129A Lecture Notes
Quantum ElectroDynamics

1 Quantum ElectroDynamics

The true quantum mechanical and relatistic theory of electromagnetism is called Quantum ElectroDynamics, or QED for shot. It combines Dirac equation to describe electron (and of course positron) and Maxwell equation for photon. The actual calculations of quantum mechanical amplitudes are based on perturbation theory, organized in terms of Feynman diagrams. Introducing QED from the first principle takes a full semester (229A in the case of Berkeley graduate curriculum) and we cannot go through it in this course. Instead, I'd like to give you rough idea on how the theory works, without getting into hardcore formalism.

2 Photons

Early development of quantum mechanics was led by the fact that electro-magnetic radiation was quantized: photons. Here we show that the Maxwell field can be regarded as a collection of harmonic oscillators. Details of this discussion are not used later in this course. This section is for only those of you who are interested in seeing this explicitly.

2.1 Classical Maxwell Field

The vector potential $\vec{A}$ and the scalar potential ϕ are combined in the four-vector potential

$$A^{\mu} = (\phi, \vec{A}). \tag{1}$$

Throughout the lecture notes, we use the convention that the metrix $g_{\mu\nu} = \text{diag}(+1, -1, -1, -1)$ and hence $A_{\mu} = g_{\mu\nu}A^{\nu} = (\phi, -\vec{A})$. The four-vector coordinate is $x^{\mu} = (ct, \vec{x})$, and correspondingly the four-vector derivative is $\partial_{\mu} = (\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla})$. The field strength is defined as $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$, and hence $F_{0i} = \partial_0 A_i - \partial_i A_0 = -\dot{\vec{A}}/c - \vec{\nabla}\phi = \vec{E}$, while $F_{ij} = \partial_i A_j - \partial_j A_i = -\vec{\nabla}_i A^j + \vec{\nabla}_j A^i = -\epsilon_{ijk}\vec{B}^k$.

In the unit we have been using where the Coulomb potential is $QQ'/4\pi r$ without a factor of $1/\epsilon_0$, the action for the Maxwell field is

$$S = -\frac{1}{4} \int dt d\vec{x} \; F^{\mu\nu} F_{\mu\nu} - A_\mu j^\mu = \frac{1}{2} \int dt d\vec{x} \left(\vec{E}^2 - \vec{B}^2 \right). \qquad (2)$$

The gauge invariance of the Maxwell field is that the vector potential A^μ and $A^\mu + \partial^\mu \omega$ (where ω is an arbitrary function of spacetime) give the same field strength and hence the same action. Using this invariance, one can always choose a particular gauge. For most purposes of non-relativistic systems encountered in atomic, molecular, condensed matter, nuclear and astrophysics, Coulomb gauge is the convenient choice, while for highly relativistic systems such as in high-energy physics. We use Lorentz gauge in this lecture note:

$$\partial_\mu A^\mu = \frac{1}{c} \frac{\partial}{\partial t} \phi + \vec{\nabla} \cdot \vec{A} = 0. \qquad (3)$$

Because of this condition, we can simplify the action. The original action is

$$
\begin{aligned}
F_{\mu\nu} F^{\mu\nu} &= (\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \\
&= 2\partial_\mu A_\nu (\partial^\mu A^\nu - \partial^\nu A^\mu) \\
&= 2\partial_\mu A_\nu \partial^\mu A^\nu - 2(\partial^\nu A_\nu)(\partial_\mu A^\mu) + \text{surface terms.} \qquad (4)
\end{aligned}
$$

Using the Lorentz gauge condition, we can drop the second term in the last line. Therefore, the Maxwell action Eq. (2) becomes

$$S = -\frac{1}{2} \int dt d\vec{x} \; \partial_\mu A_\nu \partial^\mu A^\nu. \qquad (5)$$

2.2 Quantization

In order to quantize the Maxwell field, we first determine the "canonically conjugate momentum" for the vector potential A^μ. Following the definition $p^i = \partial L / \partial \dot{q}^i$ in particle mechanics, we define the canonically conjugate momentum from the simplified action Eq. (5),

$$\pi_\mu = \frac{\partial \mathcal{L}}{\partial \dot{A}^\mu} = -\frac{1}{c^2} \dot{A}_\mu. \qquad (6)$$

Following the normal canonical commutation relation $[q^i, p^j] = i\hbar \delta^{ij}$, we set up the equal-time commutation relation

$$[A^\mu(\vec{x}, t), \pi_\nu(\vec{y}, t)] = [A^\mu(\vec{x}, t), -\frac{1}{c^2} \dot{A}_\nu(\vec{y}, t)] = i\hbar \delta^\mu_\nu \delta(\vec{x} - \vec{y}). \qquad (7)$$

To satisfy this commutation relation, we introduce the photon creation and annihilation operators

$$[a^{\mu}(p), a^{\nu\dagger}(q)] = -g^{\mu\nu}\delta_{\vec{p},\vec{q}}, \tag{8}$$

and expand the vector potential and its time derivative as

$$A^{\mu}(\vec{x}) = \sqrt{\frac{\hbar c^2}{2L^3}} \sum_{\vec{p}} \frac{1}{\sqrt{\omega_p}}(a^{\mu}(p)e^{-ip_{\mu}x^{\mu}/\hbar} + a^{\mu\dagger}(p)e^{ip_{\mu}x^{\mu}/\hbar}). \tag{9}$$

Here, $\omega_p = E_p/\hbar = c|\vec{p}|/\hbar$ is the angular frequency for the photon. L is the size of the box with periodic boundary condition. You can check that this momentum-mode expansion together with the equal-time commutation relation Eq. (8) reproduces the canonical commutation relation Eq. (7) as follows.

$$
\begin{aligned}
&[A^{\mu}(\vec{x},t), \dot{A}^{\nu}(\vec{y},t)] \\
&= \frac{\hbar c^2}{2L^3} \sum_{\vec{p},\vec{q}}(-i)[a^{\mu}(p)e^{-ip_{\mu}x^{\mu}/\hbar} + a^{\nu\dagger}(p)e^{ip_{\mu}x^{\mu}/\hbar}, a^{\mu}(p)e^{-iq_{\mu}y^{\mu}/\hbar} - a^{\nu\dagger}(q)e^{iq_{\mu}y^{\mu}/\hbar}] \\
&= \frac{\hbar c^2}{2L^3}i\sum_{\vec{p}}(-g^{\mu\nu})(e^{ip\cdot(\vec{x}-\vec{y})} + e^{-ip\cdot(\vec{x}-\vec{y})}) \\
&= -\hbar c^2 i g^{\mu\nu}\delta(\vec{x}-\vec{y}). \tag{10}
\end{aligned}
$$

At the last step, we used the correspondence in the large volume limit $\sum_{\vec{p}} = L^3 \int d\vec{p}/(2\pi\hbar)^3$.

The problem with what we have done so far is that we have not imposed the Lorentz gauge condition Eq. (3) on the vector potential yet. Acting the divergence on the momentum-mode expansion Eq. (9), we need

$$p_{\mu}a^{\mu}(p) = \frac{1}{c}\omega_p a^0(p) - \vec{p}\cdot\vec{a}^{\mu}(p) = 0. \tag{11}$$

Out of four independent annihilation operators a^{μ} for $\mu = 0, 1, 2, 3$, this constraint removes on linear combination of them. To satisfy this constraint, we introduce the polarization vectors. When $p^{\mu} = p(1, 0, 0, 1)$, the positive helicity (right-handed) circular polarization has the polarization vector $\epsilon^{\mu}_{+} = (0, 1, i, 0)/\sqrt{2}$, while the negative helicity (left-handed) circular polarization is represented by the polarization vector $\epsilon^{\mu}_{-} = (0, 1, -i, 0)/\sqrt{2}$. Two other polarization vectors are $\epsilon^{\mu}_{S} = (1, 0, 0, 1)/\sqrt{2}$ ("scalar") and $\epsilon^{\mu}_{L} = (1, 0, 0, -1)/\sqrt{2}$

3

("longitudinal"). But the longitudinal polarization vector does not satisfy the constraint Eq. (11), and this degree of freedom is removed. The scalar one is actually a gauge freedom, because under the gauge transformation $A^\mu \to A^\mu + \partial^\mu \omega$, a^μ changes by an amount proportional to p^μ and hence to ϵ_S^μ. Therefore we can also omit ϵ_S^μ because it is unphysical.

In general, for the momentum vector $p^\mu = p(1, \sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, the circular polarization vectors are given by

$$\epsilon_\pm^\mu(p) = \frac{1}{\sqrt{2}}\left(\pm\epsilon_1^\mu + i\epsilon_2^\mu\right), \tag{12}$$

where the linear polarization vectors are given by

$$\begin{align}
\epsilon_1^\mu(p) &= (0, \cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta), \tag{13}\\
\epsilon_2^\mu(p) &= (0, -\sin\phi, \cos\phi, 0). \tag{14}
\end{align}$$

We can check that these two vectors are orthogonal:

$$\epsilon_\lambda^{\mu*}(p)\epsilon_{\mu\lambda'}(p) = \delta_{\lambda,\lambda'}. \tag{15}$$

This property is satisfied for both the basis with linear polarizations $i, j = 1, 2$ and the helicity basis $i, j = \pm$.

Given the polarization vectors, we re-expand the vector potential in terms of the creation/annihilation operators

$$A^\mu(\vec{x}) = \sqrt{\frac{2\pi\hbar c^2}{L^3}}\sum_{\vec{p}}\frac{1}{\sqrt{\omega_p}}\sum_\pm(\epsilon_\pm^\mu(p)a_\pm(p)e^{-ip_\mu x^\mu/\hbar} + \epsilon_\pm^\mu(p)^*a_\pm^\dagger(p)e^{ip_\mu x^\mu/\hbar}) \tag{16}$$

With this expansion, the Lorentz gauge condition is automatically satisfied, while the creation/annihilation operators obey the commutation relations

$$[a_\lambda(p), a_{\lambda'}^\dagger(q)] = \delta_{\lambda,\lambda'}\delta_{\vec{p},\vec{q}} \tag{17}$$

for $\lambda, \lambda' = \pm$. We could also have used the linearly polarized photons

$$A^\mu(\vec{x}) = \sqrt{\frac{2\pi\hbar c^2}{L^3}}\sum_{\vec{p}}\frac{1}{\sqrt{\omega_p}}\sum_{\lambda=1}^2(\epsilon_\lambda^\mu(p)a_\lambda(p)e^{-ip_\mu x^\mu/\hbar} + \epsilon_\lambda^\mu(p)a_\lambda^\dagger(p)e^{ip_\mu x^\mu/\hbar}) \tag{18}$$

with

$$[a_\lambda(p), a_{\lambda'}^\dagger(q)] = \delta_{\lambda,\lambda'}\delta_{\vec{p},\vec{q}} \tag{19}$$

4

for $\lambda, \lambda' = 1, 2$. Clearly, two sets of operators are related by

$$a_1 = \frac{1}{\sqrt{2}}(a_+ + a_-), \qquad a_2 = \frac{i}{\sqrt{2}}(a_+ - a_-). \tag{20}$$

Once we have the mode expansion for the vector potential, one can work out the Hamiltonian

$$H = \int d\vec{x} \frac{1}{2} \left(\vec{E}^2 + \vec{B}^2 \right) = \sum_{\vec{p}} \sum_{\pm} \hbar \omega_p (a_{\pm}^{\dagger}(p) a_{\pm}(p) + \frac{1}{2}). \tag{21}$$

Because $\hbar \omega_p = c|\vec{p}|$, the dispersion relation for the photon is the familiar one $E = c|\vec{p}|$ for a massless relativistic particle.

The "vacuum" which does not have any photons is simply defined in the similar way as in the harmonic oscillator,

$$a^{\mu}(p)|0\rangle = 0. \tag{22}$$

A single photon state is given by acting a creation operator on the vacuum,

$$|\vec{p}, \lambda\rangle = a_{\lambda}^{\dagger}(p)|0\rangle. \tag{23}$$

Multi-photon states are obtained by acting the creation operator multiple times. This way, the system of photons is nothing but (an infinite collection of) harmonic oscillators.

3 Feynman Diagrams

We now have electrons and positrons described by Dirac equation, and also photons from quantized Maxwell theory. What we need next is their interaction. The inteaction is described in terms of Feynman Diagrams, which represent systematic perturbative expansion in powers of fine-structure constant $\alpha = e^2/\hbar c = 1/137$.

The basic ingredient is the emission and absorption of a photon by an electron (or a positron). This is all there is to describe all electromagnetic interactions in the QED.

Let us first look at a classic example of Rutherford scattering. You have a point positive charge situated at the origin, and you send in another charged particle (say, an electron). The point charge creates a Coulomb field around

it. The electron trajectory gets bent because of the attractive Coulomb force. Having said that the emission and absorption of a photon is all there is to describe electromagnetic interactions, we have to explain this familiar process using the language of quantized photons.

This is how it goes. The positive point charge emits a "virtual" photon, which is the quantum description of the classical Coulomb field. Then the incoming electron absorbes the virtual photon. Because the virtual photon carries a finite momentum away from the positive charge and the electron acquires it, the electron receives a kick, the quantum description of the bending of the electron.

What is a virtual photon? If you look at the kinematics of an electron absorbing a photon, it is easy to see that the photon cannot be a real photon. Take the electron initial (final) four-momentum to be $p_i^\mu = (E_i/c, 0, 0, p_i)$ ($p_f^\mu = (E_f/c, p_f \sin\theta, 0, p_f \cos\theta)$) without a loss of generality. Of course they satisfy the standard dispersion relation $p_i^2 = p_f^2 = m_e^2 c^2$. The photon brings in a four-momentum $q^\mu = p_f^\mu - p_i^\mu = ((E_f - E_i)/c, p_f \sin\theta, 0, p_f \cos\theta - p_i)$. The photon, on the other hand, has to satisfy the massless dispersion relation $q^2 = 0$. By substituting the expressions,

$$q^2 = (p_f - p_i)^2 = p_f^2 + p_i^2 - 2 p_f \cdot p_i = 2 m_e^2 c^2 - 2 \left(\frac{E_f}{c} \frac{E_i}{c} - p_i p_f \cos\theta \right). \quad (24)$$

To make the discussion simple, let us assume that the positive charge has an infinite mass (no recoil), and the initial and the final energies of the electron are the same. Then, $E_i = E_f$, and

$$q^2 = 2 m_e^2 c^2 - 2(E^2/c^2 - p^2 \cos\theta) = -2 p^2 (1 - \cos\theta) \quad (25)$$

This quantity must vanish for a photon, but it happens only if $\theta = 0$, or in other words if the electron has not been "kicked" at all. This process is impossible. At least for a "real" photon.

A virtual photon violates energy conservation to the extent the quantum mechanics allows. For the Rutherford scattering process, the initial and final electrons have the same energies (for an infinitely heavy scatterer). The virtual photon does not carry any energy but a finite momentum to "kick" the electron. Of course a real photon with a finite momentum has a finite energy, but this photon doesn't, hence virtual. But quantum mechanics allows a brief violation of energy conservation as long as the energy is recovered without getting detected. The state with energy violation ΔE is allowed to exist

only over the time interval Δt given by the energy-time uncertainty relation $\Delta E \Delta t \sim \hbar$. The more energy is violated, the shorter the state survives.

In the case of Coulomb scattering, a "kick" is due to the momentum of the virtual photon $\vec{q}^2 = -q^2 = 2p^2(1 - \cos\theta) = 4p^2 \sin^2 \theta/2$. This photon would normally have the energy $c|\vec{q}|$, but it's energy vanishes because the initial and final energies of the electron are the same. This is the violation of energy conservation by $\Delta E = 2cp \sin \theta/2$. This virtual photon is allowed to exist only for $\Delta t \sim \hbar/\Delta E = \hbar/(2cp \sin \theta/2)$, and hence can go only over the distance $c\Delta t \sim \hbar/(2p \sin \theta/2)$. This is commensurate with the classical idea that small kick (small θ) corresponds to a large impact parameter.

If two electrons scatter, called Møller scattering, they exchange a virtual photon. But two electrons are identical, and you can draw another diagram where two final state electrons are interchanged. The quantum mechanical amplitude is given by the first diagram *minus* the second diagram, where the relative minus sign represents that of the Fermi statistics.

Positrons are represented by lines going backward in time. Don't try to attach any philosophical meaning to this statement. It just turns out to be a convenient book keeping of electrons and positrons in the Feynman diagrams. Let us take the arrow of time to be upwards in the diagram. When an electron and a positron scatter, they can exchange a virtual photon sideways, but also upwards. This new diagram represents the annihilation of the electron and the positron into a virtual photon, which materializes as a pair of an electron and a positron. In the center-of-momentum frame, the virtual photon is pure energy with no momentum. Again the vanishing momentum suggests vanishing energy for the photon, but it has finite energy and hence "virtual."

The latter diagram can be used to produce other types of matter other than electron/positron. For instance, the virtual photon can become a pair of a muon and an anti-muon. The muon is similar to the electron, with the same electric charge, spin, and statistics, but 200 times heavier and decays within $2\mu \sec$.

An electron can also be "virtual." When you consider Compton scattering, it proceeds via two different Feynman diagram. In one diagram, the electron absorbs the photon, and becomes a "virtual" electron. This virtual electron emits a photon and becomes real again. In the other diagram, the electron emits a photon first and becomes virtual. Then the virtual electron absorbs the photon and becomes real. These two diagrams are added together with positive sign, and allow you to work out the rate of Comp-

7

ton scattering, its angular distribution, etc. The resulting formula is called Klein–Nishina formula, and is one of the earliest quantitative success of the QED.
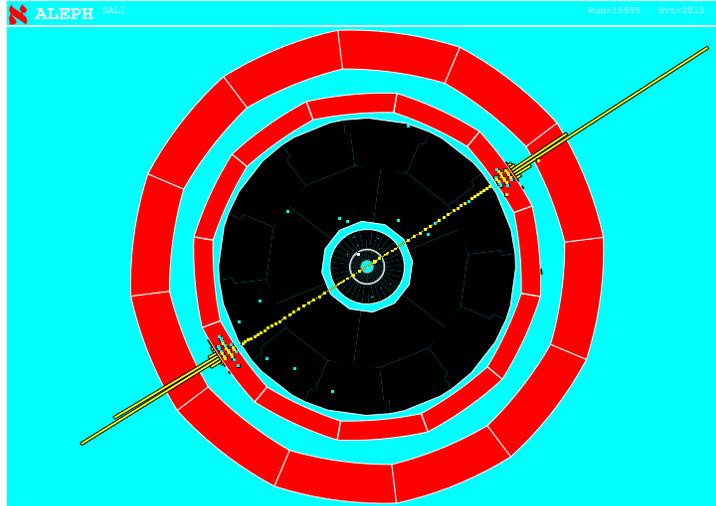

Figure 1: Bhabha scattering. The tracks are and electron and a positron because they shower in the electromagnetic calorimeter, represented in the inner red circle. The electron and positron beams are perpendicular to the event display and they collide at the center. The display is a projection on the plane transverse to the beams.

4   Casimir Effect

So far, we had not cared much about the zero-point energy of the photons in Eq. (21). It is often said that the zero-point energy just amounts to the baseline energy and all other energies are measured relative to it. In other words, pretending it doesn't exist is enough. But the zero-point energy plays a role if a change in the system affects the zero-point energy itself. (See Milonni, P. W., and Shih, M. L., 1992, *Contemp. Phys.* **33**, 313. for a review.)

To be concrete, put two conducting plates parallel to each other. The distance between the plates is d (at $z = 0$ and $z = d$), and the plates are very large of area L^2 where we take $L \to \infty$. To simplify the discussions,
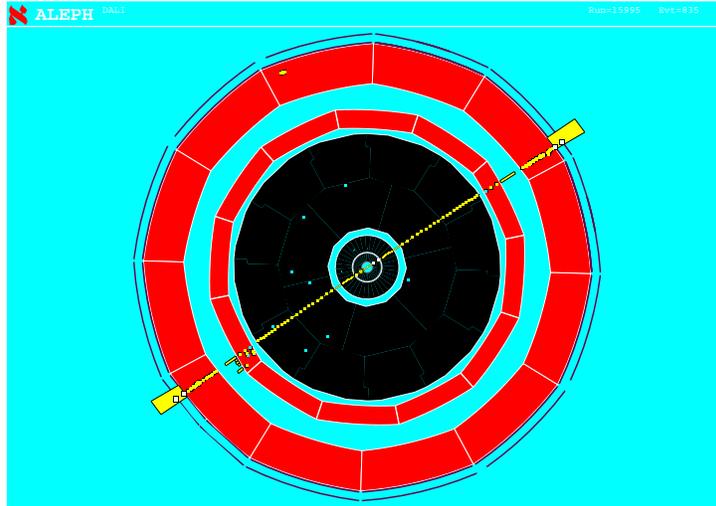
Figure 2: Pair production of muons from electron-positron annihilation. They are muons because they penetrate both layers of calorimeters and hit the "muon chamber," the outermost layers represented in purple.
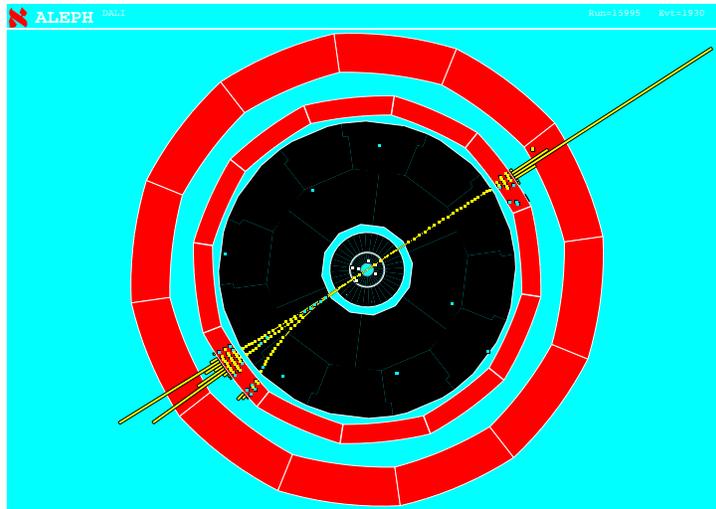

Figure 3: Pair production of tau leptons from electron-positron annihilation. One tau decays into an electron and two neutrinos, where the electrons showers in the first layer of calorimeter. The other tau decays into three pions and a neutrino where the pions are absorbed in the calorimeter.

we also place thin plates at $x = 0$, $x = L$ and $y = 0$, $y = L$ to form a box. A conducting plate imposes boundary conditions on the radiation field $E_\parallel = 0$ and $B_\perp = 0$. Given these boundary conditions, the vector potential is expanded in modes as

$$\vec{A}(\vec{x}) \sim \vec{\epsilon} \sin(n_x, n_y, n_z) \left(\frac{\pi n_x}{L} x + \frac{\pi n_y}{L} y + \frac{\pi n_z}{d} z \right), \qquad (26)$$

for n_x, n_y, and n_z non-negative integers. For each wave vector $\vec{k} = (\pi n_x/L, \pi n_y/L, \pi n_z/d)$, the polarization vector is transverse: $\vec{k} \cdot \vec{\epsilon} = 0$. If one of the n's vanishes, however, say $\vec{k} = (0, \pi n_y/L, \pi n_z/d)$, the polarization vector $\vec{\epsilon} = (1, 0, 0)$ is not allowed because that would give $A_x \neq 0$ for any x (and generic y, z) and violates the boundary condition $E_\parallel = 0$. Therefore, we have only one polarization whenever one of n's vanishes.

The sum of zero point energies in this set up is therefore given by

$$U(d) = \sum{}' \frac{1}{2} \hbar \omega \times 2 = \sum_{n_x, n_y, n_z}{}' \hbar c \left[\left(\frac{\pi n_x}{L} \right)^2 + \left(\frac{\pi n_y}{L} \right)^2 + \left(\frac{\pi n_z}{d} \right)^2 \right]^{1/2}. \qquad (27)$$

The summation $\sum'$ means that whenever one of n's vanishes, we drop the multiplicity 2 for possible transverse polarizations. Because we regard the size of the plate to be large $L \to \infty$, we can replace the sum over n_x, n_y in terms of integrals over $k_{x,y} = \pi n_{x,y}/L$,

$$U(d) = \left(\frac{L}{\pi} \right)^2 \sum_{n_z}{}' \int_0^\infty dk_x \int_0^\infty dk_y \hbar c \left[k_x^2 + k_y^2 + \left(\frac{\pi n_z}{d} \right)^2 \right]^{1/2}. \qquad (28)$$

Recall that we are interested in the *difference* in the zero-point energy when d is varied. Therefore, it is useful to compare it to the case when the plates don't exist. We can compute the zero-point energy density in the infinite volume as usual, and then calcualte the energy $U_0(d)$ multiplied by the volume $L^2 d$:

$$\begin{aligned} U_0(d) &= L d^2 \int \frac{d\vec{k}}{(2\pi)^3} \hbar c \left[k_x^2 + k_y^2 + k_z^2 \right]^{1/2} \\ &= L d^2 \int_0^\infty \frac{dk_x}{\pi} \int_0^\infty \frac{dk_y}{\pi} \int_0^\infty \frac{dk_z}{\pi} \hbar c \left[k_x^2 + k_y^2 + k_z^2 \right]^{1/2}. \end{aligned} \qquad (29)$$

What is observable is the difference

$$U(d) - U_0(d)$$

10

$$= \frac{L^2}{\pi^2} \hbar c \int_0^\infty dk_x dk_y \left\{ \sum_{n_z}' \left[k_x^2 + k_y^2 + \left(\frac{\pi n_z}{d} \right)^2 \right]^{1/2} - \frac{d}{\pi} \int_0^\infty dk_z \left[k_x^2 + k_y^2 + k_z^2 \right]^{1/2} \right\}. \tag{30}$$

Switching to the circular coordinates,

$$\int_0^\infty dk_x dk_y = \int_0^\infty k_\perp dk_\perp \int_0^{\pi/2} d\phi = \frac{\pi}{4} \int_0^\infty dk_\perp^2,$$

we find

$$U(d) - U_0(d)$$
$$= \frac{L^2}{\pi^2} \hbar c \frac{\pi}{4} \int_0^\infty dk_\perp^2 \left\{ \sum_{n_z}' \left[k_\perp^2 + \left(\frac{\pi n_z}{d} \right)^2 \right]^{1/2} - \frac{d}{\pi} \int_0^\infty dk_z \left[k_\perp^2 + k_z^2 \right]^{1/2} \right\}. \tag{31}$$

This expression appears problematic because it looks badly divergent. The divergence appears when the wave vector is large, which corresponds to high frequency photons. The point is that the conducting plates are transparent to, say, gamma rays, or in general for photons whose wavelengths are shorter than interatomic separation a. Therefore, there is naturally a damping factor $f(\omega) = f(c(k_\perp^2 + k_z^2)^{1/2})$ with $f(0) = 1$ which smoothly cuts off the integral for $\omega \gtrsim c/a$. In particular, $f(\infty) = 0$. Then the expression is safe and allows us to use standard mathematical tricks. Changing to dimensionless variables $u = (k_\perp d/\pi)^2$ and $n_z = k_z d/\pi$,

$$U(d) - U_0(d) = \frac{\pi^2 L^2 \hbar c}{4 d^3} \int_0^\infty du \left\{ \sum_{n_z}' \sqrt{u + n_z^2} - \int_0^\infty dn_z \sqrt{u + n_z^2} \right\} f(\omega). \tag{32}$$

Interchanging the sum and integral, we define the integral

$$F(n) = \int_0^\infty du \sqrt{u + n^2} f(\omega) = \int_{n^2}^\infty du \sqrt{u} f(\omega) \tag{33}$$

with $\omega = \pi c \sqrt{u}/d$ in the last expression. Using this definition, we can write

$$U(d) - U_0(d) = \frac{\pi^2 L^2 \hbar c}{4 d^3} \left\{ \frac{1}{2} F(0) + \sum_{n=1}^\infty F(n) - \int_0^\infty dn F(n) \right\}. \tag{34}$$

11

Here we can use the Euler–McLaurin formula

$$\frac{1}{2}F(0) + \sum_{n=1}^{\infty} F(n) - \int_0^{\infty} F(n) = -\frac{1}{2!}B_2 F'(0) - \frac{1}{4!}B_4 F'''(0) - \cdots . \quad (35)$$

The coefficientss B_k are Bernouilli numbers defined by the Taylor expansion

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} B_k \frac{x^k}{k!}, \quad (36)$$

and $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_4 = -1/30$, and all odd ones B_{2n-1} vanish except B_1.

Here is a physicist's proof of the Euler–McLaurin formula. Using the Taylor expansion that defines the Bernouilli numbers, we replace x to a derivative operator ∂_x

$$\frac{\partial_x}{e^{\partial_x} - 1} = \sum_{k=0}^{\infty} B_k \frac{\partial_x^k}{k!}. \quad (37)$$

We act this operator on a function $F(x)$ and integrate it over x from 0 to ∞,

$$\int_0^{\infty} dx \frac{\partial_x}{e^{\partial_x} - 1} F(x) = \int_0^{\infty} dx \sum_{k=0}^{\infty} B_k \frac{\partial_x^k}{k!} F(x). \quad (38)$$

The l.h.s. of Eq. (38) is then

$$\int_0^{\infty} \sum_{k=0}^{\infty} B_k \frac{\partial_x^k}{k!} F(x) = \int_0^{\infty} F(x) dx + \sum_{k=1}^{\infty} B_k \left[\frac{\partial_x^{k-1}}{k!} F(x) \right]_0^{\infty}$$

$$= \int_0^{\infty} F(x) dx - \sum_{k=0}^{\infty} B_k \frac{1}{k!} F^{(k-1)}(0)$$

$$= \int_0^{\infty} F(x) dx + \frac{1}{2} F(0) - \sum_{k=1}^{\infty} B_{2k} \frac{1}{(2k)!} F^{(2k-1)}(0). \quad (39)$$

On the other hand, the r.h.s. of Eq. (38) is

$$\int_0^{\infty} dx \frac{\partial_x}{e^{\partial_x} - 1} F(x). \quad (40)$$

By Taylor expanding the denominator,

$$\frac{\partial_x}{e^{\partial_x} - 1} F(x) = -\sum_{n=0}^{\infty} \partial_x e^{n\partial_x} F(x). \quad (41)$$

Note that $e^{a\partial_x} F(x) = \sum_{m=0}^{\infty} \frac{1}{m!} a^m F^{(m)}(0) = F(x + a)$ and hence the r.h.s. is

$$\int_0^{\infty} dx \frac{\partial_x}{e^{\partial_x} - 1} F(x) = -\int_0^{\infty} dx \partial_x \sum_{n=0}^{\infty} F(x + n) = \sum_{n=0}^{\infty} F(n). \quad (42)$$

12

By comparing both sides of Eq. (38), we now find

$$\sum_{n=0}^{\infty} F(n) = \int_0^{\infty} F(x)dx + \frac{1}{2}F(0) - \sum_{k=1}^{\infty} B_{2k}\frac{1}{(2k)!}F^{(2k-1)}(0). \tag{43}$$

Moving first two terms in the r.h.s. to the l.h.s, we obtain the Euler–McLaurin formula

$$\frac{1}{2}F(0) + \sum_{n=1}^{\infty} F(n) - \int_0^{\infty} F(x)dx = -\sum_{k=1}^{\infty} B_{2k}\frac{1}{(2k)!}F^{(2k-1)}(0). \tag{44}$$

Wondering why all odd B_k vanish except for B_1? It is easy to check that

$$\frac{x}{e^x - 1} + \frac{1}{2}x = \frac{x(2 + e^x - 1)}{2(e^x - 1)} = \frac{x}{2}\coth\frac{x}{2} \tag{45}$$

which is manifestly an even function of x.

Going back to the definition of our function $F(n)$ Eq. (33), we find

$$F'(n) = -2n^2 f(\pi cn/d). \tag{46}$$

As we will see below, we have d of order micron in our mind. This distance is far larger than the interatomic spacing, and hence $f(\pi cn/d)$ is constant $f = 1$ in the region of our interest. Therefore, we can ignore derivatives $f^{(n)}(0)$, and hence the only important term in the Euler–McLaurin formula Eq. (35) is $F'''(0) = -4$. We obtain

$$U(d) - U_0(d) = -\frac{\pi^2 L^2 \hbar c}{4d^3}F'''(0)\frac{1}{4!}B_4 = -\frac{\pi^2 L^2 \hbar c}{720d^3}. \tag{47}$$

In other words, there is an *attractive* force between two conducting plates

$$F = [U(d) - U_0(d)]' = \frac{\pi^2 L^2 \hbar c}{240d^4} \tag{48}$$

which is numerically

$$\frac{0.013\text{dyn/cm}^2}{(d/\mu\text{m})^3} \tag{49}$$

per unit area. This is indeed a tiny force, but Sparnay has observed it for the first time in 1958. He placed chromium steel and aluminum plates at distances between 0.3–2μm, attached to a spring. The plates are also connected to a capacitor. By measuring the capacitance, he could determine the distance, while the known sping constant can convert it to the force.

13

5 Lamb Shift

We learned with Dirac equation that states of hydrogen atom with the same principal quantum number n and the total angular momentum j remain degenerate despite the corretions from spin-orbit coupling, relativistic corrections, and Darwin term. They are, however, split as a result of full quantum interactions between the electron and photons. This is what Willis Lamb found after he worked on war-time radar technology during the WWII and came back to his lab, applied his radar technology to the hydrogen atom. He found transition spectrum between $2s_{1/2}$ and $2p_{1/2}$ states of about 1 GHz.

You heard about the Darwin term pushing the s-states up because the Zitterbewegung smears the electric field and gives rise to a delta-function potential at the origin. If there is an additional reason for the jitter of the electron, it would contribute more to the similar effect. The additional reason is the zero-point fluctuation of the radiation field. Each momentum mode of the photon has the zero-point fluctuation, and each of them jiggles the electron. That would make the electron jitter a little bit more in addition to the Zitterbewegung and pushes the s-state further up.

Let us for simplicity treat the electron non-relativistically and see how much it gets jiggled by the zero-point motion of the electric field. The classical equation of motion for the "jiggle" part of the electron position is

$$\delta\ddot{\vec{x}} = \frac{e}{m}\vec{E}. \tag{50}$$

As we discussed in the case of the Darwin term generated by the Zitterbewegung, such a "jiggle" would smear the electric field and generate additional potential term

$$\Delta V = \frac{1}{2}\langle \delta x^i \delta x^j\rangle \frac{\partial^2 e\phi}{\partial x^i \partial x^j} = \frac{1}{6}\langle(\delta\vec{x})^2\rangle\Delta(e\phi). \tag{51}$$

We are interested in the electric field caused by the zero-point motion. For each frequency mode of the photon, with frequency ω, we find then

$$\delta\vec{x}_\omega = -\frac{e}{m\omega^2}\vec{E}_\omega. \tag{52}$$

Therefore the size of the fluctuation in the electron position is

$$\langle(\delta\vec{x}_\omega)^2\rangle = \frac{e^2}{m^2\omega^4}\langle\vec{E}_\omega^2\rangle. \tag{53}$$

14

For each frequency mode, the zero-point energy of the photon is

$$\int d\vec{x} \frac{1}{8\pi} (\vec{E}_\omega^2 + \vec{B}_\omega^2) = \frac{1}{2}\hbar\omega. \tag{54}$$

Because $\vec{E}^2 = \vec{B}^2$ for a photon, we find

$$\langle \vec{E}_\omega^2 \rangle = \frac{2\pi}{L^3}\hbar\omega. \tag{55}$$

We now sum the contribution of each momentum modes incoherently (because each of them gives a random jiggling in an uncorrelated way) for both polarization states, and find

$$
\begin{aligned}
\langle (\delta\vec{x})^2 \rangle &= 2L^3 \int \frac{d\vec{k}}{(2\pi)^3} \langle (\delta\vec{x}_\omega)^2 \rangle \\
&= 2L^3 \int \frac{4\pi k^2 dk}{(2\pi)^3} \frac{e^2}{m^2\omega^4} \frac{2\pi}{L^3}\hbar\omega \\
&= \frac{2}{\pi}\frac{e^2}{m^2} \int \frac{4\pi k^2 dk}{(2\pi)^3} \frac{1}{\omega^4}\hbar\omega \\
&= \frac{2\hbar e^2}{\pi m^2 c^3} \int \frac{d\omega}{\omega}.
\end{aligned}
\tag{56}
$$

The integral over ω is logarithmically divergent both in the ultraviolet ($\omega \to \infty$) and the infrared ($\omega \to 0$).

Our calculation so far is clearly not valid when the electron is jiggled by a photon of ω larger than $mc^2/\hbar$ because the electron would become relativistic. Therefore, we assume that the correct fully relativistic treament will cut off the integral around $\omega \sim mc^2/\hbar$. Indeed, calculations using Feynman diagrams show that this is indeed the case. Because the integral depends only logarithmically on the cutoff, it should not make a big difference if the cutoff is actually $2mc^2/\hbar$ or $mc^2/2\hbar$, etc. We take it to be simply $mc^2/\hbar$. In the infrared, frequency smaller than the time for the electron to move inside the atom should not affect the result because the "jiggle" would be overshadowed by the motion of the electron itself. Therefore the integral is cutoff at around $\omega \sim cZa_B^{-1} = mcZe^2/\hbar^2$. As a result, the size of the "jiggling" is approximately

$$\langle (\delta\vec{x})^2 \rangle \simeq \frac{2\hbar e^2}{\pi m^2 c^3} \log \frac{mc^2/\hbar}{mcZe^2/\hbar^2} = \frac{2\hbar^2 \alpha}{\pi m^2 c^2} \log \frac{1}{Z\alpha}, \tag{57}$$

15

with $\alpha = e^2/\hbar c = 1/137$. Following the calculation of the energy shift of s-states from the Darwin term, we find the additional potential term Eq. (51) to be

$$
\begin{aligned}
\Delta V &= \frac{1}{6}\langle (\delta \vec{x})^2 \rangle \Delta (e\phi) \\
&\simeq \frac{1}{6}\frac{2\hbar^2 \alpha}{\pi m^2 c^2} \log \frac{1}{Z\alpha} 4\pi Z e^2 \delta(\vec{x}) \\
&= \frac{4\hbar^3 Z\alpha^2}{3m^2 c} \log \frac{1}{Z\alpha} \delta(\vec{x}).
\end{aligned}
\tag{58}
$$

The resulting energy shift for nS-states is

$$
\begin{aligned}
\Delta E_n &\simeq \frac{4\hbar^3 Z\alpha^2}{3m^2 c} \log \frac{1}{Z\alpha} |\psi_n(0)|^2 \\
&= \left[\frac{8(Z\alpha)^4 \alpha}{3\pi} \log \frac{1}{Z\alpha} \right] mc^2 \frac{1}{2n^3}.
\end{aligned}
\tag{59}
$$

Therefore, this contribution is suppressed relative to the fine-structure by α/π but is enhanced by a logarithm $\log \alpha^{-1} = 4.9$. For $n = 2$, it gives about 1 GHz for the microwave resonance frequency between $2p$ and $2s$, in rough agreement with data as we will see below.

The standard calculation uses Feynman diagrams, where the electron emits a virtual photon before it interacts with the Coulomb potential, and after the interaction it reabsorbes the virtual photon. This diagram, called the vertex correction, is actually divergent both in the ultraviolet and the infrared; reminiscent of the discussion above. It turns out, however, that the ultraviolet divergence is of a different character. The piece that corresponds to the amount of "jiggling" of the electron, more correctly called the "charge radius" of the electron, is actually ultraviolet finite in the fully relativistic calculations, supporting the rough "cutoff" at $\omega \sim mc^2/\hbar$ employed above. The ultraviolet divergence in this Feynman diagram is properly cancelled by another ultraviolet divergence called "wave function renormalization." When you use the second-order perturbation theory, your state $|\phi_n\rangle$ is modified to

$$
|\phi_n\rangle \rightarrow |\chi_n\rangle = |\phi_n\rangle + \sum_{i\neq n}\frac{|i\rangle\langle i|V|0\rangle}{E_0 - E_i} + \sum_{i,j\neq n}\frac{|j\rangle\langle j|V|i\rangle\langle i|V|0\rangle}{(E_n - E_j)(E_n - E_i)},
\tag{60}
$$

where $|0\rangle$ is the zeroth order state (not the vacuum) and $|i\rangle$ other states that mix with $|0\rangle$ due to the perturbation V. However, this state is not correctly

16

normalized, because

$$\langle \chi_n | \chi_n \rangle = 1 + \sum_i \frac{|\langle \phi_n | V | i \rangle|^2}{(E_n - E_i)^2}. \tag{61}$$

To correctly normalize the perturbed state $|\chi_n\rangle$, we need to "renormalize" it as

$$|\chi_n\rangle' = |\chi_n\rangle \frac{1}{1 + \frac{1}{2} \sum_i |\langle \phi_n | V | i \rangle|^2 / (E_n - E_i)^2} \tag{62}$$

In the case of the QED, it corresponds to the Feynman diagram where the electron emits a virtual photon and reabsorbes it without any other interactions, which is also ultraviolet divergent. This change in the normalization of the state can be shown to precisely cancel the ultraviolet divergence in the vertex correction, and hence there is no problem with the apparent divergences.

The present theoretical and experimental situation is reviewed, for example, in M.I. Eides, H. Grotch and V.A. Shelyuto, "Theory of light hydrogen-like atoms," *Phys. Rep.*, **342**, 63-261, (2001). The best experimental value of the $2s$–$2p$ splitting is

$$1.057\ 845(3)\ \text{GHz}. \tag{63}$$

The theoretical calculations depend on variety of other corrections in addition to the effect I had discussed, including the fact that the charge of the proton is not strictly point-like. The charge radius of the proton is not well determined experimentally, and limits the theoretical accuracy in calculating the level splitting. Using one particular measurement of the proton charge radius $0.862(12)$ fm, the theory gives

$$1.057\ 833(2)(4)\ \text{GHz}, \tag{64}$$

which disagrees with data at more than 2 sigma level. But other measurements of the proton charge radius disagree with this measurement, and the discrepancy becomes larger. The inconsistency among the data makes it impossible for us to draw any conclusions beyond a simple qualitative statement that theory and data agree for 6 digits.

6   Anomalous Magnetic Moment

We learned from the Dirac equation that the gyromagnetic ratio $g = 2$. This is certainly in a good agreement with data. But both experiments and
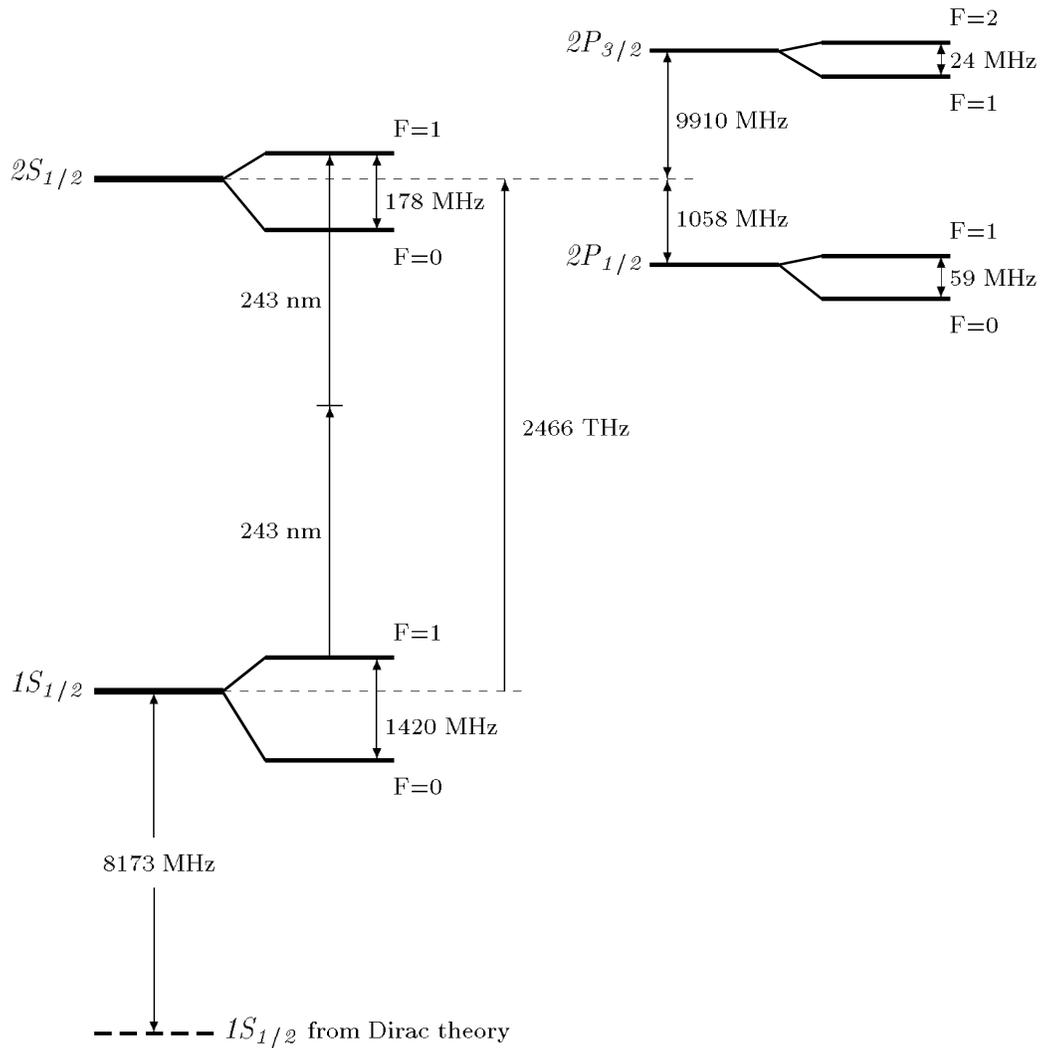
Figure 4: Low-lying hydrogen atomic levels with Lamb shifts and hyperfine splittings. Taken from M.I. Eides, H. Grotch and V.A. Shelyuto, "Theory of light hydrogenlike atoms," *Phys. Rep.*, **342**, 63-261, (2001).

theoretical calculations had progressed greatly since old days. (See V.W. Hughes and T. Kinoshita, *Rev. Mod. Phys.* **71**, S133-S139 (1999).)

The crucial theoretical progress usually attributed to Schwinger is that there is a correction at $O(\alpha)$ to the g-factor. Because of the quantum fluctuations, an electron is sometimes not just an electron but is accompanied by a photon. More precisely, the process is the following. An electron interacts with the magnetic field with $g = 2$. But before the electron interacts with the magnetic field, it can emit a (virtual) photon, violating the energy conservation. The electron then interacts with the magnetic field, and only after that, it reabsorbes the photon. Schwinger's calculation showed that the g-factor is not longer precisely 2 but is corrected as

$$g = 2\left(1 + \frac{\alpha}{2\pi}\right). \tag{65}$$

More recently, Kinoshita has calculated the correction up to $O(\alpha^4)$ with 891 Feynman diagrams. It is customary to quote $a_e = (g_e - 2)/2$. He found

$$
\begin{aligned}
a_e &= \frac{\alpha}{2\pi} - 0.328\ 478\ 965\ldots\left(\frac{\alpha}{\pi}\right)^2 + 1.181\ 241\ 456\ldots\left(\frac{\alpha}{\pi}\right)^3 \\
&\quad -1.509\ 8(384)\left(\frac{\alpha}{\pi}\right)^4 + 4.393(27) \times 10^{-12}.
\end{aligned} \tag{66}
$$

The last constant is a contribution from known heavier particles, such as μ, τ leptons, hadrons, and W, Z bosons. In order to obtain a numerical value, we need to know the fine structure constant α very precisely. The best measurement comes from the Quantum Hall Effect,

$$\alpha^{-1} = 137.036\ 003\ 7(33). \tag{67}$$

Using this value as the input, the QED predicts the electron magnetic moment to be

$$a_e = 1\ 159\ 652\ 153.5(1.2)(28.0) \times 10^{-12}. \tag{68}$$

The first uncertainty comes from the estimated error in the numerical integrations involved in the calculations, while the latter (and the dominant) error is in the input value of α.

The experiment also had made a dramatic progress. The most precise measurement was done using Penning trap. If g were exactly two, an electron moving in a magnetic field would have precisely the same values for the

cylotron frequency and the spin precession frequency. The difference between them measures $g - 2$. The best values due to Van Dyck are

$$a_{e^-} = 1\ 159\ 652\ 188.4(4.3) \times 10^{-12}, \tag{69}$$
$$a_{e^+} = 1\ 159\ 652\ 187.9(4.3) \times 10^{-12}. \tag{70}$$

The agreement between experiment and theory is truly amazing. The experimental accuracy is 4×10^{-12}, which agrees with theory at 1.3 sigma level. Incredible! This is *the* most dramatic success of quantum physics, most likely of all physical sciences.

The anomalous magnetic moment of muon is also interesting. Because muon is short-lived (lifetime is microsecond), the experimental measurement is more difficult. The trick is to actually use the decay product of the muon, $\mu^- \to e^- \bar{\nu}_e \nu_\mu$, where you can detect the electron (but not neutrinos). Luckily, parity is violated in this decay (!),[1] and the direction of the decay electron is correlated with the muon spin. By measuring the direction of the decay electrons, we can measure the muon spin precession and hence $g_\mu - 2$. Theoretically, muon is heavier and the anomalous magnetic moment is more sensitive to heavier particles than that of electron. In fact, the contribution from hadrons (pions, various mesons, protons, etc) is quite important. You may even hope that it may detect the effect of yet-undiscovered particles.

The theoretical prediction is (see, *e.g.*, U. Chattopadhyay, A. Corsetti, P. Nath, `http://arXiv.org/abs/hep-ph/0204251`)

$$a_\mu = 11\ 659\ 176.8(6.7) \times 10^{-10}. \tag{71}$$

Currently a new experiment is being conducted at Brookhaven National Laboratory and has measured the (anti-)muon magnetic moment. They reported (see H. N. Brown *et al.* [Muon $g - 2$ Collaboration], *Phys. Rev. Lett.* **86**, 2227 (2001), `http://arXiv.org/abs/hep-ex/0102017`. See also `http://phyppro1.phy.bnl.gov/g2muon/index.shtml`)

$$a_\mu = 11\ 659\ 202(14)(6) \times 10^{-10}. \tag{72}$$

---

[1] According to Leon Lederman's book "The God Particle," he conceived the parity-violation experiment in muon decay when he heard of the rumor that C.S. Wu found "large" parity violation in nuclear β-decay. Together with his collaborators, he rushed to the laboratory where a graduate student was mounting his thesis experiment. Quickly they disassembled it (!), and mounted the parity violation experiment. They result is reported in Physical Review Letters right after Wu's paper. You should feel lucky if your supervisor is not too keen on a timely success.

They agree at 1.6σ level.[2]

7 Vacuum Polarization

The Dirac sea is a collection of infinite number of electrons in negative energy states. Even though it has an infinite (negative) charge, as long as it is completely homogeneous, we will not be able to detect it because there is no preferred direction to produce an electric field. However, if the homogenity is broken, for example by the presence of a point charge, the distribution of the negative energy electrons is no longer homogeneous and can have a physically observable effect.

Suppose you place a positive point charge in the "vacuum" filled with negative energy electrons. The negative energy electrons are then attracted to the positive point charge, effectively screening the point charge. Therefore, viewed from far away, the charge of the point charge would appear less than what you originally put in. As you go closer and closer to the point charge, you see less effect due to the screening by negative energy electrons, and hence the charge appears larger and larger. In other words, the fine-structure constant α would appear larger and larger as you measure it at smaller and small distance scales. And smaller distance scales, because of the uncertainty principle, corresponds to the larger momentum transfer. This is the effect of the "vacuum polarization," causing the fine-structure constant to grow at higher momentum transfers. We often say that the fine-structure constant *runs*.

This effect, even though theoretically well worked out within the QED, had been measured for the first time in 1997, by an experiment at the electron-positron collider TRISTAN (I. Levine *et al*, "Measurement of the

---

[2]In 2001 when this result was reported, there was a big excitement in this field because the data and theory did not agree at 2.6 sigma! The best theoretical result then was $a_\mu = 11\ 659\ 159.6(6.7) \times 10^{-10}$, based on QED calculations by Kinoshita and updated by A. Czarnecki and W. J. Marciano, *Phys. Rev.* **D 64**, 013014 (2001), http://arXiv.org/abs/hep-ph/0102122. The experimental group speculated that this may be due to the quantum effect of yet-undiscovered supersymmetric partner of muon, $\tilde{\mu}$. However, it turned out that there was a sign mistake in the theoretical calculation for one small contribution. Kinoshita (Cornell) is the renowned world expert in this type of calculations. The problem was that he apparently used a software which used a differeng sign convention from his, but without reading the manual carefully. See Hayakawa and Kinoshita, http://arXiv.org/abs/hep-ph/0112102, for more details.

Electromagnetic Coupling at Large Momentum Transfer," *Phys. Rev. Lett.* **78**, 424–427 (1997)). Compared to the value of the fine-structure constant Eq. (67) at zero momentum transfer (a negligible momentum transfer compared to mc) such as in Quantum Hall Effect, they reported the fine-structure constant at momentum transfer $Q^2 = (57.77 \text{ GeV}/c)^2$ to be $\alpha^{-1}(Q^2) = 128.5 \pm 1.8 \pm 0.7$, which agrees with the theoretical prediction 129.6 ± 0.1. Note that the vacuum polarization effect is not just due to the negative energy electrons, but also due to mu and tau leptons, and all hadrons (*i.e.*, quarks).
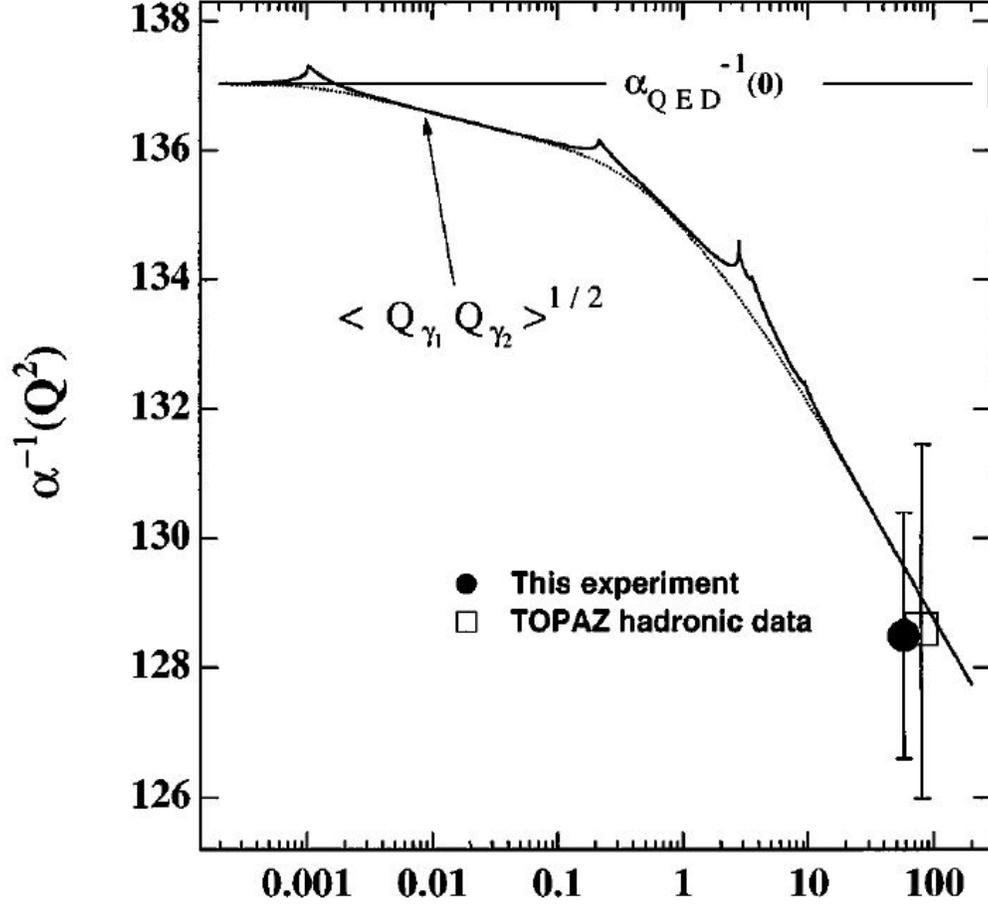
Figure 5: The measured and theoretical fine-structure constant as a function of momentum transfer Q. The solid and dotted lines correspond to positive and negative Q^2 predictions, respectively. As we probe closer to the bare charge, its effective strength increases. $\langle Q_{\gamma_1} Q_{\gamma_2} \rangle^{1/2}$ denotes the square root of the median value for the product of the photon momentum transfers in the anti-tagged $e^+e^- \to e^+e^-\mu^+\mu^-$ sample. The hadronic data point has been shifted for display. Taken from I. Levine *et al*.