INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS G: NUCLEAR AND PARTICLE PHYSICS

J. Phys. G: Nucl. Part. Phys. **29** (2003) 9–21 PII: S0954-3899(03)53696-9

Everyone makes mistakes—including Feynman

Toichiro Kinoshita

Newman Laboratory, Cornell University, Ithaca, NY 14853, USA

Received 5 September 2002
Published 17 December 2002
Online at stacks.iop.org/JPhysG/29/9

Abstract
Early stages of the theory of radiative corrections to weak interaction processes such as muon decay and $\pi - e$ decay are described based on my personal recollection. The discovery of an error in our initial paper on the muon decay played a crucial role in the realization of remarkable cancelation of mass singularities in integrated observable quantities. General formulation developed to deal with mass singularity turned out to be very handy for numerical evaluation of high-order radiative corrections to the lepton anomalous magnetic moment. This has led to the most stringent test of QED available at present. New developments of last two years are described briefly in 'Note added in proof'.

1. Introduction

This talk is dedicated to Alberto Sirlin in celebration of his seventieth birthday. I wish to convey my deep appreciation of his many important contributions to particle physics over 40 years and look forward to many more years of productive research.

Alberto arrived at Cornell as a graduate student in September 1955. I had come to Cornell as a research associate a few months earlier. Thus I have been acquainted with him for 45 years.

It was the time when experimental observations of some weak interaction processes, the so-called θ–τ puzzle in particular, began to expose the internal inconsistency in the theory of the weak interaction. Analysing this problem in great depth, Lee and Yang concluded that parity conservation, assumed to be valid in previous theories of the weak interaction, was the most likely culprit. They suggested that parity symmetry is not valid in the weak interaction and proposed ways to test it experimentally [1]. The experimental verification followed soon afterwards [2, 3]. The two-component neutrino theory (discarded previously by Pauli) became the favoured theory [4–6].

2. Radiative correction to muon decay

A detailed comparison of the two-component theory with experiment would require accounting for radiative corrections. This is because they might be as large as $\alpha\omega^2 \simeq 28.4/137$, where

$\omega = \ln(m_\mu/m_e) = 5.3316$, according to [7] on radiative corrections to the parity-conserving muon decay, on which Alberto worked before he came to Cornell. Its extension to the parity-non-conserving case is not difficult. Thus Alberto's experience enabled us to jump-start the calculation of the parity-non-conserving muon decay and finish it on a very short notice. The radiatively corrected muon decay spectrum we obtained in the two-component neutrino theory is [8]

$$\mathrm{d}N_r(x,\theta) = \frac{1}{2}A\left[3 - 2x + \frac{\alpha}{2\pi}f(x) + 6\zeta\frac{m_e}{m_\mu}\frac{1-x}{x} + \xi\cos\theta\left\{1 - 2x + \frac{\alpha}{2\pi}h(x)\right\}\right]x^2\,\mathrm{d}x\,\mathrm{d}\Omega, \tag{1}$$

where A, ξ, ζ are functions of weak coupling constants, $x = 2p_e/m_\mu$, and

$$\begin{aligned} f(x) &= (6-4x)u(x) + (6-6x)\ln x + \frac{1-x}{3x^2}[(5+17x-34x^2)(\omega+\ln x) - 22x + 34x^2], \\ h(x) &= (2-4x)u(x) + (2-6x)\ln x + \frac{1-x}{3x^2}\left[(-1-x-34x^2)(\omega+\ln x)\right. \\ &\qquad \left. - 3 + 7x + 32x^2 - \frac{4(1-x)^2}{x}\ln(1-x)\right], \end{aligned} \tag{2}$$

with

$$\begin{aligned} u(x) &= \omega^2 + \omega\left(\frac{1}{2} - 2\ln 2\right) + 2\ln 2 - 3 + \left(2\omega - 1 - \frac{1}{x}\right)\ln(1-x) \\ &\qquad + \ln x\,[3\ln(1-x) - \ln x - 2\ln 2] + L(1) - 2L(x), \end{aligned} \tag{3}$$

and

$$L(x) = \int_0^x \ln(1-t)(\mathrm{d}t/t), \qquad L(1) = -\frac{\pi^2}{6}. \tag{4}$$

The radiative correction to the decay lifetime is large, being proportional to ω^2:

$$\frac{\tau - \tau_0}{\tau_0} = -\frac{\alpha}{2\pi}(\omega^2 + \cdots) \simeq -0.029\,73. \tag{5}$$

We presented our result at the Rochester conference in the spring of 1957. Then, one day in 1958, lightening struck us. We received a preprint from Berman stating that he disagreed with our result. In particular, he mentioned that the radiative correction to the muon lifetime is linear, not quadratic, in ω, contrary to our result. When we read his preprint, however, we suspected immediately that his result, which contains a term linear in ω, must also be wrong. Alberto and I worked hard for a week or two and found that our intuition was in fact correct.

In the new result [9], spectral functions $f(x)$ and $h(x)$ have the same form as in (2), but $u(x)$ is replaced by

$$\begin{aligned} R(x) &= \omega[1.5 + 2\ln(1-x) - 2\ln x] - 2L(x) + 2L(1) - 2 - \ln x(2\ln x - 1) \\ &\qquad + \left(3\ln x - 1 - \frac{1}{x}\right)\ln(1-x). \end{aligned} \tag{6}$$

The decay spectrum still contains terms linear in $\omega(\equiv\ln(m_\mu/m_e))$. However, the muon decay lifetime now has no dependence at all on ω and the net radiative correction is very small:

$$\frac{\tau - \tau_0}{\tau_0} = -\frac{\alpha}{2\pi}\left(\frac{25}{4} - \pi^2\right) \simeq 4.17 \times 10^{-3}. \tag{7}$$

To discuss our mistake in [8] that Berman pointed out, let me focus on the inner bremsstrahlung contribution to the muon decay, which contains the factor

$$\sum_i \int_0^{k_{\max}} \frac{\mathrm{d}^3k}{\epsilon}\left[\frac{p_2\cdot e_i}{p_2\cdot\kappa} - \frac{p_1\cdot e_i}{p_1\cdot\kappa}\right]^2, \tag{8}$$

where p_1, p_2 and κ are the 4-momenta of the muon, electron and photon, and e_i are the polarization vectors of the photon.

Recall that, in a covariant calculation, the virtual photon is often treated as a vector meson of mass λ with the understanding that $\lambda \to 0$ in the physical limit. To be consistent with this, the *real photon* must also be regarded as a vector meson of mass λ. This means, in particular, that the sum in (8) must be carried out over four polarizations including time-like and longitudinal polarizations. Our mistake was that we had summed only over transverse polarizations. Furthermore, even in the limit $\lambda \to 0$, the contribution of non-transverse polarizations does not vanish if the photon has infrared divergence. As a matter of fact, this extra contribution has an ω dependence that cancels the leading ω dependence from transverse photons.

This was supposed to be well known: it is related to Feynman's famous error in matching non-relativistic and relativistic calculations of the Lamb shift that was discovered by French and mentioned in footnote 13 of Feynman's paper [10]. Unfortunately, many people, including us, had forgotten or failed to appreciate the significance of his footnote and made the same mistake again and again. (It is true that the connection with our mistake is somewhat obscure since Feynman's footnote does not deal directly with scattering states or decaying states.) In the end Berman agreed with us and revised his paper accordingly [11].

The lessons we learned from this episode are:

- Although the differential spectrum of μ–e decay diverges logarithmically as $m_e/m_\mu \to 0$, the *total* decay rate is finite in this limit.
- Cancellation of infrared divergences is a necessary but not a sufficient guarantee for the computation to be correct. (Many people were unaware of this and made the same mistake, even after our paper was published.)

We obtained a similar result for nuclear β decay under some simplifying assumptions. The radiative correction to the β-ray spectrum in the V-A theory (for $m_e \ll E$) is found to be [9]

$$\begin{aligned}\Delta P\,\mathrm{d}^3 p = \frac{\alpha}{\pi^4} G^2 E_m^5 (1-x)^2 x^2\,\mathrm{d}x \Bigg\{ & 6\ln\left(\frac{\Lambda}{m_p}\right) + 3\ln\left(\frac{m_p}{2E_m}\right) + \frac{3}{2} - \frac{2\pi^2}{3} \\ & + 4(\ln x - 1)\left[\frac{1-x}{3x} - \frac{3}{2} + \ln\left(\frac{1-x}{x}\right)\right] + \frac{(1-x)^2}{6x^2}\ln x \\ & + \Omega\left[\frac{4(1-x)}{3x} - 3 + \frac{(1-x)^2}{6x^2} + 4\ln\left(\frac{1-x}{x}\right)\right]\Bigg\},\end{aligned} \tag{9}$$

where $x = E/E_m$, $\Omega = \ln(2E_m/m_e)$, E_m is the maximum total energy of the electron and m_p is the proton mass. The spectrum diverges logarithmically for $m_e \to 0$. However, the correction to the lifetime has no $\ln m_e$ dependence:

$$\frac{\Delta\tau}{\tau_0} = -\frac{\alpha}{2\pi}\left[6\ln\left(\frac{\Lambda}{m_p}\right) + 3\ln\left(\frac{m_p}{2E_m}\right) - 2.85\right]. \tag{10}$$

Thus β and μ decays share the same $\ln m_e$ dependence, suggesting that it is more general and not limited to these decays. (See also appendix A.1.)

Feynman came to Cornell in the fall of 1958 for three months. He explained to me how he and Berman made exactly the same mistake. Feynman had asked Berman to check our calculation for his thesis work. But, actually, Feynman himself was doing this calculation independently of Berman. At the end they compared notes and were satisfied that their results agreed. When confronted with our new result which differed from theirs, they checked the notes once again and found that they made the same mistake in copying the bottom equation of a page to the top of the next page. Feynman was so disturbed by this mistake that he told me how sorry he was more than a few times while at Cornell.

3. Radiative correction to π–e decay

Feynman brought with him a new preprint of Berman on the radiative correction to π–e decay. As is well known, in the V-A theory, the ratio of π–e and π–μ decay rates is

$$R_0 = \left(\frac{m_e}{m_\mu}\right)^2 \left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2}\right)^2 \simeq 1.28 \times 10^{-4} \tag{11}$$

if the radiative correction is not included. Berman's result, including one-loop radiative correction, was of the form

$$R = R_0 \left(1 - \frac{3\alpha}{\pi} \ln(m_\mu/m_e) + \cdots\right). \tag{12}$$

This R has a rather large correction ($\sim$3%) and looked strange since R/R_0 diverged for $m_e/m_\mu \to 0$. Feynman and I were so puzzled by this result, which seemed to contradict what was discovered in μ and β decays, that we decided to check it with a fresh calculation. For the next two months we worked hard, totally independently of each other, except that we agreed to start from the same effective Lagrangian

$$g m_l \bar{\psi}_l a \psi_\nu \phi_\pi, \tag{13}$$

where l represents either a muon or an electron and $a = (1 + \mathrm{i}\gamma_5)/2$.

The radiative correction due to the virtual photon is straightforward [12]:

$$\begin{aligned} \frac{\Delta P_{\mathrm{VP}}}{P_0} &= \frac{\alpha}{\pi}\left[\frac{3}{2}\ln\frac{\Lambda}{m_\pi} - b(r)\left(\ln\frac{\lambda_{\min}}{m_\pi} - \frac{1}{2}\ln r + \frac{3}{4}\right) + \frac{r^2}{1-r^2}\ln r + \frac{1}{2}\right], \\ b(r) &= 2\left(\frac{1+r^2}{1-r^2}\ln r + 1\right), \end{aligned} \tag{14}$$

where P_0 is the uncorrected π–e decay rate, $r = m_e/m_\pi$, $\lambda_{\min}$ is the infrared cut-off mass and Λ is the ultraviolet cut-off mass.

The total probability of the inner bremsstrahlung correction is [12]

$$\begin{aligned} \frac{\Delta P_{\mathrm{IB}}}{P_0} = \frac{\alpha}{\pi}\Bigg[& b(r)\left(\ln\frac{\lambda_{\min}}{m_\pi} - \ln(1-r^2) - \frac{1}{2}\ln r + \frac{3}{4}\right) - \frac{r^2(10-7r^2)}{2(1-r^2)^2}\ln r \\ &+ \frac{2(1+r^2)}{1-r^2}L(1-r^2) + \frac{15-21r^2}{8(1-r^2)}\Bigg]. \end{aligned} \tag{15}$$

It follows from (14) and (15) that, to order α, the rate of π–e decay is

$$P = P_0(1 + \eta_e), \tag{16}$$

where

$$\begin{aligned} \eta_e = \frac{\alpha}{\pi}\Bigg[& \frac{3}{2}\ln\left(\frac{\Lambda}{m_\pi}\right) - b(r)\ln(1-r^2) - \frac{r^2(8-5r^2)}{2(1-r^2)^2}\ln r \\ &+ \frac{2(1+r^2)}{1-r^2}L(1-r^2) + \frac{19-25r^2}{8(1-r^2)}\Bigg]. \end{aligned} \tag{17}$$

Infrared divergences have cancelled out as expected.

When we finished the calculation, we compared the results and found that we agreed with each other. Unfortunately, we did not agree with Berman. In particular, our result did not have the $\ln(m_\mu/m_e)$ term. We thought for a while that Berman's calculation was wrong. But, after scrutinizing our calculation very closely, I realized that it was we who were wrong. We committed a very subtle mistake: it was in our choice of the starting Lagrangian.

Berman started from an effective Lagrangian with the derivative coupling:

$$g\bar{\psi}_l\gamma_\mu a\psi_\nu\left(\mathrm{i}\frac{\partial\phi_\pi}{\partial x_\mu} - eA_\mu\phi_\pi\right). \tag{18}$$

To simplify the algebra this is often turned into a non-derivative form

$$g(m_l - m_\nu)\bar{\psi}_l a\psi_\nu\phi_\pi \tag{19}$$

by integration by parts and use of the equation of motion

$$\mathrm{i}\frac{\partial}{\partial x_\mu}\bar{\psi}_l\gamma_\mu + eA_\mu\bar{\psi}_l + m_l\bar{\psi}_l = 0. \tag{20}$$

Our Lagrangian (13) was obtained from the equivalent Lagrangian (19) assuming that $m_\nu = 0$ and m_l is the physical mass. Unfortunately, we did not realize initially that this equation is not valid to order e^2 if m_l is the physical mass. The correct equation requires the self-mass term $\delta m_l\bar{\psi}_l$ on the right-hand side of (20). Or, equivalently, we may rewrite the corrected equation as

$$\mathrm{i}\frac{\partial}{\partial x_\mu}\bar{\psi}_l\gamma_\mu + eA_\mu\bar{\psi}_l + m_l^0\bar{\psi}_l = 0, \tag{21}$$

where m_l^0 is the *bare* mass. Berman's Lagrangian (18) may thus be replaced by

$$gm_l^0\bar{\psi}_l a\psi_\nu\phi_\pi. \tag{22}$$

This means that we can turn our incorrect result into a correct one by simply replacing the renormalized mass with the bare mass. The appearance of bare mass in this context had been noticed by Ruderman [13, 14].

The radiatively corrected decay ratio R can thus be written as [12]

$$R = R_0((1+\eta_e)/(1+\eta_\mu))(1+\delta), \tag{23}$$

where η_e and η_μ are defined by (17) and

$$\begin{aligned}\delta &= \left(m_e^0/m_e\right)^2 / \left(m_\mu^0/m_\mu\right)^2 - 1\\ &= -(3\alpha/\pi)\ln(m_\mu/m_e)\\ &\simeq -15.995\,(\alpha/\pi).\end{aligned} \tag{24}$$

This R is in exact agreement with Berman's result [15].

Measurement of the $(\pi{-}e)/(\pi{-}\mu)$ decay ratio had just started at the time of this calculation. The experimental uncertainty was still so large that the presence or absence of the $(1+\delta)$ factor could not be tested experimentally. Later more accurate measurements confirmed this large effect [16, 17]

$$R = 1.2265\,(34)\,(44)\times 10^{-4}, \qquad R = 1.2346\,(35)\,(36)\times 10^{-4}. \tag{25}$$

Our calculation of R relied on an implicit untested assumption that the UV cut-off Λ is common to both $\pi{-}e$ and $\pi{-}\mu$ decays. A justification of this assumption had to wait for the emergence of the standard model [18]. The hadronic effect was also taken into account [19]. This leads to the latest value

$$R = 1.2352\,(5)\times 10^{-4}. \tag{26}$$

Working with Feynman was a very interesting and instructive experience. Let me share one episode with you.

As is well known, the integration over three-body final states is quite non-trivial. I spent most of the two months checking my calculation of the inner bremsstrahlung term ΔP_{IB} again and again. In the end more than 30 pages of equations were needed to carry my conventional approach to the end. It also took about two months for Feynman to evaluate this integral.

Actually, I am not sure that he was working on this problem all the time. His office was next to mine so that I could hear that he was constantly practising bongo drums using the cover of the heating system as a drum. When we finally finished the work and compared notes, however, I was astounded to find that his whole calculation was written on just two sheets of paper. What he was doing during the two months was not only playing bongo but also looking for new ways of doing the integration. And he actually found a very simple and elegant method!

Since this does not seem to be widely known, let me describe it here. The decay process $\pi \to e + \bar{\nu} + \gamma$ has 4-momentum conservation:

$$p_\pi = p_e + p_\nu + k. \tag{27}$$

- Step 1. Take any final-state p_e and go to the reference frame in which the space components of p_π and p_e satisfy the relation

$$\vec{p}_\pi = \vec{p}_e. \tag{28}$$

  Then $\vec{p}_\nu$ and $\vec{k}$ are exactly back to back. Thus the angular integration becomes trivial. The result is a function of the fourth component E_e of p_e only, which can be easily converted to a covariant form.
- Step 2. Go to the pion rest frame by an E_e*-dependent* Lorentz transformation. Then the remaining integration over E_e is almost trivial. That is all.

Before we finished our work, Feynman went back to Caltech. After some exchange of letters discussing fine details of the calculation and the drafting of a report, Feynman told me to publish the paper by myself, which I did reluctantly [12]. He did not explain why he did not want to put his name on it. I can only guess some reasons. One is that he was not comfortable with the appearance of unphysical mass in observable quantities. Since the Lagrangians used in these days were just effective Lagrangians and not renormalized ones, they did not provide a proper framework to deal with such a problem. Only within the context of renormalizable theories, such as the standard model, can one treat it properly in terms of the renormalization group.

Although the presence of the $\ln(m_\pi/m_e)$ term in the π–e decay rate seemed strange at first sight, it was actually not so strange. This is because the π–e decay amplitude in the V-A theory is proportional to m_e, which multiplies $\ln(m_\mu/m_e)$ and makes the whole amplitude vanish for $m_e \to 0$. Note also that the η_e part of the radiative correction to the π–e decay behaves in the same manner as those of μ and β decays, namely, it has no $\ln m_e$ term.

4. Mass singularity

These examples convinced me that the striking cancellation of $\ln m_e$ terms in the total probability is a very general feature of quantum field theory. It seemed that the structure of general Feynman amplitudes in the massless limit deserved some attention. I spent the next few years trying to understand this problem.

The basic tool of analysis was the power counting rule to examine the behaviour of Feynman integrals in the zero-mass limit. As a function of (various) masses, the Feynman amplitude has a very complicated structure at zero-mass points. In general, its value depends on the order and direction in which the zero-mass points are approached. It is found, nevertheless, that $m \to 0$ in a propagator of mass m does not cause divergence unless it is enhanced by putting vertex-sharing propagators (including external lines) on the mass shell. The result was reported in [20]. Alberto used it in his derivation of differential equations for propagators and vertex functions in QED, and obtained results which are equivalent to the Callan–Symanzik equation [21]. (See appendix A.3.)

5. Lepton anomalous magnetic moments

It turned out that the analysis of the mass singularity was very useful in studying the m_μ/m_e dependence of the muon anomalous magnetic moment [22]. This was in fact the beginning of my active involvement in the $g-2$ problem, from which I have not yet managed to extract myself. The paper [22] was extended to the study of higher-order muon anomalous moment based on the renormalization group technique [23, 24].

The analytic tool developed in [20] to deal with general Feynman-parametric integrals also turned out to be very handy as the starting point of my work on the sixth- and eighth-order radiative corrections to the lepton $g-2$ by a numerical method [25, 26]. After converting momentum space Feynman integrals into Feynman-parametric integrals analytically, we evaluated them numerically using the iterative-adaptive Monte Carlo integration routine VEGAS [27].

In the case of the electron $g-2$, the best value of the coefficient of $(\alpha/\pi)^3$, obtained by VEGAS, is [28]

$$A_6^{(\text{num})} = 1.181\,259\,(40). \tag{29}$$

This is in good agreement with the analytic result obtained by Laporta and Remiddi several months later, after many years of hard work [29]:

$$A_6^{(\text{anal})} = 1.181\,241\,456\cdots. \tag{30}$$

At present A_8, the coefficient of $(\alpha/\pi)^4$, is known only by the VEGAS integration. The most recent reported value of A_8 is [30]

$$A_8^{(\text{num})} = -1.5098\,(384). \tag{31}$$

The project to reduce the uncertainty of A_8 by a factor of 3 or more by means of massively-parallel computers is approaching the final stage.

At present the best theoretical value of a_e, including small electroweak and hadronic terms, is

$$a_e(\text{th}) = 1159\,652\,153.5\,(1.2)\,(28.0) \times 10^{-12} \tag{32}$$

evaluated using the α obtained from the quantum Hall effect [31, 32]:

$$\alpha^{-1}(\text{qH}) = 137.036\,0037\,(33). \tag{33}$$

The value ± 1.2 in (32) is the remaining uncertainty in theory. The result (32) is to be compared with the measured values of a_e obtained in Penning trap experiments [33]:

$$a_{e^-} = 1159\,652\,188.4\,(4.3) \times 10^{-12}, \qquad a_{e^+} = 1159\,652\,187.9\,(4.3) \times 10^{-12}, \tag{34}$$

or their weighted average [32]

$$a_e = 1159\,652\,188.3\,(4.2) \times 10^{-12}. \tag{35}$$

Theory is -1.3 standard deviations away from experiment.

The QED contribution to a_μ has been computed through five loops [34, 35]

$$a_\mu(\text{QED}) = 0.5\left(\frac{\alpha}{\pi}\right) + 0.765\,857\,376\,(27)\left(\frac{\alpha}{\pi}\right)^2 + 24.050\,508\,98\,(44)\left(\frac{\alpha}{\pi}\right)^3 + 126.07\,(41)\left(\frac{\alpha}{\pi}\right)^4 + 930\,(170)\left(\frac{\alpha}{\pi}\right)^5 = 116\,584\,705.7\,(2.9)\times 10^{-11}. \quad (36)$$

The coefficients of $(\alpha/\pi)^n$ are mass dependent. Many $\ln(m_\mu/m_e)$ terms as well as some mass-independent terms can be determined analytically by renormalization group considerations [22–24]. Coefficients of α^2 and α^3 can be evaluated to any precision by expansion in mass ratios [35]. The errors in the α^2 and α^3 terms come only from measurement uncertainties of m_e/m_μ and/or m_e/m_τ. The coefficient of α^4 is known by numerical integration only. The coefficient of α^5 is only a rough estimate at present [34, 36]. The electroweak contribution has been evaluated to two-loop order [35]

$$a_\mu(\text{EW}) = 152\,(4)\times 10^{-11}. \quad (37)$$

The current best estimate of the hadronic contribution is [35]

$$a_\mu(\text{had}) = 6739\,(67)\times 10^{-11}. \quad (38)$$

The sum of (36), (37) and (38) gives the prediction of the standard model

$$a_\mu(\text{theory}) = 116\,591\,597\,(67)\times 10^{-11}. \quad (39)$$

This is in good agreement with the value

$$a_\mu(\text{exp}) = 116\,592\,050\,(460)\times 10^{-11} \quad (40)$$

obtained by combining the CERN result and the data taken through 1998 at Brookhaven National Laboratory [35, 37].

6. Fine structure constant as test of quantum mechanics

As is seen from (32), the uncertainty in $a_e(\text{th})$ is dominated by that of α given in (33). This means that this α is not accurate enough to test QED to the extent allowed by the precision of the measurement and theory of a_e. The situation is no better for other high precision values of α determined from the ac Josephson effect [32], measurement of h/m_n (m_n is the neutron mass) [38], muonium hyperfine structure [39] and caesium D_1 line[1] [40, 41]:

$$\alpha^{-1}(acJ, \gamma_{p'}) = 137.035\,9880\,(51)\,[3.7\times 10^{-8}], \quad (41)$$

$$\alpha^{-1}(m_n) = 137.036\,0119\,(51)\,[3.7\times 10^{-8}], \quad (42)$$

$$\alpha^{-1}(\mu hfs) = 137.035\,9932\,(83)\,[6.0\times 10^{-8}], \quad (43)$$

$$\alpha^{-1}(\text{Cs}D_1) = 137.035\,9924\,(41)\,[3.0\times 10^{-8}]. \quad (44)$$

Atom beam interferometry, single electron tunnelling, fine structure of the helium atom and bound electron $g-2$, may also produce very precise values of α.

This means, however, that it is the electron $g-2$ that can provide the most precise value of α at present. From the Seattle experiment and QED one obtains

$$\alpha^{-1}(a_e) = 137.035\,999\,58(14)\,(50) = 137.035\,999\,58\,(52)\,[3.8\times 10^{-9}]. \quad (45)$$

Errors on the first line are due to the α^4 term and measurement of a_e.

[1] This value will be improved very soon.

When new experiments are completed, the measurement uncertainty of a_e may be reduced by an order of magnitude [42, 43]. Further improvement in theory will enable us to determine α with an uncertainty of less than 10^{-9}.

Comparison of α cited above shows that they are in agreement with each other at the level of 10^{-7}. However, these comparisons must be regarded as testing theories underlying these measurements, rather than testing QED. Since all these determinations of α are ultimately based on quantum mechanics, they may be regarded as testing of the internal consistency of quantum mechanics itself. It will be of great interest to see whether the good agreement still holds at the level of 10^{-8} or beyond.

7. Concluding remark

Although the first paper Alberto and I wrote together [8] had an embarrassing error, it turned out to be a very productive error. If this paper did not have the mistake discussed in section 2, we would not have noticed the striking cancellation of mass singularities in integrated quantities, which we emphasized in our subsequent work [9]. I must also point out that we were very lucky to stumble upon this phenomenon. This was because we were studying the decay process rather than the scattering process. Decay processes have several mass scales (for instance m_e and m_μ) and it is thus easy to examine the limit $m_e \to 0$ while keeping m_μ finite. The same cancellation mechanism is also present in scattering processes. However, it was not noticed previously because, in a system with just one mass scale, the mass singularity is not clearly separated from singularities associated with threshold behaviour or high energy limits.

Although Alberto and I collaborated only for a couple of years and pursued separate routes afterwards, you will see that much of what we have done since then have roots in our early collaboration.

Note added in proof. Since this talk was given in the fall of year 2000, there has been new developments which render the contents of Sections 5 and 6 out-of-date unless corrected and updated. Instead of making changes in the text, however, I include them here as 'Note added in proof'. This will enable us to compare the text and Note side-by-side. It provides (hopefully) a useful example of physics as a self-correcting discipline, in which resolution of earlier mistakes serves as a stepping stone for subsequent development. After all this is in accord with the intention of this paper.

(a) New measurements of muon $g-2$ and hadronic corrections

The muon $g-2$ experiment at the Brookhaven National Laboratory, after many years of hard work, published a result based on the year 1999 data which disagreed by about 2.6 standard deviations from the prediction of the standard model. This caused a lot of excitement since it might possibly be the first concrete indication of new physics beyond the standard model. As a consequence a large number of papers have been written to explain the discrepancy. Since then the excitement has cooled down considerably by the discovery of an error in previous calculations of the hadronic light-by-light scattering contribution, reducing the discrepancy to about 1.6 standard deviations. Currently, the most accurate value, based on the analysis including the year 2000 data, is

$$a_\mu(\text{exp}) = 116\,592\,030\,(80) \times 10^{-11}.$$

This disagrees with theory, too. Unfortunately, the prediction of the standard model is not unique at this moment. It deviates from $a_\mu(\text{exp})$ by 3.0 s. d. or 1.6 s. d. depending on whether it is based on (i) the measurement of the total cross section for the e^+e^- annihilation process ($e^+e^- \to$ hadrons) or (ii) the analysis of hadronic τ-decay. Until this discrepancy is resolved, it may not be possible to confirm the existence of new physics beyond the standard model.

(b) Error in the α^4 QED term of lepton $g-2$

New measurements of the muon $g-2$ prompted reexamination not only of hadronic but also of QED contributions. While checking the previous calculation once again, we found that the α^4 contribution to the lepton $g-2$ from a gauge-invariant set of 18 Feynman diagrams containing a light-by-light scattering subdiagram internally suffered from a programming error. Correction of this error changes the coefficient of $(\alpha/\pi)^4$ in equation (36) from 126.07 (41) to about 132.23 (which is not yet final). Similarly, the value of $A_8^{(\text{num})}$ of equation (31) changes from -1.5098 (384) to about -1.7363 (which is not yet final). This also affects the value of a_e(th) given in equation (32). This correction has a negligible effect on the comparison of experiment and theory of muon $g-2$ but affects the α^4 term of the electron $g-2$ by about 16%. As a consequence it changes the value of the fine structure constant (45) obtained from theory and measurement of the electron $g-2$ to

$$\begin{aligned}\alpha^{-1}(a_e) &= 137.035\,998\,80\,(2)\,(50)\\ &= 137.035\,998\,80\,(50) \quad [3.7\times10^{-9}].\end{aligned} \tag{46}$$

Errors on the first line are due to the new α^4 term and measurement of a_e, respectively.

The value of α in (46) is still being improved and not final.

(c) New measurement of α by atom interferometry

The atom beam interferometry, in which the de Broglie wavelength of the cesium atom is measured, has now succeeded in providing a very precise value of the fine structure constant

$$\alpha^{-1}(h/M) = 137.036\,000\,3\,(10) \quad [7.4\times10^{-9}], \tag{47}$$

where M is the mass of a cesium atom.

This is almost as good as $\alpha(a_e)$ and provides the most serious test of QED and the standard model thus far at the level exceeding relative precision of 10^{-8}.

Acknowledgment

This work is supported in part by the National Science Foundation.

Appendix A. Communications with Sirlin

A.1. Sirlin to Kinoshita, 10 November 2000

Dear Tom,

I would like to thank you very much for coming to the symposium and for your interesting talk. It was also very nice to meet your wife and you after some time! (Although I saw you at the Yang Symposium last year).

Concerning your talk, Massimo Porrati gave me your transparencies and I noted some discrepancies in the citations to our papers. The problem is that we wrote a number of papers and letters and it is easy to confuse one with the other. Here are my observations:

(i), (ii) (Errors in the transparencies corrected here.)

(iii) At the time we corrected our results, I did another check on the cancellation of mass singularities. I took the corrections for muon decay for scalar, pseudoscalar and tensor interactions (which we had from the earlier paper [7]), and checked that, once the real photon contribution was corrected, the mass singularity cancels in the integrated spectrum, as well as the integrated asymmetry. So, from the fact that the cancellation occurs for the five interactions in muon decay and for the V-A interaction in beta decay, we had at the time a very strong indication that this was associated with a powerful theorem, although the proof in the general case had to wait for your subsequent analysis in [20].

(iv) Although it is true that we failed to appreciate the significance of Feynman's footnote 13, I think that on the particular issue of summing over polarizations and matching the infrared divergences Feynman was peculiarly unclear. For example, there is a short book by him, called 'Quantum Electrodynamics' (Benjamin/Cummings, 1961), which essentially contains the material of his 1953 Caltech lectures. On pages 150-151, he discusses the cancellation of infrared divergences between virtual and real soft photons in the case of electron scattering by an external potential, and he only includes the $\ln(K_m/\lambda_{\min})$ term from the inner bremsstrahlung. It seems clear that he only considered the two transverse directions of polarization in this case and made the same type of incorrect approximations in the bremsstrahlung integrals as we did in [8].

Incidentally, in 1952 Daniel Amati and I were students in a memorable course in quantum mechanics that Feynman gave in Brazil. At the end of the course, we asked his guidance about QED and he sent several copies of his Caltech lectures. While I waited in Argentina to go to UCLA, I read those notes. It is rather strange that he did not correct his notes or discuss this issue in greater detail.

(v) While I was a post-doc at Columbia, I received a letter from Feynman, dated 7 March 1958, which I have kept. The letter was about two issues.

(a) It turns out that we had briefly met at some conference and found out that we were worried about the same problem, namely the fact that experimentalists had failed to find the decay $\pi \to e + \bar{\nu}$, and that the upper bound in the branching ratio was $<1/10^5$, i.e. a factor 12 lower than the theoretical prediction of the V-A theory! I told him that I was considering a modification of the theory with the e and ν coming out at different spacetime points, which would occur, for instance, if there were a heavy intermediate particle propagating between the two. In the modern context, this will be the case if leptoquarks actually exist. But of course the concept of quark and leptoquark had not been proposed at the time. Feynman told me that he had a different idea: assuming that the self-energy of the electron is purely electromagnetic, he claimed that the most natural result for the branching ratio was about 3.5×10^{-5}. This still disagreed with experiment, but there was a factor 4 decrease in the predicted branching ratio. I thought that my approach was extremely speculative, but I did not see any way out (assuming, of course, that the experiments were correct). I wrote to T D Lee, who was on leave at Princeton, and he advised me to write a short paper, which I published in *Phys. Rev.* 1958 **111** 337. I sent a copy to Feynman, who replied in the letter of 7 March 1958. He thanked me for the paper and again mentioned his approach leading to the 3×10^{-5} branching ratio. He also added ‘Maybe experiment is wrong’.

(b) In the second paragraph he wrote that his student Sam Berman had found an error in the correction to the ρ-value that Behrends, Finkelstein and I had published. He added that the correction of this error raises ρ by about 0.01 (this is consistent with the conclusions in our paper [9], and the new Rosenson’s value became 0.68 ± 0.05. He also added that he had not seen Crowe’s data. If my memory is right, after the Berman correction, Crowe’s value was 0.68 ± 0.02. In any case, it is clear that at the moment Feynman was focusing, like us, on the corrections to the spectrum. In some sense, we were fortunate because the corrections to the lifetime became important soon afterwards (after the implications of the conserved-vector-current paper of Feynman and Gell-Mann became clear), and roughly by that time our results were corrected.

(vi) After we received the Berman paper and corrected our results, I wrote back to Feynman (on 29 April) that Berman’s point concerning the need to include all degrees of polarization associated with a ‘massive photon’ was correct, but we did not agree with the second error he had mentioned, since it violated the theorem on cancellation of mass singularities. Of course, at the time we had only a heuristic argument for such cancellation, rather than your general argument, but the cancellations were so striking that, I think, both of us were convinced that there was an underlying theorem. I did not keep a copy of my letter to Feynman (in those days there were mimeographs rather than Xerox machines), but I do have a copy of a letter that Berman sent to me, dated 6 May 1958. In the letter he said that, after reading my letter of 29 April to Feynman, he rechecked his results and found complete agreement. He also said that, in preparing the preprint, a copying error was made, resulting in the spurious term with the mass singularity. Then he thanked me for informing him of this error, 'which otherwise might have gone unnoticed'. Some time later, Berman passed through New York and asked me: what is this theorem you are talking about? If I remember correctly, I told him that at the time we did not have a general proof, but surely it was a theorem! In 1961, Berman and I overlapped at CERN, became good friends, and wrote a nice paper on a number of subtle points concerning the radiative corrections to muon and beta decays.

(vii) Sometimes I wonder what would have happened if we had not made the error. Would we have noticed the cancellation of mass singularities in integrated quantities, and consequently convinced ourselves of the existence of an underlying theorem, or would we have missed this most interesting point? Because the cancellations are so striking, I think that we would have found it, but I am not certain. Feynman, with all his genius, missed it, perhaps because of the copying error! In any case, I often mention this famous error to my students, and tell them: 'If you are going to make a mistake, make a good one and discover a theorem!'

Please, give my best regards to your wife and to Professor Salpeter.

All the best,

Alberto

A.2. Kinoshita to Sirlin, 13 November 2000

Dear Alberto,

Thanks for your informative e-mail. As you might imagine, I put together the material of my talk in a hurry and failed to detect errors in citing our papers. If there is going to be a proceeding of the symposium (which I strongly hope is the case), these errors will certainly be corrected. As a matter of fact, your e-mail contains information which will be of interest to readers and historians. If you are not going to write it by yourself, do you think it a good idea to attach it to my article as an appendix? By the way, I remember vaguely that you referred to my mass singularity paper in your paper on the renormalization group equation, but I have not found the reference. Could you give me the proper reference ? I could then include it in the proceedings.

I am impressed that you keep some letters in your file. I am rather bad in keeping letters. I have one letter from Feynman written on a scratch paper, suggesting that I should write the paper by myself. It is somewhere in my file, but I have not yet located it. I also have Feynman's original two sheet calculation in my cabinet, but do not know exactly where it actually is.

Please send my best regard to your wife.

Tom

A.3. Sirlin to Kinoshita, 20 November 2000

Dear Tom,

Thank you very much for your message. Sorry for my delay in answering: I have been sort of 'swamped' by urgent departmental matters and classes. My paper on the renormalization group equation, which is rather pedagogical and is based on considerations of mass singularities, is 'Mass divergences and Callan-Symanzik equations in quantum electrodynamics' [21].

(Several sentences are omitted here.)

The idea of appending some of the information in my e-mail as an appendix to your article is fine with me and may be of some historical interest. As I mentioned, in case anybody is interested, I kept the Feynman and Berman letters I referred to. What I did not keep, and this is a pity, is a copy of my reply to Feynman, in which I stated that Berman's additional term had to be wrong since it violated the cancellation of mass singularities in the corrections to the lifetime, although, of course, the understanding of this, in the general case, had to wait for your later work!

All the best,

Alberto

References

[1] Lee T D and Yang C N 1956 *Phys. Rev.* **104** 254

[2] Wu C S *et al* 1957 *Phys. Rev.* **105** 1413

[3] Garwin R *et al* 1957 *Phys. Rev.* **105** 1415

[4] Lee T D and Yang C N 1957 *Phys. Rev.* **105** 1671

[5] Salam A 1957 *Nuovo Cimento* **5** 299

[6] Landau L 1957 *Nucl. Phys.* **3** 127

[7] Behrends R, Finkelstein R and Sirlin A 1956 *Phys. Rev.* **101** 866

[8] Kinoshita T and Sirlin A 1957 *Phys. Rev.* **107** 593

[9] Kinoshita T and Sirlin A 1959 *Phys. Rev.* **113** 1652

[10] Feynman R P 1949 *Phys. Rev.* **76** 769

[11] Berman S M 1958 *Phys. Rev.* **112** 267

[12] Kinoshita T 1959 *Phys. Rev. Lett.* **2** 477

[13] Ruderman M A and Watson W K R 1956 *Bull. Am. Phys. Soc. Ser.* II **1** 383

[14] Gatto R and Ruderman M A 1958 *Nuovo Cimento* **8** 775

[15] Berman S M 1958 *Phys. Rev. Lett.* **1** 468

[16] Britton D I *et al* 1992 *Phys. Rev. Lett.* **68** 3000

[17] Czapek G *et al* 1993 *Phys. Rev. Lett.* **70** 17

[18] Marciano W J and Sirlin A 1976 *Phys. Rev. Lett.* **36** 1425

[19] Marciano W J and Sirlin A 1993 *Phys. Rev. Lett.* **71** 3629

[20] Kinoshita T 1962 *J. Math. Phys.* **3** 650

[21] Sirlin A 1972 *Phys. Rev.* D **5** 2132

[22] Kinoshita T 1967 *Nuovo Cimento* B **51** 140

[23] Lautrup B E and de Rafael E 1974 *Nucl. Phys.* B **70** 317

[24] Kinoshita T, Nizic B and Okamoto Y 1990 *Phys. Rev.* D **41** 593

[25] Cvitanovic P and Kinoshita T 1974 *Phys. Rev.* D **10** 4007

[26] Kinoshita T and Lindquist W B 1989 *Phys. Rev.* D **39** 2407

[27] Lepage G P 1978 *J. Comput. Phys.* **27** 192

[28] Kinoshita T 1995 *Phys. Rev. Lett.* **75** 4728

[29] Laporta S and Remiddi E 1996 *Phys. Lett.* B **379** 283

[30] Hughes V W and Kinoshita T 1999 *Rev. Mod. Phys.* **71** S133

[31] Jeffery A *et al* 1997 IEEE Trans. Instrum. Meas. **46** 264

[32] Mohr P J and Taylor B N 2000 *Rev. Mod. Phys.* **72** 351

[33] Van Dyck R S *et al* 1987 *Phys. Rev. Lett.* **59** 26

[34] Kinoshita T and Marciano W J 1990 *Quantum Electrodynamics* ed T Kinoshita (Singapore: World Scientific) pp 419–78

[35] For a recent review, see Czarnecki A and Marciano W J 2000 *Preprint* hep-ph/0010104

[36] Karshenboim S G 1993 *Yad. Fiz.* **56** 252
Karshenboim S G 1993 *Phys. At. Nucl.* **56** 857

[37] Brown H N *et al* 2000 *Phys. Rev.* D **86** 091101

[38] Krüger E *et al* 1999 *Metrologia* **36** 147

[39] Liu W *et al* 1999 *Phys. Rev. Lett.* **82** 711

[40] Udem Th *et al* 1999 *Phys. Rev. Lett.* **82** 3568

[41] Hänsch T 2000 Private communication

[42] Van Dyck R S, Schwinberg P B and Dehmelt H G 1991 Experiments with single electrons *The Electrons* ed D Hestenes and A Weingartshofer (Dordrecht: Kluwer) pp 239–93

[43] Gabrielse G and Tan J 1992 *Cavity Quantum Electrodynamics* ed P Berman (New York: Academic)