theory and experiment regarding atomic energy levels. All experimental results seemed adequately explained by the Dirac equation, without need for the extra shift $\Delta E_B^{(2)}$. That situation changed in 1947 with the announcement by Lamb and Retherford of the splitting between the $2s_{1/2}$ and $2p_{1/2}$ levels of hydrogen, measured rather accurately by microwave techniques at about 1.05 GHz. The $2p$ state makes an $E1$ transition to the $1s$ state in $\approx 10^{-9}$ sec, so it has a natural line width comparable to the Lamb shift, and the energy level diagram looks something like this:

The $2s_{1/2}$ state also has a natural line width, but it is very small on the scale of the diagram.

The suspicion immediately arose that the $\Delta E_B^{(2)}$ above, properly interpreted, would explain the Lamb shift. This was quickly done by Bethe, who used some of the renormalization ideas that had been under development by Kramers for some time. Bethe had the intuition that the phenomenon was essentially nonrelativistic, and so is his analysis. We now explain Bethe’s calculation.
First we consider the second order energy shift for a free electron (not bound in an atom) due to interactions with the \( \text{EM} \) field. This is a little easier than what we did above, since there is no atom to complicate the \( \text{EM} \). The unperturbed Hamiltonian is now

\[
H_0 = \frac{\mathbf{p}^2}{2m} + \sum_{\lambda} \hbar \omega_{\lambda} \, a_{\lambda}^\dagger a_{\lambda}
\]

with \( H_1, H_2 \) as before. The unperturbed eigenstates of the electron are normalized plane waves \( |\mathbf{p}\rangle \) in a box (later we let \( V \to \infty \)). We assume the initial state of the electron is \( |\mathbf{p}_0\rangle \). The Feynman diagram for the \( H_1^2 \) process is

![Feynman diagram](image)

where \( \mathbf{p}' \) is the momentum of the intermediate state. The energy shift is

\[
\Delta E_{\text{free}}^{(2)} = \sum_{\mathbf{p}' \lambda} \sum_{\mathbf{p}} \frac{\langle \mathbf{p}_0 | H_1 | \mathbf{p}' \lambda \rangle \langle \mathbf{p}' \lambda | H_1 | \mathbf{p}_0 \rangle}{\frac{\mathbf{p}^2}{2m} - \frac{\mathbf{p}'^2}{2m} - \hbar \omega}
\]
Doing the field matrix elements this becomes

$$\Delta E^{\text{free}}_{\text{tot}} = \left(\frac{e}{mc}\right)^2 \left(\frac{2\pi k_e^2}{\nu} \right) \sum_{\vec{\nu}'} \sum_{\mu} \sum_{\nu} \frac{1}{\omega} \left| \langle \vec{\nu}' | \hat{\mathbf{r}} \cdot \hat{\mathbf{E}}_\lambda | \vec{\nu} \rangle \right|^2 \frac{\vec{\nu}_0^2}{2m} - \frac{\vec{\nu}_0'^2}{2m} - \omega$$

The $\vec{\nu}$ inside the matrix element is the operator $-i\hbar \nabla$, so acting on $e^{-i\vec{\nu} \cdot \vec{x}}$ and the plane wave $|\vec{\nu}_0\rangle$, it gets replaced by the c-number

$$\vec{\nu} \rightarrow \vec{\nu}_0 - \hbar \hat{\mathbf{r}}.$$

However, $\hat{\mathbf{r}} \cdot \hat{\mathbf{E}}_\lambda = 0$, so the numerator becomes

$$\text{num.} = \left| \hat{\mathbf{E}}_\lambda ^* \cdot \vec{\nu}_0 - i \hbar \hat{\mathbf{r}} \right| \langle \vec{\nu}' | e^{-i\vec{\nu} \cdot \vec{x}} | \vec{\nu}_0 \rangle \right|^2$$

Now the remaining matrix element is a Kronecker $\delta$,

$$\langle \vec{\nu}' | e^{-i\vec{\nu} \cdot \vec{x}} | \vec{\nu}_0 \rangle = \delta_{\vec{\nu}', \vec{\nu}_0 - \hbar \hat{\mathbf{r}}}$$

which represents momentum conservation at the vertex of the Feynman diagram:

\[ \text{Here} \rightarrow \lambda = (\mu, \vec{p}) \]
Thus the $\vec{p}$ sum can be done, and we have

$$\Delta E_{\text{free}}^{(2)} = \frac{e^2}{m^2} \frac{2\pi \hbar}{\gamma} \sum_{\mu} \sum_{\Omega} \frac{1}{\omega} \frac{1}{\hbar^2} \frac{1}{2m} \frac{(E_k^\mu \cdot \vec{p}_0)^2}{\frac{p_0^2}{2m} - \frac{(\vec{p}_0 - \hbar \vec{k})^2}{2m} - \hbar \omega}.$$ 

Now look at the denominator:

$$\frac{p_0^2}{2m} - \frac{(\vec{p}_0 - \hbar \vec{k})^2}{2m} - \hbar \omega = \frac{p_0^2}{2m} - \frac{p_0^2}{2m} + \frac{\hbar \vec{p}_0 \cdot \vec{k}}{m} - \frac{\hbar \vec{k}^2}{2m} - \hbar \omega.$$ 

The first two terms cancel and if $\frac{p_0^2}{2m}, \hbar \omega \ll m^2$, then the next two terms are negligible, too. Then the denominator becomes simply $-\hbar \omega$. (A basic rule: at NR energies, photons contribute energy, but not much momentum.)

Now the only place the summand depends on $\mu$ or the direction of $\vec{k}$ is in $\hat{E}_k$. Take $\gamma \to \infty$, so

$$\sum_{\vec{k}} \to \frac{\gamma}{(2\pi)^3} \int_0^\infty k^2 dk \int d\Omega_k,$$

and use

$$\sum_{\mu} \int d\Omega_k \frac{1}{\hbar^2} \frac{1}{\omega} (E_k^\mu \cdot \vec{p}_0)^2 = \frac{8\pi}{3} \frac{p_0^2}{\omega}.$$ 

Then

$$\Delta E_{\text{free}}^{(2)} = \frac{4\pi}{3\hbar^2} \frac{e^2}{m^2} \hbar \frac{1}{\gamma} \int_0^\infty k^2 dk \frac{1}{\omega} \frac{p_0^2}{(-\hbar \omega)}.$$
or with \( E = \hbar \omega = \hbar ck \),

\[
\Delta E_{\text{free}}^{(2)} = -\frac{4}{3\pi} \left( \frac{e^2}{\hbar c} \right) \frac{1}{mc^2} \frac{P_0^2}{2m} \int_0^\infty dE \quad (E = \text{photon energy})
\]

This integral also diverges linearly, and has the same problem as the integral for \( \Delta E_8^{(2)} \) if we impose a cutoff on the photon energy \( E_{\text{max}} \).

Let's put an unknown cutoff \( E_{\text{max}} \) on the upper limit to this integral and write

\[
C = -\frac{4}{3\pi} \frac{\alpha}{mc^2} \int_0^{E_{\text{max}}} dE,
\]

\[\Rightarrow \]

\[
\Delta E_{\text{free}}^{(2)} = C \frac{P_0^2}{2m}.
\]

Notice that if \( E_{\text{max}} \sim mc^2 \), then \( C \sim \alpha \) = small, but if we take \( E_{\text{max}} \rightarrow \infty \), then \( C \rightarrow \infty \). \( (\text{C is dimensionless.}) \) Then the total energy of the free electron, through 2nd order perturbation theory, is

\[
E_{\text{free}}^{\text{tot}} = \frac{P_0^2}{2m} + C \frac{P_0^2}{2m} = (1 + C) \frac{P_0^2}{2m}.
\]

We cannot switch off the interaction of the EM field, so when we measure the energy of an electron with momentum...
what we get is \( (1+C) \frac{p_0^2}{2m} \), not \( \frac{p_0^2}{2m} \). Thus we must interpret

\[
(1+C) \frac{p_0^2}{2m} \quad \text{as} \quad \frac{p_0^2}{2\text{mobs}},
\]

where \( \text{mobs} \) is the observed mass of the electron, which differs from the mass parameter \( m \) that enters into the Hamiltonian:

\[
\frac{1+C}{m} = \frac{1}{\text{mobs}},
\]

\( \text{mobs} = m(1-C) \) if \( C \) small. (Note \( C < 0 \), so \( \text{mobs} > m \).

The mass parameter \( m \) is the bare mass, which is not observable.

If we model a classical electron as a particle with a mechanical mass \( m \), then the observed mass must be more because of the energy (= mass by relativity) that is contained in the electrostatic field. Since the electric field (magnitude)

\[ E = \frac{e}{r^2} \],

the total mass contained in at radii \( r > R \) is

\[
\frac{1}{c^2} \int_{4\pi} d^3 \mathbf{r} \frac{e^2}{8\pi r^4} = \frac{1}{c^2} \int_{4\pi} d^3 \mathbf{r} \frac{E^2}{8\pi} = \frac{4\pi}{c^2} \int_{R}^{\infty} d\mathbf{r} \frac{e^2}{8\pi r^4},
\]
If we set this \( m_0 \) and solve for \( R \), we get \( R = \frac{1}{2} r_e \), \( r_e = \frac{e^2}{m_0 c^2} \), the "classical radius of the electron". If we take \( R \to 0 \), we get \( \infty \) (the \( \infty \) self-mass of the electron).

Suppose we use photons of energy \( E_{\text{max}} \) to probe the field of the electron. These have minimum wavelength of order \( \frac{hc}{E_{\text{max}}} \). The total mass contained in the field outside this distance is

\[
\sim \frac{e^2}{hc} \frac{E_{\text{max}}}{m_0 c^2} = a \frac{E_{\text{max}}}{m_0 c^2}.
\]

But this is the same order of magnitude as \( C/m_0 \), so the correction between \( m \) and \( m_{\text{obs}} \) can be regarded as being due to the electron self-energy. If we take the limit \( E_{\text{max}} \to \infty \), we get the full (infinite) self-energy of the electron.

The infinite self-energy of the electron is an inconsistency in classical EM theory that cannot be cured at the classical level. It can only be avoided by not looking at problems too carefully (e.g., in books on classical EM). The difficulty can only be cured at the quantum level. Kramers
might have made greater contributions to renormalization theory, but he believed the correct strategy was to cure the classical infinities first, then quantize.
Return to the Lamb shift. The Lamb shift was discovered experimentally as a splitting between two levels that were predicted on the basis of conventional theory (the Dirac eqn) to be degenerate, but for the purposes of the following discussion let's define the Lamb shift for a single level as the difference between the Dirac prediction and the experimental value. By this definition, the Lamb shift is not the energy \( \Delta E_B^{(2)} \). Here is why.

If our theory is right, then the true (experimental) energy level of an atom must be

\[
E_B^{\text{true}} = E_B + \Delta E_B^{(2)}
\]

\[
= \langle \beta | \frac{\hbar^2}{2m} + V(\hat{r}) | \beta \rangle + \Delta E_B^{(2)}
\]

The first term, however, is not the predictions of conventional theory; rather, that is

\[
E_B^{\text{conv}} = \langle \phi | \frac{\hbar^2}{2m_0} + V(\hat{r}) | \phi \rangle.
\]

When we calculate energy levels we get a formula that contains the mass of the electron, and we must adjust the observed mass to get the conventional prediction of the energy level. We could not use the bare mass if we wanted to, because we have no value for it. But in the expression for \( E_B^{\text{true}} \) above, it is the bare mass that appears, because we are working out the predictions of quantum mechanics,
starting from the Hamiltonian which is parameterized by the bare mass. Thus

$$\Delta E^\text{lamb}_B = E^\text{int}_B - E^\text{conv}_B$$

$$= \langle B | \frac{\not{p}^2}{2m} - \frac{\not{p}^2}{2\mu_{bs}} | B \rangle + \Delta E^2_B$$

$$= - C \langle B | \frac{\not{p}^2}{2m} | B \rangle + \Delta E^2_B,$$

where we use

$$\frac{1}{m} (1 + C) = \frac{1}{m_{bs}}.$$

Both terms in $\Delta E^\text{lamb}_B$ involve divergent integrals (if $E_{\text{max}} \to \infty$); the actual value of the shift is the difference between the divergent terms.

To make the first term look like the second, want a resolution of

the identity so

$$\langle B | \frac{\not{p}^2}{2m} | B \rangle = \frac{1}{2m} \sum_I \langle B | \not{p}_I | B \rangle \cdot \langle I | \not{p}_I | B \rangle = \frac{1}{2m} \sum_I | \not{p}_I |^2.$$

Then using $C$ (a constant), we get

$$\Delta E^\text{lamb}_B = \frac{4}{3\pi} \frac{\alpha}{mc^2} \frac{1}{2m} \sum_I | \not{p}_I |^2 \int_0^{E_{\text{max}}} dE \left( 1 + \frac{E}{E_B - E_I - E} \right)$$

$$\times \left( E_B - E_I \right) \int_0^{E_{\text{max}}} \frac{dE}{E_B - E_I - E}.$$

The integral is still divergent, but now only logarithmically so, so
its value is rather insensitive to the cutoff used.

The integral must be done using the principal value prescription.

The result is

\[ \Delta E_{\text{Lamb}} = \frac{4}{3\pi} \frac{\alpha}{m c^2} \sum_i \frac{|\mathbf{p}_i|}{2m} (E_i - E_B) \ln \left( \frac{E_{\text{max}}}{1 + |E_i - E_B|} \right). \]

To further boil this down, we need to do the sum in \( I \). This is purely a problem in atomic physics. Standard tricks are used, which are summarized in Sakurai's book. At the end some numerical work is required. The result vanishes except for \( \ell = 0 \) (s-waves); in that case, for hydrogen,

\[ \Delta E_{\text{Lamb}} = \frac{4}{3\pi} \frac{\kappa^3}{n^3} \ln \left( \frac{E_{\text{max}}}{\langle |E_i - E_B| \rangle} \right) \frac{e^2}{\alpha_0}, \]

where \( n \) is the principal quantum and the average under the log is defined in a certain way by Bethe. \( (e^2/\alpha_0 = \text{atomic unit of energy}) \) With the cutoff \( E_{\text{max}} = mc^2 \), Bethe got the value 1.04 GHz, very close to the experimental value. This was the first successful application of renormalization theory.

In a proper relativistic calculation, one finds that both energy shifts (for the free and bound electrons) diverge only logarithmically in the cutoff \( E_{\text{max}} \), not linearly as here. Thus, the difference between the divergent integrals actually converges, and no cutoff is required.
The proper relativistic calculation of the Lamb shift was first carried out by French and Weisskopf, later Feynman and Schwinger.

Now we turn to relativistic quantum mechanics. We follow a roughly historical order in presenting this material, even though it involved some misconceptions at various stages. It would be possible to leapfrog directly to the modern point of view, but some important ideas would have to be badly unmotivated.

The first attempts at formulating a relativistic wave equation were made by Schrödinger. He was exploring the ideas of deBroglie, in which a classical energy and momentum $E$ and $p$ are associated with frequency $\omega$ and wavenumber $k$ by $E = \hbar \omega$, $p = \hbar k$, or as operators acting on plane waves,

$$
\begin{align*}
E & \to + i \hbar \frac{\partial}{\partial t} \\
\vec{p} & \to - i \hbar \vec{\nabla}
\end{align*}
$$

Applied to the NR free particle for which $E = \frac{p^2}{2m}$, this yields the usual NR Sch. equ,

$$
i \hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi.
$$

For a relativistic free particle, the energy and momentum are related by

$$
E = \sqrt{c^2 p^2 + m^2 c^4}.
$$
Notice that the energy includes the rest mass $mc^2$, so for low velocities $v/c \ll 1$, the relativistic formula yields

$$E \approx mc^2 + \frac{p^2}{2m}.$$ 

In any case, the relativistic energy-momentum relation yields the wave equation,

$$i\hbar \frac{\partial \psi}{\partial t} = \sqrt{-\hbar^2 c^2 \nabla^2 + m^2 c^4} \psi$$

(still for a free particle.) The square root is hard to interpret. It makes sense in momentum space, but is nonlocal in real space, and it doesn't treat space and time on an equal footing so it's hard to see how it can be covariant. To fix this, Schrödinger squared the classical energy-momentum relation,

$$E^2 = c^2 p^2 + m^2 c^4,$$

which implies a wave equ,

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = -\hbar^2 c^2 \nabla^2 \psi + m^2 c^4 \psi.$$ 

This is called the *Klein–Gordon eqn.* (Schrödinger never published it, and it was later discovered again by Klein and Gordon). An equivalent form is

$$\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = -\left( \frac{mc^2}{\hbar} \right)^2 \psi$$

which shows the d'Alembertian or wave operator on the left. On the
right appears the Compton wavelength

\[ \lambda_c = \frac{\hbar}{mc} \]

of the particle. The Compton wavelength depends on the mass and it has the following interpretation. Let a particle of mass \( m \) be confined to a box of size \( L \).

\[
\begin{array}{c}
\text{L} \\
\text{L}
\end{array}
\]

By the uncertainty principle, the momentum \( \lambda_c \) has the uncertainty \( \Delta p = \frac{\hbar}{L} \).

We ask, how small must \( L \) be such that \( \Delta p \) will reach relativistic values, say, \( mc \)? The answer is \( L \approx \lambda_c \). Note for the photon (\( m = 0 \)), \( \lambda_c = 0 \) and photons are always relativistic. For electrons, relativistic QM is necessary if the electron is examined on distance scales comparable to \( \lambda_c \), which is \( \frac{\hbar}{mc} = \alpha a_0 = \frac{1}{1836} \times \text{size of H-atom} \) (\( m = 0 \)).

Notice for a photon the Klein--Gordon eqn. just becomes the wave eqn,

\[
\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = 0
\]

which has no \( \hbar \) and which of course comes out of the classical Maxwell eqns.