial(x^1)^2} + (\alpha^2)^2\frac{\partial^2}{\partial(x^2)^2} + (\alpha^3)^2\frac{\partial^2}{\partial(x^3)^2}\right) + \beta^2m^2c^4 \\ & -i\hbar mc^3\left((\alpha^1\beta + \beta\alpha^1)\frac{\partial}{\partial x^1} + \ldots\right) \\ & -\hbar^2c^2\left((\alpha^1\alpha^2 + \alpha^2\alpha^1)\frac{\partial^2}{\partial x^1\partial x^2} + \ldots\right)\,. \end{aligned} \tag{2.2.6}$$

Thus we need to solve

$$\begin{aligned} &(\alpha^i)^2 = \mathbb{I}\,,\ i = 1, 2, 3 \\ &\alpha^i\alpha^j + \alpha^j\alpha^i = 0\,,\ i \neq j \\ &\beta^2 = \mathbb{I} \\ &\alpha^i\beta + \beta\alpha^i = 0\,,\ i = 1, 2, 3\,, \end{aligned} \tag{2.2.7}$$

where $\mathbb{I}$ denotes a unit matrix (if a subscript is added it denotes the dimensionality, e.g. $\mathbb{I}_2$ denotes a 2×2 unit matrix).

It is obviously NOT possible to solve those equations if the coefficients are simply complex numbers. So let us assume that they are $N \times N$ matrices. With some (guess)work it can be shown that the smallest value of N for which eq. (2.2.7) can be solved is $N = 4$.

This implies that the Dirac wave function is a 4-component column vector

$$\Psi(x) = \begin{pmatrix} \Psi_1(x) \\ \Psi_2(x) \\ \Psi_3(x) \\ \Psi_4(x) \end{pmatrix} \tag{2.2.8}$$

where $x \equiv (x^0, \vec{x})$ and the Dirac equation becomes a matrix equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \widehat{H}\Psi = (c\vec{\alpha} \cdot \widehat{\vec{p}} + \beta mc^2)\Psi = (-i\hbar c\vec{\alpha} \cdot \vec{\nabla} + \beta mc^2)\Psi \,. \tag{2.2.9}$$

This is a set of 4 first order linear differential equations to determine $\Psi_1, \ldots, \Psi_4$.

2.3 Representation of the Dirac Matrices

A particular set of solutions of (2.2.7) for the 4×4 matrices α^i, β can be written with the help of the 2×2 Pauli matrices σ^i, $i = 1, 2, 3$,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;,\; \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \;,\; \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.3.1}$$

which obey the following identities

$$\begin{aligned} &\left(\sigma^i\right)^2 = \mathbb{I}_2 \,,\; i = 1, 2, 3 \\ &\sigma^i \sigma^j + \sigma^j \sigma^i = 0 \,,\; i \neq j \,. \end{aligned} \tag{2.3.2}$$

We may satisfy the first two lines in eq. (2.2.7) by taking α^i to be the 4×4 matrices

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} \;,\; i = 1, 2, 3 \,. \tag{2.3.3}$$

Now we may satisfy the remaining two lines in eq. (2.2.7) by taking

$$\beta = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix} \,. \tag{2.3.4}$$

Because the σ^i are Hermitian, so are the α^i and β, *i.e.*

$$(\alpha^i)^\dagger = \alpha^i \;,\; \beta^\dagger = \beta \,. \tag{2.3.5}$$

($\dagger \leftrightarrow$ complex conjugate transposed)

2.4 Probability Density for the Dirac Equation

The Dirac equation is given by

$$i\hbar\frac{\partial\Psi}{\partial t} = -i\hbar\vec{\alpha}\cdot\vec{\nabla}\Psi + mc^2\beta\Psi \tag{2.4.1}$$

and its Hermitian conjugate is

$$-i\hbar\frac{\partial\Psi^\dagger}{\partial t} = i\hbar\vec{\nabla}\Psi^\dagger\cdot\vec{\alpha} + mc^2\Psi^\dagger\beta\,. \tag{2.4.2}$$

(Recall that α^i and β are hermitian and $(AB)^\dagger = B^\dagger A^\dagger$.)

Now take $\Psi^\dagger\times$(Dirac eqn.) and (Hermitian conjugate eqn.)$\times\Psi$ and subtract the two to obtain:

$$i\hbar\left(\Psi^\dagger\frac{\partial\Psi}{\partial t} + \frac{\partial(\Psi^\dagger)}{\partial t}\Psi\right) = -i\hbar c\left(\Psi^\dagger\vec{\alpha}\cdot\vec{\nabla}\Psi + \vec{\nabla}(\Psi^\dagger)\cdot\vec{\alpha}\Psi\right)\,. \tag{2.4.3}$$

Dividing this equation by $i\hbar$ we obtain a Continuity Equation

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0 \tag{2.4.4}$$

with the *positive definite probability density* given by

$$\rho = \Psi^\dagger\Psi = \sum_{k=1}^{4}\Psi_k^*\Psi_k = \sum_{k=1}^{4}|\Psi_k|^2 > 0\,, \tag{2.4.5}$$

and the probability current

$$\vec{J} = c\Psi^\dagger\vec{\alpha}\Psi\,. \tag{2.4.6}$$

2.5 Extreme Non-Relativistic Limit of the Dirac Equation

For a particle at rest ($\vec{p} = 0$) the Dirac equation becomes

$$\begin{aligned} i\hbar\frac{\partial\Psi}{\partial t} &= \beta mc^2\Psi \\ \Rightarrow \frac{\partial\Psi}{\partial t} &= -i\frac{mc^2}{\hbar}\beta\Psi\,. \end{aligned} \tag{2.5.1}$$

Taking $\beta = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix}$ the four equations for the components of Ψ turn into

$$\begin{aligned} \frac{\partial \Psi_1}{\partial t} &= -i\frac{mc^2}{\hbar}\Psi_1 \\ \frac{\partial \Psi_2}{\partial t} &= -i\frac{mc^2}{\hbar}\Psi_2 \\ \frac{\partial \Psi_3}{\partial t} &= +i\frac{mc^2}{\hbar}\Psi_3 \\ \frac{\partial \Psi_4}{\partial t} &= +i\frac{mc^2}{\hbar}\Psi_4 \end{aligned} \tag{2.5.2}$$

$$\Rightarrow \Psi_1 = c_1 e^{-i\frac{mc^2}{\hbar}t} \text{ e.t.c.} \tag{2.5.3}$$

where c_1 is an arbitrary constant.

Thus the general solution takes the form

$$\Psi = \begin{pmatrix} c_1 e^{-i\frac{mc^2}{\hbar}t} \\ c_2 e^{-i\frac{mc^2}{\hbar}t} \\ c_3 e^{i\frac{mc^2}{\hbar}t} \\ c_4 e^{i\frac{mc^2}{\hbar}t} \end{pmatrix} \tag{2.5.4}$$

which can be rewritten as

$$\Psi = e^{-i\frac{mc^2}{\hbar}t}\begin{pmatrix} c_1 \\ c_2 \\ 0 \\ 0 \end{pmatrix} + e^{i\frac{mc^2}{\hbar}t}\begin{pmatrix} 0 \\ 0 \\ c_3 \\ c_4 \end{pmatrix} \tag{2.5.5}$$

By acting with the Hamiltonian operator $\widehat{H} = i\hbar\frac{\partial}{\partial t}$ we find that the first term in the solution (2.5.5) carries *positive energy* ($+mc^2$) whereas the second term carries *negative energy* ($-mc^2$).

Although we found a positive probability density (contrary to the KG equation) we find that also the Dirac equation has both positive and negative energy solutions. We shall later interpret the negative energy part as due to *Anti-particles*.

2.6 Spin of the Dirac Particles

The (free) Dirac equation is

$$i\hbar\partial_t\Psi = \widehat{H}\Psi \;\text{ with }\; \widehat{H} = c\vec{\alpha}\cdot\widehat{\vec{p}} + \beta mc^2 \tag{2.6.1}$$

with $\widehat{\vec{p}} = -i\hbar\vec{\nabla}$.

Consider the total angular momentum operator

$$\vec{J} = \vec{L} + \vec{S} = \vec{x} \times \widehat{\vec{p}} + \frac{\hbar}{2}\vec{\Sigma} \text{ where } \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \tag{2.6.2}$$

Using the Uncertainty Principle

$$[x^i, \widehat{p}^j] = i\hbar\delta_{ij} \,, \tag{2.6.3}$$

it can be shown with some effort that

$$[\widehat{H}, \vec{S}] = i\hbar c(\vec{\alpha} \times \widehat{\vec{p}}) \tag{2.6.4}$$

and

$$[\widehat{H}, \vec{L}] = -i\hbar c(\vec{\alpha} \times \widehat{\vec{p}}) \,, \tag{2.6.5}$$

thus, $[\widehat{H}, \widehat{\vec{J}}]$=0. So $\vec{J}$ is a conserved quantity which we interpret as the total angular momentum.

The 3-component of Spin, S_z, is

$$S^3 = \frac{\hbar}{2}\Sigma^3 \tag{2.6.6}$$

is the matrix $\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ which has eigenvalues $\frac{\hbar}{2}$, $-\frac{\hbar}{2}$, $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$.

$\Rightarrow$ we are describing spin $\frac{1}{2}$ particles (and anti-particles).

2.7 The Covariant Form of the Dirac Equation

Multiply the Dirac equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -i\hbar c\sum_{i=1}^{3}\alpha^i\frac{\partial\Psi}{\partial x^i} + \beta mc^2\Psi \tag{2.7.1}$$

with $\frac{\beta}{c}$ to obtain

$$i\hbar\left(\beta\frac{\partial\Psi}{\partial(ct)} + \sum_{i=1}^{3}\beta\alpha^i\frac{\partial\Psi}{\partial x^i}\right) = mc\Psi \,. \tag{2.7.2}$$

Next define the matrices

$$\gamma^0 = \beta\,,\ \gamma^i = \beta\alpha^i\,,\ i = 1,2,3 \tag{2.7.3}$$

which are also called *Gamma-Matrices* and allow us to rewrite the Dirac equation as

$$i\hbar\left(\gamma^0\frac{\partial\Psi}{\partial x^0} + \sum_{i=1}^{3}\gamma^i\frac{\partial\Psi}{\partial x^i}\right)\Psi = mc\Psi \tag{2.7.4}$$

$$\to i\hbar\gamma\cdot\partial\Psi = mc\Psi \tag{2.7.5}$$

where $\partial^\mu \equiv \left(\frac{\partial}{\partial x^0}, -\vec{\nabla}\right)$ as defined in section 2.1 and

$$\gamma^\mu \equiv \left(\gamma^0, \vec{\gamma}\right) = \left(\gamma^0, \gamma^1, \gamma^2, \gamma^3\right)\,, \tag{2.7.6}$$

which makes the Dirac equation now look Lorentz covariant.

Dirac or Feynman Slash Notation

For any 4-vector A^μ, we define $\not{A} \equiv \gamma\cdot A = \gamma^0 A^0 - \vec{\gamma}\cdot\vec{A}$. Then, the Dirac equation is

$$i\hbar\not{\partial}\Psi = mc\Psi \tag{2.7.7}$$

or

$$\not{\widehat{p}}\Psi = mc\Psi \tag{2.7.8}$$

$$\Rightarrow \left(\not{\widehat{p}} - mc\right)\Psi = 0 \tag{2.7.9}$$

where $\widehat{p}^\mu = i\hbar\nabla^\mu$.

2.8 Properties of the γ-Matrices

Using the properties of the α^i and β we may show that the gamma-matrices obey the following anti-commutation identity

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\mathbb{I}_4\,, \tag{2.8.1}$$

with $g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

In particular, $(\gamma^0)^2 = \mathbb{I}_4$, $(\gamma^i)^2 = -\mathbb{I}_4$, $i = 1,2,3$.

Furthermore $(\gamma^0) = \beta^\dagger = \beta$ *i.e.* γ^0 is Hermitian, whereas $(\gamma^i)^\dagger = (\beta\alpha^i)^\dagger = (\alpha^i)^\dagger\beta^\dagger = \alpha^i\beta = -\beta\alpha^i = -\gamma^i$ *i.e.* γ^i is anti-Hermitian, $i = 1,2,3$.

Using the explicit matrices α^i and β of Section 2.4 we find

$$\gamma^0 = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix} \ , \ \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} . \tag{2.8.2}$$

2.9 The Dirac Equation and Lorentz Transformations

In the previous sections we have introduced the *covariant form* of the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\, \Psi(x) = 0\,, \tag{2.9.1}$$

without actually justifying that name. Let us clarify this point now.

We start with two comments:

1. A Lorentz transformatin takes

$$\Psi(x) \to \Psi'(x') = \Psi'(\Lambda \cdot x)\,. \tag{2.9.2}$$

   The Dirac wave function has four components but it does not transform like a 4-vector. Nevertheless we expect the transformation to be linear.

2. In the primed system $\Psi'(x')$ should obey an equation that has the same form as the Dirac equation in the un-primed system.

$$\left(i\widetilde{\gamma}^\mu \frac{\partial}{\partial x'^\mu} - m \right) \Psi'(x') = 0\,, \tag{2.9.3}$$

   where the $\widetilde{\gamma}^\mu$ are related to γ^μ by a unitary transformation $\widetilde{\gamma}^\mu = U^\dagger \gamma^\mu U$ with $U^\dagger = U^{-1}$ and obey $\{\widetilde{\gamma}^\mu, \widetilde{\gamma}^\nu\} = 2\mathbb{I}_4 \eta^{\mu\nu}$.

From the first comment we see

$$\Psi'(x') = \Psi'(\Lambda \cdot x) = S(\Lambda)\Psi(x) = S(\Lambda)\Psi(\Lambda^{-1} \cdot x') \tag{2.9.4}$$

or

$$\Psi(x) = S^{-1}(\Lambda)\Psi'(x') = S^{-1}(\Lambda)\Psi'(\Lambda \cdot x)\,. \tag{2.9.5}$$

Combining eqns. (2.9.1) and (2.9.5) we can find an equation for $\Psi'(x')$

$$(i\gamma^\mu \partial_\mu - m)\, S^{-1}(\Lambda)\Psi'(x') = 0\,. \tag{2.9.6}$$

Multiplying this equation from the left with S we find

$$\left(iS\gamma^\mu S^{-1}\partial_\mu - m\right)\Psi'(x') = 0\,. \tag{2.9.7}$$

Now $\partial_\mu = \frac{\partial x'^\nu}{\partial x^\mu}\frac{\partial}{\partial x'^\nu} = \Lambda^\nu{}_\mu\frac{\partial}{\partial x'^\nu} = \Lambda^\nu{}_\mu\partial'_\nu$ and, hence, we get

$$\left(iS\gamma^\mu S^{-1}\Lambda^\nu{}_\mu\partial'_\nu - m\right)\Psi'(x') = 0\,. \tag{2.9.8}$$

Therefore, if we are able to find an $S(\Lambda)$ such that

$$\Lambda^\nu{}_\mu S\gamma^\mu S^{-1} = \gamma^\nu \tag{2.9.9}$$

then we would have proven form invariance (or commonly called convariance) of the Dirac equation! We can write this condition also as

$$S^{-1}\gamma^\nu S = \Lambda^\nu{}_\mu\gamma^\mu\,. \tag{2.9.10}$$

So let us consider infinitesimal Lorentz Transformations

$$\Lambda^\nu{}_\mu = \delta^\nu{}_\mu - i\delta\omega^\nu{}_\mu\,, \tag{2.9.11}$$

and write the transformation matrix acting on the Dirac spinor also in infinitesimal form

$$S = \mathbb{I}_4 - i\delta\sigma\,. \tag{2.9.12}$$

Its inverse is

$$S^{-1} = \mathbb{I}_4 + i\delta\sigma\,. \tag{2.9.13}$$

Plugging these expressions into (2.9.10) we find

$$\begin{aligned}&\left(\mathbb{I}_4 + i\delta\sigma\right)\gamma^\nu\left(\mathbb{I}_4 - i\delta\sigma\right) = \left(\delta^\nu{}_\mu - i\delta\omega^\nu{}_\mu\right)\gamma^\mu\\ &\Rightarrow \left[\gamma^\nu,\sigma\right] = \omega^\nu{}_\mu\gamma^\mu\end{aligned} \tag{2.9.14}$$

Solutions for this equation are easy to find and well known. For $\omega^\nu{}_\mu = K_{x,y,z}; J_{x,y,z}$ we find $\sigma(K_x), \sigma(K_y), \ldots, \sigma(J_z)$ with

$$\sigma(K_i) = \frac{i}{4}\left[\gamma^0,\gamma^i\right] = \frac{i}{2}\begin{pmatrix}0 & \sigma_i\\ \sigma_i & 0\end{pmatrix}, \tag{2.9.15}$$

$$\sigma(J_i) = \frac{i}{8}\epsilon_{ijk}\left[\gamma^j,\gamma^k\right] = \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix} = \frac{1}{2}\Sigma_i\,. \tag{2.9.16}$$

Note that the $\sigma(K_i)$ and $\sigma(J_i)$ obey the same commutation relations as the K_i and J_i. They correspond to a particular representation of the Lorentz group called the spinor representation. The $\sigma(K_i)$ and $\sigma(J_i)$ correspond to the vector representation which acts by generalised rotations on Lorentz 4-vectors.

Finite LTs for Dirac spinors are again obtained by exponentiating the infinitesimal LTs:

$$S = \exp(-i\delta\sigma)\,, \tag{2.9.17}$$

so that under a LT

$$\Psi \to S\Psi\,. \tag{2.9.18}$$

Rotation of Dirac fermion:

$$\begin{aligned} S &= \exp(-i(\phi_x\sigma(J_x) + \phi_y\sigma(J_y) + \phi_z\sigma(J_z))) \\ &= \exp(-i\vec{\phi}\cdot\vec{\Sigma}/2) = \begin{pmatrix} \exp(-\frac{i}{2}\vec{\phi}\cdot\vec{\sigma}) & 0 \\ 0 & \exp(-\frac{i}{2}\vec{\phi}\cdot\vec{\sigma}) \end{pmatrix} \end{aligned} \tag{2.9.19}$$

For a rotation around the z-axis $\phi_x = \phi_y = 0$ and $\phi_z = \phi$ this becomes

$$\begin{pmatrix} \exp(-i\phi/2) & 0 & 0 & 0 \\ 0 & \exp(+i\phi/2) & 0 & 0 \\ 0 & 0 & \exp(-i\phi/2) & 0 \\ 0 & 0 & 0 & \exp(+i\phi/2) \end{pmatrix}. \tag{2.9.20}$$

In particular if $\phi = 2\pi$, $S = -\mathbb{I}_4$, i.e. under a 2π-rotation $\Psi \to -\Psi$ and a Dirac fermion is only invariant under a 4π rotation. This is in contrast to a 4-vector that comes back to itself after a 2π rotation. One consequence of this is that physical observables always contain an even number of fermions which is invariant under 2π rotations.

Boost of Dirac fermion: Consider a boost in the z direction

$$\begin{aligned} S &= \exp(-i\omega\sigma(K_z)) \\ &= \exp\begin{pmatrix} 0 & \frac{\omega}{2}\sigma_z \\ \frac{\omega}{2}\sigma_z & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cosh(\omega/2) & 0 & \sinh(\omega/2) & 0 \\ 0 & \cosh(\omega/2) & 0 & -\sinh(\omega/2) \\ \sinh(\omega/2) & 0 & \cosh(\omega/2) & 0 \\ 0 & -\sinh(\omega/2) & 0 & \cosh(\omega/2) \end{pmatrix} \end{aligned} \tag{2.9.21}$$

Acting with this S on a plane wave solution with $\vec{p} = (0,0,0)$ we should obtain a plane wave solution with $\vec{p} = (0,0,p_z)$. In order to show this it is helpful to use the identities $\cosh(\omega/2) = \sqrt{\frac{E+m}{2m}}$ and $\sinh(\omega/2) = \frac{p_z}{\sqrt{2m(E+m)}}$ with $E = \sqrt{m^2 + p_z^2}$. (Homework)

2.10 Plane Wave Solutions of the Dirac Equation

From now on we will work in *natural units* $\hbar = c = 1$. We look for plane wave solutions of the Dirac equation of the form

$$\Psi = e^{\mp ip\cdot x} \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.10.1}$$

where ϕ are the upper two components and χ the lower two components of Ψ and $p \cdot x = Et - \vec{p}\cdot\vec{x}$, with $E > 0$.

The factor $e^{-ip\cdot x}$ gives solutions with positive energy E and momentum $\vec{p}$, and the factor $e^{+ip\cdot x}$ gives solutions with negative energy $-E$ and momentum $-\vec{p}$. Substituting back into the Dirac equation

$$\left(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)\Psi = i\frac{\partial\Psi}{\partial t}\,, \tag{2.10.2}$$

and using $\partial_t\Psi = \mp iE\Psi$ and $\partial_x\Psi = \pm ip_x\Psi$ e.t.c., we obtain

$$\begin{aligned} (-i(\pm i)\vec{\alpha}\cdot\vec{p} + \beta m)\Psi &= i(\mp iE)\Psi \\ \rightarrow (\pm\vec{\alpha}\cdot\vec{p} + \beta m)\Psi &= \pm E\Psi\,. \end{aligned} \tag{2.10.3}$$

If we use the standard representation for the α^i and β from Section 2.4 we obtain

$$\begin{pmatrix} m\mathbb{I} & \pm\vec{\sigma}\cdot\vec{p} \\ \pm\vec{\sigma}\cdot\vec{p} & -m\mathbb{I} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \pm E \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.10.4}$$

which gives the coupled set of equations

$$\begin{aligned} (m \mp E)\phi \pm \vec{\sigma}\cdot\vec{p}\chi &= 0 \\ \pm\vec{\sigma}\cdot\vec{p}\phi - (m \pm E)\chi &= 0\,. \end{aligned} \tag{2.10.5}$$

We will construct the solutions in such a way that they have a straighforward $\vec{p} = 0$ limit.

<u>Positive Energy Solutions</u>

$$\begin{aligned} (m - E)\phi + \vec{\sigma}\cdot\vec{p}\chi &= 0 \\ \vec{\sigma}\cdot\vec{p}\phi - (m + E)\chi &= 0\,. \end{aligned} \tag{2.10.6}$$

For $\vec{p} = 0$, $E = m$ and $\chi = 0$ (in agreement with Section 2.6).

For $\vec{p} \neq 0$ it is convenient to solve for χ in terms of ϕ using the second equation in (2.10.6), *i.e.*

$$\chi = \frac{\vec{\sigma} \cdot \vec{p}}{(E+m)} \phi \tag{2.10.7}$$

then,

$$\Psi = e^{-ip \cdot x} \begin{pmatrix} \phi \\ \frac{\vec{\sigma} \cdot \vec{p}}{(E+m)} \phi \end{pmatrix} . \tag{2.10.8}$$

Note that the first equation in (2.10.6) only gives the on-(mass)shell condition $E^2 = \vec{p}^2 + m^2$.

We can write ϕ in terms of $\phi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\phi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

There are thus two independent positive energy solutions

$$\Psi = e^{-ip \cdot x} U(p, s) \ , \ \ s = 1, 2 \tag{2.10.9}$$

where

$$U(p, s) = \sqrt{E+m} \begin{pmatrix} \phi_s \\ \frac{\vec{\sigma} \cdot \vec{p}}{(E+m)} \phi_s \end{pmatrix} \tag{2.10.10}$$

is a positive energy Dirac spinor. In the last expression a convenient normalization factor has been introduced.

It can be checked that the first equation in (2.10.6) is automatically satisfied by using the identity $(\vec{\sigma} \cdot \vec{p})^2 = \vec{p}^2 \mathbb{I}_2$.

Negative Energy Solutions

$$\begin{aligned} (m+E)\phi - \vec{\sigma} \cdot \vec{p} \chi &= 0 \\ -\vec{\sigma} \cdot \vec{p} \phi - (m-E)\chi &= 0 \, . \end{aligned} \tag{2.10.11}$$

For $\vec{p} = 0$, $E = m$ and $\phi = 0$ (in agreement with Section 2.6).

For $\vec{p} \neq 0$ it is convenient to solve for ϕ in terms of χ using the first equation in (2.10.11), *i.e.*

$$\phi = \frac{\vec{\sigma} \cdot \vec{p}}{(E+m)} \chi \tag{2.10.12}$$

then,

$$\Psi = e^{+ip \cdot x} \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{(E+m)} \chi \\ \chi \end{pmatrix} . \tag{2.10.13}$$

Note that the second equation in (2.10.11) only gives the on-(mass)shell condition $E^2 = \vec{p}^2 + m^2$.

We can write χ in terms of $\chi_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\chi_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

There are thus two independent negative energy solutions

$$\Psi = e^{+ip\cdot x} V(p,s) \ , \quad s = 1,2 \tag{2.10.14}$$

where

$$V(p,s) = \sqrt{E+m} \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{(E+m)} \chi_s \\ \chi_s \end{pmatrix} \tag{2.10.15}$$

is a negative energy Dirac spinor. In the last expression a convenient normalization factor has been introduced.

It can be checked that the second equation in (2.10.11) is automatically satisfied by using the identity $(\vec{\sigma}\cdot\vec{p})^2 = \vec{p}^2 \mathbb{I}_2$.

<ins>Interpretation</ins>

To find the physical interpretation for the four independent solutions we consider the rest frame $\vec{p} = 0$. Then:

$$\begin{aligned} U(p,1) &= \sqrt{2m} \begin{pmatrix} \phi^1 \\ 0 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ U(p,2) &= \sqrt{2m} \begin{pmatrix} \phi^2 \\ 0 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\ V(p,1) &= \sqrt{2m} \begin{pmatrix} 0 \\ \chi^1 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ V(p,2) &= \sqrt{2m} \begin{pmatrix} 0 \\ \chi^2 \end{pmatrix} = \sqrt{2m} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \end{aligned} \tag{2.10.16}$$

Furthermore, for $\vec{p} = 0$ we have $\vec{L} = \vec{x} \times \vec{p} = 0$ so that the total angular momentum operator becomes

$$\begin{aligned} \vec{J} &= \vec{L} + \vec{S} = \vec{S} = \frac{1}{2}\vec{\Sigma} \\ \Rightarrow & \; S_z = \begin{pmatrix} \frac{1}{2}\sigma_3 & 0 \\ 0 & \frac{1}{2}\sigma_3 \end{pmatrix} \end{aligned} \tag{2.10.17}$$

Thus, $U(p,1)$ and $U(p,2)$ are *positive energy solutions* with S_z eigenvalues $s_z = +1/2$ and $s_z = -1/2$ respectively, whereas $V(p,1)$ and $V(p,2)$ are *negative energy solutions* with S_z eigenvalues $s_z = -1/2$ and $s_z = +1/2$ respectively.

In general, U and V are the Lorentz boosts of these solutions to a frame where $\vec{p} \neq 0$. Interpret the negative energy solutions later.

2.11 Properties of Solutions

Since $e^{-ip\cdot x}U(p,s)$ is a solution of the Dirac equation

$$\begin{aligned} & (i\gamma\cdot\partial - m)(e^{-p\cdot x}U(p,s)) = 0 \\ \to & \left(i\gamma^0\frac{\partial}{\partial t} + i\vec{\gamma}\cdot\vec{\nabla} - m\right)e^{-i(Et-\vec{p}\cdot\vec{x})}U(p,s) = 0 \\ \to & (E\gamma^0 - \vec{p}\cdot\vec{\gamma} - m)U = 0 \\ \to & (p\cdot\gamma - m)U = 0 \\ \Rightarrow & (\not{p} - m)U(p,s) = 0 \end{aligned} \tag{2.11.1}$$

i.e. $U(p,s)$ obeys the Dirac equations with $\widehat{p}^\mu$ simply replaced by p^μ. Similarly, we find

$$(\not{p} + m)V(p,s) = 0\,. \tag{2.11.2}$$

We will often make use of the *adjoint spinors* which are defined as

$$\overline{U}(p,s) \equiv U^\dagger(p,s)\gamma^0\,, \;\; \overline{V}(p,s) \equiv V^\dagger(p,s)\gamma^0\,. \tag{2.11.3}$$

By taking the Hermitian adjoint of the equation obeyed by U and V we find

$$\overline{U}(p,s)(\not{p} - m) = 0\,, \;\; \overline{V}(p,s)(\not{p} + m) = 0\,. \tag{2.11.4}$$

One may also check directly (using again $(\vec{\sigma}\cdot\vec{p})^2 = \vec{p}^2\mathbb{I}_2$) that

$$U^\dagger(p,s)U(p,s) = V^\dagger(p,s)V(p,s) = 2E\,, \;\; s = 1,2\,, \tag{2.11.5}$$

and

$$\overline{U}(p,s)U(p,s) = 2m\,, \;\; \overline{V}(p,s)V(p,s) = -2m\,, \;\; s = 1,2\,. \tag{2.11.6}$$

2.12 Anti-Particles — Hole Theory

Since the Dirac equation has negative energy solutions, why do positive energy electrons not radiate energy and fall into a negative energy state? Dirac: Negative energy states are completely *filled* and the Pauli exclusion principle (which applies to fermions) forbids the transition. Consequently, the Vacuum is a state with all positive energy states empty *but* all negative energy states filled.

(picture goes here)

If a photon excites a negative energy e^- of energy $-|E_2|$ into a positive energy e^- of energy $|E_1|$, we observe the production of an e^- of mass m, charge $-|e|$ and energy $|E_1|$, and a *Hole* in the negative energy sea (Pair production). Note that there is a *gap* of $2mc^2$ between the negative and positive energy states and, hence, the photon energy $h\nu = |E_1| + |E_2|$ must be larger than $2mc^2$ for this to happen. The hole appears as a particle of mass m, charge $+|e|$ and energy $+|E_2|$.
$\Rightarrow$ The existence of the Positron (and Anti-particles in general) is predicted!

(picture goes here)

The absence of a spin-up electron of energy $-|E|$ and momentum $-\vec{p}$ is equivalent to the presence of a spin-down positron of energy $+|E|$ and momentum $+\vec{p}$. (Think about time running backwards or the arrow in a Feynman diagram reversed)

(picture goes here)

Thus, the electron wavefunction $e^{ip\cdot x}V(p,s)$ corresponding to energy $-E$ and momentum $-\vec{p}$ describes a positron of energy $+E$ and momentum $+\vec{p}$. Also, $V(p,1)$ and $V(p,2)$ which describe spin down and spin up negative energy electrons must describe spin up and spin down positrons.

2.13 Vacuum Polarization

In general the infinite negative charge of the vacuum produces no effect because the distribution of charge is homogeneous.

However, consider the effect of a positive energy electron with charge $-|e|$ on the vacuum. It repels the negative energy electrons and electrically polarises the vacuum. Thus the physical charge $-|e|$ seen by a test charge at a large distance from the electron is numerically smaller than the bare charge $-|e_0|$, i.e. $|e| < |e_0|$.

(picture goes here)

However, if the test charge comes very close it will see the bare charge $-|e_0|$. For S-wave electrons ($l = 0$) in an atom, the proton sees a charge numerically greater than the ordinary electric charge $|e|$. Note that for $l > 0$ the wavefunction vanishes at the origin and the proton feels a numerically smaller charge. This effect leads to measurable shifts of the energy levels of atoms.

2.14 Charge Conjugation Symmetry $\mathcal{C}$

We construct an operator acting on the Dirac wave function

$$\mathcal{C} : \Psi \to \Psi_C \tag{2.14.1}$$

which turns a positive energy electron wavefunction (e^-) into a negative energy wave function (e^+) with the same momentum and spin state. If $\Psi = e^{-ip\cdot x}U(p,s)$ then $\Psi_C = e^{ip\cdot x}V(p,s)$. The required operation turns out to be

$$\mathcal{C} : \Psi \to \Psi_C = C\gamma^0\Psi^*\,, \tag{2.14.2}$$

where $C = i\gamma^2\gamma^0$. Useful properties: $C^\dagger = -C$, $C^2 = -\mathbb{I}$, $C^{-1} = -C$ and $C\gamma^\mu C = (\gamma^\mu)^T$.

A symmetry of a wave equation is an operation on a wave function $\Psi \to \Psi'$ and on the space-time coordinates $x \to x'$ such that Ψ' obeys the same equation as Ψ, with x replaced by x'.

$\Psi \to \Psi_C$, $x \to x'$ can be shown to be a symmetry of the Dirac equation as follows:

Proof: We claim that the Dirac equation is charge conjugation invariant. The Dirac equation may be written as

$$(i\gamma^\mu\partial_\mu - m)\Psi = 0, \tag{2.14.3}$$

taking the complex conjugate gives

$$(-i(\gamma^\mu)^*\partial_\mu - m)\Psi^* = 0\,. \tag{2.14.4}$$

Now multiply from the left with $C\gamma^0$

$$\begin{aligned}
&(-iC\gamma^0(\gamma^\mu)^*\partial_\mu - m)\Psi^* = 0\\
\to\quad &(i\gamma^\mu(C\gamma^0))\partial_\mu - mc\gamma^0)\Psi^* = 0\\
\to\quad &(i\gamma^\mu\partial_\mu - m)\Psi_C = 0\,,
\end{aligned} \tag{2.14.5}$$

where we have used the identity $C\gamma^0(\gamma^\mu)^* = -\gamma^\mu(C\gamma^0)$. This shows that Ψ_C obeys the same equation as Ψ.

2.15 Space Inversion $\mathcal{P}$

The Dirac equation is also invariant under reflection of space coordinates in the origin

$$\mathcal{P} : \vec{x} \to \vec{x}' = -\vec{x}\,, \; t \to t' = t\,. \tag{2.15.1}$$

The corresponding operation[2] on Dirac spinors is

$$\mathcal{P} : \Psi \to \Psi' = P\Psi \tag{2.15.2}$$

with $P = \gamma^0$. It can be checked by direct calculation that $PU(\vec{p}, s) = U(-\vec{p}, s)$ i.e. $\vec{p} \to -\vec{p}$ as expected for space inversion, but the spin state is unchanged. Also $PV(\vec{p}, s) = -V(-\vec{p}, s)$. The -1 factor indicates that anti-particles have opposite Parity to particles.

In this case, to check invariance of the Dirac equation, it is necessary to replace ∂_x, ∂_y and ∂_z by $-\partial_x$, $-\partial_y$ and $-\partial_z$, as well as replacing Ψ by Ψ', i.e. Ψ' obeys the same equation as Ψ with ∂_x, ∂_y and ∂_z replaced by $-\partial_x$, $-\partial_y$ and $-\partial_z$.

2.16 Time Reversal $\mathcal{T}$

The Dirac equation also has a symmetry under time reversal

$$\mathcal{T} : t \to t' = -t\,, \; \vec{x} \to \vec{x}' = \vec{x}\,, \tag{2.16.1}$$

the appropriate transformation of Ψ is

$$\mathcal{T} : \Psi \to \Psi' = T\psi^* \tag{2.16.2}$$

with $T = -\gamma^1\gamma^3$. It can be checked directly that this is the correct transformation by showing that

$$TU^*(\vec{p}, 1) = +U(-\vec{p}, 2)\,, \; TV^*(\vec{p}, 1) = -V(-\vec{p}, 2) \tag{2.16.3}$$

Thus, the transformation changes a solution of the Dirac equation with momentum $\vec{p}$ and spin up into a solution with momentum $-\vec{p}$ and spin down.

This is as expected for time reversal, since $\vec{p} = m\vec{v}/\sqrt{1 - \vec{v}^2/c^2}$ and $\vec{L} = \vec{x} \times \vec{p}$, and thus under time reversal $\vec{p} \to -\vec{p}$ and $\vec{L} \to -\vec{L}$ and in particular $L_z \to -L_z$. We assume that this applies to any AM operator, so that in particular $S_z \to -S_z$. In this case, Ψ' obeys the same equation as Ψ with ∂_t replaced by $-\partial_t$.

[2] Since $\mathcal{P}$ also transforms the space time coordinates this operation should be written more properly as $\Psi(t, \vec{x}) \to \Psi'(t', \vec{x}') = \Psi'(t, -\vec{x}) = P\Psi(t, \vec{x})$. Hence $\Psi'(t, \vec{x}) = P\Psi(t, -\vec{x})$; a similar comment applies to time reversal T.

2.17 Dirac Covariants

It is important in the study of the Weak Interactions to know the properties of objects like $\overline{\Psi}\gamma^\mu\Psi$, $\overline{\Psi}\gamma^\mu\gamma_5\Psi$, etc, where we introduced

$$\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} . \tag{2.17.1}$$

We list here the behaviour of some of the Dirac covariants under Lorentz transformations, $\mathcal{P}$, $\mathcal{C}$:

Covariant	LT's	$\mathcal{P}$	$\mathcal{C}$
$\overline{\Psi}\Psi$	scalar	$+\overline{\Psi}\Psi$	$-\overline{\Psi}\Psi$
$\overline{\Psi}\gamma_5\Psi$	pseudoscalar	$-\overline{\Psi}\gamma_5\Psi$	$-\overline{\Psi}\gamma_5\Psi$
$\overline{\Psi}\gamma^\mu\Psi$	4-vector	$+\overline{\Psi}\gamma^0\Psi$ $-\overline{\Psi}\gamma^i\Psi$	$+\overline{\Psi}\gamma^\mu\Psi$
$\overline{\Psi}\gamma^\mu\gamma_5\Psi$	(pseudo) 4-vector	$-\overline{\Psi}\gamma^0\gamma_5\Psi$ $+\overline{\Psi}\gamma^i\gamma_5\Psi$	$-\overline{\Psi}\gamma^\mu\gamma_5\Psi$

(γ_5 has the properties $\{\gamma_5, \gamma^\mu\} = 0$, $\gamma_5^\dagger = \gamma_5$)

For example under P the behaviour of the vector current is

$$\begin{aligned} \overline{\Psi}\gamma^\mu\Psi \quad \rightarrow \quad & \overline{\Psi}'\gamma^\mu\Psi' \\ &= (\Psi')^\dagger\gamma^0\gamma^\mu\Psi' \\ &= \Psi^\dagger(\gamma^0)^\dagger\gamma^0\gamma^\mu\gamma^0\Psi \\ &= \Psi^\dagger\gamma^0\gamma^0\gamma^\mu\gamma^0\Psi \\ &= \overline{\Psi}\gamma^0\gamma^\mu\gamma^0\Psi \end{aligned} \tag{2.17.2}$$

Thus,

$$\overline{\Psi}\gamma^0\Psi \rightarrow \overline{\Psi}\gamma^0\gamma^0\gamma^0\Psi = \overline{\Psi}\gamma^0\Psi \tag{2.17.3}$$

$$\begin{aligned} \overline{\Psi}\gamma^i\Psi \quad \rightarrow \quad & \overline{\Psi}\gamma^0\gamma^i\gamma^0\Psi \\ &= -\overline{\Psi}(\gamma^0)^2\gamma^i\Psi \\ &= -\overline{\Psi}\gamma^i\Psi \end{aligned} \tag{2.17.4}$$

In the Relativistic version of Time Dependent Perturbation Theory (Feynman Diagrams) the probability amplitudes for Electromagnetic Scattering of 2 particles via photon exchange contains a factor

$$
\begin{aligned}
&\overline{U}(p_2,s_2)\gamma_\mu U(p_1,s_1)\overline{U}(p_4,s_4)\gamma^\mu U(p_3,s_3) \\
&= \overline{U}\gamma^0 U\overline{U}\gamma^0 U - \sum_{i=1}^{3}\overline{U}\gamma^i U\overline{U}\gamma^i U\,,
\end{aligned}
\tag{2.17.5}
$$

which is *invariant under both $\mathcal{P}$ and $\mathcal{C}$* because the two negative signs for the $\overline{U}\gamma^i U$ cancel for space inversion. Thus the Electromagnetic interactions are both space reflection invariant and charge conjugation invariant.

(picture goes here)

The corresponding probability amplitude for the Weak Interaction has a factor

$$
\begin{aligned}
&\overline{U}(p_2,s_2)\gamma_\mu(\mathbb{I}-\gamma_5)U(p_1,s_1)\overline{U}(p_4,s_4)\gamma^\mu(\mathbb{I}-\gamma_5)U(p_3,s_3) \\
&= \overline{U}\gamma_\mu U\overline{U}\gamma^\mu U + \overline{U}\gamma\mu\gamma_5 U\overline{U}\gamma^\mu\gamma_5 U \\
&-\overline{U}\gamma_\mu\gamma_5 U\overline{U}\gamma^\mu U - \overline{U}\gamma_\mu U\overline{U}\gamma^\mu\gamma_5 U\,.
\end{aligned}
\tag{2.17.6}
$$

The term

$$
\begin{aligned}
&\overline{U}\gamma_\mu\gamma_5 U\overline{U}\gamma^\mu U \\
&= \overline{U}\gamma^0\gamma_5 U\overline{U}\gamma^0 U - \sum_{i=1}^{3}\overline{U}\gamma^i\gamma_5 U\overline{U}\gamma^i U
\end{aligned}
\tag{2.17.7}
$$

changes sign under both $\mathcal{P}$ and $\mathcal{C}$ transformations, because $\overline{\Psi}\gamma^\mu\Psi$ and $\overline{\Psi}\gamma^\mu\gamma_5\Psi$ transform with opposite signs both for $\mu = 0$ and $\mu = i$. Thus, the weak interactions *break both space inversion and charge conjugation invariance.* This manifests itself in the angular dependence of scattering processes (e.g. $\cos\theta$ changes sign under space inversion: $0 \to \pi - 0$). Note however that the combined action of $\mathcal{C}$ and $\mathcal{P}$, $\mathcal{CP}$, is a symmetry of this interaction.

Note, that in a general theory $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{T}$ are not preserved, but the combination of the three transformations $\mathcal{CPT}$ is always a symmetry.

2.18 Neutrinos

Some modification of RQM is needed in the physically important case of massless spin-1/2 particles — Neutrinos (with todays experimental evidence of Neutrino oscillations this is not quite true, nevertheless it is a very good approximation.)

First, define the Helicity of a particle as the component of its AM $\vec{J} = \vec{L} + \vec{\Sigma}/2$ in its direction of motion. For a Dirac particle,

$$\text{Helicity} = \vec{J} \cdot \frac{\vec{p}}{|\vec{p}|} = \frac{\vec{\Sigma}}{2} \cdot \frac{\vec{p}}{|\vec{p}|} \tag{2.18.1}$$

because $\vec{L} \cdot \vec{p} = (\vec{x} \times \vec{p}) \cdot \vec{p} = 0$.

Experimental observation shows that whereas an e^- can have Helicity $+1/2$ or $-1/2$, a Neutrino (which is massless) can only have Helicity $-1/2$ and an Anti-Neutrino can only have Helicity $+1/2$.

Thus, whereas we need four degrees of freedom to describe the 2 spin states of an electron or positron, we need only 2 degrees of freedom to describe the spin states of the neutrino and anti-neutrino. We need to discard 2 spin states of the Dirac particle.

Now we return to the Dirac equation for a positive energy solution of energy E and momentum $\vec{p}$. However, we choose a different representation of the Dirac matrices (and hence a different representation of the gamma-matrices). This does not effect the physics but makes the proof much easier. It may be checked that

$$\beta = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix} \tag{2.18.2}$$

also obey the Dirac Algebra (2.2.7).

For $\Psi = e^{-ip\cdot x} \begin{pmatrix} \phi \\ \chi \end{pmatrix}$ the Dirac equation reduces to $(\vec{\alpha} \cdot \vec{p} + \beta m)\Psi = E\Psi$ (see Section 2.11), from which we get

$$\begin{pmatrix} \vec{\sigma} \cdot \vec{p} & m\mathbb{I} \\ m\mathbb{I} & -\vec{\sigma} \cdot \vec{p} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = E \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \tag{2.18.3}$$

which gives the coupled equations

$$\begin{aligned} \vec{\sigma} \cdot \vec{p}\phi + m\chi &= E\phi \\ m\phi - \vec{\sigma} \cdot \vec{p}\chi &= E\chi \,. \end{aligned} \tag{2.18.4}$$

Now taking $m = 0$ for a massless neutrino decouples the two equations,

$$\begin{aligned} \vec{\sigma} \cdot \vec{p}\phi &= E\phi \\ \vec{\sigma} \cdot \vec{p}\chi &= -E\chi \,, \end{aligned} \tag{2.18.5}$$

and since $E = |\vec{p}|$ for $m = 0$,

$$\begin{aligned} \frac{\vec{\sigma} \cdot \vec{p}}{2|\vec{p}|}\phi &= \frac{1}{2}\phi \\ \frac{\vec{\sigma} \cdot \vec{p}}{2|\vec{p}|}\chi &= -\frac{1}{2}\chi \,. \end{aligned} \tag{2.18.6}$$

Thus the upper 2 components of Ψ describe Helicity 1/2, and the lower two describe helicity $-1/2$ when Ψ has positive energy. To obtain an appropriate Ψ to describe a Neutrino we perform a projection that removes the upper 2 components.

This may be achieved by using $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. With the above choice of $\vec{\alpha}$ and β,

$$\gamma^0 = \beta = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}, \quad \gamma^i = \beta\alpha^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} \tag{2.18.7}$$

and

$$\gamma^5 = i \begin{pmatrix} -\sigma^1\sigma^2\sigma^3 & 0 \\ 0 & \sigma^1\sigma^2\sigma^3 \end{pmatrix} = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \tag{2.18.8}$$

If we form $\frac{1}{2}(\mathbb{I}-\gamma_5) = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I} \end{pmatrix}$, then we may use it to project the upper 2 components and leave the Helicity $-1/2$ components. Thus,

$$\Psi_L \equiv \frac{1}{2}(\mathbb{I} - \gamma_5)\Psi\,, \tag{2.18.9}$$

with Ψ a positive energy spinor, may be used to describe the neutrino.

If instead we start from a negative energy solution Ψ, the from Section 2.11

$$(-\vec{\alpha}\cdot\vec{p} + \beta m)\Psi = -E\Psi\,. \tag{2.18.10}$$

For $\Psi = e^{ip\cdot x}\begin{pmatrix} \phi \\ \chi \end{pmatrix}$ we then have

$$\begin{pmatrix} -\vec{\sigma}\cdot\vec{p} & m\mathbb{I} \\ m\mathbb{I} & \vec{\sigma}\cdot\vec{p} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = -E \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \tag{2.18.11}$$

which gives the coupled equations

$$\begin{aligned} -\vec{\sigma}\cdot\vec{p}\phi + m\chi &= -E\phi \\ m\phi + \vec{\sigma}\cdot\vec{p}\chi &= -E\chi\,. \end{aligned} \tag{2.18.12}$$

Now taking $m = 0$ for a massless anti-neutrino

$$\begin{aligned} \vec{\sigma}\cdot\vec{p}\phi &= E\phi \\ \vec{\sigma}\cdot\vec{p}\chi &= -E\chi\,, \end{aligned} \tag{2.18.13}$$

and since $E = |\vec{p}|$ for $m = 0$,

$$\begin{aligned} \frac{\vec{\sigma}\cdot\vec{p}}{2|\vec{p}|}\phi &= \frac{1}{2}\phi \\ \frac{\vec{\sigma}\cdot\vec{p}}{2|\vec{p}|}\chi &= -\frac{1}{2}\chi\,. \end{aligned} \tag{2.18.14}$$

This is the same as for the positive energy solution. Thus, the upper 2 components of Ψ still describe Helicity $+1/2$ and the lower 2 components describe Helicity $-1/2$. To obtain an appropriate Ψ to describe an Anti-Neutrino with Helicity $+1/2$ we need a negative energy state with Helicity $-1/2$.

Thus, $\Psi_L = \frac{1}{2}(\mathbb{I} - \gamma_5)\Psi$ with Ψ a negative energy spinor, may be used to describe the anti-neutrino with Helicity $+1/2$.

2.19 Feynman's Interpretation of the Klein-Gordon Equation

In Section 1.2 we abandoned the KG equation because the Probability density

$$\rho = i\left(\phi^* \partial_t \phi - \phi \partial_t \phi^*\right) \tag{2.19.1}$$

could give negative values (we have renamed the wavefuntion Ψ by ϕ).

It can be checked by direct substitution that the KG equation

$$(\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + m^2)\phi = 0 \tag{2.19.2}$$

has positive energy solutions

$$\phi = N e^{-ip\cdot x} = N e^{-i(Et - \vec{p}\cdot\vec{x})} \tag{2.19.3}$$

and negative energy solutions

$$\phi = N e^{ip\cdot x} = N e^{i(Et - \vec{p}\cdot\vec{x})}\,. \tag{2.19.4}$$

The probability density of such solutions is

$$\rho = |N|^2(\pm 2E)\,. \tag{2.19.5}$$

Thus, negative probabilities come from negative energy solutions. These are (as usual) the problem.

We need an interpretation for the negative energy solutions of the KG equation. Dirac Hole theory will NOT work for the spin-0 Bosons described by the KG equation, because they do not obey the Dirac exclusion principle to give a filled negative energy sea.

Feynman gave an alternative way of interpreting negative energy solutions which works for both bosons and fermions!

The emission/absorption of an anti-particle with 4-momentum p^μ is equivalent to the absorption/emission of a negative energy particle with 4-momentum $-p^\mu$.

In Feynman diagrams, which are the rules of calculating scattering and decay amplitudes in RQM, when Anti-Particles are involved we draw lines for negative energy particles propagating backwards in time and use Feynman's interpretation. E.g. for electromagnetic Electron-Positron scattering via photon exchange there are two diagrams that contribute:

(picture goes here)

2.20 Dirac Equation in an Electromagnetic Field

In classical relativistic mechanics the interaction of a particle carrying charge q in an external electromagnetic field can be obtained by substituting the momentum as

$$p^\mu \to p^\mu + qA^\mu \,, \tag{2.20.1}$$

where A^μ is the 4-vector potential

$$A^\mu \equiv (A^0, \vec{A}) = (\phi, \vec{A}) \tag{2.20.2}$$

with ϕ the scalar potential and $\vec{A}$ the vector potential. (Remember: $\vec{E} = -\vec{\nabla}\phi - \partial_t \vec{A}$, $\vec{B} = \vec{\nabla} \times \vec{A}$).

This works also for RQM

$$\widehat{p}^\mu \to \widehat{p}^\mu + qA^\mu \tag{2.20.3}$$

or equivalently

$$\partial^\mu \to \partial^\mu - iqA^\mu \,. \tag{2.20.4}$$

The free particle Dirac equation is $(i\gamma \cdot \partial - m)\Psi = 0$. Making the above substitution

$$\gamma \cdot \partial \to \gamma \cdot \partial - iq\gamma \cdot A \tag{2.20.5}$$

the Dirac equation in an electromagnetic field is

$$(i\gamma \cdot \partial - m)\Psi = -q\gamma \cdot A \tag{2.20.6}$$

or

$$(i\partial\!\!\!/ - m)\Psi = -qA\!\!\!/ \,. \tag{2.20.7}$$

It is sometimes convenient to write the equation in terms of the Dirac matrices i.e. in Hamiltonian form. Begin with the equation

$$(i\gamma^0 \frac{\partial}{\partial t} + i\vec{\gamma} \cdot \vec{\nabla} - m)\Psi = -q(A^0\gamma^0 - \vec{A} \cdot \vec{\gamma})\Psi \tag{2.20.8}$$

and multiply from left with $\gamma^0 = \beta$. Since $(\gamma^0)^2 = \mathbb{I}$ and $\gamma^0\gamma^i = \beta\beta\alpha^i = \alpha^i$ we obtain

$$\begin{aligned} i\frac{\partial\Psi}{\partial t} &= \left((-i\vec{\nabla} + q\vec{A}) \cdot \alpha + \beta m\right) \Psi - qA^0\Psi \\ \rightarrow \quad i\frac{\partial\Psi}{\partial t} &= \left(\alpha \cdot \widehat{\vec{\Pi}} + \beta m\right) \Psi - qA^0\Psi\,, \end{aligned} \tag{2.20.9}$$

where

$$\widehat{\vec{\Pi}} = -i\vec{\nabla} + q\vec{A} = \widehat{\vec{p}} + q\vec{A}\,. \tag{2.20.10}$$

Gauge Invariance

Remember that the EM field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is invariant under gauge transformations $A_\mu \rightarrow A_\mu + \partial_\mu\chi \equiv \tilde{A}_\mu$. However the Dirac equation is not invariant under gauge transformation

$$\begin{aligned} (i\not{\partial} + q\not{A} - m)\Psi &\rightarrow (i\not{\partial} + q\tilde{\not{A}} - m)\Psi \\ &= (i\not{\partial} + q\not{A} + q\not{\partial}\chi - m)\Psi\,. \end{aligned} \tag{2.20.11}$$

This can be fixed if we transform the Dirac wavefunction by a spacetime dependent phase factor

$$\Psi \rightarrow e^{i\alpha(x)}\Psi = \tilde{\Psi}\,. \tag{2.20.12}$$

So we find

$$\begin{aligned} (i\not{\partial} + q\not{A} - m)\Psi = 0 \quad &\rightarrow \quad (i\not{\partial} + q\tilde{\not{A}} - m)\tilde{\Psi} = 0 \\ &\rightarrow \quad e^{i\alpha(x)}\left(i\not{\partial}\Psi - (\not{\partial}\alpha)\Psi + (q\not{A} + q\not{\partial}\chi - m)\Psi\right) = 0\,. \end{aligned} \tag{2.20.13}$$

If we now require that the 2nd term cancels the 4th term in the last line of that equation, which simply implies

$$\alpha(x) = q\chi(x)\,, \tag{2.20.14}$$

then we get

$$e^{iq\chi(x)}\left(i\not{\partial} + q\not{A} - m\right)\Psi = 0\,. \tag{2.20.15}$$

So up to an irrelevant overall phase factor we have recovered the original Dirac equation! Hence the Dirac equation is *forminvariance under gauge transformations*:

$$\begin{aligned} A_\mu &\rightarrow \tilde{A}_\mu = A_\mu + \partial_\mu\chi(x)\,, \\ \Psi &\rightarrow \tilde{\Psi} = e^{iq\chi(x)}\Psi\,. \end{aligned} \tag{2.20.16}$$

Charge Conjugation

It is interesting to see what happens to the Dirac equation coupled to the EM field under charge conjugation:

$$\mathcal{C} : \Psi \to \Psi_C = C\gamma^0\Psi^* \,, \tag{2.20.17}$$

with $C = i\gamma^2\gamma^0$.

It turns out that

$$(i\partial\!\!\!/ + qA\!\!\!/ - m)\,\Psi = 0 \to (i\partial\!\!\!/ - qA\!\!\!/ - m)\,\Psi_C = 0 \tag{2.20.18}$$

i.e. under charge conjugation the charge of the particle flips sign. (Proof in homework)

2.21 The Magnetic Moment of the Electron

In the non-relativistic limit the rest mass mc^2 is the largest energy in the problem (since $|\vec{v}|^2 << c^2$) and we can write for a positive energy solution

$$\Psi = e^{-imt}\begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.21.1}$$

where ϕ and χ vary slowly with time and will be called large and small components for reasons that will become clear in a moment.

Substituting in the Dirac equation (with the Dirac representation for β and α^i, which is more appropriat for studying non-relativistic limits) in an electromagnetic field

$$me^{-imt}\begin{pmatrix} \phi \\ \chi \end{pmatrix} + ie^{-imt}\begin{pmatrix} \partial_t\phi \\ \partial_t\chi \end{pmatrix} = e^{-imt}\begin{pmatrix} (-qA^0+m)\mathbb{I} & \vec{\sigma}\cdot\widehat{\vec{\Pi}} \\ \vec{\sigma}\cdot\widehat{\vec{\Pi}} & (-qA^0-m)\mathbb{I} \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix}, \tag{2.21.2}$$

multiplying with e^{+imt} and subtracting the first term on the left hand side from both sides we obtain

$$\begin{pmatrix} i\partial_t\phi \\ i\partial_t\chi \end{pmatrix} = \begin{pmatrix} -qA^0\mathbb{I} & \vec{\sigma}\cdot\widehat{\vec{\Pi}} \\ \vec{\sigma}\cdot\widehat{\vec{\Pi}} & (-qA^0-2m)\mathbb{I} \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix}. \tag{2.21.3}$$

The lower equtions is

$$i\partial_t\chi = \vec{\sigma}\cdot\widehat{\vec{\Pi}}\phi - (qA^0+2m)\chi \,. \tag{2.21.4}$$

For χ varying slowly with time and under the assumption $2m >> qA_0$

$$\chi \sim \frac{\vec{\sigma}\cdot\widehat{\vec{\Pi}}}{2m}\phi \tag{2.21.5}$$

where $\widehat{\vec{\Pi}} = \widehat{\vec{p}} + q\vec{A}$. Hence, for momenta and EM fields small compared to the rest mass

$$\chi << \phi\,. \tag{2.21.6}$$

Now using the top equation

$$i\partial_t \phi = -qA^0\phi + \vec{\sigma}\cdot\widehat{\vec{\Pi}}\chi\,, \tag{2.21.7}$$

we get, using eqn. (2.21.5)

$$\Rightarrow i\partial_t \phi = -qA^0\phi + \frac{(\vec{\sigma}\cdot\widehat{\vec{\Pi}})^2}{2m}\phi\,. \tag{2.21.8}$$

To simplify this further we may use the identity

$$(\vec{\sigma}\cdot\vec{a})(\vec{\sigma}\cdot\vec{b}) = \vec{a}\cdot\vec{b}\mathbb{I} + i\vec{\sigma}\cdot(\vec{a}\times\vec{b})\,, \tag{2.21.9}$$

which follows from

$$\sigma^i\sigma^j = \delta^{ij}\mathbb{I} + i\epsilon^{ijk}\sigma^k \tag{2.21.10}$$

where summation over k is understood. The Kronecker delta is defined as $\delta^{ij} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$; and $\epsilon^{123} = \epsilon^{231} = \epsilon^{312} = +1$, $\epsilon^{132} = \epsilon^{213} = \epsilon^{321} = -1$ and otherwise $\epsilon^{ijk} = 0$. In terms of the ϵ-tensor

$$\left(\vec{a}\times\vec{b}\right)^i = \epsilon^{ijk}a^j b^k \tag{2.21.11}$$

and, hence,

$$(\vec{\sigma}\cdot\widehat{\vec{\Pi}})^2\phi = (\vec{\sigma}\cdot\widehat{\vec{\Pi}})(\vec{\sigma}\cdot\widehat{\vec{\Pi}})\phi = \widehat{\vec{\Pi}}\cdot\widehat{\vec{\Pi}}\mathbb{I}\phi + i\vec{\sigma}\cdot(\widehat{\vec{\Pi}}\times\widehat{\vec{\Pi}})\phi\,. \tag{2.21.12}$$

Now ($\hbar = 1$),

$$(\widehat{\vec{\Pi}}\times\widehat{\vec{\Pi}})\phi = (\widehat{\vec{p}}+q\vec{A})\times(\widehat{\vec{p}}+q\vec{A})\phi = (-i\vec{\nabla}+q\vec{A})\times(-i\vec{\nabla}+q\vec{A})\phi = -(\vec{\nabla}+iq\vec{A})\times(\vec{\nabla}+iq\vec{A})\phi\,. \tag{2.21.13}$$

The x component of this expression is

$$\begin{aligned} & -(\partial_y + iqA_y)(\partial_z + iqA_z)\phi + (\partial_z + iqA_z)(\partial_y + iqA_y)\phi \\ = \; & -\left[\partial_y(iqA_z\phi) - \partial_z(iqA_y\phi) + iqA_y\partial_z\phi - iqA_z\partial_y\phi\right] \\ = \; & -iq[\partial_y A_z - \partial_z A_y]\phi \\ = \; & -iqB_x\phi\,, \end{aligned} \tag{2.21.14}$$

and hence

$$(\widehat{\vec{\Pi}}\times\widehat{\vec{\Pi}})\phi = -iq\vec{B}\phi\,. \tag{2.21.15}$$

Now

$$(\vec{\sigma} \cdot \widehat{\vec{\Pi}})^2 \phi = \widehat{\vec{\Pi}} \cdot \widehat{\vec{\Pi}} \phi + q \vec{\sigma} \cdot \vec{B} \phi \,, \tag{2.21.16}$$

hence, the non-relativistic limit of the Dirac equation in an EM field (also called Pauli equation) may be written as

$$i \frac{\partial \phi}{\partial t} = \left(-qA^0 + \frac{(\widehat{\vec{p}} + q\vec{A})^2}{2m} + \frac{q \vec{\sigma} \cdot \vec{B}}{2m} \right) \phi \tag{2.21.17}$$

or writing $\vec{S} = \frac{\hbar}{2} \vec{\sigma}$ for the spin of the electron (setting $\hbar = 1$) we obtain

$$i \frac{\partial \phi}{\partial t} = \left(-qA^0 + \frac{(\widehat{\vec{p}} + q\vec{A})^2}{2m} + \frac{q \vec{S} \cdot \vec{B}}{m} \right) \phi \tag{2.21.18}$$

where ϕ is a two-component wave function for the non-relativistic spin-1/2 particle.

Comparing with the usual form of the non-rel. Schrödinger equation

$$i \frac{\partial \Psi}{\partial t} = \left(-\frac{1}{2m} \vec{\nabla}^2 + V \right) \Psi \,, \tag{2.21.19}$$

we can interpret the last term in eqn. (2.21.18) as a potential energy $-\vec{\mu}_{spin} \cdot \vec{B}$ due to the spin magnetic moment of the electron in an external magnetic field. Thus the spin magnetic moment is

$$\vec{\mu}_{spin} = -\frac{q\vec{S}}{m} \equiv -g \frac{q\vec{S}}{2m} \tag{2.21.20}$$

where g is the so-called gyromagnetic ration. Hence, the Dirac equation predicts $g = 2$ whereas classically we would expect $g = 1$. This prediction was confirmed experimentally and is one of the spectacular successes of the Dirac equation! Including radiative correction from Quantum Electrodynamics (QED) yields a more precise value of $g = 2(1.0011\ldots)$ which agrees up to nine digits after the dot with experiment!

2.22 Hydrogen Atom Spectrum

In the presence of an electrostatic potential $V(r)$ the Dirac equation becomes

$$\widehat{H} \Psi = (\vec{\alpha} \cdot \widehat{\vec{p}} + \beta m + V(r)) \Psi = i \frac{\partial \Psi}{\partial t} \,, \tag{2.22.1}$$

for positive energy solutions with energy eigenvalue $E > 0$ we make an separation ansatz

$$\Psi = e^{-iEt} \Psi_0(r, \theta, \phi) \tag{2.22.2}$$

so that

$$i\partial_t \Psi = E\Psi \tag{2.22.3}$$

which gives a time independent equation

$$(\vec{\alpha} \cdot \widehat{\vec{p}} + \beta m + V(r))\Psi = E\Psi \,. \tag{2.22.4}$$

For a Hydrogen-like atom we take

$$V(r) = -\frac{Z\alpha}{r} \ , \ \ \alpha = \frac{e^2}{4\pi} \,. \tag{2.22.5}$$

The total AM operator $\vec{J}$ commutes with $(\vec{\alpha} \cdot \widehat{\vec{p}} + \beta m)$ as in section 2.7. Also $\vec{J}$ commutes with $V(r)$ because $V(r)$ is independent of Spin and as in section 1.3, orbital AM operator $\vec{L}$ commutes with $V(r)$. Thus $[\vec{J}, \widehat{H}] = 0$.

The problem can be solved using simultaneous eigenstates $\psi^l_{j,m}$ of $\vec{J}^2$, J_z and the parity operator P (which takes $\vec{x} \to -\vec{x}$). The corresponding quantum numbers are $j(j+1)$, m and $(-1)^l$, where l is orbital angular momentum.

For the spin-1/2 electron, the allowed values of j are $j = l \pm 1/2$. The gory details of the calculation can be found in section 2.3.2 of [6], with the result for the energy levels

$$E_{n,j} = m_e \left[1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2} - \frac{1}{2}\frac{Z^4\alpha^4}{n^3}\left(\frac{1}{j+1/2} - \frac{3}{4n} \right) + \mathcal{O}((Z\alpha)^6) \right] \tag{2.22.6}$$

This result predicts correctly the splitting of the energy levels with the same principle quantum number n but different j (Fine Splitting); It does not predict the observed splitting of energy levels with the same n and j but different parity $(-1)^l$ (Lamb Shift). This requires the quantization of the EM field $A^\mu$and, hence, the use of Quantum Electrodynamics (QED). Other important quantum corrections are discussed in [6].

(draw the energy levels with $n = 1$ and $n = 2$)

Bibliography

[1] J. Bjorken and S. Drell, *Relativistic Quantum Mechanics* and *Relativistic Quantum Fields*, McGraw-Hill.

[2] M.E. Peskin and D.V. Schröder, *An Introduction to Quantum Field Theory*, Addison-Wesley.

[3] S. Weinberg, *The Quantum Theory of Fields, Vol 1*, Cambridge University Press.

[4] L.H. Ryder, *Quantum Field Theory*, Cambridge University Press.

[5] A. Zee, *Quantum Field Theory in a Nutshell*, Princeton University Press.

[6] C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, McGraw-Hill.

[7] I.J. Aitchison and A.J. Hey, *Gauge Theories in Particle Physics*, Dover Publications.

[8] F. Mandl and G. Shaw, *Quantum Field Theory*, John Wiley & Sons.

[9] W. Greiner, *Relativistic Quantum Mechanics: Wave Equations*, Springer-Verlag.