TRIANGLE ANOMALY AND THE MUON $g-2$*

Andrzej Czarnecki

Department of Physics, University of Alberta
Edmonton, AB T6G 2J1, Canada

William J. Marciano

Physics Department, Brookhaven National Laboratory
Upton, NY 11973, USA

and Arkady Vainshtein

William I. Fine Theoretical Physics Institute, University of Minnesota
116 Church St. SE, Minneapolis, MN 55455, USA

(Received October 24, 2003)

Hadronic electroweak corrections to the muon anomalous magnetic moment $(g-2)$ are reviewed. Emphasis is on clarification of discrepancies among various published studies. A theorem on non-renormalization of the transversal part of a correlator of two vector currents and an axial current is reviewed and its consequences in the form of superconvergent sum rules are discussed.

PACS numbers: 13.40.Em, 14.60.Ef

1. Introduction

The subject of this paper is a class of two-loop electroweak contributions to the muon $g-2$ containing a fermion triangle along with a virtual photon and Z boson, as shown in Fig. 1. In the standard model all charged fermions contribute to the triangle loop. Individual contributions of fermions lighter than the Z boson are enhanced by large logarithms $\ln(M_Z/m)$, where m denotes the mass of the fermion in the loop or of the muon, whichever is heavier. Those large logarithms were first found in [1] where the diagrams

* Presented at the XLIII Cracow School of Theoretical Physics, Zakopane, Poland, May 30–June 8, 2003.

in Fig. 1 where evaluated only for leptons in the triangle loop. However, it was pointed out in [2] and explicitly shown in [3] that such logarithms cancel in sums over all fermions of a given generation (as long as $m \ll M_Z$ for all fermions in the generation). The source of this cancellation can be traced back to the cancellation of anomalies, given by the same fermionic triangles, within a given generation. It is required for renormalizability.

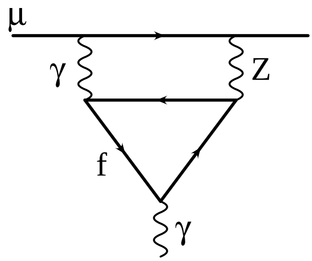

Fig. 1. Effective $Z\gamma\gamma^*$ coupling induced by a fermion triangle, contributing to $a_\mu^{\rm EW}$.

Some subtlety here is that the anomaly cancellation refers to the longitudinal part of the axial Z boson current, while both transversal and longitudinal parts of the fermionic triangle contribute to $g-2$. At the level of free quarks the cancellation is, of course, explicit, so the only question is whether it can be spoiled by strong interactions. Since $\ln M_Z$ arises from the loop momenta much larger than the hadronic scale $\Lambda_{\rm QCD}$ perturbative QCD should provide the answer. For the longitudinal part the Adler–Bardeen theorem [4] guarantees the absence of gluonic corrections governed by $\alpha_{\rm s}(M_Z)$. We will describe below why this is also true for the transversal part confirming the cancellation of $\ln M_Z$ within a generation.

For the light quark loops strong interactions become essential at the range of loop momenta on the order of hadronic scale. Not only is $\alpha_{\rm s} \sim 1$ but also nonperturbative phenomena defining hadron masses are crucial in this range. Namely, these dynamical phenomena — not the current masses of light quarks — define effective infrared cutoff in the logarithmic integrals. While an exact calculation in this range is not possible, in the paper [3] (CKM) a crude approximation was used: "constituent masses" of 300 MeV assigned to quarks u, d and 500 MeV to the quark s played the role of the infrared cutoff. The contribution of the first two generations of fermions was found to be

$$\begin{aligned}\Delta a_\mu^{\rm CKM}(e,u,d;\mu,c,s) &= \frac{G_{\rm F}}{\sqrt{2}}\frac{m_\mu^2}{8\pi^2}\frac{\alpha}{\pi}\left[-6\ln\frac{(m_u m_c)^{4/3}}{(m_d m_s)^{1/3} m_\mu^2} - \frac{49}{3} + \frac{8\pi^2}{9}\right] \\ &\simeq -8.3\times 10^{-11} \qquad (\text{with } m_c = 1.3~\text{GeV})\,. \qquad (1)\end{aligned}$$

Some uncertainty was assigned to this result, but the cancellation of $\ln M_Z$, present in that model, was believed to be rigorously valid.

Around the same time the paper [5] (PPR) appeared, in which the fermion triangle contributions of Fig. 1 were also studied. The result of the PPR paper for the first two generations was

$$\begin{aligned}\Delta a_\mu^{\rm PPR}(e,u,d;\mu,c,s) &= \frac{G_{\rm F}}{\sqrt{2}}\frac{m_\mu^2}{8\pi^2}\frac{\alpha}{\pi}\left[-\frac{14}{3}\ln\frac{M_Z^2}{m_\mu^2}+4\ln\frac{M_Z^2}{m_c^2}-\frac{107}{9}+\frac{8\pi^2}{9}\right]\\ &\simeq -8.7\times 10^{-11}. \qquad (2)\end{aligned}$$

This calculation was done in the chiral limit, so that the u,d,s masses were replaced by m_μ in (2), just like the electron mass. The M_Z dependence of the logarithms would cancel if the first coefficient $(-14/3)$ were replaced by $-12/3=-4$. We have argued that -4 is indeed the correct coefficient and that the factor $-14/3$ arises due to an incomplete accounting of u,d,s contributions in Eq. (20) in [5].

Since the numerical values of (1) and (2) are very similar, and since analytical results in [5] were given for individual flavors and not for their sum, the presence of a residual $\ln M_Z$ in the final result of [5] was inconspicuous. The difference was first pointed out by Mingxing Luo in a private communication to W. Marciano.

More recently, the authors of [5] together with M. Knecht have revisited that issue in a detailed study [6]. They maintained the finding of [5] and argued that the cancellation of $\ln M_Z$ in [3] is a "spurious" result of the naive constituent quark model. If the conclusion of [6] were correct, it would call into question the validity of QCD studies in a variety of contexts, since it suggests that low energy strong dynamics may influence very short distance phenomena, a violation of the basic tenet of asymptotic freedom.

On the other hand, with respect to large distances, the analyses in [5, 6] correctly emphasize differences in the hadronic dynamics of longitudinal and transversal parts of quark triangles. These differences are not properly accounted for by approximation with constituent masses of light quarks. Moreover, constituent masses break the chiral structure of the fermion loops.

The above challenge prompted us to reanalyze the triangle contributions to $g-2$ in [7]. We believe to have located the source of the error in [6]; our arguments are summarized in this text. We also address some questions raised after [7] was published. Let us stress that the discrepancy in question is not relevant for an interpretation of the present $g-2$ experiment, whose design accuracy is $\pm 40\times 10^{-11}$. This is an academic dispute *par excellence.* However, its clarification is important, especially if our community is to determine the much more difficult hadronic three-loop light-by-light effects that limit the accuracy of the standard model prediction for $g-2$.

2. Hadronic triangle diagrams

Some details of the virtual fermion triangle are shown in Fig. 2. For the determination of the muon anomalous magnetic moment, we are interested in the $Z^* \to \gamma^*$ transition between virtual Z and γ in the presence of the external magnetic field to first order in this field. Moreover, the magnetic field is constant so Z^* and γ^* carry the same momentum q. In this approximation, but including all effects of strong interactions in the quark loop (such as gluon exchanges and non-perturbative effects), the amplitude of the $Z^* - \gamma^*$ transition $T_{\mu\nu}(q)$ can be parametrized using two Lorentz-invariant functions of the external momentum q,

$$T_{\mu\nu}(q) = \frac{e}{4\pi^2}\left[w_{\rm T}(q^2)\left(-q^2\tilde{F}_{\mu\nu}+q_\mu q^\sigma\tilde{F}_{\sigma\nu}-q_\nu q^\sigma\tilde{F}_{\sigma\mu}\right)+w_{\rm L}(q^2)\,q_\nu q^\sigma\tilde{F}_{\sigma\mu}\right]. \tag{3}$$

Here, $\tilde{F}_{\mu\nu} = (1/2)\varepsilon_{\mu\nu\gamma\delta}F^{\gamma\delta}$ denotes the dual of the external electromagnetic field tensor. The structure (3) is obtained in the following way. We construct $T_{\mu\nu}$ which is a pseudo-tensor under Lorentz transformations using the vector q_μ and the pseudo-tensor $\tilde{F}_{\mu\nu}$. From these objects we can construct three structures: $\tilde{F}_{\mu\nu}$, $q_\mu q^\sigma\tilde{F}_{\sigma\nu}$, and $q_\nu q^\sigma\tilde{F}_{\sigma\mu}$. In addition, $T_{\mu\nu}$ must be consistent with electromagnetic current conservation, $q^\mu T_{\mu\nu} = 0$. This leaves us with two independent structures and we group them in such a way that one is transversal and the other longitudinal with respect to q_ν.

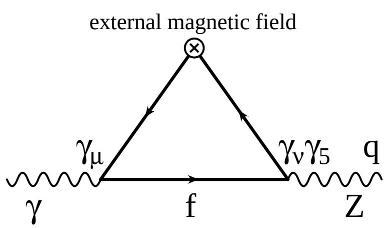

Fig. 2. Virtual fermion loop with structure of couplings and external momentum.

The existence of the longitudinal part stems from the axial ABJ anomaly [9] which in fact exactly determines $w_{\rm L}(q^2)$ in the chiral limit of massless fermions f,

$$q^\nu T_{\mu\nu} = \frac{e}{4\pi^2}w_{\rm L}(q^2)q^2q^\sigma\tilde{F}_{\sigma\mu}\,. \tag{4}$$

Comparing this with the axial anomaly we find

$$w_{\rm L}(q^2) = -\frac{2}{q^2}\,2I_f^3N_fQ_f^2 \equiv \frac{2}{Q^2}\,2I_f^3N_fQ_f^2\,, \tag{5}$$

where Q_f and I_f^3 are the electric charge and the third component of weak isospin projection while N_f accounts for the number of colors in the case of quarks and equals one for leptons. There are no corrections to this result in the limit $m_f = 0$, neither perturbative — due to Adler–Bardeen theorem [4], nor nonperturbative — due to 't Hooft consistency condition [10].

There is no such strong result for w_{T}, but there is a very useful theorem proven by Vainshtein [8]. It states that for massless fermions, the following relation holds to all orders of perturbation theory:

$$2w_{\mathrm{T}}(q^2) = w_{\mathrm{L}}(q^2) \qquad (\text{for } m_f = 0)\,. \tag{6}$$

There are nonperturbative corrections to this result which are given by powers of $\Lambda_{\mathrm{QCD}}^2/Q^2$ but no perturbative ones. One way to prove Vainshtein's theorem is to consider the imaginary part of $T_{\mu\nu}$. The crucial point is that $\mathrm{Im}\,T_{\mu\nu}$ is symmetric under $\mu \leftrightarrow \nu$, $q \leftrightarrow -q$. Indeed, since $\mathrm{Im}\,T_{\mu\nu}$ is given by convergent diagrams, we can freely use the anti-commutation of γ_5 to move it from the axial vertex $\gamma_\nu\gamma_5$ to the vector one, γ_μ. In the limit $m_f = 0$, this involves commuting γ_5 with an even number of γ matrices, no matter how many gluon emissions occur on the way.

Thus, we find

$$\begin{aligned} &\mathrm{Im}\,\left[w_{\mathrm{T}}(q^2)\left(-q^2\tilde{F}_{\mu\nu} + q_\mu q^\sigma \tilde{F}_{\sigma\nu} - q_\nu q^\sigma \tilde{F}_{\sigma\mu}\right) + w_{\mathrm{L}}(q^2)\,q_\nu q^\sigma \tilde{F}_{\sigma\mu}\right] \\ &\propto q_\mu q^\sigma \tilde{F}_{\sigma\nu} + q_\nu q^\sigma \tilde{F}_{\sigma\mu}\,. \end{aligned} \tag{7}$$

This is possible only if

$$2\,\mathrm{Im}\,w_{\mathrm{T}} = \mathrm{Im}\,w_{\mathrm{L}} = \mathrm{const}\,\delta(q^2)\,. \tag{8}$$

Only with such a form for $\mathrm{Im}\,w_{\mathrm{T,L}}(q^2)$, $\mathrm{Im}\,T_{\mu\nu}$ has no terms antisymmetric in $\mu \leftrightarrow \nu$. The coefficient of the delta-function in Eq. (8) is fixed to be $2\pi \times 2I_f^3 N_f Q_f^2$ by the exact form of w_{L} in (5). Both $w_{\mathrm{T}}(q^2)$ and $w_{\mathrm{L}}(q^2)$ are analytical functions decreasing at large q^2. It means that they satisfy unsubtracted dispersion relations

$$w_{\mathrm{T,L}}(q^2) = \frac{1}{\pi}\int_0^\infty ds\,\frac{\mathrm{Im}\,w_{\mathrm{T,L}}(s)}{s-q^2}\,, \tag{9}$$

which when combined with Eq. (8) for imaginary parts implies relation (6) for the real parts as well.

Note that the real part of $T_{\mu\nu}$ contains an antisymmetric term $\tilde{F}_{\mu\nu}$. In terms of diagrammatic calculations it arises from counterterms fixed by the conservation of the vector current. When the Pauli-Vilars regularization is used these counterterms are given by heavy regulator loops.

To summarize, we now know that in the chiral limit $m_f = 0$, the longitudinal function $w_{\rm L}$ is exactly given by Eq. (5) and the transversal function $w_{\rm T}$ is known up to nonperturbative corrections,

$$w_{\rm T} = \frac{2I_f^3 N_f Q_f^2}{Q^2} + \text{non-perturbative corrections}\,. \tag{10}$$

At this point we can apply our findings to an analysis of logarithmic contributions $\ln M_Z$ to $g-2$. We will use the fact that Eqs. (5) and (10) rigorously determine the large Q^2 asymptotics of $w_{L,T}$; the non-perturbative corrections to $w_{\rm T}$, to be discussed below, are suppressed by extra powers of $\Lambda^2_{\rm QCD}/Q^2$ and do not influence $\ln M_Z$ terms. Corrections due to deviations from the chiral limit are also suppressed by powers of m_f/Q.

How do $w_{\rm L}$ and $w_{\rm T}$ enter the $g-2$ contribution of the diagrams in Fig. 1? It was shown in [7] that for a determination of $\ln M_Z$ terms we can use the following simple representation:

$$\Delta a_\mu^{\rm EW} \simeq \frac{\alpha}{\pi}\,\frac{G_\mu\, m_\mu^2}{8\pi^2\sqrt{2}} \int\limits_{m_\mu^2}^{\infty} dQ^2 \left(w_{\rm L} + \frac{m_Z^2}{m_Z^2 + Q^2}\, w_{\rm T} \right) . \tag{11}$$

Two points can be made regarding this integral:

- $\int^\infty dQ^2 w_{\rm L}$ diverges. The theory is inconsistent unless the anomaly cancellation condition is satisfied,

$$\sum_f I_f^3\, N_f\, Q_f^2 = 0 \tag{12}$$

- $\int^\infty dQ^2 M_Z^2\, w_{\rm T}/(m_Z^2 + Q^2) \simeq \ln M_Z$. With $w_{\rm T} = w_{\rm L}/2$ at large Q the anomaly cancellation condition (12) leads to the cancellation of $\ln M_Z$ within a generation where $m_f \ll M_Z$ for all fermions.

In the analyses [5,6] the authors missed the leading $1/Q^2$ part in $w_{\rm T}(u,d,s)$ but not in $w_{\rm L}(u,d,s)$ and $w_{\rm T,L}(e,\mu,c)$. Naturally, they arrived then at the expression (2) where $\ln M_Z$ terms do not cancel. Technically the mistake in [5, 6] arose from an incorrect construction of the Operator Product Expansion (OPE) in the part related to $w_{\rm T}$. Referring to [7] for details note that the authors missed the leading operator which reflects interaction with the soft electromagnetic field (external magnetic field) at short distances. As we demonstrated above this leads to an inconsistency.

On the other hand, the authors of [5, 6] were correct in their analysis of nonperturbative corrections, they were the first to point out that these corrections start with $\Lambda^4_{\rm QCD}/Q^6$ in $w_{\rm T}$. So, below we will try to explain following [7] what can be done when we account for both the leading perturbative term $1/Q^2$ in $w_{\rm T}$ and subleading nonperturbative terms $\Lambda^4_{\rm QCD}/Q^6$ and higher.

3. Large Q^2 expansion, sum rules and models for $\mathrm{Im}\, w_{\rm T,L}$

For simplicity we limit ourselves in this section to the first generation, *i.e.* u and d quarks. As we mentioned above, the first nonperturbative corrections in the chiral limit $m_{u,d} = 0$ are of order $\Lambda^4_{\rm QCD}/Q^6$ in $w_{\rm T}$, as has been shown in detail in [6, 7]. A particular example illustrating how the $d = 6$ operators appear in the OPE is given in Fig. 3. The diagrams allow for a perturbative calculation of the OPE coefficient while averaging of the four-fermion operator in the external magnetic field involves nonperturbative physics.

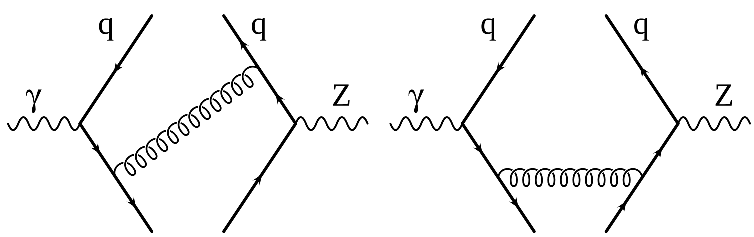

Fig. 3. Diagrams for four-fermion operators responsible for leading nonperturbative corrections to $w_{\rm T}$.

Of course, the expansion of $w_{\rm T}$ in powers of $1/Q^2$ continues further and we are not able to find the sum. Still the analysis gives some exact relations: the first term is $1/Q^2$ with the coefficient 1 (it is normalized to $\sum_{u,d} 2I_3 N_f Q_f^2 = 1$), while the coefficient of $1/Q^4$ is zero. These relations can be rewritten in the form of sum rules for $\mathrm{Im}\, w_{\rm T}$. Indeed, using the dispersion representation (9) at large Q^2 we find

$$\int_0^\infty ds\, \mathrm{Im}\, w_{\rm T}(s) = \pi\,, \qquad \int_0^\infty ds\, s\, \mathrm{Im}\, w_{\rm T}(s) = 0\,. \tag{13}$$

The existence of such sum rules has been pointed out in [11], where their exactness was considered puzzling. It does not look puzzling to us, as there are many examples of such exact sum rules like Weinberg sum rules or sum rules for the DIS structure functions. Moreover, even if $w_{\rm T}$ would not have the leading $1/Q^2$ term, as assumed in [6], these sum rules would be no less exact (except the right hand side would be zero in the first one).

It is instructive to compare these relations with similar sum rules for the longitudinal function $w_{\rm L}$ where all powers of $1/Q^2$ are known,

$$\int_0^\infty ds\, {\rm Im}\, w_{\rm L}(s) = 2\pi\,, \qquad \int_0^\infty ds\, s^k\, {\rm Im}\, w_{\rm L}(s) = 0 \quad (k=1,\ldots,\infty)\,. \tag{14}$$

This set of relations implies a unique solution ${\rm Im}\, w_{\rm L}(s) = 2\pi\delta(s)$ showing that the massless pion is the only contributing intermediate state.

For the transversal function $w_{\rm T}$ the intermediate hadronic states have to be 1^+ mesons with isospin 1 and 0 or 1^- mesons with isospin 1. The lightest ones ρ, ω and a_1 are massive even in the chiral limit. Representing ${\rm Im}\, w_{\rm T}$ as

$${\rm Im}\, w_{\rm T} = \pi \sum_i g_i \delta(s - m_i^2) \tag{15}$$

we get from Eq. (13)

$$\sum_i g_i = 1\,, \qquad \sum_i g_i\, m_i^2 = 0\,. \tag{16}$$

The analysis of [6] assumed $\sum_i g_i = 0$.

Until now we did not make any specific assumptions. Let us now assume a saturation by the lightest states and consider ρ and ω to be degenerate. Then the solution of relations (16) gives

$$w_{\rm T} = \frac{1}{m_{a_1}^2 - m_\rho^2}\left[\frac{m_{a_1}^2}{Q^2+m_\rho^2} - \frac{m_\rho^2}{Q^2+m_{a_1}^2}\right]. \tag{17}$$

This model for $w_{\rm T}$ by construction satisfies exact relations (13) and was used in [7]. Of course, one can modify the model adding extra resonances but numerically a_μ is not much sensitive to these modifications.

One more comment on the chiral symmetry breaking by quark masses. Accounting for nonvanishing $m_{u,d}$ leads to negligible effects for $w_{\rm T}$. The main (and important) effect for $w_{\rm L}$ is a shift of the pole from 0 to m_π^2, *i.e.* to $w_{\rm L} = 2/(Q^2+m_\pi^2)$.

These refinements of $w_{\rm T,L}$ in the long distance hadronic triangle effects for the u,d quark loops, together with similar ones for the s quark, were made in our study [7]. Numerically it changes the result in Eq. (1) to $\Delta a_\mu^{\rm CMV}(e,u,d;\mu,c,s) = -6.7(1.0)\times 10^{-11}$.

4. Summary

Although the combined hadronic and leptonic triangle loop effects in electroweak corrections are not that significant for the muon $g-2$, it is reassuring that a theory based on the Operator Product Expansion and its matching with hadronic phenomenology allows for quite accurate calculations. We resolved a conceptual controversy existing in literature concerning cancellation of short distances between fermionic triangles. It seems to be possible now to develop an analogous approach to improve theoretical accuracy for the light-by-light part of hadronic effects.

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the DOE grants DE-AC03-76SF00515 and DE-FG02-94ER408.

REFERENCES

[1] E.A. Kuraev, T.V. Kukhto, A. Schiller, *Sov. J. Nucl. Phys.* **51**, 1031 (1990); T.V. Kukhto, E.A. Kuraev, A. Schiller, Z.K. Silagadze, *Nucl. Phys.* **B371**, 567 (1992).

[2] T. Kinoshita, W.J. Marciano, in *Quantum Electrodynamics*, edited by T. Kinoshita, World Scientific, Singapore 1990, p. 419.

[3] A. Czarnecki, B. Krause, W. Marciano, *Phys. Rev.* **D52**, 2619 (1995).

[4] S.L. Adler, W.A. Bardeen, *Phys. Rev.* **182**, 1517 (1969).

[5] S. Peris, M. Perrottet, E. de Rafael, *Phys. Lett.* **B355**, 523 (1995).

[6] M. Knecht, S. Peris, M. Perrottet, E. de Rafael, *J. High Energy Phys.* **11**, 003 (2002).

[7] A. Czarnecki, W.J. Marciano, A. Vainshtein, *Phys. Rev.* **D67**, 073006 (2003).

[8] A. Vainshtein, *Phys. Lett.* **B569**, 187 (2003).

[9] S.L. Adler, *Phys. Rev.* **177**, 2426 (1969); J.S. Bell, R. Jackiw, *Nuovo Cim.* **A60**, 47 (1969).

[10] G. 't Hooft, in *Recent Developments In Gauge Theories,* eds. G. 't Hooft *et al.*, Plenum Press, New York 1980.

[11] E. de Rafael, talk given at the International Symposium on Lepton Moments, Cape Cod, June'03, `http://g2pc1.bu.edu/~leptonmom/talks/deRafael.pdf` and private communication.