5-loop QED calculation of the electron anomalous magnetic moment

Sergey Volkov

MPIK Heidelberg

Progress in algorithms and numerical tools for QCD
Orsay, 2026

Electron's *(g-2)/2*: the current status

Experiment:

a_e=0.00115965218073(28) [**2011**, D. Hanneke, S. Fogwell Hoogerheide, G. Gabrielse, Phys. Rev. A 83, 052122]

a_e=0.00115965218059(13) [**2022**, X. Fan, T. G. Myers, B. A. D. Sukra, G. Gabrielse, arXiv:2209.13084]

Theory:

$$a_e = a_e(QED) + a_e(hadronic) + a_e(electroweak),$$

$$a_e(QED) = \sum_{n\geqslant 1} \left(\frac{\alpha}{\pi}\right)^n a_e^{2n},$$

$$a_e^{2n} = A_1^{(2n)} + A_2^{(2n)}(m_e/m_\mu) + A_2^{(2n)}(m_e/m_\tau) + A_3^{(2n)}(m_e/m_\mu, m_e/m_\tau)$$

a_e=0.00115965218154 9(4)(12)(299)

α^{-1}=137.035999046(27) [**2018**, R. H. Parker et al., Science 360, 191-195]

a_e=0.00115965218019 7(4)(12)(94)

α^{-1}=137.035999206(11) [**2020**, L. Morel et al., Nature 588, 61-65]

Uncertainties come from: $A_1^{(10)}$, hadronic+electroweak, α

(recent hadronic data give a different value and are contradictive)

$A_1^{(10)}$=5.887(55) – S. Volkov & AHKN (T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio), **2025** consensus value

The history of the universal QED contributions $A_1^{(2n)}$ calculations

$$a_e(QED) = \sum_{n \geqslant 1} \left(\frac{\alpha}{\pi}\right)^n a_e^{2n},$$

$$a_e^{2n} = A_1^{(2n)} + A_2^{(2n)}(m_e/m_\mu) + A_2^{(2n)}(m_e/m_\tau) + A_3^{(2n)}(m_e/m_\mu, m_e/m_\tau)$$

- J.Schwinger [**1948**], analytically: $A_1^{(2)}$=**0.5**
- R. Karplus, N. Kroll [**1949**] – $A_1^{(4)}$ with a mistake
  A.Petermann [**1957**], C. Sommerfield [**1958**], analytically: $A_1^{(4)}$=**-0.328478966...**
- ~**1970**...~**1975**, $A_1^{(6)}$, numerically:
  1. M.Levine, J. Wright.
  2. R. Carroll, Y. Yao.
  3. T. Kinoshita, P. Cvitanović.
     T. Kinoshita, P. Cvitanović [**1974**]: $A_1^{(6)}$=**1.195±0.026**
- E. Remiddi, S. Laporta et al., ~**1965**..**1996**, analytically: $A_1^{(6)}$=**1.181241456**...
- T. Kinoshita, M. Nio et al., numerically, **2015**: $A_1^{(8)}$=**-1.91298(84)**
  (first estimations in 1980-x)
- S. Laporta, semianalytically, **2017**: $A_1^{(8)}$=**-1.9122457649...**
- T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio, numerically, **2012**: first numerical values of $A_1^{(10)}$
- S. Volkov, numerically **2019**: discrepancy of **4.8σ** in $A_1^{(10)}$
- **2025**: the discrepancy was resolved: $A_1^{(10)}$=**5.887(55)**

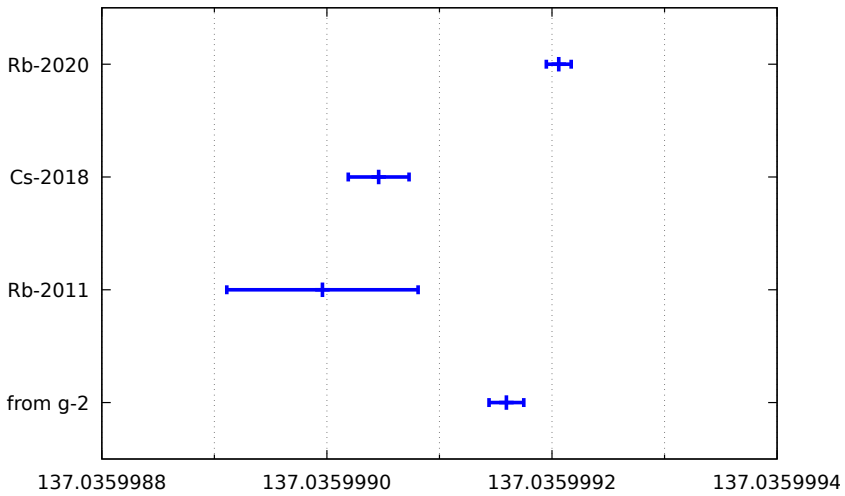
The history of the mass-dependent QED contributions calculations

$$a_e(QED) = \sum_{n \geqslant 1} \left(\frac{\alpha}{\pi}\right)^n a_e^{2n},$$

$$a_e^{2n} = A_1^{(2n)} + A_2^{(2n)}(m_e/m_\mu) + A_2^{(2n)}(m_e/m_\tau) + A_3^{(2n)}(m_e/m_\mu, m_e/m_\tau)$$

- H.H.Elend [1966], analytically (**recently obtained masses are substituted into the analytical results**):
  for electron $A_2^{(4)}(m_e/m_\mu)=0.519738676(24)\cdot10^{-6}$, $A_2^{(4)}(m_e/m_\tau)=0.183790(25)\cdot10^{-8}$,
  for muon $A_2^{(4)}(m_\mu/m_e)=1.0942583093(76)$, $A_2^{(4)}(m_\mu/m_\tau)=0.000078076(11)$.
- ~**1970**...~**1990**, numerically, different research groups, increasing precision:
  J. Aldins, S. J. Brodsky, C. Chlouber, A. J. Dufner, T. Kinoshita, W. J. Marciano, B. Nizic, Y. Okamoto, M. A. Samuel...
  T. Kinoshita, W. J. Marciano [**1990**]: for muon $A_2^{(6)}(m_\mu/m_e)=$**22.8671(22)**.
- M. A. Samuel, G. Li, S. Laporta, E. Remiddi [**1991-1993**], analytically:
  for electron $A_2^{(6)}(m_e/m_\mu)=$**-0.737394164(24)·10^{-5}**, $A_2^{(6)}(m_e/m_\tau)=$**-0.658273(79)·10^{-7}**,
  for muon $A_2^{(6)}(m_\mu/m_e)=$**22.86837998(20)**, $A_2^{(6)}(m_\mu/m_\tau)=$**0.000360671(94)**.
- S. Laporta [**1993**], semianalytically: for muon $A_3^{(6)}(m_\mu/m_e,m_\mu/m_\tau)=$**0.0005238(19)**.
- A. Czarnecki, M. Skrzypek, B. Krause [**1999**], analytically: for muon $A_3^{(6)}(m_\mu/m_e,m_\mu/m_\tau)=$**0.000527738(75)**.
- T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio [**2012**], numerically:
  for electron $A_2^{(8)}(m_e/m_\mu)=$**0.0009222(66)**, $A_2^{(8)}(m_e/m_\tau)=$**8.24(12)·10^{-6}**,
  for muon $A_2^{(8)}(m_\mu/m_e)=$**132.6852(60)**, $A_2^{(8)}(m_\mu/m_\tau)=$**0.04234(12)**, $A_3^{(8)}(m_\mu/m_e,m_\mu/m_\tau)=$**0.06272(4)**,
  $A_2^{(10)}(m_\mu/m_e)=$**742.18(87)** [**NOT DOUBLE-CHECKED**], $A_2^{(10)}(m_\mu/m_\tau)=$**-0.068(5)**, $A_3^{(10)}(m_\mu/m_e,m_\mu/m_\tau)=$**2.011(10)**.
- A. Kurz, T. Liu, P. Marquard, M. Steinhauser [**2013**], semianalytically:
  for electron $A_2^{(8)}(m_e/m_\mu)=$**0.00916197070 3(372)**, $A_2^{(8)}(m_e/m_\tau)=$**7.42924(118)·10^{-6}**,
  for muon $A_2^{(8)}(m_\mu/m_\tau)=$**0.0424941(53)**.
- A. Kurz, T. Liu, P. Marquard, V. A. Smirnov, A. V. Smirnov, M. Steinhauser [**2016**], semianalytically:
  for muon $A_2^{(8)}(m_\mu/m_e)=$**132.86(48)**, $A_3^{(8)}(m_\mu/m_e,m_\mu/m_\tau)=$**0.0627220(100)**.

a_e and α



α^{-1} obtained by different methods

- $\alpha^{-1}[a_e] = 137.035999159(16)$
- $\alpha^{-1}[\text{Rb-2011}] = 137.035998996(85)$
  (R. Bouchendira, P. Cladé, S. Guellati-Khélifa, F. Nez, F. Biraben, PRL 106, 080801, 2011 + CODATA-2014)
  $\alpha^{-1}[\text{Cs-2018}] = 137.035999046(27)$
  (R. H. Parker, C. Yu, W. Zhong, B. Estey, H. Müller, Science 360, 191, 2018)
  $\alpha^{-1}[\text{Rb-2020}] = 137.035999206(11)$
  (L. Morel, Z. Yao, P. Cladé, S. Guellati-Khélifa, Nature 588, 61, 2020)

10-th order: my and AHKN results

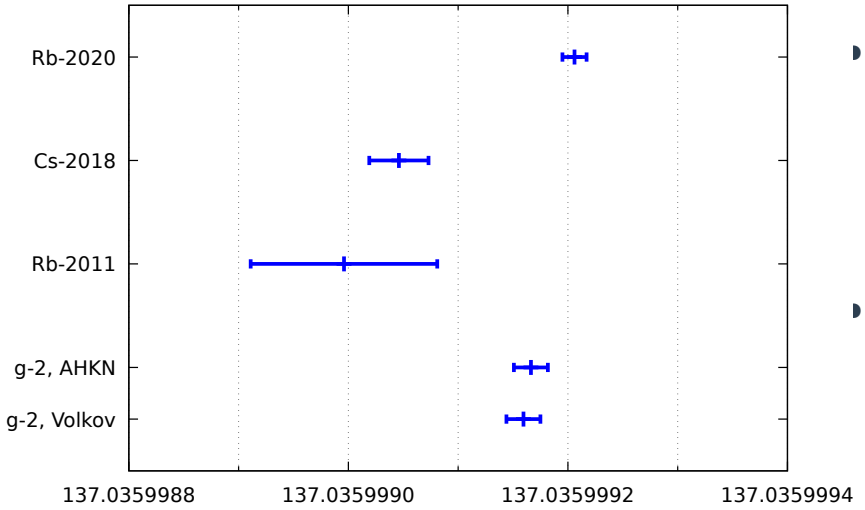
- AHKN = T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio

  $A_1^{(10)}$[AHKN] = 5.870(128) [**2025**]
  [T. Aoyama, M. Hayakawa, A. Hirayama, M. Nio, Phys. Rev. D 111, L031902]

- My result

  $A_1^{(10)}$[Volkov] = 5.891(61) [**2024**]
  [S. Volkov, Phys. Rev. D 110, 036001]

a_e and α: before resolving the discrepancy in $A_1^{(10)}$

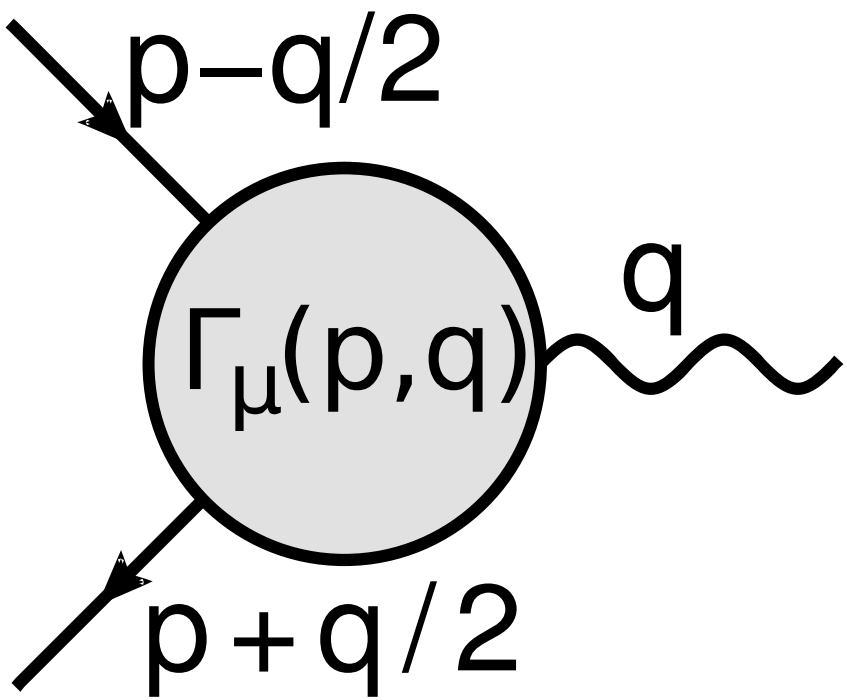


α^{-1} obtained by different methods

- $\alpha^{-1}[a_e, \text{Volkov}] = 137.035999159(16)$
  (S. Volkov, Phys. Rev. D 110, 036001, 2024)
  $\alpha^{-1}[a_e, \text{AHKN}] = 137.035999166(16)$ ← OLD VALUE
  (T. Aoyama, T. Kinoshita, M. Nio, Atoms 7, 28, 2019)

- $\alpha^{-1}[\text{Rb-2011}] = 137.035998996(85)$
  (R. Bouchendira, P. Cladé, S. Guellati-Khélifa, F. Nez, F. Biraben, PRL 106, 080801, 2011 + CODATA-2014)
  $\alpha^{-1}[\text{Cs-2018}] = 137.035999046(27)$
  (R. H. Parker, C. Yu, W. Zhong, B. Estey, H. Müller, Science 360, 191, 2018)
  $\alpha^{-1}[\text{Rb-2020}] = 137.035999206(11)$
  (L. Morel, Z. Yao, P. Cladé, S. Guellati-Khélifa, Nature 588, 61, 2020)

Anomalous magnetic moment from Feynman diagrams



In the “spinor sandwich”, the Feynman amplitude $\Gamma_\mu(p,q)$ can be expressed in terms of three form-factors:

$$\overline{u}\Gamma_\mu(p,q)v = \overline{u}(f(q^2)\gamma_\mu + g(q^2)p_\mu + h(q^2)q_\mu)v$$

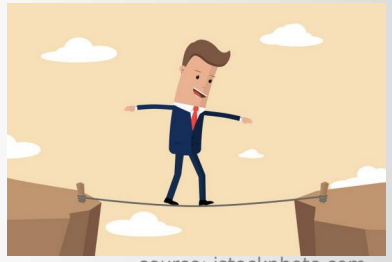
The anomalous magnetic moment is extracted from

$$\lim_{q^2\to 0} g(q^2)$$

Approaches to handle 5 loops

With infinitesimal regularization parameters (dimensional regularization is usually used):

- less troubles with a complicated structure of divergences;
- this path is very narrow: it is well known how much computer resources it takes;
- enormous amount of symbolic manipulations at 5 loops.



source: istockphoto.com

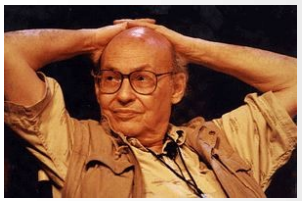
Without infinitesimal regularization parameters (point-by-point subtraction of divergences):

- finite integrals are required from the beginning;
- more freedom for optimizations;
- feasible at higher orders.

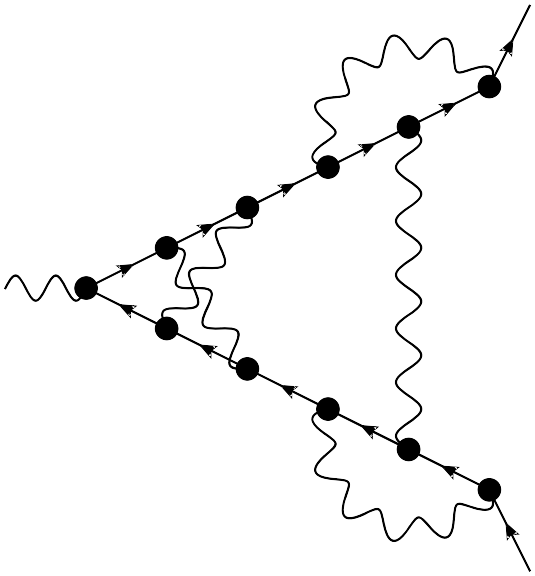
My calculation: general approach

You don't understand anything until you learn it more than one way.

Marwin Minsky
(a mathematician, computer scientist, cognitive scientist, father of AI)



source: www.bcp.psych.ualberta.ca

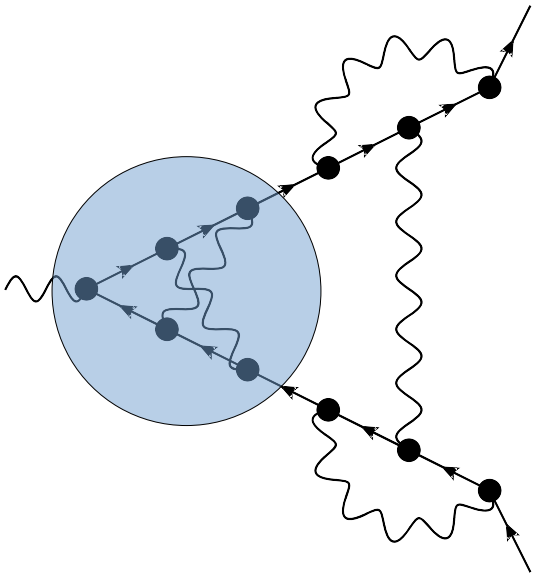


an example of a diagram of $A_1^{(10)}$

Diagrams: 12672
(undirected: 5536)

- We construct an integral in Feynman-parameteric space for each diagram:
  $\int_{z\geq 0} f(z_1,z_2,\ldots,z_n)\delta(z_1+\ldots+z_n-1)d^n z$
  A Feynman parameter z_j is assigned to each line ($n=15$).
- Divergences (UV, IR, mixed) are removed under the integral sign (before integration).
- Things like dimensional regularization, expansions of infinities or master integrals are never used.

Handling UV divergences in QED Feynman diagrams



To obtain subtraction counterterms for removing UV divergences, the Feynman amplitudes of subdiagrams are modified.

The procedure is equivalent to the renormalization (of the **mass**, **charge**, **wave function**).

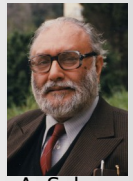
General idea how to work at any order:
F. J. Dyson, Phys. Rev. 75, 1736 (1949)
A. Salam, Phys. Rev. 82, 217 (1951)

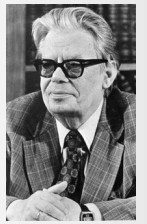
First attempt of rigorous proof, the hope that QFT can be treated rigorously, inspiration:
N. N. Bogoliubov and O. S. Parasiuk, Acta Math. 97, 227 (1957)

Rigorous proof of finiteness (UV only), Schwinger parameters:
K. Hepp, Commun. Math. Phys. 2, 301 (1966)

Rigorous proof of finiteness (UV only), momentum space:
W. Zimmermann, Commun. Math. Phys. 15, 208 (1969)


F. Dyson


A. Salam


N. Bogoliubov


O. Parasiuk


K. Hepp


W. Zimmermann

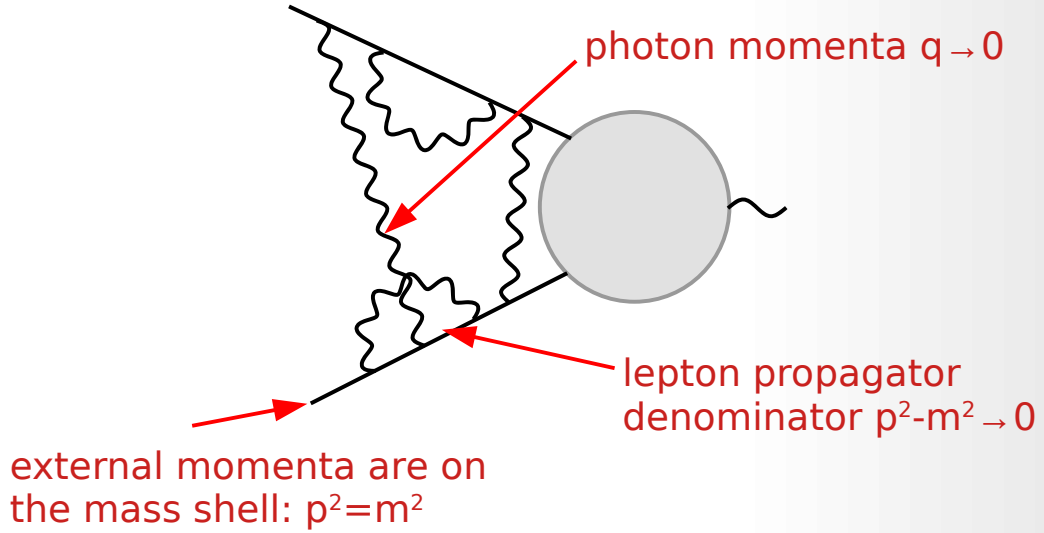
Picture sources: wikipedia.org, ru.ruwiki.ru, www.ias.edu, twas.org, pro-physik.de

Handling UV divergences is a **beautiful story**:

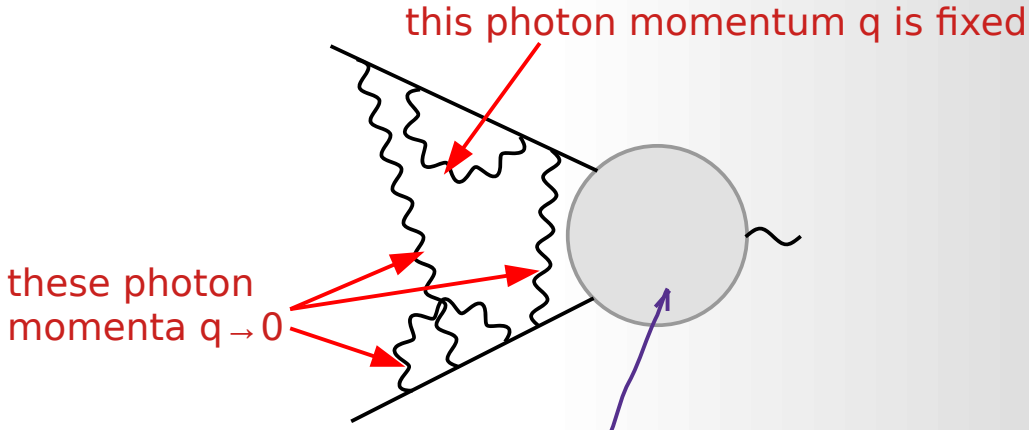
- rigorous proofs;
- It is universal.

IR divergences in QED: how they work?

Logarithmic IR divergences



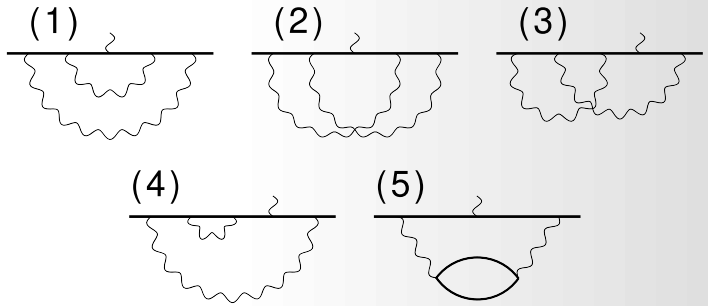
Power-type IR divergences



IR and UV divergences can mix with each other! (if it is UV divergent)

What if we apply BPHZ-like in-place on-shell renormalization?

- Two types of IR divergences:
  - contained in original diagrams;
  - coming from renormalization.
  They cancel each other: this is why the electron *g-2* is IR-finite.
- The cancellation is inter-diagram.

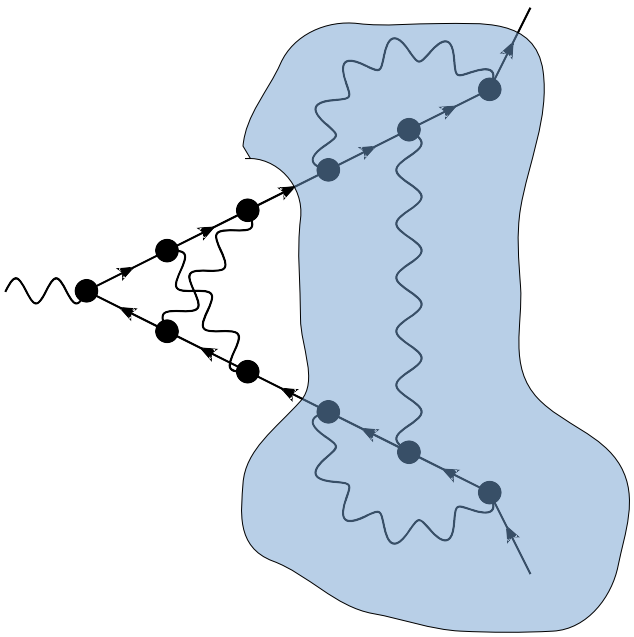


Diagrams for $A_1^{(4)}$

№	Value
1	0.77747802
2	-0.46764544
3	0.564021-(1/2)log(λ^2/m^2)
4	-0.089978+(1/2)log(λ^2/m^2)
5	0.0156874

Contributions to $A_1^{(4)}$ (Petermann, 1957)

Handling IR divergences in Feynman diagrams of the electron *g-2*



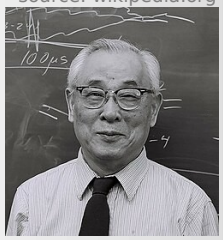
To obtain subtraction counterterms for removing IR divergences, we typically modify the expression part related to the part outside of a subdiagram.

Different approaches to the removal of IR divergences (together with the UV and mixed ones):

- M. J. Levine, J. Wright, Phys. Rev. D 8, 3171 (**1973**);
- R. Carroll, Y.-P. Yao, Phys. Lett. 48B, 125 (**1974**);
- P. Cvitanović, T. Kinoshita, Phys. Rev. D 10, 3991 (**1974**);
- T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio, Nucl. Phys. B 796, 184 (**2008**);
- **S. Volkov**, J. Exp. Theor. Phys. 122, 1008 (**2016**);
- **S. Volkov**, Phys. Part. Nuclei 53, 805 (**2022**).

simplest

source: wikipedia.org



Toichiro Kinoshita
first went beyond $A_1^{(6)}$

Handling IR divergences is **not a beautiful story at all**:

- not universal (for specific observables only, like the anomalous magnetic moment);
- no rigorous proofs.

My divergence elimination method: general idea, linear operators

- For the in-place on-shell renormalization the linear operators L (for vertex-like subdiagrams) and B (for electron self-energy) are used:

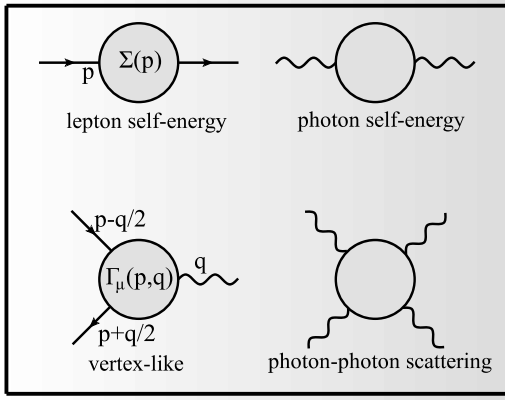
$$L\Gamma_\mu(p,q) = (a(m^2) + b(m^2)m + c(m^2)m^2)\gamma_\mu,$$
$$\text{where } \Gamma_\mu(p,0) = a(p^2)\gamma_\mu + b(p^2)p_\mu + c(p^2)\not{p}p_\mu + d(p^2)(\not{p}\gamma_\mu - \gamma_\mu\not{p}),$$
$$\not{p} = p^\nu\gamma_\nu$$

UV-divergent IR-finite

no overall UV div. IR-divergent

no overall UV div. IR-finite

$$B\Sigma(p) = r(m^2) + s(m^2)m + 2m(r'(m^2) + ms'(m^2))(\not{p} - m),$$
$$\text{where } \Sigma(p) = r(p^2) + s(p^2)\not{p}$$



They follow the rules of QED.
L and B produce IR divergences, it is important for the IR-divergence cancellation in the final value.

- We **redistribute divergences** between diagrams and integration regions using special operators U:

$$U\Gamma_\mu(p,q) = a(M^2)\gamma_\mu, \qquad U\Sigma(p) = r(m^2) + s(m^2)m + s(M^2)(\not{p} - m)$$

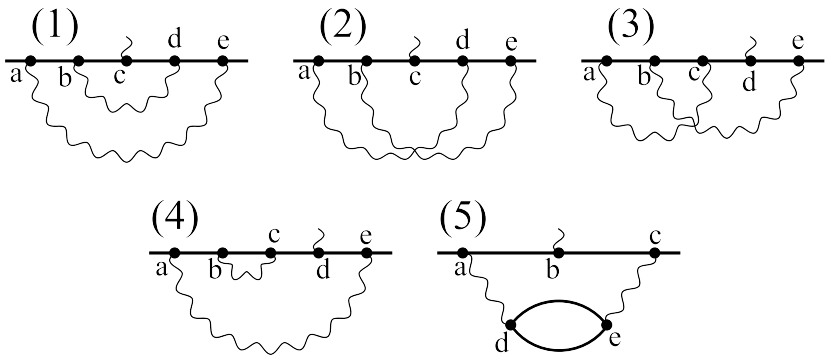
They preserve the Ward identity. They are IR-safe. They extract the overall (superficial) UV divergence. The mass part is extracted completely.

m is the lepton mass. M^2 is an arbitrary parameter.

Calculation **2019**: $M^2=m^2$

Calculation **2024**: $M^2=-m^2$ (1.5 times **less cancellations** in individual diagrams contributions)

My divergence elimination nethod:
2-loop example



$A_1^{(4)}$[Petermann]=-0.328478966…

$A_1^{(4)}$[Volkov]=-0.3284777(34)

- Linear operators:
  A = anomalous magnetic moment projector (multiplied by γ_μ)
  L, *B* = on-shell renormalization for vertex-like and lepton self-energy diagrams
  U = intermediate operators for vertex-like and lepton self-energy
  P = on-shell renormalization for photon self-energy
- The method removes all divergences.
- The subtraction is equivalent to the on-shell renormalization after summation over diagrams.

№	My expression	On-shell renorm.	Difference
1	A_G-$A_G U_{bcd}$-(L_G-U_G)A_{bcd}	A_G-$A_G L_{bcd}$	(L_G-U_G)A_{bcd}-A_G(L_{bcd}-U_{bcd})
2	A_G	A_G	0
3	2(A_G-$A_G U_{abc}$)	2(A_G-$A_G L_{abc}$)	2A_G(U_{abc}-L_{abc})
4	2(A_G-$A_G U_{bc}$)	2(A_G-$A_G B_{bc}$)	2A_G(U_{bc}-B_{bc})
5	A_G-$A_G P_{de}$	A_G-$A_G P_{de}$	0

№	My value	Petermann, 1957
1	0.77747774(18)	0.77747802
2	-0.4676475(17)	-0.46764544
3	-0.0640193(19)	-0.564021-(1/2)log(λ^2/m^2)
4	-0.5899758(14)	-0.089978+(1/2)log(λ^2/m^2)
5	0.0156895(25)	0.0156874

How to integrate?

- We have to integrate
  $$\int_{z\geq 0} f(z_1,\ldots,z_M)\,\delta(z_1+\ldots+z_M-1)\,d^M z$$
  with *M=14* (reduced by 1 due to some trick) for each of 5536 Feynman diagrams.
- Many variables: Monte Carlo is the only feasible solution.
- Uniform probability distribution for random samples works extremely bad.
- Importance sampling: more samples for regions with large $|f(z)|$.
- $f(z)$ is singular at the boundary, has acute peaks.
- Adaptive algorithms work not so well, because the number of qualitatively different regions in the integration domain grows exponentially (or even faster) with M.

Monte Carlo integration: general ideas

- We employ a **nonadaptive** approach for the Monte Carlo integration.
- The ideas come from the theory of renormalization:
  E. Speer, J. Math. Phys. 9, 1404 (1968)
  (the asymptotic behavour of the integrand can be explained by ultraviolet degrees of divergence of **all** subdiagrams).
- If there is no infrared cut-off, these ideas are not enough. We need more:
  S. Volkov, Nucl. Phys. B 961, 115232 (2020).
- The algorithm is still partially adaptive:
  - all diagrams are evaluated simultaneously, the time distribution is being adjusted adaptively;
  - several parameters are preadjusted (by lower-order experiments).

Nonadaptive Monte Carlo integration: diagram-specific probability density functions

- Integral: $\int\limits_{z_1,\ldots,z_M>0} f(z_1,\ldots,z_M)\,\delta(z_1+\ldots+z_M-1)\,dz$
- Hepp sectors: $z_{j_1}\geq z_{j_2}\geq\ldots\geq z_{j_M}$
  (14!=**87178291200** sectors for each diagram)
- Probability density: $C\times\dfrac{\prod\limits_{j=2}^{M}\left(z_{j_l}/z_{j_{l-1}}\right)^{Deg(\{j_l,j_{l+1},\ldots,j_M\})}}{z_1 z_2\ldots z_M}$

  Deg is defined on subsets of *{1,2,…,M}*
  (idea: E. Speer, J. Math. Phys. 9, 1404 (1968); 2^{16}=**16384** subsets for each diagram)

- My ideas are:
  1) how to calculate *Deg(s)* for each set *s*
     based on a theory (S. Volkov, Nucl. Phys. B 961, 115232 (2020))
     and adjustment
  2) how to generate random samples fast
  S. Volkov, Phys. Rev. D 96, 096018 (2017); S. Volkov, Phys. Rev. D 98, 076018 (2018)

Integration: huge computer resources

- The integrand code size: **1 TB**
  (the compilation required **1 CPU core-month**)
- Nvidia GPUs were used for integration.
- Two calculations:
  - [**2019**] a part of $A_1^{(10)}$, supercomputer “Govorun” (Dubna), Nvidia V100, **4.5 GPU-years**;
  - [**2024**] full $A_1^{(10)}$, supercomputer “HoreKA” (Karlsruhe), Nvidia A100, **11.5 GPU-years**.
- Total number of Monte Carlo samples $\approx 10^{15}$.

Summary

- I have my own calculation of $A_1^{(10)}$ based on point-by-point subtraction of divergences under the integral sign, combinatorial-inspired nonadaptive Monte Carlo integration, parallel realization on GPUs.
- At this moment, $A_1^{(10)}$ is double-checked (the current precision is 1.5% due to a statistical uncertainty of the Monte Carlo integration).
- The methods based on point-by-point subtraction of divergences before integration are very promising at high orders: they do not require huge amounts of symbolic manipulations, they give more freedom for optimizations.