Final Solutions

1. *Consider the decay of the $2p$ state of hydrogen atom to the $1s$ state. Calculate the amplitude of the decay for $m = +1$ state using plane waves for photons, and explain the θ dependence of the amplitude for each helicity ± 1 of the final-state photon in terms of the angular momentum conservation. Show that the rate is the same as the decay rate of the $m = 0$ state.*

The only differences from the calculation in the lecture note "Quantum Field Theory III (Radiation Field)" are: the matrix elements $\langle 1s|\vec{D}|2p, m = \pm 1\rangle$ and the polarization vectors $\vec{\epsilon}_\lambda(\vec{q})^*$. (The complex conjugation is very important for helicity eigenstates.) The matrix elements are

$$\langle 1s|\vec{D}|2p, m = 1\rangle = e\int_0^\infty r^2 dr d\Omega \frac{1}{a^3} 2e^{-r/a} Y_0^{0*} \frac{\sqrt{6}}{12}\frac{r}{a} e^{-r/2a} Y_1^1 \vec{x}. \tag{1}$$

Because $Y_1^1 = -\sqrt{3}8\pi\cos\theta e^{i\phi}$, only x and y components of the dipole operator survive integration over ϕ. The result of the integration is

$$\langle 1s|\vec{D}|2p, m = 1\rangle = -e\frac{128}{243}a(1, i, 0). \tag{2}$$

On the other hand, the polarization vectors for helicity ± 1 states are (Eqs. (16,17,18) in the lecture note):

$$\epsilon_\pm(\vec{p}) = \frac{1}{\sqrt{2}}(\pm\cos\theta\cos\phi - i\sin\phi, \pm\cos\theta\sin\phi + i\cos\phi, \mp\sin\theta). \tag{3}$$

Therefore the transition matrix element Eq. (56) is

$$\begin{aligned}
\langle f|V|i\rangle &= \frac{i}{\hbar}|\vec{q}|\sqrt{\frac{2\pi\hbar c^2}{L^3}}\frac{1}{\sqrt{\omega_q}}\vec{\epsilon}_\lambda^*(\vec{q})\cdot\langle 1s|\vec{D}|2p, m = 1\rangle \\
&= \frac{i}{\hbar}|\vec{q}|\sqrt{\frac{2\pi\hbar c^2}{L^3}}\frac{1}{\sqrt{\omega_q}}(-e)\frac{128}{243}a \\
&\quad (1, i, 0)\cdot\frac{1}{\sqrt{2}}(\pm\cos\theta\cos\phi + i\sin\phi, \pm\cos\theta\sin\phi - i\cos\phi, \mp\sin\theta) \\
&= \frac{i}{\hbar}|\vec{q}|\sqrt{\frac{2\pi\hbar c^2}{L^3}}\frac{1}{\sqrt{\omega_q}}(-e)\frac{1}{\sqrt{2}}\frac{128}{243}a(1\pm\cos\theta)e^{i\phi}.
\end{aligned} \tag{4}$$

The θ dependence has a simple interpretation in terms of angular momentum conservation. The initial state has the angular momentum $J_z = +1$. When a positive helicity photon is emitted along the negative z-axis, it carries away the angular momentum $J_z = -1$, and hence the final state must have $J_z = +2$.

However, our final state is $1s$ and hence $J_z = 0$. Such a transition must be forbidden. Indeed, $(1 + \cos\theta)$ factor does that precisely. Only the other hand, if the positive helicity photon is emitted along the positive z-axis, the angular momentum conservation is satisfied. Since unit angular momentum should give an amplitude linear in $\cos\theta$ (in general higher j gives j-th order polynomials), $(1+\cos\theta)$ is the only possible θ dependence. The argument for a negative helicity photon is similar.

The rate is calculated in the same way as in Eq. (60),

$$\begin{aligned} W_i &= \int \frac{d\Omega_q}{(2\pi\hbar)^3}|\vec{q}|^2\frac{(2\pi)^2|\vec{q}|}{\hbar}\sum_{\pm}\left|(-e)\frac{1}{\sqrt{2}}\frac{128}{243}a(1\pm\cos\theta)e^{i\phi}\right|^2 \\ &= \int \frac{d\Omega_q}{2\pi\hbar^4}|\vec{q}|^3\sum_{\pm}\left(e\frac{1}{\sqrt{2}}\frac{128}{243}a\right)^2(1\pm 2\cos\theta+\cos^2\theta) \\ &= \int \frac{d\Omega_q}{2\pi\hbar^4}|\vec{q}|^3\left(e\frac{1}{\sqrt{2}}\frac{128}{243}a\right)^2 2(1+\cos^2\theta) \\ &= \frac{8}{3}\left(\frac{128}{243}\right)^2 e^2\frac{q^3a^2}{\hbar^4} \\ &= \frac{2}{3}\left(\frac{256}{243}\right)^2 e^2\frac{q^3a^2}{\hbar^4}. \end{aligned} \tag{5}$$

Note that you need to sum over helicities of the photon to obtain the total decay rate of the $2p$ state. This result agrees completely with the decay orate of the $|2p, m = 0\rangle$ state calculated in the lecture note, confirming the rotational invariance of the result.

2. *How can the* $2s$ *state decay to the* $1s$ *state? Do not calculate the rate, but discuss it.*

In multipole expansion, we saw that a photon carries angular momentum of *at least* one. On the other hand, the initial and final state here both have zero angular momentum. Therefore, the $2s$ state *cannot* decay to the $1s$ state by emitting a single photon. When you were told that Lyman series of hydrogen spectrum shows transitions between states with principal angular momentum n and 1, you were cheated; $n = 1$ case shows only transitions between $2p$ and $1s$, but *not* $2s$! Then how does the $2s$ state decay to the ground state? It has to emit *two* photons. Note that an emission of two photons at the same time does not give you a discrete spectrum. Only the sum of two photon energies is constrained. The dominant contribution is the emission of two photons both in the E1 multipole (electric dipoles) where the angular momenta of both photons cancel and hence they are in the $J = 0$ configuration. Angular momentum consideration suggests the combination $|k10, k10\rangle + 2|k11, k1-1\rangle$.

Then the next question is how can two photons be emitted. There are two possibilities. One is to use the term

$$V = \left(\frac{e}{c}\right)^2 \vec{A}(\vec{x})^2 \tag{6}$$

in the Hamiltonian. However, using the expression for the electric dipole photon Eq. (86) in the lecture note, the matrix is $\langle 1s|2s\rangle$ which vanishes identically. Then we have to go to higher order in kr for the electric dipole mode function $\vec{u}^E_{k10}$. Another possibility is to use the operator we've been using

$$V = -\frac{e}{c}\vec{p}\cdot\vec{A}(\vec{x}) \tag{7}$$

twice. The transition element is then

$$\langle 1s + \text{two photons}|V\frac{1}{E_{2s}-H_0}V|2s\rangle = \sum_i \langle 1s + \text{two photons}|V|i\rangle\frac{1}{E_{2s}-E_i}\langle i|V|2s\rangle. \tag{8}$$

The perturbation Hamiltonian V on the left creates the intermediate states, $|2p+\text{photon}\rangle$ (or any other $|np\rangle$ state), while that on the right creates another photon making transition to the $1s$ state. You then have to sum over all intermediate states.

Order of magnitude of the process can be easily be estimated by neglecting all numerical factors but by keeping dependences on the physical constants. We have seen that dipole transition matrix elements $\langle f|V|i\rangle$ go as $\sim ea_0E_\gamma^{1/2}L^{-3/2}$. Then the transition element Eq. (8) goes as

$$\sum_i \langle 1s + \text{two photons}|V|i\rangle\frac{1}{E_{2s}-E_i}\langle i|V|2s\rangle \sim (ea_0E_\gamma^{1/2}L^{-3/2})^2\frac{1}{e^2/a_0} = a_0^3E_\gamma L^{-3}. \tag{9}$$

Similarly, the contribution from the term Eq. (6) can be shown to be of the same order of magnitude. The rate goes as square of this: $a_0^6E_\gamma^2L^{-6}$. The phase space integral summing over two final state photons goes as $(L^3q^3/\hbar^3)^2 \sim L^6E_\gamma^6/(\hbar c)^6$. Here I made an approximation that two photon energies are comparable. This is indeed the case because the phase space $E_1^3E_2^3$, subject to the constraint that E_1+E_2 is fixed, is maximized when $E_1 \sim E_2$. Together with the delta function in energy and $1/\hbar$ in Fermi's golden rule, the estimate of the decay rate is $a_0^6E_\gamma^7/\hbar^7c^6 \simeq \alpha^8mc^2/\hbar$. On the other hand, the dipole transition rate for $2p \to 1s$ is $\alpha^5mc^2/\hbar$ up to numerical factors. Therefore, the decay rate of $2s \to 1s$ is roughly $\alpha^3 \simeq 4\times10^{-7}$ smaller.

A detailed calculation shows that the decay rate is very small: 8.229 sec^{-1}, even another order of magnitude smaller than the above rough estimate due to numerical factors. This is smaller than the dipole transition from $2p$ to $1s$ by eight orders of magnitude! The $2s$ state is hence said to be *metastable*.

You may think that the $2s$ state decays into $2p$ state first, which is lower than the $2s$ state because of the Lamb shift, and then decays into $1s$ state, rather than going through quantum intermediate states as in Eq. (8) or emitting two photons directly from the operator Eq. (6). Recalling that the decay rate due to electric dipole transition is proportional to the energy of the photon to the cube power, and knowing that the decay rate we calculated for $2p \to 1s$ was $6.27 \times 10^8 \text{sec}^{-1}$, we can estimate the order of magnitude of the decay rate of $2s$ to $2p$ due to the Lamb shift. The level splitting is about 1 GHz in frequency. The energy of the photon then is $h\nu \simeq 4.1 \times 10^{-6}$ eV. On the other hand, the energy of the photon in the $2p \to 1s$ decay is $13.6 \times (1 - 1/4)$ eV. Therefore the $2s \to 2p$ decay gives the energy of the photon 4.1×10^{-7} time smaller than $2p \to 1s$, and hence the decay rate is suppressed by $(4.1 \times 10^{-7})^3 = 6.7 \times 10^{-20}$. This is indeed very small! What it means is that the decay of $2s$ state going through the $2p$ state is possible, but this decay is limited by the slowness of $2s \to 2p$ transition and the two photon transition discussed above is far more important.

Finally, once you consider the spin of the electron (not required in this problem), the decay $2s_{1/2} \to 1s_{1/2}$ allows even parity $j = 1$ photon: an M1 transition. However, as we see in Problem 3, the M1 transition causes only the spin flip and does not change the spatial wave function at the leading order in kr. Hence the amplitude picks up higher order in the Taylor expansion in the spherical Bessel function $j_1(kr)$, which is $(kr)^3$ compared to kr of the leading term. Recall that the M1 is already one order higher in kr compared to the E1. Therefore there is an overall $(kr)^3$ suppression in the amplitude, and hence $(kr)^6$ in the rate compared to the E1 transition rate. Because $k \sim \alpha^2 mc/\hbar$ for the typical photon wave vector from the atomic transitions and $r \sim a_0 = \hbar/(\alpha mc)$, $kr \sim \alpha$. This leads to a factor of α^6 suppression in the rate relative to the E1 case, and hence negligible compared to the two-photon process.

In this decay, therefore, two photons are emitted promptly, giving vanishing total angular momentum. This is the ideal system for testing Einstein–Podolsky–Rosen paradox and Bell's inequality and had been used for that purpose.

Table 1: Comparison of Theoretical and Experimental Total Decay Rates of the $2s_{1/2}$ State (in s^{-1}). Taken from G.W.F. Drake, in "The Spectrum of Atomic Hydrogen Advances," edited by G. W. Series, World Scientific, 1988.

Ion	Theory	Experiment
He^+	526.61	525 ± 5
O^{7+}	2.1552×10^6	$(2.21 \pm 0.22) \times 10^6$
F^{8+}	4.3699×10^6	$(4.22 \pm 0.28) \times 10^6$
S^{15+}	1.3964×10^8	$(1.37 \pm 0.13) \times 10^8$
Ar^{17+}	2.8590×10^8	$(2.868 \pm 0.029) \times 10^8$

I could not find experimental data on the $2s$ lifetime for hydrogen, except the statement that it is very difficult to measure because of its too small decay rate (or too long lifetime). Instead, I found comparison between theory and data for hydrogen-like atoms. Following the analysis above with $Z > 1$, the two E1 photon emission process scales as Z^6, while one M1 photon emission as Z^{10}. At some point, the M1 emission catches up. Here is a table that compares theory and experiment. For Ar, the data is sufficiently accurate to be sensitive to the small M1 contribution of 0.0908×10^8 s^{-1}.

3. *The coupling of the magnetic moment to the magnetic field $V = -\vec{\mu} \cdot \vec{B}$ can also cause transitions. (One such example is the hyperfine transition in hydrogen atom.) By expanding the Hamiltonian in multipoles, show that emission or absorption of a photon can change the spin state by M1 transitions.*

The expression for $\vec{u}^M_{k10}$ is given in Eq. (84). Other $m = \pm$ states can also easily be calculated.

$$\begin{aligned}
\vec{u}^M_{k11} &= \frac{i}{\hbar}\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} 2j_1(kr)Y_1^1 \frac{1}{\sqrt{1(1+1)}} \\
&= \frac{i}{\hbar}\begin{pmatrix} (L_+ + L_-)/2 \\ (L_+ - L_-)/2i \\ L_z \end{pmatrix} 2j_1(kr)Y_1^1 \frac{1}{\sqrt{2}} \\
&= i\begin{pmatrix} Y_1^0/2 \\ -Y_1^0/2i \\ Y_1^1 \end{pmatrix} 2j_1(kr)\frac{1}{\sqrt{2}} \\
&= \sqrt{\frac{3}{4\pi}}\begin{pmatrix} \cos\theta \\ i\cos\theta \\ -\sin\theta e^{i\phi} \end{pmatrix} j_1(kr) \\
&\simeq \sqrt{\frac{k}{12\pi}}\begin{pmatrix} z \\ iz \\ -(x+iy) \end{pmatrix}
\end{aligned} \tag{10}$$

Similarly,

$$\begin{aligned}
\vec{u}^M_{k1-1} &= \frac{i}{\hbar}\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} 2j_1(kr)Y_1^{-1} \frac{1}{\sqrt{1(1+1)}} \\
&= \frac{i}{\hbar}\begin{pmatrix} (L_+ + L_-)/2 \\ (L_+ - L_-)/2i \\ L_z \end{pmatrix} 2j_1(kr)Y_1^{-1} \frac{1}{\sqrt{2}}
\end{aligned}$$

$$
\begin{aligned}
&= i \begin{pmatrix} Y_1^0/2 \\ Y_1^0/2i \\ -Y_1^{-1} \end{pmatrix} 2j_1(kr)\frac{1}{\sqrt{2}} \\
&= \sqrt{\frac{3}{4\pi}} \begin{pmatrix} \cos\theta \\ -i\cos\theta \\ -\sin\theta e^{-i\phi} \end{pmatrix} j_1(kr) \\
&\simeq \sqrt{\frac{k}{12\pi}} \begin{pmatrix} z \\ -iz \\ -(x-iy) \end{pmatrix}. \qquad (11)
\end{aligned}
$$

The next step is to calculate the magnetic field $\vec{B}$ for the magnetic dipole component. Using Eq. (76), the magnetic dipole component is

$$
\vec{A}_{M1} = \sum_m \int \frac{dk}{2\pi} k^2 \sqrt{\frac{2\pi\hbar c^2}{\omega}} (a^M_{k1m}\vec{u}^M_{k1m} + a^{m\dagger}_{k1m}\vec{u}^{M*}_{k1m}), \qquad (12)
$$

and hence the magnetic field is

$$
\begin{aligned}
\vec{B}_{M1} &= \sum_m \int \frac{dk}{2\pi} k^2 \sqrt{\frac{2\pi\hbar c^2}{\omega}} (a^M_{k1m}\nabla\times\vec{u}^M_{k1m} + a^{M\dagger}_{k1m}\nabla\times\vec{u}^{M*}_{k1m}) \\
&\simeq \int \frac{dk}{2\pi} k^2 \sqrt{\frac{2k}{3\pi}} \left(a^M_{k11}\frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix} + a^M_{k10}\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + a^M_{k1-1}\frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix} + h.c. \right), \qquad (13)
\end{aligned}
$$

where $h.c.$ stands for hermitian conjugate.

Finally, we substitute the magnetic field in the operator

$$
H = -\vec{\mu}\cdot\vec{B} = -g\frac{e\hbar}{2mc}\vec{s}\cdot\vec{B}. \qquad (14)
$$

we find

$$
\begin{aligned}
H &= -g\frac{e\hbar}{2mc}\int \frac{dk}{2\pi} k^2 \sqrt{\frac{2k}{3\pi}} \\
&\quad \left(a^M_{k11}\frac{1}{\sqrt{2}}(-is_x + s_y) + a^M_{k10}(-s_z) + a^M_{k1-1}\frac{1}{\sqrt{2}}(is_x + s_y) + h.c. \right) \\
&= -g\frac{e\hbar}{2mc}\int \frac{dk}{2\pi} k^2 \sqrt{\frac{2k}{3\pi}} \\
&\quad \left(\frac{-i}{\sqrt{2}}a^M_{k11}s_+ - a^M_{k10}s_z + \frac{i}{\sqrt{2}}a^M_{k1-1}s_- + \frac{i}{\sqrt{2}}a^{M\dagger}_{k11}s_- - a^{M\dagger}_{k10}s_z + \frac{-i}{\sqrt{2}}a^{M\dagger}_{k1-1}s_+ \right). \qquad (15)
\end{aligned}
$$

Clearly, one can, for instance, emit a M1 photon with $m = 1$ with the creation operator $a^{M\dagger}_{k11}$ while lowering the spin by s_- conserving the angular momentum of the system. This operator can cause a transition from the triplet state $|e^\uparrow p^\uparrow\rangle$ to the singlet state $|e^\downarrow p^\uparrow\rangle$ in the hyperfine transition, giving rise to 21cm line important for astronomy. Because of the small level splittings in the hyperfine transition $\Delta E_{hf} = (hc/\lambda) = 5.9 \times 10^{-6}$ eV, even a very cold gas can emit it. In fact, there are 2.725K comsic microwave background *anywhere* in the Universe, which can excite the atom to the triplet state because $kT = 2.3 \times 10^{-4}$ eV $\gg \Delta E_{hf}$. For instance, the rotational speed of hydrogen gas in the outer region of galaxies is determined using the Dopper shifts in 21cm line. The fact that the rotational speed is much larger than the estimate based on the luminous matter (stars) within the given radius is the best evidence for the existence of dark matter in galactic halo.

Note that, at the leading order in kr we had kept, the operator does not depend on the electron variables except spin, and hence does not affect the spatial wave function of the electron.

The expansion of $(\vec{p} - \frac{e}{c}\vec{A})^2/2m$ also gives the coupling of the magnetic field to the orbital angular momentum. Therefore the M1 transition can cause transitions among different $L_z = m\hbar$ states in a given l multiplet. Transitions among split Zeeman levels can occur this way.

4. *The relativistic field equation for a spinless particle in the presence of the Maxwell field is*

$$\left[\left(-i\hbar\frac{1}{c}\frac{\partial}{\partial t} - \frac{e}{c}A^0\right)^2 - \left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2 - m^2c^2\right]\phi = 0. \tag{16}$$

Answer the following questions.

(a) *We would like to determine energy eigenvalue E in the presence of Coulomb potential $eA^0 = \frac{Ze^2}{r}$. Show that time-independent field equation for the radial wave function $\phi = R(r)Y^m_l e^{-iEt/\hbar}$ has the form*

$$\left[\frac{\hbar^2}{2\mu}\left(-\frac{d^2}{dr^2} - \frac{2}{r}\frac{d}{dr} + \frac{\lambda(\lambda+1)}{r^2}\right) - \frac{Ze^2}{r}\right]R = \epsilon R. \tag{17}$$

Write μ, λ, ϵ in terms of E, m, and l.

The field equation for $\phi = R(r)Y^m_l e^{-iEt/\hbar}$ is

$$\left[\frac{1}{c^2}\left(-E - \frac{Ze^2}{r}\right)^2 - \hbar^2\left(\frac{1}{r^2}\frac{d^2}{dr^2}r^2 - \frac{l(l+1)}{r^2}\right) - m^2c^2\right]R = 0. \tag{18}$$

By reorganizing terms, we find

$$\left[\frac{\hbar^2 c^2}{2E}\left(-\frac{d^2}{dr^2}-\frac{2}{r}\frac{d}{dr}+\frac{l(l+1)-Z^2\alpha^2}{r^2}\right)-\frac{Ze^2}{r}\right]R=\frac{E^2-(mc^2)^2}{2E}R. \tag{19}$$

$\alpha = e^2/\hbar c$ is the fine-structure constant. By comparing to the Schrödinger-like equation Eq. (17), we find

$$\mu = E/c^2 \tag{20}$$

$$\lambda = \sqrt{\left(l+\frac{1}{2}\right)^2 - Z^2\alpha^2} - \frac{1}{2} \tag{21}$$

$$\epsilon = \frac{E^2-(mc^2)^2}{2E}. \tag{22}$$

4. (b) Eq. (17) has exactly the same form as the Schrödinger equation for the hydrogen atom, except that λ is not an integer. Therefore the boundstate eigenvalues are given by

$$\epsilon = -\frac{1}{2}\frac{Z^2\alpha^2\mu c^2}{\nu^2},$$

where the "principal quantum number" ν takes values $\nu = \lambda+1, \lambda+2, \lambda+3, \cdots$. Solve for E.

Using the result from the previous problem,

$$\frac{E^2-(mc^2)^2}{2E} = -\frac{1}{2}\frac{Z^2\alpha^2 E}{\nu^2}. \tag{23}$$

Solving for E, we find

$$E = \frac{mc^2}{\sqrt{1+Z^2\alpha^2/\nu^2}}. \tag{24}$$

4. (c) Expand E up to $O(Z^2\alpha^2)$ and show that it agrees with the result of conventional Schrödinger equation including the rest energy.

By expanding Eq. (24) up to $O(Z^2\alpha^2)$, we find

$$E = mc^2\left(1-\frac{1}{2}\frac{Z^2\alpha^2}{\nu^2}+O(Z^4\alpha^4)\right). \tag{25}$$

Note that $\lambda = l + O(Z^2\alpha^2)$. Therefore, $\nu = \lambda + k$ (k is a non-negative integer) and hence ν is also an integer up to an $O(Z^2\alpha^2)$ correction. Neglecting $O(Z^4\alpha^4)$ terms, we find the principal quantum number $n = \nu + O(Z^2\alpha^2)$ and hence

$$E = mc^2 - \frac{1}{2}\frac{Z^2\alpha^2 mc^2}{n^2} + O(Z^4\alpha^4). \tag{26}$$

The result agrees with conventional Schrödinger equation at this order.

4. *(e) Expand E up to $O(Z^4\alpha^4)$, and discuss the interpretation of the correction.*

We this expand expand E in Eq. (24) up to $O(Z^4\alpha^4)$, and find

$$E = mc^2\left(1 - \frac{1}{2}\frac{Z^2\alpha^2}{\nu^2} + \frac{3}{8}\frac{Z^4\alpha^4}{\nu^4} + O(Z^6\alpha^6)\right). \tag{27}$$

The difference between ν and n at $O(Z^2\alpha^2)$ cannot be ignored in the second term because it gives rise to a term of $O(Z^4\alpha^4)$. By expanding λ up to $O(Z^2\alpha^2)$,

$$\lambda = l - \frac{Z^2\alpha^2}{2l+1} + O(Z^4\alpha^4), \tag{28}$$

we can write

$$\nu = n - \frac{Z^2\alpha^2}{2l+1} + O(Z^4\alpha^4), \tag{29}$$

and hence

$$E = mc^2\left(1 - \frac{1}{2}\frac{Z^2\alpha^2}{n^2} - \frac{Z^4\alpha^4}{(2l+1)n^3} + \frac{3}{8}\frac{Z^4\alpha^4}{n^4} + O(Z^6\alpha^6)\right). \tag{30}$$

As before, the second term is the term we obtain in non-relativistic Schrödinger equation.

The question is what are the next two terms. They are the so-called "relativistic correction," obtained by expanding the relativistic kinetic energy

$$\sqrt{\vec{p}^2c^2 + (mc^2)^2} = mc^2 + \frac{\vec{p}^2}{2m} - \frac{1}{8}\frac{(\vec{p}^2)^2}{m^3c^2} + O(\vec{p}^6). \tag{31}$$

Because $|\vec{p}|/m = v = Z\alpha$ in hydrogen-like atoms, $O(\vec{p}^6) \sim O(Z^6\alpha^6)$ and these terms are beyond our interest. We can rewrite

$$\vec{p}^2|nlm\rangle = 2m\left(\frac{Ze^2}{r} - \frac{1}{2}\frac{Z^2\alpha^2mc^2}{n^2}\right), \tag{32}$$

and hence

$$\langle nlm| - \frac{1}{8}\frac{(\vec{p}^2)^2}{m^3c^2}|nlm\rangle = -\frac{1}{2mc^2}\langle nlm|\left(\frac{Ze^2}{r} - \frac{1}{2}\frac{Z^2\alpha^2mc^2}{n^2}\right)^2|nlm\rangle. \tag{33}$$

Using (see below for the derivation of these expectation values)

$$\langle nlm|\frac{1}{r}|nlm\rangle = \frac{1}{n^2a}, \qquad \langle nlm|\frac{1}{r^2}|nlm\rangle = \frac{2}{(2l+1)n^3a^2}, \tag{34}$$

with $a = \hbar^2/mZe^2 = \hbar/mcZ\alpha$, we find

$$\langle nlm| - \frac{1}{8}\frac{(\vec{p}^2)^2}{m^3c^2}|nlm\rangle = -\frac{Z^4\alpha^4}{(2l+1)n^3} + \frac{3}{8}\frac{Z^4\alpha^4}{n^4}. \tag{35}$$

This precisely reproduces the $O(Z^4\alpha^4)$ terms in Eq. (24), and hence the relativistic correction $-\frac{1}{8}\frac{\vec{p}^4}{m^3c^4}$ is their origin. Obviously, there is no spin-orbit coupling because the Klein–Gordon field does not have spin. What is more interesting is that there is no Darwin term; the Klein–Gordon particle does not do Zitterbewegung! In fact, if you take the square root of the Klein–Gordon equation and consider the Hamiltonian to be $H = \sqrt{c^2\vec{p}^2 + m^2c^4}$, the Heisenberg equation would give the velocity $\dot{\vec{x}} = [\vec{x}, H]/i\hbar = c^2\vec{p}/\sqrt{c^2\vec{p}^2 + m^2c^4}$ which is perfectly normal, showing no sign of Zitterbewegung.

The energy levels of the Klein–Gordon equation in the Coulomb potential is the starting point for the study of π-mesic atoms, *i.e*, the bound states of negative pions π^- to nuclei.

The following derivations are by Ed. Thank you, Ed! We can derive Eq. (34) without suffering through generating functions for Laguerre polynomials by using the Feynman-Hellman theorem which states

$$\langle\psi|\frac{\partial H}{\partial\lambda}|\psi\rangle = \frac{\partial E}{\partial\lambda}, \tag{36}$$

quite generally when a Hamiltonian H, its eigenstates $|\psi\rangle$, and its eigenvalues E depend on a parameter λ. (The eigenstates if degenerate must be diagonalized not to mix under infinitessimal changes in λ.) To show equation (36) start with

$$\frac{\partial}{\partial\lambda}(H|\psi\rangle) = \frac{\partial}{\partial\lambda}(E|\psi\rangle) \tag{37}$$

$$\frac{\partial H}{\partial\lambda}|\psi\rangle + H\frac{\partial}{\partial\lambda}|\psi\rangle = \frac{\partial E}{\partial\lambda}|\psi\rangle + E\frac{\partial}{\partial\lambda}|\psi\rangle, \tag{38}$$

and act on the left with $\langle\psi|$. Then $\langle\psi|H = \langle\psi|E$ so that the unwanted terms drop out:

$$\langle\psi|\frac{\partial H}{\partial\lambda}|\psi\rangle + E\langle\psi|\frac{\partial}{\partial\lambda}|\psi\rangle = \langle\psi|\frac{\partial E}{\partial\lambda}|\psi\rangle + E\langle\psi|\frac{\partial}{\partial\lambda}|\psi\rangle \Longrightarrow \tag{39}$$

$$\langle\psi|\frac{\partial H}{\partial\lambda}|\psi\rangle = \frac{\partial E}{\partial\lambda}. \tag{40}$$

Now for the non-relativistic hydrogen atom,

$$H = \frac{\hbar^2}{2m}\left(-\frac{d^2}{dr^2} - \frac{2}{r}\frac{d}{dr} + \frac{l(l+1)}{r^2}\right) - \frac{Ze^2}{r}, \tag{41}$$

$$E = -\frac{Z^2\alpha^2mc^2}{2n^2}. \tag{42}$$

Mathematically, we can consider Z to be a continuous parameter and apply the Feynman-Hellman theorem,

$$\langle nlm|\frac{1}{r}|nlm\rangle = -\frac{1}{e^2}\langle nlm|\frac{\partial H}{\partial Z}|nlm\rangle = -\frac{1}{e^2}\frac{\partial E}{\partial Z} = \frac{1}{e^2}\frac{Z\alpha^2 mc^2}{n^2} = \frac{1}{n^2 a}, \tag{43}$$

which is the first of (34). To find the second relation we can basically repeat the above argument with l in place of Z, but there is one subtlety. For the Hamiltonian (41), the radial eigenvalue problem is well-defined even for non-integer l. But when solving for the radial wavefunction, we find a principle quantum number $n = n_r + l + 1$ where $n_r = 0, 1, 2, \ldots$ must be an integer for the hypergeometric series to terminate and give a normalizable radial wavefunction. When we differentiate with respect to l we must hold n_r, not n, fixed. In other words, $\frac{\partial n}{\partial l} = 1$. Then

$$\begin{aligned}
\langle nlm|\frac{1}{r^2}|nlm\rangle &= \frac{2m}{\hbar^2(2l+1)}\langle nlm|\frac{\partial H}{\partial l}|nlm\rangle & (44)\\
&= \frac{2m}{\hbar^2(2l+1)}\frac{\partial E}{\partial l} & (45)\\
&= \frac{2m}{\hbar^2(2l+1)}\frac{2Z^2\alpha^2 mc^2}{2n^3}\frac{\partial n}{\partial l} & (46)\\
&= \frac{2}{(2l+1)n^3 a^2}. & (47)
\end{aligned}$$

Now that you have seen how to obtain the energy levels for the Klein–Gordon equation, you must be wondering what we do for the Dirac equation. Here is how you do it. Starting from the Dirac equation

$$\left[E + \frac{Ze^2}{r} - c\vec{\alpha}\cdot\vec{p} - mc^2\beta\right]\psi = 0, \tag{48}$$

multiply by

$$\left[E + \frac{Ze^2}{r} + c\vec{\alpha}\cdot\vec{p} + mc^2\beta\right] \tag{49}$$

from the left. Then you find

$$\left[\left(E + \frac{Ze^2}{r}\right)^2 - c^2\vec{p}^2 - (mc^2)^2 + c\vec{\alpha}\cdot\left(-i\hbar\vec{\nabla}\frac{Ze^2}{r}\right)\right]\psi = 0. \tag{50}$$

The anti-commutation relation $\{\alpha^i, \alpha^j\} = 2\delta^{ij}$, $\{\alpha^i, \beta\} = 0$ had been used in simplifying the expression. Writing out the derivative acting on the Coulomb potential, we find

$$\left[\left(E + \frac{Ze^2}{r}\right)^2 - c^2\vec{p}^2 - (mc^2)^2 + i\hbar c\vec{\alpha}\cdot\hat{\vec{r}}\frac{Ze^2}{r^2}\right]\psi = 0, \tag{51}$$

using the notation $\hat{\vec{r}} = \vec{r}/r$. At this point, we also rewrite $\vec{p}^2$ using the spherical coordinates,

$$\left[\left(E + \frac{Ze^2}{r}\right)^2 + c^2\hbar^2\left(\frac{1}{r}\frac{d^2}{dr^2}r - \frac{l(l+1)}{r^2}\right) - (mc^2)^2 + i\hbar c\vec{\alpha}\cdot\hat{\vec{r}}\frac{Ze^2}{r^2}\right]\psi = 0. \quad (52)$$

We can block-diagonalize the matrix $\vec{\alpha}$ as

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix}. \quad (53)$$

Then depending on upper or lower two components, we have $\vec{\alpha}\cdot\hat{\vec{r}} = \pm\vec{\sigma}\cdot\hat{\vec{r}}$. Then the equation becomes

$$\left[E^2 - (mc^2)^2 + 2E\frac{Ze^2}{r} + c^2\hbar^2\left(\frac{1}{r}\frac{d^2}{dr^2}r - \frac{l(l+1) + Z^2\alpha^2 \pm iZ\alpha\vec{\sigma}\cdot\hat{\vec{r}}}{r^2}\right)\right]\psi = 0. \quad (54)$$

The non-trivial point with this equation is to deal with the numerator $l(l+1) + Z^2\alpha^2 \pm iZ\alpha\vec{\sigma}\cdot\hat{\vec{r}}$. The trick is to note that it commutes with $\vec{J} = \vec{L} + \vec{\sigma}/2$. Therefore, we can look at the subspace of the Hilbert space with fixed j and hence $l = j \pm 1/2$. On this space, the numerator has the form

$$l(l+1) + Z^2\alpha^2 \pm iZ\alpha\vec{\sigma}\cdot\hat{\vec{r}} = \begin{pmatrix} (j+\frac{1}{2})(j+\frac{3}{2}) + Z^2\alpha^2 & \mp iZ\alpha \\ \mp iZ\alpha & (j-\frac{1}{2})(j+\frac{1}{2}) + Z^2\alpha^2 \end{pmatrix}. \quad (55)$$

The eigenvlaues of this matrix are easily obtained, but we intentionally write the eigenvalues as $\lambda(\lambda+1)$. The motiation to do so must be clear from what we did with the Klein–Gordon equation. The two solutions are

$$\lambda_+ = \left[\left(j+\frac{1}{2}\right)^2 - Z^2\alpha^2\right]^{1/2}, \qquad \lambda_- = \left[\left(j+\frac{1}{2}\right)^2 - Z^2\alpha^2\right]^{1/2} - 1. \quad (56)$$

Using λ, the Dirac equation is now

$$\left[E^2 - (mc^2)^2 + 2E\frac{Ze^2}{r} + c^2\hbar^2\left(\frac{1}{r}\frac{d^2}{dr^2}r - \frac{\lambda(\lambda+1)}{r^2}\right)\right]\psi = 0. \quad (57)$$

It has the same form as the Klein–Gordon equation except λ. By following the same arguments, we find the energy eigenvalues

$$E = \frac{mc^2}{\sqrt{1 + Z^2\alpha^2/\nu^2}}, \quad (58)$$

with $\nu = \lambda + 1, \lambda + 2, \cdots$. The solutions with both λ_+ and λ_- give the same set of ν's, except that the smallest ν is obtained only from λ_- with $j = 1/2$. This corresponds to the fact that $n = 1$ state has only $l = 0$ which does not mix with an $l = 1$ state. The degeneracy of the eigenvalues for two solutions is split only by Lamb shift. The principal quantum number is ν at the lowest order in $Z\alpha$, and hence

$$\nu = n + \left[\left(j + \frac{1}{2}\right)^2 - Z^2\alpha^2\right]^{1/2} - \left(j + \frac{1}{2}\right). \tag{59}$$

We finally find the energy levels of the Dirac equation

$$E = mc^2 \left[1 + \frac{Z^2\alpha^2}{n - (j + 1/2) + [(j + 1/2)^2 - Z^2\alpha^2]^{1/2}}\right]^{-1/2}. \tag{60}$$