Dirac particles and electromagnetic interactions

Bill Celmaster

May 8, 2021

1 Classical theory of a charged particle in an EM field

1.1 Classical equation of motion

- The fields of electrodynamics are the 4-tuplet of the vector potential $A^\mu(x)$. We will refer to A_0 as the scalar potential ϕ.
- The familiar fields of electromagnetism are

$$\begin{aligned} \mathbf{E} &= -\nabla\phi - \frac{1}{c}\partial_0 \mathbf{A} \\ \mathbf{B} &= \nabla \times \mathbf{A} \end{aligned} \tag{1}$$

  where the nabla operator ∇ is defined by $\nabla_j = \frac{\partial}{\partial q_j}$. (When we work in the non-relativistic limit, we don't distinguish between upper and lower indices.)
- For particles in the non-relativistic limit, we have

$$\frac{d}{dt}(m\mathbf{v}) = e\left(\mathbf{E} + \frac{1}{c}\mathbf{v} \times \mathbf{B}\right) \tag{2}$$

1.2 The classical Hamiltonian

In this section, the takeaway is

$$H == \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + e\phi$$

- Recall that the connection between classical and quantum mechanics, is made via the Lagrangian and Hamiltonian. We first show that eq.(2) can be derived from the Euler-Lagrange equations for the Lagrangian $L = \frac{m\dot{\mathbf{q}}^2}{2} + e\left(\frac{\dot{\mathbf{q}}\cdot\mathbf{A}}{c} - \phi\right)$. We treat the electromagnetic field as non-dynamic. That is, we regard it as external and unaffected by the charged particle so that the EM fields don't appear in the Euler-Lagrange equations.

  - Let the classical action $S = \int dtL$. The E-L equations are equivalent to the variational equation $\frac{\delta S}{\delta q_i} = 0$.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \tag{3}$$

Notice the total derivative on the LHS. We will expand that as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \left(\frac{\partial}{\partial t} + \frac{\partial q_j}{\partial t}\frac{\partial}{\partial q_j} + \frac{\partial \dot{q}_j}{\partial t}\frac{\partial}{\partial \dot{q}_j}\right)\frac{\partial L}{\partial \dot{q}_i}.$$

Then plugging the Lagrangian into the Euler-Lagrange equations, and noting that $\frac{\partial L}{\partial q_i} = e\left(-\frac{\partial \phi}{\partial q_i} + \frac{1}{c}(\dot{\mathbf{q}}\cdot\nabla)A_i\right)$ and $p_i \equiv \frac{\partial L}{\partial \dot{q}_i} = m\dot{q}_i + \frac{e}{c}A_i$ (we've introduced the definition of the canonical momentum p_i) we get[1]

$$\begin{aligned} e\left(-\frac{\partial \phi}{\partial q_i} + \frac{1}{c}(\dot{\mathbf{q}}\cdot\nabla)A_i\right) &= \left(\frac{\partial}{\partial t} + \frac{\partial q_j}{\partial t}\frac{\partial}{\partial q_j}\right)\left(m\dot{q}_i + \frac{e}{c}A_i\right) \\ &= \frac{d}{dt}(m\dot{q}_i) + \frac{e}{c}\left(\frac{\partial A_i}{\partial t} + \dot{q}_j\frac{\partial A_i}{\partial q_j}\right) \\ &= \frac{d}{dt}(m\dot{q}_i) + \frac{e}{c}\frac{\partial A_i}{\partial t} + \frac{e}{c}\left(-(\dot{\mathbf{q}}\times(\nabla\times\mathbf{A}))_i + (\dot{\mathbf{q}}\cdot\nabla)A_i\right) \end{aligned} \tag{4}$$

After an easy simplification we then end up with

$$\frac{d}{dt}m\mathbf{v} = e\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right).$$

This confirms our choice of Lagrangian.

[1] The Lagrangian has no explicit time dependence, so the $\frac{\partial}{\partial t}$ term makes no contribution.

- The Hamiltonian H is defined as $\dot{\mathbf{q}} \cdot \mathbf{p} - L$, where all instances of $\dot{\mathbf{q}}$ are replaced by $\frac{1}{m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)$. Then

$$\begin{aligned} H &= \frac{1}{m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right) \cdot \mathbf{p} - \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 - \frac{e}{c}\left(\mathbf{p} - \frac{e}{mc}\mathbf{A}\right) \cdot \mathbf{A} + e\phi \\ &= \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + e\phi \end{aligned} \tag{5}$$

1.3 The magnetic moment

For a uniform (i.e., constant) magnetic field, the field A_i can be written (the form is known as the 'symmetric gauge') as

$$\mathbf{A} = \frac{1}{2}\mathbf{r} \times \mathbf{B}.$$

Then

$$\frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 = \frac{\mathbf{p}^2}{2m} - \frac{e}{2mc}\mathbf{L} \cdot \mathbf{B} + \frac{e^2}{8mc^2}\left(r^2\mathbf{B}^2 - (\mathbf{r} \cdot \mathbf{B})^2\right)$$

2 Non-relativistic quantum theory of a charged particle in an EM field

In this section, I follow the prescription that we find in the earliest expositions of quantum mechanics.

- In the classical theory, identify the Hamiltonian, and write it in terms of the coordinate and momentum variables. We did this in the last section. To repeat:

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + e\phi \tag{6}$$

- Let $\xi(t, \mathbf{x})$ be the electron wave-function and interpret H and $\mathbf{P}$ as operators, then write

$$H\xi(t, \mathbf{x}) = \left(\frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + e\phi\right)\xi(t, \mathbf{x}). \tag{7}$$

- Finally make the identifications

$$H \to i\frac{\partial}{\partial t}, \quad p_j \to -i\frac{\partial}{\partial x^j} \tag{8}$$

so the Schrodinger equation becomes

$$i\frac{\partial \xi(t, \mathbf{x})}{\partial t} = \left(\frac{1}{2m}\left(-i\nabla - \frac{e}{c}\mathbf{A}\right)^2 + e\phi\right)\xi(t, \mathbf{x}). \tag{9}$$

3 Dirac particles review

- The Dirac action for a <u>free</u> spin-$\frac{1}{2}$ particle:

$$\mathcal{S} = \int d^4x \mathcal{L} = \int d^4x \psi^\dagger(x)\gamma_0 \left(i\partial\!\!\!/ - m\right)\psi(x) \tag{10}$$

  - ψ is a 4-spinor – a vector with 4 complex components.
  - $\gamma^0, \gamma^1, \gamma^2, \gamma^3$ are 4x4 matrices.
  - $\partial\!\!\!/ \equiv \sum_\mu \gamma^\mu \partial_\mu$

- Equation of motion:

$$\frac{\delta \mathcal{S}}{\delta \psi^\dagger} = 0$$

  implies the Dirac equation

$$\boldsymbol{\left(i\partial\!\!\!/ - m\right)\psi(x) = 0.}$$

- General solution of Dirac equation:

$$\psi(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}^m}} \left(a_p^s u_{\mathbf{p}}^s e^{-ipx} + b_p^{s*} v_{\mathbf{p}}^s e^{ipx}\right), \tag{11}$$

  - $\omega_{\mathbf{p}}^m = p_0 = \sqrt{m^2 + \mathbf{p}^2}$.
  - s is called the spin and ranges from 1 to 2.
  - $u_{\mathbf{p}}^s$ and $v_{\mathbf{p}}^s$ are complex 4-vectors chosen with appropriate orthonormality conditions

$$\begin{aligned} u_{\mathbf{p}}^{s\dagger} u_{\mathbf{p}}^{s'} &= 2\omega_{\mathbf{p}}^m \delta_{ss'} = v_{\mathbf{p}}^{s\dagger} v_{\mathbf{p}}^{s'} \\ v_{\mathbf{-p}}^{s\dagger} u_{\mathbf{p}}^{s'} &= u_{-\mathbf{p}}^{s\dagger} v_{\mathbf{p}}^{s'} = 0 \\ \sum_s u_{\mathbf{p}i}^s u_{\mathbf{p}j}^{s*} &= \left(p\!\!\!/\gamma^0 + m\gamma^0\right)_{ij} \\ \sum_s v_{\mathbf{p}i}^s v_{\mathbf{p}j}^{s*} &= \left(p\!\!\!/\gamma^0 - m\gamma^0\right)_{ij} \end{aligned} \tag{12}$$

    In these expressions, $p\!\!\!/ = \gamma^0\omega_{\mathbf{p}}^m + \gamma^i p_i$.
  - a_p^s and b_p^s are coefficients.

- Express a_p^s and b_p^s in terms of ψ by employing inverse Fourier transform:

$$
\begin{aligned}
a_p^{s\dagger} &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}^m}} \int d^3\mathbf{x} \psi^\dagger(t,\mathbf{x}) u_{\mathbf{p}}^s e^{-ipx} \\
b_p^{s\dagger} &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}^m}} \int d^3\mathbf{x} v_{\mathbf{p}}^{s\dagger} \psi(t,\mathbf{x}) e^{-ipx}
\end{aligned}
\tag{13}
$$

In the free theory, these operators are time-independent.

- Promote ψ, a_p^s and b_p^s to operators on a Hilbert space TBD. The canonical commutation relations for the fields in the Lagrangian are:

$$
\begin{aligned}
\{\psi^\alpha(x), \frac{\partial \mathcal{L}}{\partial \dot{\psi}_\beta}(y)\} &= \{\psi^\alpha(x), i\psi^{\dagger\beta}(y)\} \\
&= i\delta(\mathbf{x}-\mathbf{y})\delta^{\alpha\beta} \\
\{\psi^\alpha(x), i\psi^\beta(y)\} &= 0 \\
\{\psi^{\dagger\alpha}(x), i\psi^{\dagger\beta}(y)\} &= 0
\end{aligned}
\tag{14}
$$

and these imply from eq. (13)

$$
\begin{aligned}
\{a_p^s, a_{p'}^{\dagger s'}\} &= (2\pi)^3\delta^3(\mathbf{p}-\mathbf{p}')\delta^{ss'} \\
\{b_p^s, b_{p'}^{\dagger s'}\} &= (2\pi)^3\delta^3(\mathbf{p}-\mathbf{p}')\delta^{ss'} \\
\{a_p^s, a_{p'}^{s'}\} = \{a_p^{\dagger s}, a_{p'}^{\dagger s'}\} &= \{b_p^s, b_{p'}^{s'}\} = \{b_p^{\dagger s}, b_{p'}^{\dagger s'}\} = 0
\end{aligned}
\tag{15}
$$

- We can now generate the Hilbert space, as a Fock space constructed from one-particle states $|s,p\rangle = a_p^{s\dagger}|\Omega\rangle$ and one-antiparticle states $|s,p\rangle' = b_p^{s\dagger}|\Omega\rangle$, where $|\Omega\rangle$ denotes the vacuum state. We call $a_p^{s\dagger}$ and $b_p^{s\dagger}$ *creation operators.*

4 Quantum Mechanics review

- Start with the Heisenberg picture since QFT is in the Heisenberg picture. Operators evolve in time as

$$
\mathcal{O}(t) = e^{iHt}\mathcal{O}(0)e^{-iHt}. \tag{16}
$$

States don't evolve in time, in the Heisenberg picture. We describe states using bra-ket notation. e.g. $|\xi\rangle$ or (the dual) $\langle\xi|$. **Don't get confused by the notation ξ. We'll end up talking about this as a wavefunction and <u>not</u> a field.**

- We'll also end up using the Schrodinger picture. In that picture, operators don't evolve, but states do. The Schrodinger state $|\xi(t)\rangle$ is defined as

$$|\xi(t)\rangle = e^{-iHt}|\xi\rangle. \tag{17}$$

  We see from this, that the Heisenberg-picture state $|\xi\rangle$ is related to the Schrodinger-picture state by

$$|\xi\rangle = |\xi(0)\rangle.$$

- If we take derivatives in the Schrodinger picture, we get

$$i\frac{d}{dt}|\xi(t)\rangle = H|\xi(t)\rangle.$$

  You should recognize this as the so-called time-dependent Schrodinger equation. [2]

- A more familiar form of the Schrodinger equation is the wavefunction form. We take the inner product of both sides by applying the ket $\langle x|$ which is an eigenvector of the position operator $\hat{x}$.

$$i\frac{d}{dt}\langle x|\xi(t)\rangle = \langle x|H|\xi(t)\rangle.$$

  We usually write $\langle x|\xi(t)\rangle \equiv \xi(t,x)$.

  In QFT, the position operator has a diminished role relative to quantum mechanics. Remember, in QFT, time (which is not regarded as an operator in ordinary quantum mechanics) and position are both parameters of the fields and have similar roles.

- To connect QM to QFT, let's focus on the 4-vectors of states $|\mathbf{x}\rangle_i$ and $|\mathbf{x}\rangle_i'$ rather than the operator $\hat{\mathbf{x}}$. We do that by defining $|\mathbf{x}\rangle_i$ as a superposition of states $|s,\mathbf{p}\rangle = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}a_{\mathbf{p}}^{s\dagger}|\Omega\rangle$ and similarly with $|\mathbf{x}\rangle_i'$ as

[2] There is also a time-independent Schrodinger equation

$$H|\xi\rangle = E|\xi\rangle,$$

also known as the energy eigenvalue equation. **Note that the time-dependent and time-independent Schrodinger equations have nothing to do with one another!**. The time-independent Schrodinger equation can be useful for solving the time-dependent Schrodinger equation.

a superposition of states $|s,p\rangle' = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} b_{\mathbf{p}}^{s\dagger}|\Omega\rangle$. We'll define these as Heisenberg states, so will set $t = 0$.

$$
\begin{aligned}
{}_i\langle \mathbf{x}| &= \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \langle s,\mathbf{p}|(u_{\mathbf{p}}^s)_i = \langle\Omega|\psi_i(0,\mathbf{x}) \\
{}_i\langle \mathbf{x}|' &= \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \langle s,\mathbf{p}|'(v_{\mathbf{p}}^{s*})_i = \langle\Omega|\psi_i^\dagger(0,\mathbf{x})
\end{aligned} \tag{18}
$$

where $|s,\mathbf{p}\rangle$ is defined as an eigenstate of the momentum operator $\hat{\mathbf{p}}$ whose eigenvalue[3] is $\mathbf{p}$. Similarly with $|s,\mathbf{p}\rangle'$.

- We can now define spinor particle and antiparticle wavefunctions by

$$
\begin{aligned}
\xi_i(\mathbf{x}) &= {}_i\langle\mathbf{x}|\xi\rangle = \langle\Omega|\psi_i(0,\mathbf{x})|\xi\rangle \\
\tilde{\xi}_i(\mathbf{x}) &= {}_i\langle\mathbf{x}|'\xi\rangle = \langle\Omega|\psi_i^\dagger(0,\mathbf{x})|\xi\rangle
\end{aligned} \tag{19}
$$

- As an example of how to make contact with something familiar from QM, consider the action of the momentum operator.

$$
\begin{aligned}
{}_i\langle\mathbf{x}|\hat{\mathbf{p}}|\xi\rangle &= \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \langle s,\mathbf{p}|\hat{\mathbf{p}}|\xi\rangle (u_{\mathbf{p}}^s)_i \\
&= \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} \mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} \langle s,\mathbf{p}|\xi\rangle (u_{\mathbf{p}}^s)_i \\
&= \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} (-i\nabla)\, e^{i\mathbf{p}\cdot\mathbf{x}} \langle s,\mathbf{p}|\xi\rangle (u_{\mathbf{p}}^s)_i \\
&= (-i\nabla) \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \langle s,\mathbf{p}|\xi\rangle (u_{\mathbf{p}}^s)_i \\
&= (-i\nabla)\ {}_i\langle\mathbf{x}|\hat{\mathbf{p}}|\xi\rangle = (-i\nabla)\,\xi_i(\mathbf{x}).
\end{aligned} \tag{20}
$$

Similarly ${}_i\langle\mathbf{x}|'\hat{\mathbf{p}}|\xi\rangle = (-i\nabla)\,\hat{\xi}_i(\mathbf{x})$.

5 Field equations with an external electromagnetic field

The full theory of the universe doesn't involve any *external* fields. However, it's useful in practice, to isolate "large" sources of electromagnetic fields, from

[3] Strictly speaking, what I mean is that $\mathbf{p}$ is a triplet of eigenvalues.

the particles that interact with them. In Appendix A, I discuss a systematic way of deriving from a full dynamic theory, the external-field theory which approximates the effect of large sources on electrons, etc. In what follows, I simply assume the result summarized in the following

- The (external) fields of electrodynamics are the 4-tuplet of the vector potential $A^\mu(x)$. Commonly, A_0 is written as the scalar potential ϕ. The external fields are **NOT** operators on a Hilbert space. They are just real functions (sometimes called **c-numbers**).

- The electric and magnetic fields are

$$\begin{aligned} E_i &= -\partial_i A_0 + \partial_0 A_i \\ B_i &= \epsilon^{ijk}\partial_j A_k \end{aligned} \tag{21}$$

- The Lagrangian for a Dirac particle in the presence of an external electromagnetic field is

$$\mathcal{L} = \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi - e\bar{\psi}\gamma^\mu A_\mu\psi \tag{22}$$

  where A_μ is an external field – by which I mean that it is a real-valued 4-vector whose values are 'given', rather than derived.

- Recall from months ago, our exploration of quantum field theory from the perspective of path integrals. The basic idea was that all physical quantities of interest could be derived from an integral over variables (fields) of the exponential of the action – the action being a function of those variables. One of the things we learned was that if the action was quadratic in those variables, the integral could be (in principle) computed exactly. What that tells us, is that the above Lagrangian (and its integral, the action) leads to an exactly solvable quantum field theory. It will turn out that the solutions can't be expressed in closed form, but that they can be expanded in powers of $|\mathbf{p}|/m$ – i.e. a relativistic expansion.

- The variation equation

$$\frac{\delta\mathcal{S}}{\delta\psi^\dagger} = 0$$

  implies the equation of motion

$$\left(i\partial\!\!\!/ - m - e\gamma^\mu A_\mu(x)\right)\psi(x) = 0. \tag{23}$$

- For future reference, expand and rewrite this equation as

$$i\partial_0\psi(x) = -i\gamma^0\gamma^i\partial_i\psi(x) + \left(m\gamma^0 + e\gamma^0\gamma^\mu A_\mu(x)\right)\psi(x) \tag{24}$$

  where I've multiplied both sides by γ_0 in order to rewrite terms.

- Promote ψ to the operator $\hat{\psi}$. Because A_μ is an external field, we don't promote that to an operator and treat it instead as a complex-valued vector.

- Now we proceed by a sequence of steps which bring this into the form of the time-dependent Schrodinger equation.

  – We will be interested in the spinor particle wavefunctions $\xi_i(t, \mathbf{x})$ defined as $\langle\Omega|\psi_i(0, \mathbf{x})|\xi(t)\rangle$ where the Schrodinger state, defined in eq. (17) is $|\xi(t)\rangle = e^{-iHt}|\xi\rangle$. We see from eq. (19) that

$$\begin{aligned} \xi_i(t, \mathbf{x}) &= \langle\Omega|\psi_i(0, \mathbf{x})e^{-iHt}|\xi\rangle \\ &= \langle\Omega|e^{-iHt}e^{iHt}\psi_i(0, \mathbf{x})e^{-iHt}|\xi\rangle \\ &= \langle\Omega|e^{iHt}\psi_i(0, \mathbf{x})e^{-iHt}|\xi\rangle \\ &= \langle\Omega|\psi_i(t, \mathbf{x})|\xi\rangle \end{aligned} \tag{25}$$

  where we've used the fact that the vacuum is invariant under the time-translation operation, and we've also used the fact that the field transforms as the Heisenberg operator in eq. (16),

$$\psi_i(t, \mathbf{x}) = e^{iHt}\psi_i(0, \mathbf{x})e^{-iHt}.$$

  Similarly, we can see that

$$\tilde{\xi}_i(t, \mathbf{x}) = \langle\Omega|\psi_i^\dagger(t, \mathbf{x})|\xi\rangle.$$

  Note that in the non-interacting theory $\langle\Omega|\psi_i(0, \mathbf{x})|\xi(t)\rangle =_i \langle\mathbf{x}|\xi(t)\rangle$, as explained in eq. (18).

  – Next, we'll take the time-derivative of the 4-spinor $\xi(t, \mathbf{x})$, and then invoke eq. (24).

$$\begin{aligned} i\partial_t\xi(t, \mathbf{x}) &= \langle\Omega|i\partial_t\psi(t, \mathbf{x})|\xi\rangle \\ &= \langle\Omega| - i\gamma^0\gamma^j\partial_j\psi(t, \mathbf{x}) + \left(m\gamma^0 + e\gamma^0\gamma^\mu A_\mu(t, \mathbf{x})\right)\psi(t, \mathbf{x})|\xi\rangle \\ &= \gamma^0\gamma^j\left(-i\partial_j\right)\xi(t, \mathbf{x}) + \left(m\gamma^0 + e\gamma^0\gamma^\mu A_\mu(t, \mathbf{x})\right)\xi(t, \mathbf{x}) \end{aligned} \tag{26}$$

  Similarly,

$$i\partial_t\tilde{\xi}(t, \mathbf{x}) = -\gamma^{0*}\gamma^{j*}\left(i\partial_j\right)\tilde{\xi}(t, \mathbf{x}) - \left(m\gamma^{0*} + e\gamma^{0*}\gamma^{\mu*}A_\mu(t, \mathbf{x})\right)\tilde{\xi}(t, \mathbf{x}) \tag{27}$$

- Before leaving the antiparticle equation, we'll transform it into a form more similar to the particle equation, following pp. 192 and 193 of Schwartz. Define $\xi_c = -i\gamma^2\tilde{\xi}$. This is known as the charge-conjugate of ξ and continues to describe the antiparticle wavefunction. Define new γ-matrices by $\hat{\gamma^\mu} = \gamma^2\gamma^{\mu *}\gamma^2$. It isn't hard to see that these new γ-matrices satisfy the Dirac algebra and therefore can be used in the Dirac equation. The previous equation becomes (by inserting $\gamma^2\gamma^2 = -\mathbf{I}$ in appropriate places)

$$i\partial_t \xi_c(t, \mathbf{x}) = \hat{\gamma}^0\hat{\gamma}^j\left(-i\partial_j\right)\xi_c(t, \mathbf{x}) + \left(m\hat{\gamma}^0 - e\hat{\gamma}^0\hat{\gamma}^\mu A_\mu(t, \mathbf{x})\right)\xi_c(t, \mathbf{x}) \tag{28}$$

This looks just like the particle equation except that e is replaced by $-e$. In other words, it describes the motion of a particle whose mass is the same as the original particle but whose charge is opposite.

- To proceed, I will follow the outline of Problem 10.1 in Schwartz's text applied to the particle equation. It will be useful to write the spinor ξ as $\begin{pmatrix}\xi^L \\ \xi^R\end{pmatrix}$ where ξ^L and ξ^R are each 2-spinors. Similarly, we will write $\tilde{\xi}$ as $\begin{pmatrix}\tilde{\xi}^L & \tilde{\xi}^R\end{pmatrix}$.

  * Rewrite eq. (26), suppress the $(t, \mathbf{x})$ arguments and expand the matrices.

$$\begin{aligned}\left(i\partial_t - eA_0\right)\xi &= \gamma^0\gamma^j\left(-i\partial_j\right)\xi + \left(m\gamma^0 + e\gamma^0\gamma^j A_j\right)\xi \\ &= \begin{pmatrix} -\left(-i\partial_j + eA_j\right)\sigma_j & m\mathbf{I}_{2\text{x}2} \\ m\mathbf{I}_{2\text{x}2} & \left(-i\partial_j + eA_j\right)\sigma_j\end{pmatrix}\xi\end{aligned} \tag{29}$$

  * We can then re-apply the operator on the left hand side to obtain

$$\left(i\partial_t - eA_0\right)^2\xi = \begin{pmatrix}\alpha & 0 \\ 0 & \alpha\end{pmatrix}\xi \tag{30}$$

    where

$$\alpha = \left(-i\partial_j + eA_j\right)\left(-i\partial_k + eA_k\right)\sigma_j\sigma_k + m^2\mathbf{I}_{2\text{x}2}.$$

    We simplify α by noting a couple of things about the terms with $\sigma_j\sigma_k$. First, there is the Pauli matrix identity $\sigma_j\sigma_k = i\epsilon^{jkl}\sigma_l + \delta_{jk}$. Second, note that for an arbitrary function f, we have $-i\partial_j(A_k f) - iA_j(\partial_k f) = -i\left(\partial_j A_k\right)f - i\left(A_k\partial_j + A_j\partial_k\right)f$.

Apply these observations to α.[4]

$$\begin{aligned}\alpha &= \left(-\partial_j\partial_j + e^2 A_j A_j - 2iA_j\partial_j - i\left(\partial_j A_j\right) + m^2\right)\mathbf{I}_{2\text{x}2} + e\epsilon^{jkl}\partial_j A_k \sigma_l \\ &= \left((-i\nabla - e\mathbf{A})^2 + m^2\right)\mathbf{I}_{2\text{x}2} - e\mathbf{B}\cdot\sigma\end{aligned} \tag{31}$$

where $\mathbf{B}$ is the magnetic field. All of this can be written in slightly more conventional notation by recognizing, in the Schrodinger equation, that the operator $i\partial_t$ is the Hamiltonian H, and the operator $-i\nabla$ is the momentum $\mathbf{p}$ (see for example eqs. (20)). Then our Schrodinger equation can be written as

$$(H - eA_0)^2\,\xi = \begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}\xi \tag{32}$$

where

$$\begin{aligned}\alpha &= \left((\mathbf{p} - e\mathbf{A})^2 + m^2\right)\mathbf{I}_{2\text{x}2} - e\mathbf{B}\cdot\sigma \\ &= \left((c\mathbf{p} - e\mathbf{A})^2 + m^2c^4\right)\mathbf{I}_{2\text{x}2} - e\hbar c\mathbf{B}\cdot\sigma\end{aligned} \tag{33}$$

In the last line, we re-introduced units that had been suppressed.

* Now we'll take the square-root of the operators on both sides and expand to leading order in c.[5] We set up the solution and expansion by first examining eq. (26). From this, we can see that the LHS operator is just $(H - eA_0)\,\mathbf{I}_{4\text{x}4}$ and the RHS operator is $\begin{pmatrix}0 & \sqrt{\alpha}\\ \sqrt{\alpha} & 0\end{pmatrix}$. We define the square-root of α by

[4] In trying to reconcile signs, I've come across a rather bothersome set of conventions. Hopefully I've understood them. In classical pre-relativistic EM, there was no notion of covariant and contravariant tensors, so if $\mathbf{V}$ is a non-relativistic 3-vector, then $V_i = V^i$. However, if V is a 4-vector, then $V_i = -V^i$. Where we get into trouble, is when we use the same letter for both the 3-vector and the 4-vector. Here's an example. Consider the 3-vector $\mathbf{x}$. This is meant to denote the triplet (x, y, z). What about the 4-vector x? $x_i = -x^i$. Which one is y (for example)? Is it x_2 or is it x^2? We need to pick a convention, in this case x^2. In other words, our convention is that $x^i = \mathbf{x}_i = \mathbf{x}^i$, where we use the boldface to denote the 3-vector. For momentum, we use the opposite convention, $p_i = \mathbf{p}_i$. Other conventions required here are $\nabla_i = \partial_i = \frac{\partial}{\partial^i}$, $\mathbf{A}_i = \mathbf{A}^i = -A_i$ (see Cambridge EM notes of Tong) and $\mathbf{B} = \nabla_i \times \mathbf{A}$.

[5] **WARNING**: These are 4x4 matrices, some of whose components are operators (like the gradient), so the square-root has many branches. At the very least, each of the diagonal entries can be plus or minus. More significantly, the square-root of a diagonal matrix can be non-diagonal. However, the original equation was linear, so the ambiguities are actually fictitious. We'll choose the square-root by solving the linear equation to leading order in c and then expand around that solution. This will give a unique square-root.

again examining eq. (26) and noting that the leading term in c is $mc^2\mathbf{I}_{2\text{x}2}$.

$$\begin{aligned}\sqrt{\alpha} &= mc^2\sqrt{\left(\mathbf{I}_{2\text{x}2} + \frac{1}{m^2c^4}\left(c\mathbf{p} - e\mathbf{A}\right)^2\mathbf{I}_{2\text{x}2} - \frac{e\hbar}{m^2c^3}\mathbf{B}\cdot\sigma\right)} \\ &= mc^2\mathbf{I}_{2\text{x}2} + \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2\mathbf{I}_{2\text{x}2} - \frac{e\hbar}{2mc}\mathbf{B}\cdot\sigma + \mathcal{O}(\frac{1}{c^2})\end{aligned} \tag{34}$$

We could leave things explicitly in terms of the Pauli matrices σ, but often we prefer to express equations using the spin $\mathbf{S} = 2\hbar\sigma$. To remind you, the spin is defined to be the angular momentum – in units of $\hbar$ – for a spin-1/2 particle, or in our group-theoretic language, the spin is the operator that generates the 2D representation of rotations. Putting everything together, our Schrodinger-style equation looks like

$$\begin{aligned}\left(H - eA_0\right)\xi^L &= \left(mc^2\mathbf{I}_{2\text{x}2} + \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2\mathbf{I}_{2\text{x}2} - \frac{e}{mc}\mathbf{B}\cdot\mathbf{S} + ...\right)\xi^R \\ \left(H - eA_0\right)\xi^R &= \left(mc^2\mathbf{I}_{2\text{x}2} + \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2\mathbf{I}_{2\text{x}2} - \frac{e}{mc}\mathbf{B}\cdot\mathbf{S} + ...\right)\xi^L.\end{aligned} \tag{35}$$

We can simplify further by defining $\xi^+ = \frac{\xi^L + \xi^R}{\sqrt{2}}$. Then

$$\left(H - eA_0\right)\xi^+ = \left(mc^2\mathbf{I}_{2\text{x}2} + \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2\mathbf{I}_{2\text{x}2} - \frac{e}{mc}\mathbf{B}\cdot\mathbf{S} + ...\right)\xi^+. \tag{36}$$

Other than the overall constant mc^2, this equation is the Schrodinger-Pauli equation. Compare this to the non-relativistic Schrodinger equation, eq. (7) which looks similar except for the explicit spin term.

* The antiparticle equation looks the same except that the charge is reversed.

– As mentioned before, a further simplification can be made when the magnetic field is uniform (i.e. constant). In that case, the field A_i can be written as

$$\mathbf{A} = \frac{1}{2}\mathbf{r}\times\mathbf{B}.$$

We use this in our Schrodinger-Pauli equation by expanding $\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 =$

$\left(-i\hbar\nabla - \frac{e}{c}\mathbf{A}\right)^2$.

$$\begin{aligned}\left(-i\hbar\nabla - \frac{e}{c}\mathbf{A}\right)^2 &= -\hbar^2\nabla^2 + \frac{e^2}{4c^2}\left(\mathbf{r}^2\mathbf{B}^2 - (\mathbf{r}\cdot\mathbf{B})^2)\right) - i\hbar\frac{e}{c}\mathbf{r}\times\mathbf{B}\cdot\nabla \\ &= \mathbf{p}^2 - \frac{e}{c}(\mathbf{r}\times\mathbf{p})\cdot\mathbf{B} + \frac{e^2}{4c^2}\left(\mathbf{r}^2\mathbf{B}^2 - (\mathbf{r}\cdot\mathbf{B})^2)\right) \\ &= \mathbf{p}^2 - \frac{e}{c}\mathbf{L}\cdot\mathbf{B} + 2m\mathcal{D}\end{aligned} \tag{37}$$

where $\mathcal{D} = \frac{e^2}{4c^2}\left(\mathbf{r}^2\mathbf{B}^2 - (\mathbf{r}\cdot\mathbf{B})^2)\right)$ is known as the diamagnetic term, and is a very small effect.

Now we are ready to rewrite our Schrodinger-Pauli equation.

$$(H - eA_0)\,\xi^+ = \left(mc^2 + \frac{1}{2m}\left(\mathbf{p}^2 - \frac{e}{c}\mathbf{B}\cdot[\mathbf{L}+2\mathbf{S}]\right) + \mathcal{D} + ...\right)\xi^+. \tag{38}$$

An alternative derivation, based on the Dirac representation of gamma matrices and spinors, is provided in an Appendix.

- In a purely classical theory, the spin term is absent. The energy has a term proportional to the electron's angular momentum dotted into the magnetic field. This causes electrons, along certain trajectories, to be deflected in a magnetic field. The coefficient $\frac{e}{2mc}$ of $\mathbf{B}\cdot\mathbf{L}$ is known as the classical magnetic moment. In the quantum theory, there is an additional energy term – and hence deflection term – proportional to $\mathbf{B}\cdot\mathbf{S}$. The coefficient twice the classical magnetic moment. **We say that the electron quantum magnetic moment equals $\frac{g_e e}{2mc}$, where g_e is called the electron Landé g-factor**. According to the Dirac equation, we find that $g_e = 2$. It's common in the literature, to simply refer to g, but in a specific context. So, for example, in recent experiments, we hear that measurements are being made of $g-2$ for the muon. The muon, like the electron, is governed by the Dirac equation, so we might expect $g-2=0$. However, remember that we made a number of approximations. The most significant of these, was that we completely ignored the interplay between electrons and photons. In fact, photons played no role whatsoever in the Dirac equation, since we assumed the electric field was external – or in particular, that its dynamics were not affected in any way by the electrons (we were essentially ignoring Newton's third law). Once we take these interactions into account, there are corrections to g and it is no longer exacty 2.

A Field equations with quantum electrodynamics

In this appendix, I discuss the validity of treating the electromagnetic field as an external function in the Lagrangian. This subject does not get discussed in most texts or online lecture notes. In fact, I've only found one treatment and it's rather involved. For those interested, you can find this in Steven Weinberg "The Quantum Theory of Fields Volume I" section 13.6 (pages 556-562). This section depends critically on several prior specialized sections on Feynman diagrams etc. Here is a sketch of the key points

- The electric and magnetic fields can be expressed directly in terms of the tensor
$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{39}$$
- The Lagrangian for Dirac and electromagnetic fields was derived over the space of many years, based on a requirement of Lorentz invariance together with a connection to Planck's theory of quantum electromagnetism. For the problem of interest to us here, we assume a theory with two kinds of charged particles. The first charged particle is a light spin-1/2 fermion such as the electron. The second charged particle is a heavy fermion such as a proton (but the method will generalize even to charged scalars). Ultimately, we will treat the heavy particles as being the source(s) of the external electromagnetic field.
$$\begin{aligned} \mathcal{L} = &-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi} \left(i\gamma^\mu \partial_\mu - m \right) \psi - e\bar{\psi}\gamma^\mu A_\mu \psi \\ &+ \bar{\Psi} \left(i\gamma^\mu \partial_\mu - m \right) \Psi - q\bar{\Psi}\gamma^\mu A_\mu \Psi \end{aligned} \tag{40}$$
where the field Ψ represents the heavy particle.
- The variation equation
$$\frac{\delta \mathcal{S}}{\delta \psi^\dagger} = 0$$
implies the equation of motion
$$\left(i\partial\!\!\!/ - m - e\gamma^\mu A_\mu(x) \right) \psi(x) = 0. \tag{41}$$
This looks a bit like the Dirac equation. However, there are separate equations of motion for A_μ and Ψ in terms of the other fields. The resulting coupled set of equations can't be solved analytically.

- Our goal is to show that we can approximate all calculations involving the light fermion by assuming a Lagrangian of the form eq. (22).
- Start the analysis by considering the scattering of an electron off a heavy fermion. This can be treated in perturbation theory by adding up Feynman diagrams depicted on page 561 of Weinberg – namely, diagrams which include a heavy fermion with some number of photon lines emitted from it. A photon line represents a propagator with a 4-momentum that is the difference of the 4-momenta of the heavy fermion lines at the photon-fermion-antifermion vertex.
- Those propagators can be shown to be dominated by intermediate momenta that approach 0 (at which point the propagator acquires a pole). When the diagrams are added up, including only contributions of the pole-residues, the result is the same as if one had a Lagrangian with an external coulomb field.
- An arbitrary external field then represents a functional smearing of the initial state of the heavy fermion.
- The approximation obtains corrections of order k^2/p^2 where k is the typical electron energy and p is the typical heavy-particle energy. These corrections are smaller as the heavy-particle mass is larger. Of course, in practical situations, the heavy particles actually represent something like a magnet, which consists of large numbers of bound charges. But the Feynman-diagram argument shows how the formalism permits a separation into dynamical particles and external fields.

B Dirac representation version of the equations

- Again we assume A_μ is time-independent. For this section, I will work in the Dirac representation.

$$\gamma^0 = \begin{pmatrix} \mathbf{I}_2 & 0 \\ 0 & -\mathbf{I}_2 \end{pmatrix}, \gamma^i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{42}$$

It will be useful to write the spinor ξ as $\begin{pmatrix} \xi^+ \\ \xi^- \end{pmatrix}$ where ξ^+ and ξ^- are each 2-spinors. Similarly, we will write $\tilde{\xi}$ as $\begin{pmatrix} \tilde{\xi}^+ & \tilde{\xi}^- \end{pmatrix}$. It turns out that the relationship between the Weyl (chiral) representation that we've used in the main text, and this Dirac representation, implies that $\xi^+ = \frac{\xi^L + \xi^R}{\sqrt{2}}$ which is consistent with our definition of ξ^+ in the main text.

– Rewrite eq. (26), suppress the $(t, \mathbf{x})$ arguments and expand the matrices.

$$(i\partial_t - eA_0)\,\xi = \gamma^0\gamma^j\,(-i\partial_j)\,\xi + \left(m\gamma^0 + e\gamma^0\gamma^j A_j\right)\xi$$
$$= \begin{pmatrix} m\mathbf{I}_{2\text{x}2} & (-i\partial_j + eA_j)\,\sigma_j \\ (-i\partial_j + eA_j)\,\sigma_j & -m\mathbf{I}_{2\text{x}2} \end{pmatrix}\xi \tag{43}$$

– Following identifications made in the main text, we obtain

$$(H - eA_0)\,\xi^+ = mc^2\xi^+ + c\sigma\cdot\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)\xi^- \tag{44}$$

$$(H - eA_0)\,\xi^- = -mc^2\xi^- + c\sigma\cdot\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)\xi^+. \tag{45}$$

– The next steps follow the Wikipedia article "Electron magnetic moment" and seem to me to lean heavily on physical intuition rather than rigorous logic. We start by assuming the equations can be solved for eigenvectors of H, so we take $H\xi = E\xi$. The assumption seems reasonable, since the momentum and Hamiltonian ought to commute. Furthermore, we also assume that for slow particles, we have $E - eA_0 = mc^2\left(1 + \mathcal{O}\left(\frac{v}{c}\right)^2\right)$.[6] (I believe it would be possible to solve these equations starting with the assumption that $E - eA_0 \approx -mc^2$ but that's not an intuitively reasonable solution.)

– From eq. (44), our approximation implies that ξ^- is of order $\mathcal{O}\left(\frac{v}{c}\right)\xi^+$. So one conclusion already, is that the bottom part of the 4-spinor is suppressed relative to the top part of the 4-spinor. Then, from eq. (45) we obtain

$$2mc^2\xi^- = c\sigma\cdot\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)\xi^+ + \mathcal{O}\left(\frac{v}{c}\right)^2 mc^2 \tag{46}$$

from which we conclude

$$\xi^- = \frac{1}{2mc}\sigma\cdot\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)\xi^+ + ... \tag{47}$$

[6] It's also possible to start with the ansatz that $E = mc^2\left(1 + \mathcal{O}\left(\frac{v}{c}\right)^2\right)$ and after expansion in orders of $\frac{v}{c}$ and e, we'd come to the same conclusion.

We can then substitute back into eq. (44), and perform expansions as we did in the main text.

$$
\begin{aligned}
(E - eA_0 - mc^2)\xi^+ &= \frac{1}{2m}\left(\sigma \cdot \left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)\right)^2 \xi^+ + ... \\
&= \frac{1}{2m}\left(\mathbf{p}^2 - \frac{e}{c}\mathbf{B}\cdot[\mathbf{L} + 2\mathbf{S}] + \mathcal{D} + ...\right)\xi^+
\end{aligned}
\tag{48}
$$