The hydrogen atom and the Dirac equation – Part I: Massaging the Schrodinger equation for an electron

Bill Celmaster

June 6, 2021

This accompanies the QM youtube lecture 4.3 by Burton Zwiebach, `https://www.youtube.com/watch?v=tl7q_VZ3eIQ`, which is also covered in online notes `https://ocw.mit.edu/courses/physics/8-06-quantum-physics-iii-spring-2018/lecture-notes/MIT8_06S18ch2.pdf`.

1 Motivating the spin magnetic moment term in the Schrodinger equation

- Zwiebach defines the magnetic moment of a current loop as

$$\boldsymbol{\mu} = I\mathbf{A}$$

  where I've set $c = 1$ and I is the current in the loop. Also, the area vector $\mathbf{A}$ has magnitude A which is the area enclosed by the loop and has direction normal to the plane of the loop.

- Zwiebach then goes on to say that in the presence of a magnetic field $\mathbf{B}$ the energy of that current-loop is

$$H = -\boldsymbol{\mu} \cdot \mathbf{B}.$$

  This is something that can be derived from the Lorentz force of classical electromagnetism.

- The definition of the magnetic moment can be rewritten in terms of the classical angular momentum.

– The current in a circular loop is just the charge per unit time through a point in the loop.

$$
\begin{aligned}
I &= q\frac{v}{\text{loop-circumference}} \\
&= q\frac{v}{2\pi r} \\
&= q\frac{mv}{2m\pi r}
\end{aligned}
\tag{1}
$$

We recognize mv as the magnitude p of the electron's momentum. Then we note that $A = \pi r^2$ to get the magnitude

$$
\begin{aligned}
\mu &= q\frac{p}{2m\pi r}\pi r^2 \\
&= q\frac{rp}{2m} \\
&= \frac{qL}{2m}
\end{aligned}
\tag{2}
$$

where we recognize the classical angular momentum as $L = rp$, in the case that the radius vector is perpendicular to the velocity vector, as happens in the loop.

– This derivation leads us to

$$
\boldsymbol{\mu} = \frac{q}{2m}\mathbf{L}.
$$

2 Introducing Pauli spinors

Zwiebach, in lecture 4.3, evidently builds on some earlier lectures where he has introduced the idea that an electron is represented by a 2-component vector called a Pauli spinor.

I frankly don't know what was the concept or motivation behind his introduction of Pauli spinors so rather than guess, I'll make my own introduction. Previously, we encountered representations of the rotation group. Here's a reminder.

2.1 Representations and Pauli spinors

- A finite linear representation – up to a phase – of the rotation group is a set, $M(R)$ of matrix functions of rotations R so that $M(R_1 R_2) = M(R_1)M(R_2)e^{i\theta}$.

- In order to construct quantum theories with rotational symmetry, we deal with multiplet states that transform under rotations by the action of a matrix represention (up to a phase) of the rotation group.

- When the states form a doublet, they transform under 2D matrix representations (up to a phase) of the rotation group. The doublets are called **(Pauli) spinors**. Reminder:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3}$$

2.2 Lie Algebra of the Rotation Group

Consider a rotation around the x-axis $\mathbf{R_x}(\theta_\mathbf{x})$. Expand it as a Taylor expansion for small values of θ_x.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x \\ 0 & -\sin\theta_x & \cos\theta_x \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + i\theta_x \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} + \dots \tag{4}$$

Define $\mathbf{I} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\mathbf{J_x} \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$.

We call J_x the angular momentum around the x-axis. It is the quantum operator corresponding to classical angular momentum.

We can rewrite Equation (4) as $\mathbf{R_x}(\theta_\mathbf{x}) = \mathbf{I} + i\theta_x \mathbf{J_x} + \dots$. It turns out that if we continue the Taylor expansion, we find

$$\mathbf{R_x}(\theta_\mathbf{x}) = \mathbf{I} + i\theta_x \mathbf{J_x} + \frac{(i\theta_x^2)}{2}\mathbf{J_x}^2 + \dots = \exp(i\theta_x \mathbf{J_x}) \tag{5}$$

Note that

$$\mathbf{R_y}(\theta_\mathbf{y})\mathbf{R_x}(\theta_\mathbf{x}) \neq \mathbf{R_x}(\theta_\mathbf{x})\mathbf{R_y}(\theta_\mathbf{y})$$

The operations don't commute.

It follows that for the infinitesimal generators

$$\mathbf{J_y J_x} \neq \mathbf{J_x J_y}$$

and in general we can show by direct computation, that

$$[\mathbf{J_i}, \mathbf{J_j}] = i\epsilon_{ijk}\mathbf{J_k} \tag{6}$$

where we've equated the indices $(1, 2, 3)$ with (x, y, z) and we've defined the commutator $[\mathbf{J_i}, \mathbf{J_j}] \equiv \mathbf{J_i J_j} - \mathbf{J_j J_i}$. The generators J_i together with the commutation relations of equation (6) are known as the **LIE ALGEBRA of the rotation group.**

2.3 Spin

- Just like we had matrix representations (up to a phase) of the rotations, we also have matrix representations of their generators – i.e., the angular momenta J_i as illustrated in eq. (5).

- For spinor representations, the angular momenta J_i are represented by the 2 x 2 Pauli spin matrices $\sigma_i/2$.

- Since these matrices represent the angular momentum operators, they must satisfy the same Lie Algebra commutation relations – namely,

$$[\frac{\sigma_i}{2}, \frac{\sigma_j}{2}] = i\epsilon_{ijk}\frac{\sigma_k}{2} \tag{7}$$

We can verify this by explicitly computing matrix products with the Pauli spin matrices.

- **EXERCISE**. Show, by explicit computation with the Pauli spin matrices, that we also have the anticommutation relations

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}.$$

HINT: Verify the relation for only two cases – the case when $i = 1$ and $j = 2$, then the case when $i = 1$ and $j = 1$. If you get the right answers, declare victory.

- We say that $\frac{\sigma}{2}$ is the *spin* of the electron, representing the angular momentum operator for the electron doublet and we sometimes denote is as **S**.

3 Magnetic moment and spin

- Now that we've identified the spin **S** as an angular momentum, we can try picturing it as something characterizing something like the internal angular momentum of the electron (hence the name *spin*).

- This is, at best, a rather crude picture. In fact, since the spin is a matrix (or more precisely a triplet of matrices) and the classical angular momentum is just a real vector, the two concepts are only related abstractly. When we eventually deal with quantum mechanics, then we treat both **S** and **L** as operators (albeit on different spaces) and it becomes more natural to compare the two.

- In the meantime, we can put together various things we've discussed. We had

$$H = -\boldsymbol{\mu} \cdot \mathbf{B}$$

and for a current-loop we had

$$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{L}$$

so altogether we have

$$H = -\frac{q}{2m}\mathbf{L} \cdot \mathbf{B}.$$

- Although the electron itself can't be regarded as a current loop, it seems reasonable to think that we could replace the loop angular momentum $\mathbf{L}$ with the intrinsic electron angular momentum $\mathbf{S}$ leading to a contribution – independent of current – of

$$H = -\frac{q}{2m}\mathbf{S} \cdot \mathbf{B}?$$

- The way Zwiebach portrays this, we need to multiply by a fudge factor g to reflect the crudeness of the spin picture. I don't think this is the way things happened. But for now, that suffices for motivation.

4 Simple one-component Schrodinger equation

- The Schrodinger equation is

$$i\partial_t \psi(t, \mathbf{x}) = H\psi(t, \mathbf{x})$$

where ψ is the Schrodinger wavefunction (**NOT** an operator or field) and H is the Hamiltonian.

- Also, the energies of quantum states are given by the equation

$$H\psi_E = E\psi_E$$

where ψ_E is an eigenfunction and E is its energy eigenvalue.

- The Hamiltonian for a free particle is given clasically by

$$H = \frac{\mathbf{p} \cdot \mathbf{p}}{2m}$$

and in QM, we use the *correspondence principle* to set

$$\mathbf{p} = -i\boldsymbol{\nabla}.$$

Recall that $\nabla_i = \frac{\partial}{\partial x^i}$.

- Substituting the correspondence principle into the Hamiltonian, we get

$$H = -\frac{\boldsymbol{\nabla}^2}{2m}.$$

5 Extending the Schrodinger equation to include Pauli spinors

- Again, what follows should be regarded as motivation. I'm not sure it's historically correct but even if it is, there's a lot of context missing.

- Since we are going to extend the Schrodinger equation to deal with doublets, let's find an alternative expression for the Schrodinger Hamiltonian for a free theory. We had $H = -\frac{\boldsymbol{\nabla}^2}{2m}$. Now we want to act on a 2D space of functions, so we simply generalize the Hamiltonian to

$$H = -\mathbf{I}_{2\text{x}2}\frac{\boldsymbol{\nabla}^2}{2m}.$$

We could obtain this by writing $H = \frac{\mathbf{p^2}}{2m}$ and substituting $\mathbf{p} = -i\boldsymbol{\nabla}$. But let's see what happens if instead we write

$$\hat{\mathbf{p}} = -i\boldsymbol{\sigma} \cdot \boldsymbol{\nabla} = -i\sum_j \boldsymbol{\sigma}_j \partial_j.$$

THIS IS PURELY A GUESS FOR MOTIVATIONAL PURPOSES!!!

- **EXERCISE:** Show that

$$\left(-i\boldsymbol{\sigma} \cdot \boldsymbol{\nabla}\right)\left(-i\boldsymbol{\sigma} \cdot \boldsymbol{\nabla}\right) = -\mathbf{I}_{2\text{x}2}\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}.$$

Hint: Use the following identity which you can derive from the Pauli spin matrix commutation and anticommutation relations above (or by brute force).

$$\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j = \delta_{ij}\mathbf{I}_{2\text{x}2} + i\epsilon_{ijk}\boldsymbol{\sigma}_j. \tag{8}$$

- From the exercise, we see that we could write

$$H = \frac{\hat{\mathbf{p}}^2}{2m}$$

so that demonstrates there are at least two definitions of the free momentum which lead to the same Hamiltonian.

- Now we're ready to extend Schrodinger's equation. You may remember from a previous set of notes, that in classical theory, the Hamiltonian for a charged particle in an EM field is

$$H = \frac{(\mathbf{p} - \mathbf{eA})^2}{2m} + e\phi.$$

So just as we generalized $\mathbf{p}$ to $\hat{\mathbf{p}} = -i\boldsymbol{\sigma} \cdot \boldsymbol{\nabla}$, let's generalize $(\mathbf{p} - \mathbf{eA})$ to $\boldsymbol{\Pi} = \boldsymbol{\sigma} \cdot (-i\boldsymbol{\nabla} - e\mathbf{A})$ so that

$$H = \frac{(\boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{eA}))\,(\boldsymbol{\sigma} \cdot (\mathbf{p} - \mathbf{eA}))}{2m} + e\phi\mathbf{I}_{2\text{x}2}$$

- **EXERCISE:** From the σ multiplication identity eq. (8) above, show Zweibach's result

$$(\boldsymbol{\sigma} \cdot a)(\boldsymbol{\sigma} \cdot b) = (\mathbf{a} \cdot \mathbf{b})\mathbf{I}_{2\text{x}2} + \mathbf{i}\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}). \qquad (9)$$

  Hint: Expand $(\boldsymbol{\sigma} \cdot a)$ as $\boldsymbol{\sigma}_i a_i$ and $(\boldsymbol{\sigma} \cdot b)$ as $\boldsymbol{\sigma}_j b_j$ and then apply eq.(8).

- **SOMETHING MYSTERIOUS FROM QM**: Before we can properly expand the above Hamiltonian, we have to recognize the mysterious fact from QM, that $p_i A_j \neq A_j p_i$. This is explained in what follows. In regular math, $a_i b_j = b_j a_i$. This not generally true in QM because the variables of interest are operators rather than numbers. For example, suppose a_i is the operator $\mathbf{x}_i$ and b_j is the operator $\mathbf{p}_j$. (Respectively the i component of the position vector and the j component of the momentum vector.)(By the way, the combination $\mathbf{x_i p_j}$ should remind you of a component of the angular momentum.) In QM, we have the commutation relationship $\mathbf{x_i p_j} = \mathbf{p_j x_i} + \mathbf{i}\delta_{\mathbf{ij}}$ so that $a_j b_j \neq b_j a_j$.

  The notation tends to obscure the distinction between 'operator' and 'eigenvalue'. For example, suppose we have a wavefunction $\psi(x)$. We use the language of operators and states by saying

  – ψ is a state – that is, a map from real 3-tuplets $((x_1, x_2, x_3))$ to the complex number $\psi(\mathbf{x})$.

– x_j is an operator which transforms the state ψ to the state which maps the 3-tuplet $\mathbf{x}$ to $x_j\psi(\mathbf{x})$. (Often we write $\hat{x}_j$ to denote the operator, but I'm not going to do that right now.)

In that language of operators and states, the electromagnetic field $\mathbf{A_i}(\mathbf{x})$ is an operator and not a state. Specifically, it is the operator which transforms the state ψ to the state which maps the 3-tuplet $\mathbf{x}$ to $A_i(\mathbf{x})\psi(\mathbf{x})$.

What about the operator p_j? When applied to the state ψ, we obtain a new state which maps $\mathbf{x}$ to $-i\partial_j\psi(\mathbf{x})$.

We can then apply this formalism to the joint operator $p_j A_k(x)$ as applied to the state $\psi(x)$. First one applies the operator $A_k(\mathbf{x})$ to the state ψ. This is a new state. Then apply the operator p_j to that new state. The result is different than if we first apply p_j to ψ and then afterwards apply the operator A_k to that state.

$$-i\partial_j[A_k(x)\psi(x)] = \left(-i\partial_j A_k(x)\right)\psi(x) + A_k(x)\left(-i\partial_j\psi(x)\right)$$

compared to

$$A_k(x)\left(-i\partial_j\psi(x)\right).$$

The extra term is $\left(-i\partial_j A_k(x)\right)\psi(x)$. **Abstractly, we would write**

$$[p_j, A_k(x)] = -i\partial_j A(x).$$

If $A(x) = x$, we see that the commutator is $-i$ which should be a familiar commutation relation.

- **EXERCISE:**   Use the above mysterious result to show that

$$\epsilon_{ijk}\left(\mathbf{p_j} - \mathbf{eA_j}\right)\left(\mathbf{p_k} - \mathbf{eA_k}\right) = -2ie\epsilon_{ijk}\partial_j A_k.$$

- If we use Zwiebach's identity eq. (9) and the last exercise above, we see that one of the terms in the Hamiltonian will be

$$-ie\frac{\boldsymbol{\sigma}_i\epsilon_{ijk}\partial_j A_k}{2m} = i\frac{e}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}$$

as Zwiebach shows.