

Advanced Quantum Field Theory : Renormalization, Non-Abelian Gauge Theories and Anomalies

*Lecture notes
Amsterdam-Brussels-Geneva-Paris Doctoral School
“Quantum Field Theory, Strings and Gravity”
Brussels, October 2026*

Adel Bilal

Laboratoire de Physique de l'École Normale Supérieure
PSL Research University, CNRS
24 rue Lhomond, F-75231 Paris Cedex 05, France

Abstract

These are lecture notes of an advanced quantum field theory course intended for graduate students in theoretical high energy physics who are already familiar with the basics of QFT. After a recap on the functional integral in quantum field theory, the first part discusses the formulation and renormalization of non-abelian gauge theories, as well as general renormalization theory. The second part gives an introduction to anomalies. Starting from the simple abelian anomaly, we mainly discuss the chiral (gauge) anomalies and end with a brief discussion of gauge and gravitational anomalies in arbitrary dimensions, as well as their possible cancellation.

Contents

1	Some elements of quantum field theory and the functional integral	1
1.1	Functional integrals	1
1.1.1	Path integrals in QM	1
1.1.2	Functional integrals in QFT	2
1.2	Propagators	3
1.3	Interactions and Feynman diagrams	5
1.3.1	Feynman diagrams from functional integrals	6
1.3.2	Relation with Dyson's formula and Wick's theorem	7
1.4	Generating functional for connected Green functions	8
1.5	Quantum effective action	9
1.6	Loop diagrams: regularization and renormalization	10
1.7	An example : vacuum polarization in QED	12
1.8	Exercises	15
1.8.1	Linear operators on $L^2(\mathbf{R}^4)$	15
1.8.2	Green's function and generating functional in $0 + 1$ dimensions	16
1.8.3	Effective action from integrating out fermions in QED	17
1.8.4	An example of asymptotic series	18
2	General renormalization theory	19
2.1	Degree of divergence	19
2.2	Structure of the divergences	21
2.3	Bogoliubov-Parasiuk-Hepp-Zimmermann prescription and theorem	24
2.4	Summary of the renormalization program and proof	25
2.5	The criterion of renormalizability	26
2.6	Exercices	27
2.6.1	Yukawa theory	27
2.6.2	Quantum gravity	27
2.6.3	Non-renormalizable interactions from integrating out a massive field	28
3	Non-abelian gauge theories: formulation and quantization	29
3.1	Non-abelian gauge transformations and gauge invariant actions	29
3.2	Quantization	33
3.2.1	Faddeev-Popov method	33
3.2.2	Gauge-fixed action, ghosts and Feynman rules	36
3.2.3	BRST symmetry	37
3.3	BRST cohomology	40
3.4	Exercices	43
3.4.1	Some one-loop vertex functions	43
3.4.2	More general gauge-fixing functions	44
3.4.3	Two ghost - two antighost counterterms	44
4	Renormalization of non-abelian gauge theories	45
4.1	Slavnov-Taylor identities and Zinn-Justin equation	45
4.1.1	Slavnov-Taylor identities	45
4.1.2	Zinn-Justin equation	46
4.1.3	Antibracket	47
4.1.4	Invariance of the measure under the BRST transformation	47
4.2	Renormalization of gauge theories theories	49
4.2.1	The general structure and strategy	49
4.2.2	Constraining the divergent part of Γ	51
4.2.3	Conclusion and remarks	56
4.3	Background field gauge	58
4.4	One-loop β -functions for Yang-Mills and supersymmetric Yang-Mills theories	60
4.4.1	β -function for Yang-Mills theory	61

4.4.2	β -functions in supersymmetric gauge theories	64
5	Overview	66
6	Various manifestations of an anomaly	67
6.1	Noether's theorem and a Ward identity	67
6.2	Non-invariance of the fermionic measure under chiral transformations (abelian anomaly) . . .	69
6.2.1	Transformation of the measure	71
6.2.2	Defining and computing the determinant	72
6.2.3	Relation of the abelian anomaly with the index theorem	75
6.3	Matter effective action and more Ward identities	76
6.3.1	Anomalous Ward identity for the abelian anomaly	77
6.3.2	Anomalous Ward identity for the chiral anomaly	78
7	Computing the triangle diagrams	79
7.1	Computations for the abelian anomaly	79
7.2	Computations for the chiral anomaly	82
7.2.1	Chiral fermions	82
7.2.2	Matter Lagrangian for chiral fermions	83
7.2.3	Feynman diagram computation of the triangle for chiral fermions	84
8	General characterization and properties of the anomalies	88
8.1	Group structure	88
8.2	Particles or antiparticles	88
8.3	Locality	89
8.4	How many points necessary for an anomalous diagram?	89
8.5	Finiteness	90
8.6	One-loop exactness	91
8.7	Relevant and irrelevant anomalies	91
8.8	Anomaly cancellation	92
8.9	Wess-Zumino consistency condition	93
8.10	Reformulation in terms of differential forms	93
9	BRST cohomology and descent equations	95
9.1	BRST transformations	95
9.2	Reformulating relevant anomalies as BRST cohomologies	96
9.3	Solving the WZ-condition	97
9.4	Descent equations : characteristic classes and Chern-Simons forms	98
9.4.1	Characteristic classes	98
9.4.2	Chern-Simons forms and descent equations	100
9.5	BRST cohomology at ghost number one	102
9.6	Gauge anomaly in $d = 2n$ dimensions	103
10	Gauge and gravitational anomalies and index theorems	105
10.1	Anomalies and index theorems	105
10.2	Gravitational anomalies	108
10.3	Anomaly cancellation in ten-dimensional IIB supergravity	111
11	Conclusions of part II	114

Part I : Renormalization and non-abelian gauge theories

1 Some elements of quantum field theory and the functional integral

This section aims to give an overview of the functional integral formulation of quantum field theory we will be using throughout these notes, and how it is related to the (perturbative) Feynman diagram expansion. While this is standard material and can be found in many textbooks, the presentation here will follow mostly [7].

1.1 Functional integrals

Functional integrals are the generalization to QFT of path integrals in ordinary QM. Although not necessarily mathematically rigorous, functional integrals are a useful tool for many computations whose accuracy can be confirmed by independent “proper” Feynman diagram computations. Our signature is $(-+++)$ and, of course, $\hbar = c = 1$.

1.1.1 Path integrals in QM

Let us first briefly introduce path integrals. We will be working in the Heisenberg picture and $|q, t\rangle$ denotes the eigenstate of the position operator $Q_H(t)$ at time t , corresponding to the eigenvalue q , and similarly $|p, t\rangle$ for the momentum operator $P_H(t)$ and eigenvalue p . For a small time interval Δt , we can approximate the scalar product by the following integral ¹:

$$\begin{aligned}\langle q', t + \Delta t | q, t \rangle &= \langle q', t | e^{-i\Delta t H(Q, P)} | q, t \rangle \simeq \langle q', t | (1 - i\Delta t H(Q, P)) | q, t \rangle \\ &= \int dp \langle q', t | (1 - i\Delta t H(Q, P)) | p, t \rangle \langle p, t | q, t \rangle = \int dp \langle q', t | (1 - i\Delta t H(q', p)) | p, t \rangle \langle p, t | q, t \rangle \\ &\simeq \int \frac{dp}{2\pi} e^{-i\Delta t H(q', p)\Delta t + ip(q' - q)},\end{aligned}\tag{1.1}$$

where we assumed a certain operator ordering in H such that $\langle q', t | H(Q, P) | p, t \rangle = H(q', p) \frac{e^{ipq'}}{\sqrt{2\pi}}$. For a finite time interval $t_f - t_i$, we proceed by dividing it into small chunks $t_f - t_i = N\Delta t$ and let $t_k = t_i + k\Delta t$ ($t_0 = t_i$, $t_N = t_f$). Repeating the previous process at each step yields the following formula, where $\prod dq_k \simeq \mathcal{D}q$ and $\prod \frac{dp_k}{2\pi} \simeq \mathcal{D}p$:

$$\begin{aligned}\langle q_f, t_f | q_i, t_i \rangle &= \int \prod dq_1 \dots dq_{N-1} \langle q_f, t_f | q_{N-1}, t_{N-1} \rangle \dots \langle q_1, t_1 | q_i, t_i \rangle \\ &\simeq \int \prod dq_1 \dots dq_{N-1} \prod \frac{dp_1}{2\pi} \dots \frac{dp_N}{2\pi} \exp \left\{ -i \sum_{k=1}^N H(q_k, p_k) \Delta t + i \sum_{k=1}^N p_k (q_k - q_{k-1}) \right\} \\ &\simeq \int_{q(t_i)=q_i, q(t_f)=q_f} \mathcal{D}q \mathcal{D}p \exp \left\{ -i \int_{t_i}^{t_f} d\tau H(q(\tau), p(\tau)) + i \int_{t_i}^{t_f} d\tau p(\tau) \dot{q}(\tau) \right\},\end{aligned}\tag{1.2}$$

¹Note that $H = H(Q_H(t), P_H(t)) = H(Q, P)$ does not depend on t .

where in the last step we have taken the limit $N \rightarrow \infty$, so that one integrates over all paths $q(t)$ and $p(t)$ in phase space with the prescribed initial and final conditions.

This can easily be generalized to obtain the matrix elements of the time-ordered product of operators $O_a(Q(t_a), P(t_a))$ constructed from the Heisenberg operators Q and P at times $\{t_a\}_a$:

$$\langle q_f, t_f | T \prod_a O_a(Q(t_a), P(t_a)) | q_i, t_i \rangle = \int_{q(t_i)=q_i}^{q(t_f)=q_f} \mathcal{D}q \mathcal{D}p \prod_a O_a(q(t_a), p(t_a)) e^{i \int_{t_i}^{t_f} d\tau (p(\tau) \dot{q}(\tau) - H(q(\tau), p(\tau)))} . \quad (1.3)$$

This formulation is called the Hamiltonian formulation. If the operators $\mathcal{O}_a$ do not depend on momenta and the Hamiltonian is at most quadratic in the latter, the integration over $\mathcal{D}p$ is Gaussian and can be easily performed. This is the only case we will consider from here on out. Thus, one gets:

$$\langle q', t' | T \prod_a O_a(Q(t_a)) | q, t \rangle = \mathcal{N} \int_{q(t_i)=q_i}^{q(t_f)=q_f} \mathcal{D}q \prod_a O_a(q(t_a)) (\text{Det } \hat{A})^{-1/2} e^{i \int_{t_i}^{t_f} d\tau (p_*(\tau) \dot{q}(\tau) - H(q(\tau), p_*(\tau)))} , \quad (1.4)$$

where $\hat{A}$ denotes the integral kernel of the Hamiltonian involving the momentum,² $p_*(t)$ the extremal argument of the exponent and $\mathcal{N}$ a potentially infinite normalization constant. Eventually, we will be interested in normalized quantities, i.e. quotients, and then this constant will drop out. It should be noted that performing this Gaussian integration corresponds to doing the Legendre transformation from the Hamiltonian to the Lagrangian, and yields the Lagrangian formulation of path integrals:

$$\langle q', t' | T \prod_a O_a(Q(t_a)) | q, t \rangle = \int_{q(t_i)=q_i, q(t_f)=q_f} \mathcal{D}q \prod_a \mathcal{O}_a(q(t_a)) (\text{Det } \hat{A})^{-1/2} e^{i \int_{t_i}^{t_f} d\tau L(q(\tau), \dot{q}(\tau))} . \quad (1.5)$$

If $\hat{A}$ is independent of Q , $\text{Det } \hat{A}$ will be absorbed into the normalization constant. If not, we will rewrite this quantity as $(\text{Det } \hat{A})^{-1/2} = \exp(-\frac{1}{2} \log \text{Det } \hat{A}) = \exp(-\frac{1}{2} \text{Tr } \log \hat{A})$ and consider it an additional contribution to the Lagrangian.

1.1.2 Functional integrals in QFT

In QFT, the quantities of interest are the vacuum expectation values of time-ordered products of field operators (or functions of the field operators) and not so much transition amplitudes between eigenstates of the field operators. This induces two major modifications in the integral formalism: prescribed initial and final conditions become obsolete, and one needs to take the “vacuum wave function” into account, which results in a purely imaginary, infinitesimally small contribution to the Lagrangian. These are the $i\epsilon$ -terms which will tell us that the propagators are the Feynman propagators.

When generalizing path integrals to functional integrals, we will start with the Lagrangian formulation and not preoccupy ourselves with its derivation from the Hamiltonian formulation. In particular, when dealing with (non-abelian) gauge theories, the necessary gauge-fixing and ghost terms will be added in when necessary. The Lagrangian (more precisely Lagrangian density) $\mathcal{L}$ at a given space-time point x depends only on the fields $\psi_l(x)$ and their derivatives $\partial_\mu \psi_l(x)$ at the same point. For brevity, we will simply denote it by $\mathcal{L}(\psi(x))$. For bosonic fields, this yields the following

²More precisely, if the eigenvalues of the integral kernel are a_k , one gets $\prod_k \frac{1}{2\pi} \sqrt{\frac{2\pi}{ia_k}} = (\text{Det } (2\pi i \hat{A}))^{-1/2}$. But we include all the numerical factors into the normalization constant $\mathcal{N}$ and simply write $(\text{Det } \hat{A})^{-1/2}$.

generic formula for expectation values between the in and out vacua ³ of time-ordered products of Heisenberg-picture operators $O_a(\Psi)$ (that only depend on the Heisenberg fields $\Psi_l(x)$ and not on their momenta $\Pi_l(x)$):

$$\begin{aligned} \langle \text{vac, out} | T \left\{ \prod_a O_a(\Psi(x_a)) \right\} | \text{vac, in} \rangle \\ = \mathcal{N} \int \prod_l \mathcal{D}\psi_l (\text{Det} \hat{\mathcal{A}}(\psi))^{-1/2} \prod_a O_a(\psi(x_a)) \exp \left\{ i \int d^4x \mathcal{L}(\psi(x)) \right\}, \end{aligned} \quad (1.6)$$

where the notation ψ (and Ψ) is used to indicate a dependence in all fields ψ_l (and Ψ_l), and $\hat{\mathcal{A}}$ is again the integral kernel of the term in the Hamiltonian that is quadratic in the momentum fields. In most theories, $\hat{\mathcal{A}}$ does not depend on the fields ψ , and its determinant is an irrelevant constant.

For anti-commuting Dirac fields ψ and $\bar{\psi}$, the analogue of the bosonic formula (1.6) becomes the following:

$$\begin{aligned} \langle \text{vac, out} | T \left\{ \prod_a O_a(\Psi(x_a), \bar{\Psi}(x_a)) \right\} | \text{vac, in} \rangle \\ = \mathcal{N} \int \prod_l \mathcal{D}\psi_l \mathcal{D}\bar{\psi}_l \prod_a O_a(\psi(x_a), \bar{\psi}(x_a)) \exp \left\{ i \left[\int d^4x \mathcal{L}(\bar{\psi}, \psi) \right] \right\}. \end{aligned} \quad (1.7)$$

Note that the normalization constant $\mathcal{N}$ depends on how exactly one defines the functional integral measure, but does not depend on which operator insertions are present.

1.2 Propagators

The free propagators are defined as:

$$-i\Delta_{lk}(x, y) = \langle \text{vac, out} | T(\Psi_l(x) \Psi_k(y)) | \text{vac, in} \rangle |_{\text{no interactions}}. \quad (1.8)$$

In a free theory, Heisenberg and interaction pictures coincide. Free propagators are determined by the free part of the action, i.e. the part of the Lagrangian density that is quadratic in the fields. Hence, the computation of the propagators reduces, yet again, to computing a Gaussian integral. In the bosonic case, the free action is of the form ⁴:

$$\int d^4x \mathcal{L}_0 = -\frac{1}{2} \int d^4x d^4y \sum_{l,l'} \psi_l(x) \mathcal{D}_{l,l'}(x, y) \psi_{l'}(y). \quad (1.9)$$

For a hermitian scalar field, $\mathcal{L}_0 = -\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{m^2}{2} \phi^2$, so that $\mathcal{D}(x, y) = (-\partial^\mu \partial_\mu + m^2) \delta^{(4)}(x - y)$. In the free theory the generic formula (1.6) becomes:

$$-i\Delta_{lk}(x, y) = \mathcal{N} \int \prod_{l'} \mathcal{D}\psi_{l'} \psi_l(x) \psi_k(y) \exp \left\{ i \int d^4x \mathcal{L}_0 \right\}. \quad (1.10)$$

³States with particle labels are hard to interpret in an interacting theory. This is why we define “in” and “out” states: these are eigenstates of the Hamiltonian of the interacting theory which asymptotically (as $t \rightarrow -\infty$ for in-states and $t \rightarrow +\infty$ for out-states) behave like a collection of non-interacting particles with the same labels. This “asymptotic behavior” is, of course, to be understood with regards to time-dependent operators as the states themselves are time-independent in the Heisenberg picture. We will define in and out states more precisely below in subsection 1.3.2.

⁴One should not confuse the kernel $\mathcal{D}(x, y)$ acting on ψ with the functional integral measure $\mathcal{D}\psi$.

In a free theory, the “in” and “out” states coincide and one has:

$$1 = \langle \text{vac, out} | \text{vac, in} \rangle |_{\text{no interactions}} = \mathcal{N} \int \prod_{l'} \mathcal{D}\psi_{l'} \exp \left\{ i \int d^4x \mathcal{L}_0 \right\} , \quad (1.11)$$

which allows us to eliminate the normalization constant by rewriting (1.10) as

$$-i\Delta_{lk}(x, y) = \frac{\int \prod_{l'} \mathcal{D}\psi_{l'} \psi_l(x) \psi_k(y) \exp \left\{ i \int d^4x \mathcal{L}_0 \right\}}{\int \prod_{l'} \mathcal{D}\psi_{l'} \exp \left\{ i \int d^4x \mathcal{L}_0 \right\}} . \quad (1.12)$$

Let us now introduce the generating functional for the free theory:

$$Z_0[J] = \int \prod_{l'} \mathcal{D}\psi_{l'} \exp \left\{ i \int d^4x [\mathcal{L}_0(x) + J_l(x) \psi^l(x)] \right\} . \quad (1.13)$$

One can then easily see that, in a free theory, the n -point function can be rewritten using the functional integral and derivative formalism, as well as the generating functional, as ⁵:

$$\begin{aligned} \langle \text{vac, out} | T(\Psi_{l_1}(x_1) \dots \Psi_{l_n}(x_n)) | \text{vac, in} \rangle \Big|_{\text{no interaction}} &= \frac{\int \prod_{l'} \mathcal{D}\psi_{l'} \psi_{l_1}(x_1) \dots \psi_{l_n}(x_n) \exp \left\{ i \int d^4x \mathcal{L}_0 \right\}}{\int \prod_{l'} \mathcal{D}\psi_{l'} \exp \left\{ i \int d^4x \mathcal{L}_0 \right\}} \\ &= (Z_0[0])^{-1} (-i)^n \frac{\delta}{\delta J_{l_1}(x_1)} \dots \frac{\delta}{\delta J_{l_n}(x_n)} Z_0[J] \Big|_{J=0} , \end{aligned} \quad (1.14)$$

One can then compute $Z_0[J]$ as a Gaussian integral since $\mathcal{L}_0$ is quadratic. This is done by “completing the square” in the exponent, and yields:

$$Z_0[J] = Z_0[0] \exp \left(\frac{i}{2} \int d^4x d^4y J_l(x) \mathcal{D}_{lk}^{-1}(x, y) J_k(y) \right) , \quad (1.15)$$

where $Z_0[0] = \left(\text{Det} \left[\frac{i\hat{\mathcal{D}}}{2\pi} \right] \right)^{-1/2}$. This formula yields a simple expression of the free propagator:

$$-i\Delta_{lk}(x, y) = (-i)^2 (Z_0[0])^{-1} \frac{\delta}{\delta J_l(x)} \frac{\delta}{\delta J_k(y)} Z_0[J] \Big|_{J=0} = -i\mathcal{D}_{lk}^{-1}(x, y) . \quad (1.16)$$

We can now use the Fourier transform to determine this expression explicitly. Translation invariance reduces the number of variables and allows us to write:

$$\mathcal{D}_{lk}(x, y) \equiv \mathcal{D}_{lk}(x - y) = \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \mathcal{D}_{lk}(p) \quad (1.17)$$

The inverse operator $(\mathcal{D}^{-1})_{lk}(x, y)$ is given by the Fourier transform of $(\mathcal{D}^{-1})_{lk}(p)$, which is the inverse matrix of $\mathcal{D}_{lk}(p)$:

$$\Delta_{lk}(x, y) \equiv \Delta_{lk}(x - y) = \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} (\mathcal{D}^{-1})_{lk}(p) . \quad (1.18)$$

For the example of the hermitian scalar field, with the replacement $m^2 \rightarrow m^2 - i\epsilon$, we get

$$\mathcal{D}(x, y) = (-\partial^\mu \partial_\mu + m^2 - i\epsilon) \delta^{(4)}(x - y) = \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} (p^2 + m^2 - i\epsilon) ,$$

⁵For anticommuting, i.e. fermionic fields one has to be careful about the ordering as they may differ by a sign.

with inverse

$$\Delta_{\text{scalar}}(x-y) = \int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} \frac{1}{p^2 + m^2 - i\epsilon} , \quad (1.19)$$

so that the momentum space propagator is

$$\frac{-i}{(2\pi)^4} \Delta_{\text{scalar}}(p) = \frac{-i}{(2\pi)^4} \frac{1}{p^2 + m^2 - i\epsilon} . \quad (1.20)$$

Let us note that for fermions, i.e. for anti-commuting fields, Gaussian functional integrals can be defined similarly as

$$\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp \left(\int (\bar{\psi} K \psi + \bar{\eta} \psi + \bar{\psi} \eta) \right) = \mathcal{N} \text{Det } K \exp \left(- \int \bar{\eta} K^{-1} \eta \right) , \quad (1.21)$$

which shows very analogously that the fermionic propagator is also given by the inverse of the kernel of the quadratic part of the Lagrangian, i.e. for Dirac fields by $(\not{\partial} + m - i\epsilon)^{-1}$, or more explicitly using the Fourier transform by

$$\begin{aligned} \langle 0 | T(\psi(x) \bar{\psi}(y)) | 0 \rangle = -i \Delta_{\text{Dirac}}(x-y) &= \int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} (i\not{p} + m - i\epsilon)^{-1} \\ &= -i \int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} \frac{-i\not{p} + m}{p^2 + m^2 - i\epsilon} , \end{aligned} \quad (1.22)$$

so that the momentum space propagator is

$$\frac{-i}{(2\pi)^4} \Delta_{\text{Dirac}}(p) = \frac{-i}{(2\pi)^4} (i\not{p} + m - i\epsilon)^{-1} = \frac{-i}{(2\pi)^4} \frac{-i\not{p} + m}{p^2 + m^2 - i\epsilon} . \quad (1.23)$$

Note that in the fermionic Gaussian integral (1.21) the determinant of the kernel K appears with a positive power.

1.3 Interactions and Feynman diagrams

We now go back to the full interacting theory and explain how to compute the vacuum expectation values (1.6) of time-ordered Heisenberg-picture operators. Without much loss of generality, we will assume that the products of the $O_a(\Psi(x_a))$ are simply products of the $\Psi(x_b)$ (and $\text{Det } \hat{\mathcal{A}} = \text{const}$). Then, much as in (1.14), these vacuum expectation values $\langle \text{vac}, \text{out} | T(\Psi_{l_1}(x_1) \dots \Psi_{l_p}(x_p)) | \text{vac}, \text{in} \rangle$ can be obtained by multiple functional derivatives with respect to a source function $J(x)$ of a generating functional

$$Z[J] = \int \mathcal{D}\psi \exp \left\{ i \int d^4 x [\mathcal{L}(x) + J_l(x) \psi^l(x)] \right\} , \quad (1.24)$$

where now the full Lagrangian is present in the exponent. Moreover, to eliminate the unknown normalization constant $\mathcal{N}$ we take again a ratio with $\langle \text{vac}, \text{out} | \text{vac}, \text{in} \rangle$. Note that, contrary to the free theory, the latter no longer equals one, since the in- and out-vacua now differ by a non-trivial phase. Then the “normalized p -point function” is given by

$$\frac{\langle \text{vac}, \text{out} | T(\Psi_{l_1}(x_1) \dots \Psi_{l_p}(x_p)) | \text{vac}, \text{in} \rangle}{\langle \text{vac}, \text{out} | \text{vac}, \text{in} \rangle} = (Z[0])^{-1} (-i)^p \frac{\delta}{\delta J_{l_1}(x_1)} \dots \frac{\delta}{\delta J_{l_p}(x_p)} Z[J] \Big|_{J=0} . \quad (1.25)$$

1.3.1 Feynman diagrams from functional integrals

In order to relate this quantity to the sum of Feynman diagrams, we would like to do a perturbative expansion in the coupling constant. Let us begin by separating the Lagrangian into its free - the quadratic - part, and the interaction Lagrangian: $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\text{int}}$ and proceed to develop $e^{i \int \mathcal{L}_{\text{int}}}$ in a Taylor series:

$$\begin{aligned}
Z[J] &= \int \mathcal{D}\psi \, e^{i \int d^4x \left(\mathcal{L}_0(x) + \mathcal{L}_{\text{int}}(x) + J_l(x) \psi^l(x) \right)} \\
&= \int \mathcal{D}\psi \sum_{n=0}^{\infty} \frac{i^n}{n!} \left[\int d^4x \, \mathcal{L}_{\text{int}}(\psi(x)) \right]^n e^{i \int d^4x \left(\mathcal{L}_0(x) + J_l(x) \psi^l(x) \right)} \\
&= \sum_{n=0}^{\infty} \frac{i^n}{n!} \left[\int d^4x \, \mathcal{L}_{\text{int}} \left(-i \frac{\delta}{\delta J(x)} \right) \right]^n \int \mathcal{D}\psi \, e^{i \int d^4x \left(\mathcal{L}_0(x) + J_l(x) \psi^l(x) \right)} \\
&= \sum_{n=0}^{\infty} \frac{i^n}{n!} \left[\int d^4x \, \mathcal{L}_{\text{int}} \left(-i \frac{\delta}{\delta J(x)} \right) \right]^n Z_0[J] , \tag{1.26}
\end{aligned}$$

where $Z_0[J]$ is the generating functional of the free theory as computed above in (1.15) with $\mathcal{D}_{lk}^{-1} \equiv \Delta_{lk}$:

$$Z_0[J] = Z_0[0] \exp \left(\frac{1}{2} \int d^4x \, d^4y \, (iJ_l(x)) (-i\Delta_{lk}(x, y)) (iJ_k(y)) \right) . \tag{1.27}$$

This expression allows us to give a diagrammatic interpretation of the generating functional $Z[J]$. Each functional derivative $-i\delta/\delta J_l(x)$ acting on $Z_0[J]$ yields a propagator $-i\Delta_{lk}(x, y)$ attached to a vertex at x and multiplied with a $J_k(y)$, integrated over y (which reads $\int d^4y (-i\Delta_{lk}(x, y)) (iJ_k(y))$). The number of propagators attached to each vertex is equal to the number of fields in the interaction Lagrangian, which translates into the number of functional derivatives $-i\delta/\delta J(x)$ applied to $Z_0[J]$. The number of vertices of a diagram coincides with the order of expansion n . Every propagator is either attached to a vertex or a $iJ_k(z_i)$. Thus, $Z[J]$ can be interpreted as the sum of all Feynman diagrams, with any number of external lines attached to a factor $iJ(z_i)$ (integrated d^4z_i), times $Z_0[J]$. In particular, the contribution from each vertex can be read from the “numerical factors” contained in the corresponding interaction Lagrangian. These factors typically include some coupling constant g , and the expansion then is an expansion in powers of g .

Furthermore, the additional functional derivatives in (1.25) act similarly, but replace a $J_i(z_i)$ by an external point x_i . Since in the end one is instructed to set $J = 0$, there cannot be left any propagators with J 's attached. This means that all propagators have their ends attached either to a vertex or to an external point. Very similarly, the denominator $Z[0]$ in (1.25) has all propagators connected to vertices with no external points. This equals one plus the contribution from all so-called vacuum bubble diagrams. Thus, the ratio (1.25) appears as the ratio of all Feynman diagrams with p external lines and of all Feynman diagrams without external lines (one plus all vacuum bubbles). Since the sum of all diagrams in the numerator can be organized as the product of (i) the sum of the diagrams without vacuum bubbles and (ii) one plus all vacuum bubbles, we see that these vacuum bubble diagrams cancel between the numerator and the denominator, and the normalized p -point function (1.25) is given exactly by the sum of all Feynman diagrams with p external propagators and containing *no* vacuum bubbles.

We should note that in (1.26) we have interchanged the order of the (convergent) sum for the exponential series and of the functional integration. In general this cannot be justified and in most

cases results in a divergent series. This means that $Z[J] \equiv Z[J](g)$ considered as a function of the coupling constant g is not Taylor expandable, and the corresponding series is a diverging one. However, in many theories, if g is small enough, the sum truncated at low order still yields very good approximations. Such series are known as asymptotic series.⁶ This is the case in particular in QED where perturbation theory at order e^9 (four loops) yields results for the magnetic moment of the muon that are in impressive agreement with the experimental values.

1.3.2 Relation with Dyson's formula and Wick's theorem

We will now re-derive the previous result from a somewhat more conventional perspective, using Dyson's formula and Wick's theorem. We begin with the relation between the Heisenberg-picture fields $\Psi_l(x)$ and the interaction-picture fields⁷ $\psi_l(x)$ which is

$$\Psi_l(t, \vec{x}) = e^{iHt} e^{-iH_0 t} \psi_l(t, \vec{x}) e^{iH_0 t} e^{-iHt} = \Omega(t) \psi_l(t, \vec{x}) \Omega^{-1}(t) \quad , \quad \Omega(t) \equiv e^{iHt} e^{-iH_0 t} . \quad (1.28)$$

The unitary operators $\Omega(t)$ provide the desired relation between the two pictures. They also allow us to define the in- and out-states. Here we only need the in- and out-vacuum states :

$$|\text{vac, in}\rangle = \lim_{\theta \rightarrow -\infty} \Omega(\theta) |0\rangle \quad , \quad |\text{vac, out}\rangle = \lim_{\theta \rightarrow +\infty} \Omega(\theta) |0\rangle . \quad (1.29)$$

Then, assuming $x_1^0 > x_2^0 > \dots x_p^0$ without loss of generality, we have

$$\begin{aligned} \langle \text{vac, out} | T \left\{ \prod_a \Psi_{l_a}(x_a) \right\} | \text{vac, in} \rangle &= \langle \text{vac, out} | \Psi_{l_1}(x_1) \Psi_{l_2}(x_2) \dots \Psi_{l_p}(x_p) | \text{vac, in} \rangle \\ &= \langle 0 | \Omega^{-1}(+\infty) \Omega(t_1) \psi_{l_1}(x_1) \Omega^{-1}(t_1) \dots \psi_{l_p}(t_p) \Omega^{-1}(t_p) \Omega(-\infty) | 0 \rangle \\ &= \langle 0 | \mathcal{U}(+\infty, t_1) \psi_{l_1}(x_1) \mathcal{U}(t_1, t_2) \psi_{l_2}(x_2) \dots \psi_{l_p}(x_p) \mathcal{U}(t_p, -\infty) | 0 \rangle , \end{aligned} \quad (1.30)$$

where $\mathcal{U}$ is the interaction picture evolution operator :

$$\mathcal{U}(t, t') = \Omega^{-1}(t) \Omega(t') = e^{iH_0 t} e^{-iH(t-t')} e^{-iH_0 t'} . \quad (1.31)$$

It satisfies a simple differential equation and initial condition

$$i \frac{d}{dt} \mathcal{U}(t, t') = e^{iH_0 t} (-H_0 + H) e^{iH(t-t')} e^{-iH_0 t'} = e^{iH_0 t} V e^{-iH_0 t} \mathcal{U}(t, t') = V_I(t) \mathcal{U}(t, t') , \quad \mathcal{U}(t, t) = \mathbf{1} , \quad (1.32)$$

$V_I(t)$ being the interaction Hamiltonian $H - H_0$ in the interaction-picture. In quantum field theory, V_I is given by a space-integral of a local interaction Hamiltonian density $V_I(t) = \int d^3x \mathcal{H}_{\text{int}}(t, \vec{x})$. The latter is given by a sum of products of interaction-picture fields (and possibly their derivatives). The solution of (1.32) is then given by the time-ordered exponential

$$\mathcal{U}(t, t') = T \exp \left(-i \int_{t'}^t d\tau V_I(\tau) \right) = T \exp \left(-i \int_{t'}^t d\tau \int d^3x \mathcal{H}_{\text{int}}(\tau, \vec{x}) \right) . \quad (1.33)$$

⁶Probably the best-known asymptotic series is Stirling's approximation for the Γ -function for large arguments.

⁷We use the same symbol ψ_l for the interaction picture fields as we used for the functional integration variables, which should hopefully not cause any confusion.

This evolution operator obeys the composition law $\mathcal{U}(t_1, t_2)\mathcal{U}(t_2, t_3) = \mathcal{U}(t_1, t_3)$. It is then easy to see that (1.30) can be very simply rewritten as

$$\langle \text{vac, out} | T \left\{ \prod_a \Psi_{l_a}(x_a) \right\} | \text{vac, in} \rangle = \langle 0 | T \left\{ \psi_{l_1}(x_1) \dots \psi_{l_p}(x_p) \exp \left(-i \int d^4 z \mathcal{H}_{\text{int}}(z) \right) \right\} | 0 \rangle, \quad (1.34)$$

where the integral in the exponent is over all space and all times (from $-\infty$ to $+\infty$).

This can be evaluated in perturbation theory by expanding the exponential into a Taylor series⁸

$$\begin{aligned} \langle \text{vac, out} | T \left\{ \prod_a \Psi_{l_a}(x_a) \right\} | \text{vac, in} \rangle \\ = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} T \left(\int d^4 z_1 \dots d^4 z_n \langle 0 | \psi_{l_1}(x_1) \dots \psi_{l_p}(x_p) \mathcal{H}_{\text{int}}(z_1) \dots \mathcal{H}_{\text{int}}(z_n) | 0 \rangle \right), \end{aligned} \quad (1.35)$$

and using Wick's theorem. Wick's theorem allows us to rewrite any product of interaction-picture fields in terms of a sum of terms with an arbitrary number of pairs of fields "contracted" and the remaining fields being in normal order. A normal-ordered expression of interaction-picture fields (i.e. free fields) has all its annihilation operators a to the right and all its creation operators $a^\dagger$ to the left, so that its vacuum expectation value vanishes (since $a|0\rangle = 0$ and $\langle 0|a^\dagger = 0$). As a result, the only terms that have a non-vanishing contribution are the ones with *all* fields contracted. Each contraction between a field at y_A and a field at y_B (the y 's being either a x_l or a z_l) is given by the time-ordered two-point function of interaction-picture fields. Since these interaction-picture fields evolve with the free Hamiltonian H_0 they are free fields and the two-point function

$$\langle 0 | T \psi_l(y_A) \psi_k(y_B) | 0 \rangle = -i \Delta_{lk}(y_A - y_B) \quad (1.36)$$

coincides with the propagator as determined above. These propagators are all connected either to an external point x_i or to an internal point z_j which are again the vertices of the corresponding Feynman diagram. In the simplest quantum field theories, the interaction Hamiltonian and interaction Lagrangian are simply related as $\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}}$, and then it is not too difficult to convince oneself that Wick's theorem applied to Dyson's formula (1.35) exactly generates the same Feynman diagrams as the expansion from the functional integral above (1.26).⁹

1.4 Generating functional for connected Green functions

The p -point Green function $\hat{G}_{(p)}(x_1, \dots, x_p)$ without vacuum-bubbles is a sum of connected and disconnected pieces. The *connected* p -point Green functions $G_{(p)}^C(x_1, \dots, x_p)$ can be recursively defined as : $G_{(1)}^C(x) \equiv \hat{G}_{(1)}(x)$ and $G_{(2)}^C(x_1, x_2) \equiv \hat{G}_{(2)}(x_1, x_2) - G_{(1)}^C(x_1)G_{(1)}^C(x_2)$, etc.¹⁰ $G_{(p)}^C(x_1, \dots, x_p)$ then is the sum of the corresponding connected Feynman diagrams.

⁸Note that we again interchange the order of the sum and of computing the matrix elements.

⁹If the interaction Hamiltonian contains derivatives of the fields, there are some subtleties, but we will not discuss this any further.

¹⁰For the general p -point function one needs to consider all non-trivial partitions $\mathcal{P}$ of $(x_1, \dots, x_p)$ and subtract $\sum_{\mathcal{P}} (\pm) G^C \dots G^C$ from $G_{(p)}(x_1, \dots, x_p)$, where the possible minus signs take into account the exchange of fermionic labels.

This recursive relation can also be rewritten as a relation between generating functionals. We denote $iW[J]$ the generating functional of connected Green functions:

$$iW[J] = iW[0] + \sum_{p=1}^{\infty} \frac{1}{p!} \int d^4x_1 \dots d^4x_p G_{(p)}^{C, l_1 \dots l_p}(x_1, \dots, x_p) iJ_{l_1}(x_1) \dots iJ_{l_p}(x_p) . \quad (1.37)$$

The term of order 0, namely $iW[0]$, corresponds to connected 0-point Green function, i.e. to connected vacuum-bubbles. It is worth noting that, for any $p \geq 1$, G_p^C cannot contain vacuum-bubbles.

One can show that this generating functional $W[J]$ of all connected Green's functions relates to the generating functional $Z[J]$ of all Green's functions in the fully interacting theory as follows:

$$Z[J] = \exp(iW[J]) . \quad (1.38)$$

1.5 Quantum effective action

We define the functional $\Gamma[\varphi]$ as the Legendre transform of $W[J]$. For simplicity, we suppress again all labels on the fields. First, let

$$\phi_J(x) \equiv \frac{\delta}{\delta J(x)} W[J] = -i \frac{\delta}{\delta J(x)} \log Z[J] = \frac{1}{Z[J]} \left(-i \frac{\delta}{\delta J(x)} Z[J] \right) . \quad (1.39)$$

The expression on the r.h.s. is similar to the one-point Green function without vacuum bubbles (which is the connected one-point function) $\hat{G}_{(1)}(x) \equiv G_{(1)}^C(x)$ but computed with $J \neq 0$, thus keeping the additional interaction terms ϕJ in the Lagrangian. We see that $\phi_J(x)$ is the connected one-point function in the presence of the additional interactions generated by the sources.

The relation $\phi_J(x) = \frac{\delta}{\delta J(x)} W[J]$ can be inverted to get $J(x)$ as a function of $\phi(x)$. We denote $j_\varphi(x)$ the function such that $\phi_J(x) = \varphi(x)$ if $J(x) = j_\varphi(x)$. We can now use φ as variable and define the Legendre transform of W as

$$\Gamma[\varphi] = W[j_\varphi] - \int d^4x \varphi(x) j_\varphi(x) . \quad (1.40)$$

It can be shown (cf the remark below) that this is the generating functional of one-particle-irreducible diagrams, also called 1PI, which are connected diagrams that do not become disconnected when a single line is cut. More specifically, if one develops $\Gamma[\varphi]$ in powers of the $\varphi(x_i)$, just as we developed $W[J]$ in powers of the $J(x_i)$, the coefficients $\Gamma_{(p)}(x_1, \dots, x_p)$ are the 1PI p -point functions which are given by the sum of all corresponding 1PI Feynman diagrams with p external points at x_i , $i = 1, \dots, p$.

$\Gamma[\varphi]$ also bears the name of *quantum effective action* because its tree level contribution coincides with the classical action. As a matter of fact, let us introduce a loop-counting parameter λ by replacing the action as $S \rightarrow \frac{1}{\lambda} S$ and the sources as $J \rightarrow \frac{1}{\lambda} J$. This means that all vertices are multiplied by $\frac{1}{\lambda}$ and all propagators by λ . External lines get an overall factor λ^0 since they get a factor λ from the propagator and a $\frac{1}{\lambda}$ from the sources. The overall factor of a diagram is λ^{I-V} , where I denotes the number of internal lines and V the number of vertices. We also denote L the number of loops and C the number of connected components. Using the diagrammatic identity

$$I - V = L - C , \quad (1.41)$$

the overall factor can also be written as λ^{L-C} . Since C is fixed, λ is a loop-counting parameter. We look at connected components, i.e. $C = 1$. Thus, taking the limit $\lambda \rightarrow 0$ amounts to isolating the tree level contribution. One can then write the following loop expansions

$$W[J, \lambda] = \sum_{L \geq 0} \lambda^{L-1} W_L[J] \quad , \quad \Gamma[\varphi, \lambda] = \sum_{L \geq 0} \lambda^{L-1} \Gamma_L[\varphi] \quad , \quad (1.42)$$

where $W_0[J]$ and $\Gamma_0[\varphi]$ correspond to the classical limit ¹¹ (tree-level contributions), and the $W_L[J]$ and $\Gamma_L[\varphi]$ for $L \geq 1$ to L -loop contributions, which can be interpreted as quantum corrections.

We will now proceed by approximating the generating functional $W[J, \lambda]$ in the classical limit $\lambda \rightarrow 0$ using the saddle point method. We have, cf (1.38),

$$\exp \left\{ iW[J, \lambda] \right\} = Z[J, \lambda] = \int \mathcal{D}\psi \exp \left\{ \frac{i}{\lambda} \left(S[\psi] + \int d^4x J(x) \psi(x) \right) \right\} \quad . \quad (1.43)$$

If we let $\lambda \rightarrow 0$, the integral is dominated ¹² by the ψ_J that solve $\frac{\delta S}{\delta \psi} + J = 0$. It follows that, at tree level,

$$W_0[J] = S[\psi_J] + \int d^4x J(x) \psi_J(x) \quad , \quad (1.44)$$

is the Legendre transform of the classical action. On the other hand, since $W[J]$ is the inverse Legendre transform of $\Gamma[\varphi]$, we conclude that in the classical limit the quantum effective action equals the classical action :

$$\Gamma_0[\varphi] = S[\varphi] \quad . \quad (1.45)$$

Since the $\Gamma_L[\varphi]$ for $L \geq 1$ correspond to loop corrections, Γ represents indeed a (quantum corrected or) quantum effective action.

One can show [7, 23] by a similar saddle-point argument, that doing the functional integral, with $S[\psi]$ replaced by $\Gamma[\varphi]$, generates *at tree-level* the full generating functional $Z[J] = e^{iW[J]}$. This shows that all Feynman diagrams are generated as tree diagrams with (full) propagators that follow from the quadratic part of Γ and with effective vertices that follow from the non-quadratic parts of Γ . Since a given diagram is obtained in only one way from 1PI-vertices and full propagators this means that these effective vertices are 1PI, and this is why Γ is the generating functional of 1PI diagrams. We will not go into further detail.

1.6 Loop diagrams: regularization and renormalization

We have shown above how to obtain the Feynman rules to compute a generic n -point Green function $\widehat{G}_{(n)}(x_1, \dots, x_n)$ without vacuum bubbles, or the corresponding connected n -point function $G_{(n)}^C(x_1, \dots, x_n)$. If one further restricts to 1PI n -point diagrams, one obtains the 1PI n -point function $\Gamma_{(n)}(x_1, \dots, x_n)$. In practice one is more interested in the (four-dimensional) Fourier transforms, e.g.

$$\Gamma_{(n)}(p_1, \dots, p_n) = \int d^4x_1 \dots d^4x_n \exp \left(i \sum_{r=1}^n (p_r)_\mu x_r^\mu \right) \Gamma_{(n)}(x_1, \dots, x_n) \quad , \quad (1.46)$$

¹¹If we did not use units such that $\hbar = 1$, we would have $e^{\frac{i}{\hbar} S}$ and $\hbar$ would act as the loop-counting parameter λ . This is why $\lambda \rightarrow 0$ is the classical limit.

¹²For $\lambda \rightarrow 0$ we have the saddle-point approximation $e^{iW_0/\lambda} = (\text{Det } \frac{i\mathcal{D}}{2\pi})^{-1/2} e^{i(S(\psi_J) + \int J\psi_J)/\lambda}$, or taking the logarithm, $\frac{W_0}{\lambda} = \frac{1}{\lambda} (S(\psi_J) + \int J\psi_J) + \frac{i}{2} \log \text{Det } \frac{i\mathcal{D}}{2\pi}$. As expected, the determinant only contributes at order λ^0 which corresponds to one loop.

where all momenta are counted as flowing into the diagram. It follows from invariance under translations $x_i^\mu \rightarrow x_i^\mu + a^\mu$ that one of the integrations simply yields a $\delta^{(4)}(\sum p_i)$ of overall 4-momentum conservation.

Tree diagrams are always well-defined finite expressions of the external momenta p_r . In particular, all momenta carried by the external *and* internal propagators are perfectly determined by four-momentum conservation at each vertex.

On the other hand, for a diagram containing L loops, I internal lines and V vertices, one has I internal momenta, constrained by conservation at V vertices. One of these constraints only reproduces the overall conservation of the external momenta, and hence there are $I - (V - 1) = L$ unconstrained momenta : there is exactly one unconstrained momentum per loop, also referred to as the loop momenta. The integrations over these loop four-momenta k_i^μ may be divergent at large values of $|k_i^\mu|$. For example, if there is a single loop containing N internal scalar lines, i.e. N scalar propagators, and if the vertices contain no momentum factors (interactions without derivatives) then the loop integrand will behave for large k as $d^4k (k^2)^{-N}$. There is a standard procedure to replace k^0 by ik_E^0 , called the Wick rotation, which makes the loop momentum a Euclidean 4-momentum and provides a factor of i from the change of variables. For large Euclidean k one can then use 4-dimensional spherical coordinates, so that $d^4k = k^3 dk d^3\Omega$ and the integrand behaves as $dk k^{3-2N}$. One sees that this is logarithmically divergent for $N = 2$ and convergent for $N \geq 3$. On the other hand, if the N internal lines are fermionic lines with Dirac propagators that behave as $\frac{1}{k}$, the large k behavior now is $d^4k k^{-N} \sim dk k^{3-N}$. This is then divergent for $N \leq 4$ and convergent for $N \geq 5$.

These divergent integrals must be regularized. There are various ways to proceed. One may artificially “improve” the large momentum behavior of the propagators in one way or the other. The Pauli-Villars regularization to be discussed in section 4 essentially does this. Another very convenient regularization is the dimensional regularization which formally continues the integrand to an arbitrary complex value d of the dimension. The integral then appears as a meromorphic function of the space-time dimension $d = 4 - \epsilon$ which has a pole at $d = 4$ (if the original integral was divergent), i.e. at $\epsilon = 0$. Unfortunately, dimensional regularization does not work in the presence of structures that are specific to 4 dimensions, like the $\epsilon^{\mu\nu\rho\sigma}$ -tensor or the Dirac matrix γ_5 .

Of course, once one removes the regulator, the loop expressions will diverge again. But in general, one has a clear separation between the diverging piece, i.e. the polar part $\sim \frac{1}{\epsilon^r}$, $r = 1, 2, \dots$, and the finite part. One can then add the so-called counterterms to the Lagrangian that are of higher order in the coupling constant(s) and that are $\sim \frac{1}{\epsilon^r}$ with exactly the right coefficients, so that they produce a tree-level vertex that cancels the diverging piece of the corresponding loop diagram. Allowed counterterms must be local expressions in the fields involving finitely many derivatives, corresponding to counterterm vertices that are polynomial in the momenta. It is thus important that the divergences from the loop diagrams have a corresponding structure. Instead of adding these counterterms “by hand”, one can also consider that the original Lagrangian was made from so-called bare fields and bare masses and bare coupling constants, and then split this Lagrangian into a finite Lagrangian plus these same counterterms. Rather than discussing the procedure in general, we will work out an example in the next subsection.

1.7 An example : vacuum polarization in QED

The QED Lagrangian is given by

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{L}_{\text{matter}} + \mathcal{L}_{\text{gauge fixing}} . \quad (1.47)$$

Of course, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $\mathcal{L}_{\text{matter}} = -\bar{\psi}(\not{D} + m)\psi$ with $D_\mu = \partial_\mu + ieA_\mu$ being the covariant derivative.¹³ The gauge-fixing Lagrangian can be chosen to be $\mathcal{L}_{\text{gauge fixing}} = -\frac{\alpha}{2}(\partial_\mu A^\mu)^2$, but this is not important for the present computation, as we will not need the photon propagator.

It is interesting to briefly discuss the simplest one-loop diagram in QED which is the vacuum polarization diagram. This is the 1PI two photon function¹⁴ $\Pi_{\mu\nu}(q)$. At one loop, there is a single Feynman diagram that contributes at order e^2 as shown in Fig. 1. The Feynman rule for the Dirac

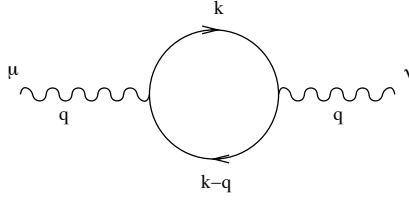


Figure 1: The one-loop vacuum polarization diagram in QED

propagator was already given in (1.23). The vertex follows from the expression of the covariant derivative in $\mathcal{L}_{\text{matter}} = -\bar{\psi}(\not{D} + m)\psi + \mathcal{L}_{\text{int}}$ with $\mathcal{L}_{\text{int}} = -ie\bar{\psi}\gamma^\mu\psi A_\mu$. Thus, the electron-electron-photon vertex contributes $i(2\pi)^4 \times (-ie\gamma^\mu) = (2\pi)^4 e\gamma^\mu$. Then we get for this diagram

$$i(2\pi)^4 \Pi_{\text{loop}, e^2}^{\mu\nu}(q) = - \int d^4k \text{tr} \left\{ \frac{(-i)}{(2\pi)^4} \frac{-i\not{k} + m}{k^2 + m^2 - i\epsilon} (2\pi)^4 e\gamma^\mu \frac{(-i)}{(2\pi)^4} \frac{-i(\not{k} - \not{q}) + m}{(k - q)^2 + m^2 - i\epsilon} (2\pi)^4 e\gamma^\nu \right\} , \quad (1.48)$$

where the overall minus sign is due to the fermion loop.

Before doing the actual computation of this diagram, let us comment on a few properties. First of all, one sees that for large loop-momentum k the integrand behaves as $d^4k k^{-2}$, so that one expects this diagram to diverge quadratically, i.e. as Λ^2 , if one cuts off the loop-momentum at $|k| \leq \Lambda$. One also says it has superficial degree of divergence 2. One must then regularize this expression in one way or another. For example, in dimensional regularization to be discussed below, we let $d = 4 - \epsilon$, so that for a finite regulator $\epsilon \neq 0$ the integral is actually finite. One can then split this expression into a diverging part which is $\sim \frac{1}{\epsilon}$ and a finite part, which has a finite limit when $\epsilon \rightarrow 0$. If one takes a derivative of this regularized integral (1.48) with respect to the external momentum q this changes the large-momentum behavior of the integrand into $d^4k k^{-3}$, so that the superficial degree of divergence has been lowered from 2 to 1. Taking 2 more derivatives yields a degree of divergence -1 , i.e. a convergent integral. But this means that 3 derivatives have annihilated the diverging part, so the latter must be a polynomial of degree 2 in the external momentum. Indeed, the explicit computation below will yield

$$\Pi_{\text{loop}, e^2}^{\mu\nu}(q) = (q^2 \eta^{\mu\nu} - q^\mu q^\nu) \left(-\frac{e^2}{6\pi^2} \frac{1}{\epsilon} + f(q^2) + \mathcal{O}(\epsilon) \right) , \quad (1.49)$$

¹³The covariant derivative is $D_\mu = \partial_\mu - iqA_\mu$. But for the electron, $q = -e$, where $e > 0$ is the elementary charge.

¹⁴Somewhat more precisely, the 1PI 2-point function is $\Gamma_{(2)}^{\mu\nu}(q, q') = \delta^{(4)}(q - q') \Pi^{\mu\nu}(q)$.

where $f(q^2)$ is some finite, non-polynomial function of q^2 (and m^2 and some additional scale one needs to introduce in dimensional regularization to keep the coupling constant e dimensionless). The overall factor $(q^2\eta^{\mu\nu} - q^\mu q^\nu)$ actually is dictated by gauge-invariance which implies that one must have $q_\mu \Pi^{\mu\nu}(q) = 0 = q_\nu \Pi^{\mu\nu}(q)$.

As mentioned above, the diverging piece can be cancelled by a corresponding (polynomial) counterterm which one can either add to the initial Lagrangian “by hand” or more elegantly obtain in the following way. One considers that the fields appearing in the Lagrangian are so-called bare fields A_B^μ and ψ_B , and that the coupling constant and mass also are a bare coupling e_B and bare mass m_B . The bare fields, bare couplings and bare masses are related to the renormalized A^μ , ψ , e and m by so-called renormalization constants Z_2 and Z_3 (and δm)

$$\psi_B = \sqrt{Z_2}\psi \quad , \quad A_B^\mu = \sqrt{Z_3}A^\mu \quad , \quad e_B = \frac{1}{\sqrt{Z_3}}e \quad , \quad m_B = m - \delta m \quad . \quad (1.50)$$

In the absence of interactions, there is no need to renormalize and then the Z factors are all 1 and $\delta m = 0$. Hence $(Z_2 - 1)$, $(Z_3 - 1)$ and δm must all be at least of order e . Actually, they are power series in e^2 with the lowest order being of order e^2 . The fact that the Z -factors of A^μ and e are opposite implies that the covariant derivative $\partial_\mu + ie_B A_B^\mu = \partial_\mu + ie A_\mu$ remains unchanged which is again a consequence of gauge invariance. In particular, we can then rewrite the original Lagrangian as ($F_{\mu\nu}^B = \partial_\mu A_\nu^B - \partial_\nu A_\mu^B$)

$$-\frac{1}{4}F_{\mu\nu}^B F_B^{\mu\nu} - \bar{\psi}_B(\gamma^\mu(\partial_\mu + ie_B A_B^\mu) + m_B)\psi_B = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \bar{\psi}(\gamma^\mu(\partial_\mu + ie A_\mu) + m)\psi + \mathcal{L}_{\text{ct}} \quad , \quad (1.51)$$

where the counterterm Lagrangian is

$$\mathcal{L}_{\text{ct}} = -\frac{1}{4}(Z_3 - 1)F_{\mu\nu}F^{\mu\nu} - (Z_2 - 1)\bar{\psi}(\not{\partial} + m)\psi + Z_2\delta m\bar{\psi}\psi - (Z_2 - 1)ie\bar{\psi}\not{A}\psi \quad . \quad (1.52)$$

While this looks like a trivial rewriting of the Lagrangian, the point is to consider the terms on the right-hand side of (1.51) which are not in $\mathcal{L}_{\text{ct}}$ as the “classical” Lagrangian which defines the Feynman rules, and consider $\mathcal{L}_{\text{ct}}$ as a higher-order term that provides extra interactions, called the counterterms. These extra interactions may correspond to vertices with only two legs as is the case for the first three terms in $\mathcal{L}_{\text{ct}}$, or a vertex with three legs as is the case for the last term. In particular, one easily figures out that the term $-\frac{1}{4}(Z_3 - 1)F_{\mu\nu}F^{\mu\nu}$ contributes

$$\Pi_{\text{ct},e^2}^{\mu\nu} = -(Z_3 - 1)(q^2\eta^{\mu\nu} - q^\mu q^\nu) \quad (1.53)$$

to the vacuum polarization at order e^2 .

The full order e^2 contribution to the vacuum polarization then is

$$\Pi_{e^2}^{\mu\nu}(q) = \Pi_{\text{loop},e^2}^{\mu\nu}(q) + \Pi_{\text{ct},e^2}^{\mu\nu} = (q^2\eta^{\mu\nu} - q^\mu q^\nu) \left(-\frac{e^2}{6\pi^2\epsilon} + f(q^2) - (Z_3 - 1) \right) \quad . \quad (1.54)$$

The a priori unknown constant Z_3 is fixed by some physical requirement, referred to as a renormalization condition. If we require that the photon propagator has its pole at $q^2 = 0$ so that the photon remains massless, and that the residue at the pole has the same form as expected at tree level, then the term multiplying the $(q^2\eta^{\mu\nu} - q^\mu q^\nu)$ must vanish at $q^2 = 0$. This fixes

$$Z_3 - 1 = -\frac{e^2}{6\pi^2\epsilon} + f(0) \quad , \quad (1.55)$$

so that

$$\Pi_{e^2}^{\mu\nu}(q) = (q^2 \eta^{\mu\nu} - q^\mu q^\nu) \left(f(q^2) - f(0) \right), \quad (1.56)$$

which is now a well-defined finite function of the momentum.

Let us now explicitly do the computation of the loop diagram. We choose dimensional regularization. This means that we work with Feynman rules that would follow from the corresponding d -dimensional Lagrangian. In particular all momenta become d -dimensional, and so does the loop integration. The non-trivial fact is that one may “analytically continue” to complex values of d with the integrals being meromorphic functions of d with poles at various integer values of d . While in 4 dimensions the coupling e is a dimensionless quantity, this is not so in arbitrary d . Indeed, in d dimensions, the Dirac fields have engineering dimension $\frac{d-1}{2}$ while the photon field has dimension $\frac{d-2}{2}$ (just as scalars). Then the interaction $\bar{\psi}\gamma^\mu\psi A_\mu$ has dimension $\frac{3}{2}d - 2 = d + \frac{d-4}{2}$ and only if we replace the coupling e by $e\tilde{\mu}^{(4-d)/2}$ the coupling e is still dimensionless. This means that the vacuum polarization gets an additional factor $\tilde{\mu}^{4-d}$.

Once we have dimensionally regularized our integral, it is a finite (converging) expression, and we can perform the usual manipulations such as shifting the integration variable. Applying Feynman’s formula $\frac{1}{AB} = \int_0^1 dx \frac{1}{[xA + (1-x)B]^2}$ allows to rewrite

$$\Pi_{\text{loop}, e^2}^{\mu\nu}(q) = \frac{-i e^2}{(2\pi)^4} \tilde{\mu}^{4-d} \int_0^1 dx \int d^d k \frac{\text{tr} [\dots]}{[(k-xq)^2 + m^2 + x(1-x)q^2 - i\epsilon]^2}. \quad (1.57)$$

The Dirac trace can still be evaluated as in 4 dimensions¹⁵ so that

$$\text{tr} [\dots] = 4 \left[-k^\mu (k-q)^\nu - k^\nu (k-q)^\mu + k(k-q)\eta^{\mu\nu} + m^2 \eta^{\mu\nu} \right]. \quad (1.58)$$

Shifting the integration variable $k - xq \rightarrow k$ this becomes

$$\begin{aligned} \Pi_{\text{loop}, e^2}^{*\mu\nu}(q) = \frac{-4i e^2}{(2\pi)^4} \tilde{\mu}^{4-d} \int_0^1 dx \int d^d k & \left[-(k+xq)^\mu (k-(1-x)q)^\nu - (k+xq)^\nu (k-(1-x)q)^\mu \right. \\ & \left. + (k+xq)(k-(1-x)q)\eta^{\mu\nu} + m^2 \eta^{\mu\nu} \right] \frac{1}{[k^2 + m^2 + x(1-x)q^2]^2}. \end{aligned} \quad (1.59)$$

Now one can do the Wick rotation which amounts to the change of variables $k^0 = ik_E^0$, so that $k^\mu k_\mu = k_E^2$ (recall that our Lorentz signature is $(-+++)$) and $d^d k = id^d k_E$. In order for this to make sense one must also continue the external q^μ to a Euclidean q_E^μ . Furthermore, due to rotational symmetry, terms in the numerator that are linear in k^μ don’t contribute to the integral, while terms $k^\mu k^\nu$ can be replaced by $\frac{1}{d} \eta^{\mu\nu} k^2$ which become $\frac{1}{d} \eta^{\mu\nu} k_E^2$ after the continuation. Hence

$$\Pi_{\text{loop}, e^2}^{\mu\nu}(q_E) = \frac{-4i e^2}{(2\pi)^4} \tilde{\mu}^{4-d} \int_0^1 dx \int d^d k_E \frac{[m^2 - x(1-x)q_E^2 + (1-\frac{2}{d})k_E^2] \eta^{\mu\nu} + 2x(1-x)q_E^\mu q_E^\nu}{[k_E^2 + m^2 + x(1-x)q_E^2]^2}. \quad (1.60)$$

Next we use the formula

$$\int \frac{d^d k_E (k_E^2)^s}{(k_E^2 + R^2)^N} = \pi^{d/2} \frac{\Gamma(\frac{d}{2} + s) \Gamma(N - \frac{d}{2} - s)}{\Gamma(\frac{d}{2}) \Gamma(N)} (R^2)^{\frac{d}{2} + s - N}, \quad (1.61)$$

for $s = 0$ and $s = 1$ with $R^2 = m^2 + x(1-x)q_E^2$. This yields

$$\begin{aligned} \Pi_{\text{loop}, e^2}^{\mu\nu}(q_E) = \frac{4e^2}{(2\pi)^4} \tilde{\mu}^{4-d} \int_0^1 dx & \left\{ \left[[m^2 - x(1-x)q_E^2] \eta^{\mu\nu} + 2x(1-x)q_E^\mu q_E^\nu \right] \pi^{d/2} \Gamma\left(2 - \frac{d}{2}\right) (R^2)^{\frac{d}{2}-2} \right. \\ & \left. + \left(1 - \frac{2}{d}\right) \eta^{\mu\nu} \pi^{d/2} \frac{\Gamma(\frac{d}{2} + 1) \Gamma(1 - \frac{d}{2})}{\Gamma(\frac{d}{2})} (R^2)^{\frac{d}{2}-1} \right\}. \end{aligned} \quad (1.62)$$

¹⁵All one needs is that $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \mathbf{1}$ and the only issue concerns $\text{tr} \mathbf{1}$ which can be chosen to be $2^{d/2} = 4 \times 2^{-\epsilon/2}$ or just 4. The ambiguity consisting in an overall $2^{-\epsilon/2}$ is of the same type as the ambiguity in choosing to continue also the factors $\frac{1}{(2\pi)^4}$ to $\frac{1}{(2\pi)^d}$. In the end this only changes the renormalization constants Z_i by finite amounts and does not affect the renormalized Π^* or Σ^* .

Observing that $(1 - \frac{2}{d}) \frac{\Gamma(\frac{d}{2}+1)\Gamma(1-\frac{d}{2})}{\Gamma(\frac{d}{2})} = (1 - \frac{2}{d}) \frac{d}{2} \Gamma(1 - \frac{d}{2}) = (\frac{d}{2} - 1)\Gamma(1 - \frac{d}{2}) = -\Gamma(2 - \frac{d}{2})$ we get

$$\begin{aligned}\Pi_{\text{loop},e^2}^{\mu\nu}(q_E) &= \frac{4e^2}{(2\pi)^4} \Gamma\left(2 - \frac{d}{2}\right) \pi^{d/2} \int_0^1 dx \left[\underbrace{[m^2 - x(1-x)q_E^2 - R^2]}_{-2x(1-x)q_E^2} \eta^{\mu\nu} + 2x(1-x)q_E^\mu q_E^\nu \right] \left(\frac{R^2}{\tilde{\mu}^2}\right)^{\frac{d}{2}-2} \\ &= \frac{8e^2}{(2\pi)^4} \Gamma\left(2 - \frac{d}{2}\right) \pi^{d/2} (q_E^\mu q_E^\nu - q_E^2 \eta^{\mu\nu}) \int_0^1 dx x(1-x) \left[\frac{m^2 + x(1-x)q_E^2}{\tilde{\mu}^2} \right]^{\frac{d}{2}-2} .\end{aligned}\quad (1.63)$$

Next one sets $d = 4 - \epsilon$ and expands the result in ϵ . Recall that

$$\Gamma\left(2 - \frac{d}{2}\right) = \Gamma\left(\frac{\epsilon}{2}\right) = \frac{2}{\epsilon} - \gamma + \mathcal{O}(\epsilon) , \quad (1.64)$$

and also $a^\epsilon = e^{\epsilon \log a} = 1 + \epsilon \log a + \mathcal{O}(\epsilon^2)$, so that

$$\begin{aligned}\Gamma\left(2 - \frac{d}{2}\right) \pi^{d/2-2} [\dots]^{\frac{d}{2}-2} &= \left(\frac{2}{\epsilon} - \gamma + \mathcal{O}(\epsilon)\right) \left(1 - \frac{\epsilon}{2} \log \pi - \frac{\epsilon}{2} \log[\dots] + \mathcal{O}(\epsilon^2)\right) \\ &= \frac{2}{\epsilon} - \gamma - \log \pi - \log[\dots] + \mathcal{O}(\epsilon) .\end{aligned}\quad (1.65)$$

Since $\int_0^1 dx x(1-x) = \frac{1}{6}$, and reverting back to Minkowski signature, we finally get

$$\Pi_{\text{loop},e^2}^{\mu\nu}(q^2) = (q^2 \eta^{\mu\nu} - q^\mu q^\nu) \frac{e^2}{6\pi^2} \left(-\frac{1}{\epsilon} + \frac{\gamma + \log \pi}{2} + 3 \int_0^1 dx x(1-x) \log \left[\frac{m^2 + x(1-x)q^2}{\tilde{\mu}^2} \right] + \mathcal{O}(\epsilon) \right) . \quad (1.66)$$

We indeed see the diverging term $(q^2 \eta^{\mu\nu} - q^\mu q^\nu) \left(-\frac{e^2}{6\pi^2} \frac{1}{\epsilon} \right)$, as anticipated in eq. (1.49), and we can read off the finite function $f(q^2)$. In particular,

$$f(q^2) - f(0) = \frac{e^2}{2\pi^2} \int_0^1 dx x(1-x) \log \left[1 + x(1-x) \frac{q^2}{m^2} \right] . \quad (1.67)$$

1.8 Exercises

1.8.1 Linear operators on $L^2(\mathbf{R}^4)$

Quite often one has to manipulate linear operators $\hat{A}$ acting on functions $f \in L^2(\mathbf{R}^4)$. These are *not* operators acting in the Hilbert space of quantum field theory but in a “much smaller” Hilbert space. A typical example are differential operators. In general one can write

$$(\hat{A}f)(x) = \int d^4y A(x,y) f(y) . \quad (1.68)$$

It is very convenient to use the notation $A(x,y) = \langle x | \hat{A} | y \rangle$ where $|x\rangle$ is the ket of the non-normalisable basis of $L^2(\mathbf{R}^4)$ such that $\hat{X}^\mu |x\rangle = x^\mu |x\rangle$ (eigenstate of the position operator). One also uses the non-normalisable eigenstates $|p\rangle$ of $\hat{P}^\mu$. Recall that

$$\mathbf{1} = \int d^4x |x\rangle \langle x| = \int d^4p |p\rangle \langle p| \quad , \quad \langle x | p \rangle = \frac{e^{ipx}}{(2\pi)^2} . \quad (1.69)$$

a) Show that an n th order differential operator such that $(\hat{A}f)(x) = \frac{\partial}{\partial x^{\mu_1}} \dots \frac{\partial}{\partial x^{\mu_n}} f(x)$, corresponds to $A(x,y) = (-)^n \frac{\partial}{\partial y^{\mu_1}} \dots \frac{\partial}{\partial y^{\mu_n}} \delta^{(4)}(x-y) = \frac{\partial}{\partial x^{\mu_1}} \dots \frac{\partial}{\partial x^{\mu_n}} \delta^{(4)}(x-y)$. Show that for this same operator one also has $A(p,\tilde{p}) \equiv \langle p | \hat{A} | \tilde{p} \rangle = i^n p_{\mu_1} \dots p_{\mu_n} \delta^{(4)}(p - \tilde{p})$.

b) Show that for any $\hat{A}$, the matrix elements $A^{-1}(x, y) \equiv \langle x | \hat{A}^{-1} | y \rangle$ of the inverse operator satisfy

$$\int d^4 z A^{-1}(x, z) A(z, y) = \delta^{(4)}(x - y) . \quad (1.70)$$

Write the analogous relation for $A^{-1}(p, \tilde{p}) = \langle p | \hat{A}^{-1} | \tilde{p} \rangle$ and $A(\tilde{p}, p')$.

c) Show that if $A(x, y)$ only depends on the difference $x - y$ (translation invariance), then

$$\langle p | \hat{A} | \tilde{p} \rangle = A(p) \delta^{(4)}(p - \tilde{p}) \quad \text{with} \quad A(x, y) = \int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} A(p) , \quad (1.71)$$

and that in this case the matrix elements of the inverse operator simply are

$$\langle p | \hat{A}^{-1} | \tilde{p} \rangle = \frac{1}{A(p)} \delta^{(4)}(p - \tilde{p}) \quad \Rightarrow \quad A^{-1}(x, y) = \int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} \frac{1}{A(p)} . \quad (1.72)$$

More generally, show that for such operators one has for a “general” function F of $\hat{A}$:

$$\langle x | F(\hat{A}) | y \rangle = \int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} F(A(p)) . \quad (1.73)$$

d) Use these results to show that the inverse of the Feynman propagator of a scalar field $A(x, y) = \Delta_F(x - y)$ is given by $A^{-1}(x, y) = \left[-\frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x_\mu} + m^2 - i\epsilon \right] \delta^{(4)}(x - y)$. Deduce the differential equation satisfied by the Feynman propagator.

f) We will also have to compute traces of certain operators on $L^2(\mathbf{R}^4)$. By definition

$$\text{Tr } \hat{A} = \int d^4 x \langle x | \hat{A} | x \rangle = \int d^4 x A(x, x) . \quad (1.74)$$

Explicitly show that this is equivalent to $\text{Tr } \hat{A} = \int d^4 p \langle p | \hat{A} | p \rangle = \int d^4 p A(p, p)$. Show that for an operator invariant under translations (cf (1.71)) formally one has

$$\text{Tr } F(\hat{A}) = \delta^{(4)}(0) \int d^4 p F(A(p)) . \quad (1.75)$$

Interpret $\delta^{(4)}(0)$ as $\langle p | p \rangle$ and argue that if one replaces $\mathbf{R}^4$ by a hypercube of volume L^4 , this factor would be replaced according to $\delta^{(4)}(0) \rightarrow \left(\frac{L}{2\pi}\right)^4$.

1.8.2 Green's function and generating functional in 0 + 1 dimensions

Consider the 0 + 1 dimensional quantum “field” theory (this is actually quantum mechanics) with action

$$S_0[\phi] = \int_{-\infty}^{\infty} dt \frac{1}{2} \left[\dot{\phi}^2 - \omega^2 \phi^2 \right] , \quad (1.76)$$

where $\dot{\phi} = \frac{\partial \phi}{\partial t} \equiv \partial_t \phi$.

a) Determine the corresponding Green's functions $G(t)$ as the solutions of

$$(\partial_t^2 + \omega^2) G(t) = \delta(t) , \quad (1.77)$$

by first solving this equation for all $t \neq 0$ and then matching solutions to get the correct discontinuity at $t = 0$ to reproduce the δ -singularity.

b) Obtain Feynman's Green's function from the Fourier transform with the $i\epsilon$ prescription:

$$G_F(t) = \int_{-\infty}^{\infty} \frac{dp^0}{2\pi} \frac{e^{-ip^0 t}}{-(p^0)^2 + \omega^2 - i\epsilon} , \quad (1.78)$$

and compare with the results from a).

c) Define the generating functional as

$$Z_0[J] = \int \mathcal{D}\phi \exp \left[iS_0[\phi] + i \int dt J(t)\phi(t) \right] \quad (1.79)$$

and compute $Z_0[J]/Z_0[0]$. Explicitly convince yourself that

$$\frac{1}{Z_0[0]} \frac{\delta}{\delta J(t_1)} \cdots \frac{\delta}{\delta J(t_4)} Z_0[J] \Big|_{J=0} \quad (1.80)$$

gives the functional integral representation of the (free) four-point function $\langle T[\phi(t_1) \dots \phi(t_4)] \rangle_0$.

d) Now add to the action S_0 an interaction term $S = S_0 + S_{\text{int}}$ with $S_{\text{int}} = \int_{-\infty}^{\infty} dt \frac{g}{6} \phi^3$. Define $Z[J]$ as $Z_0[J]$ but with S_0 replaced by S . Compute $Z[J]$ and $Z[0]$ to second order in g . Give the graphical representation of the terms you find.

e) Deduce the full 2-point function as

$$\langle T[\phi(t_1)\phi(t_2)] \rangle = -\frac{1}{Z[0]} \frac{\delta}{\delta J(t_1)} \frac{\delta}{\delta J(t_2)} Z[J] \Big|_{J=0} . \quad (1.81)$$

Discuss the graphical representation of the terms you get and observe that the vacuum bubbles have cancelled.

1.8.3 Effective action from integrating out fermions in QED

a) Show that in spinor quantum electrodynamics, the functional integral representation of time-ordered expectation values of operators that do not involve the matter fields can be done in two steps, where the first step involves doing the functional integral over the fermions only. This amounts to computing

$$\begin{aligned} S_{\text{vac,vac}}[A] &= \langle \text{vac, out} | \text{vac, in} \rangle_A \\ &= |\mathcal{N}|^2 \int \prod_l \mathcal{D}\psi_l \mathcal{D}\bar{\psi}_l \exp \left\{ -i \int d^4x \bar{\psi}(x) (\not{\partial} + m + ieA(x)) \psi(x) \right\} . \end{aligned} \quad (1.82)$$

b) Compute this integral exactly in terms of an appropriate functional determinant. Of course, this determinant is ill-defined and needs regularization. (One could e.g. replace $\text{Det}[\dots] = \prod_n \lambda_n$ by $\prod_{n=0}^N \lambda_n$ and let $N \rightarrow \infty$ in the end, or replace it by $\prod_n \lambda_n e^{-\epsilon n^2}$ and let $\epsilon \rightarrow 0$ in the end.) Once regularized, it can be manipulated as an ordinary determinant. Show that

$$\frac{S_{\text{vac,vac}}[A]}{S_{\text{vac,vac}}[0]} = \text{Det} \left[1 + (\not{\partial} + m)^{-1} ieA(x) \right] . \quad (1.83)$$

Use $\text{Det}(1 + M) = \exp(\text{Tr} \log(1 + M))$ and the power series expansion of $\log(1 + x)$ to get

$$\frac{S_{\text{vac,vac}}[A]}{S_{\text{vac,vac}}[0]} = \exp \left\{ -\text{Tr} \sum_{n=1}^{\infty} \frac{1}{n} \left[(\not{\partial} + m)^{-1} (-ie) \not{A}(x) \right]^n \right\} , \quad (1.84)$$

and give a Feynman diagram interpretation of each term in the sum.

Hint : Recall $\text{Tr}(\dots) = \int d^4p \langle p | \text{tr}_D(\dots) | p \rangle$ where tr_D indicates a trace over the Dirac matrices and $\langle p | (\not{\partial} + m)^{-1} | x \rangle = (i\not{p} + m)^{-1} \langle p | x \rangle = (i\not{p} + m)^{-1} \frac{e^{-ip_\mu x^\mu}}{(2\pi)^2}$, as well as the definition of the Fourier transform $\tilde{f}(q) = \int d^4x e^{-iq_\mu x^\mu} f(x)$ and $f(x) = \int \frac{d^4p}{(2\pi)^4} e^{iq_\mu x^\mu} \tilde{f}(q)$.

c) One can show that each term in the sum in the exponent is purely imaginary which allows to write

$$S_{\text{vac,vac}}[A] = S_{\text{vac,vac}}[0] \exp \left\{ i \int d^4x \mathcal{L}_{\text{eff}}^{1\text{-loop}}[A] \right\} , \quad (1.85)$$

where $\mathcal{L}_{\text{eff}}^{1\text{-loop}}[A]$ is the effective action obtained by integrating out the fermions. Show that by taking functional derivatives with respect to $A_{\mu_i}(x_i)$ of $S_{\text{vac,vac}}[A]$ one generates time-ordered (connected) correlation functions of the electromagnetic current operators $J^{\mu_i}(x_i) = -ie\bar{\psi}(x_i)\gamma^{\mu_i}\psi(x_i)$.

1.8.4 An example of asymptotic series

a) Let

$$f(g) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx e^{-x^2 - gx^4} . \quad (1.86)$$

Compute $f(0)$ and show that f is a well-defined monotonously decreasing function for $g > 0$. What happens for $g < 0$?

b) In order to try to compute f in perturbation theory, expand e^{-gx^4} in a (converging !) power series. Truncate this series at order N . We let

$$f_N(g) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx e^{-x^2} \sum_{n=0}^N \frac{1}{n!} (-gx^4)^n . \quad (1.87)$$

Since this is a finite sum, exchange the order of the sum and the integration and explicitly compute the integrals for each term so as to obtain the coefficients a_n in

$$f_N(g) = \sum_{n=0}^N a_n g^n . \quad (1.88)$$

c) Compute $\frac{a_n}{a_{n+1}}$ and deduce that the limit $N \rightarrow \infty$ does not exist or, in other words, that the corresponding infinite series has a vanishing radius of convergence. Explain this fact in terms of the behaviour of $f(g)$ for negative g .

d) Numerically evaluate $f(g)$ for small $g > 0$, say $g = 0.01$, and compute $f_N(g)$ for $N = 0, 1, 2, 3, 4$. What do you observe ? Repeat the same procedure for $g = 0.3$.

2 General renormalization theory

As we have seen in the example of the vacuum polarization, but also more generally, the notion of renormalization arises in order to keep correctly normalized full propagators, with their poles at the physical masses. This requirement is independent of any divergences arising in loop integrals. However, the one-loop contributions typically are divergent and have to be regulated. Then, at least in QED, after performing the renormalization, according to the renormalization conditions, the renormalized quantities turn out to be finite (i.e. have finite limits even if the regulator is removed). The purpose of the general renormalization theory is to show that this is no accident but remains valid to all orders in perturbation theory, at least for so-called renormalizable theories as QED or scalar ϕ^4 -theory in four dimensions.

2.1 Degree of divergence

The *superficial* degree of divergence D characterizes the behaviour of the momentum integral $\int d^4k_1 \dots d^4k_L [\dots]$ when all $|k_i| \rightarrow \infty$ with a common $k \rightarrow \infty$. More precisely: if the integral behaves as $\sim \int k^{D-1} dk$ then D is the superficial degree of divergence. An integral with $D \geq 0$ is called superficially divergent, and one with $D < 0$ is called superficially convergent. This does not necessarily mean the integral really diverges or converges, but we will see that the superficial degree of divergence D nevertheless plays an important role. Often, we will talk somewhat loosely about the superficial degree of divergence D of a diagram, meaning the D of the associated integral. To determine D let

- I_f be the number of internal lines of the field f ,
- E_f be the number of external lines of the field f ,
- N_i the number of vertices of type i with d_i derivatives and n_{if} fields f attached.

Now, for large momenta the propagators behave as $\Delta_f(k) \sim k^{-2+2s_f}$, where

- for scalars: $s_f = 0$,
- for spin 1/2: $s_f = 1/2$,
- for massive spin 1: $s_f = 1$ since $\Delta(q)^{\mu\nu} \sim \frac{1}{q^2} \left(\eta^{\mu\nu} - \frac{q^\mu q^\nu}{m^2} \right)$
- and for photons or gravitons: $s_f = 0$, since: $\Delta(q)^{\mu\nu} \sim \frac{1}{q^2} \left(\eta^{\mu\nu} - \xi \frac{q^\mu q^\nu}{q^2} \right)$.

It follows that

$$D = 4L + \sum_i N_i d_i + \sum_f I_f (-2 + 2s_f) . \quad (2.1)$$

Now use the relation $L = I - V + 1 = \sum_f I_f - \sum_i N_i + 1$ to obtain

$$D = 4 + \sum_i N_i (d_i - 4) + \sum_f I_f (2 + 2s_f) . \quad (2.2)$$

Use further $\sum N_i n_{if} = 2I_f + E_f$ to get

$$D = 4 + \sum_i N_i \left((d_i - 4 + \sum_f n_{if} (s_f + 1)) \right) - \sum_f E_f (s_f + 1) , \quad (2.3)$$

or

$$D = 4 - \sum_f E_f (s_f + 1) - \sum_i N_i \Delta_i , \quad (2.4)$$

with

$$\Delta_i = 4 - d_i - \sum_f n_{if} (s_f + 1) . \quad (2.5)$$

One can repeat this argument in an arbitrary space-time dimension d . One simply has to replace $4L \rightarrow dL$ in (2.1). The final formula then is modified as

$$D = d - \sum_f E_f \left(s_f + \frac{d-2}{2} \right) - \sum_i N_i \Delta_i \quad (2.6)$$

$$\Delta_i = d - d_i - \sum_f n_{if} \left(s_f + \frac{d-2}{2} \right) . \quad (2.7)$$

Note that scalar theories in $d = 2$ have the particularity that the superficial degree of divergence does not depend on the number of external lines. Indeed, in this case one simply has $D = 2 - \sum_i N_i \Delta_i$ and $\Delta_i = 2 - d_i$. Nonlinear sigma models have $d_i = 2$ and hence $\Delta_i = 0$, so all diagrams have $D = 2$, independently of the external lines. On the other hand, a simple ϕ^4 -theory in 2 dimensions has $\Delta = 2$ and only loop diagrams with a single vertex can be (superficially) divergent.

There is an alternative way to determine the Δ_i of a given interaction, which works easily as long as only scalars, spin $\frac{1}{2}$ particles, photons or gravitons are involved (no massive spin 1). The engineering (mass) dimension $\mathcal{D}_f$ of a given field can be read from the kinetic term in the Lagrangian. In d dimensions one gets $\mathcal{D}_f = \frac{d-2}{2}$ for scalars, photons or gravitons, and $\mathcal{D}_f = \frac{d-1}{2}$ for spin $\frac{1}{2}$ fields. This coincides with $\mathcal{D}_f = s_f + \frac{d-2}{2}$. (This does not work for massive spin 1 fields that have $s_f = 1$ but engineering dimension 1 - it is the explicit appearance of the mass m which messes up the argument.) Then any interaction of type i involving n_{if} fields f and d_i derivatives has dimension $d_i + \sum_f n_{if} \mathcal{D}_f$. Since the Lagrangian must have dimension d , the coupling constant g_i must have dimension $[g_i] = d - d_i - \sum_f n_{if} \mathcal{D}_f$ or

$$[g_i] = \Delta_i , \quad (2.8)$$

so that the value of Δ_i can be directly read from the (mass) dimension of the coupling constant. As an example, in Table 2, we consider the various interactions in QED in four dimensions, where one has $\mathcal{D}_\psi = \frac{3}{2}$ and $\mathcal{D}_{A_\mu} = 1$ and one indeed verifies that the Δ_i coincide with the dimensions of the coupling constants.

One can similarly assign Δ_a^{comp} to composite operators $\mathcal{O}_a(x)$ inserted in a diagram. Δ_a^{comp} is the same as one would define for the corresponding interaction, and eq. (2.4) would be replaced by $D = 4 - \sum_f E_f \mathcal{D}_f - \sum_i N_i \Delta_i - \sum_a N_a \Delta_a^{\text{comp}}$.

The importance of the notion of superficial degree of divergence D resides in the following remarks:

$-ie\psi A\psi$	$\Delta = 4 - 2 \times \frac{3}{2} - 1 = 0$
$(Z_3 - 1)F_{\mu\nu}F^{\mu\nu}$	$\Delta = 4 - 2 - 2 = 0$
$(Z_2 - 1)\psi\bar{\psi}$	$\Delta = 4 - 2 \times \frac{3}{2} - 1 = 0$
$[-(Z_2 - 1)m + Z_2\delta m]\psi\psi$	$\Delta = 4 - 2 \times \frac{3}{2} = 1$

Table 1: The values of Δ for the various interactions in QED

- If $\Delta_i \geq 0$ for all i then only Green's functions with $d - \sum_f E_f \mathcal{D}_f \geq 0$ can have $D \geq 0$ and, since $\mathcal{D}_f > 0$ (except for $s_f = 0$ in $d = 2$ dimensions) there are only finitely many superficially divergent Green's functions (finitely many values of E_f). Note that, unless $\Delta_i > 0$, $\forall i$, these are still infinitely many diagrams (arbitrary N_i for fixed E_f).
- Interactions with $\Delta_i \geq 0$ are called *renormalizable interactions*. Theories with only renormalizable interactions are called *renormalizable theories*. In such theories only finitely many Green's functions are superficially divergent. We will see that in such theories *all* Green's functions can be rendered finite by a finite number of counterterms corresponding to the redefinition of a finite number of physical constants (couplings and masses) and the (re)normalizations of the fields.
- If $\Delta_i < 0$ for some interaction, then an infinite number of Green's functions (arbitrary numbers E_f of external fields) are superficially divergent. In general, one then needs an infinite number of counterterms to make them finite. Such theories are called *non-renormalizable*.

The importance of the notion of superficial divergence is partly due to the following **Theorem** (Weinberg): If $D < 0$ for the complete integration and *any* sub-integration (i.e. holding some linear combination of momenta fixed) then the integral is really convergent.

In particular, at one-loop, there are no sub-integrations and, hence, by this theorem, any one-loop integral with $D < 0$ is convergent. Note also that this theorem does not say anything about the divergence of integrals with $D > 0$. We have seen examples in QED where the integrals are less divergent than expected. The vacuum polarization diagram e.g. had $D = 2$, but gauge invariance allows one to “pull out” two factors of external momentum by writing $\Pi_{\mu\nu} = (q^2\eta_{\mu\nu} - q_\mu q_\nu)\pi(q^2)$ with $\pi_{1\text{-loop}}(q^2)$ only logarithmically divergent ($D = 0$). More generally, symmetries may result in cancellations between the divergences arising from individual diagrams yielding less divergent or even finite Green's functions.

2.2 Structure of the divergences

Suppose an integral has $D < 0$, i.e. it is superficially convergent. This means that if all k are simultaneously taken to be large the integral converges. Thus the only possibility for this integral not to converge is that it diverges if some combination of momenta is held fixed and the divergence is due to the sub-integration over the other momenta. This must necessarily correspond to a subdiagram. Hence a superficially convergent diagram can only be divergent due to a diverging subdiagram (with $D' \geq 0$), as shown e.g. in Fig. 2. However, such subdiagrams are of lower order in perturbation theory. If they have already been rendered finite by the addition of appropriate counterterms at the lower order in perturbation theory, one no longer has to worry about such diverging subdiagrams

any more. Henceforth, we will assume that such diverging subdiagrams have been taken care of and that superficially convergent diagrams are convergent.

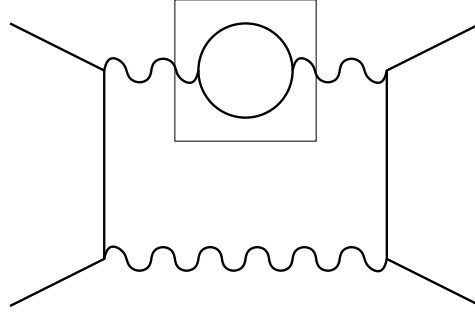


Figure 2: A superficially convergent diagram containing a divergent vacuum-polarization subdiagram.

In the following we will argue a lot about the diverging and finite parts of a diagram. One should keep in mind that we always work with regularized expressions, and by a diverging expression we mean one that does diverge once we remove the regulator. In dimensional regularization in particular, one can also make a clear distinction between the diverging and the finite parts of an expression, the former involving inverse powers of ϵ (the “polar part”) and the latter involving no or positive powers of ϵ .

If an integral $I(p_i)$ depending on external momenta p_i has superficial degree of divergence $D \geq 0$, then by differentiating with respect to the external momenta lowers D by one unit since

$$\frac{\partial}{\partial p_i^\mu} \frac{1}{(k + p_i)^2 + m^2 - i\epsilon} = \frac{-2(k + p_i)_\mu}{[(k + p_i)^2 + m^2 - i\epsilon]^2} . \quad (2.9)$$

(There is a caveat to this argument to be discussed soon.) Differentiating $D + 1$ times results in an integral having degree of divergence -1 , i.e. which is superficially convergent. According to the above remarks it is then convergent. Hence

$$\frac{\partial}{\partial p_{i_1}^{\mu_1} \dots \partial p_{i_{D+1}}^{\mu_{D+1}}} I(p_i) = J(p_i) , \quad (2.10)$$

where $J(p_i)$ is convergent. Upon integrating with respect to the external momenta p_i we get

$$I(p_i) = \mathcal{P}_D(p_i) + \hat{I}(p_i) , \quad (2.11)$$

with $\mathcal{P}_D(p_i)$ a polynomial of order D in the external momenta with (a priori) divergent coefficients and $\hat{I}(p_i)$ a convergent integral. Obviously, the diverging part $\mathcal{P}_D(p_i)$ is entirely contained in the first $D + 1$ terms in a Taylor expansion of $I(p_i)$. We may then rephrase this by saying that $I(p_i)$ minus the first $D + 1$ terms in a Taylor expansion in the external momenta is finite. We have seen this in our one-loop computations of the vacuum polarization, see e.g. (1.49) and (1.54) which one can rewrite as

$$\Pi_{1\text{-loop}}^{\mu\nu}(q) = (q^2 \eta^{\mu\nu} - q^\mu q^\nu) \pi_{1\text{-loop}}(0) + (q^2 \eta^{\mu\nu} - q^\mu q^\nu) \pi(q^2) . \quad (2.12)$$

The first terms $\sim \pi_{1\text{-loop}}(0)$ indeed constitute a second-order polynomial in the external momentum which are the first terms in a Taylor expansion of $\Pi_{1\text{-loop}}^{\mu\nu}(q)$ around $q = 0$, while the last term $\sim \pi(q^2)$ is the renormalized finite vacuum polarization.

It is very important that the divergent part of the integral $I(p_i)$ is a polynomial in the external momenta since it is precisely such divergences that can be cancelled by *local* counterterms. Indeed, a local counterterm is of the form

$$\prod_f \prod_{i=1}^{E_f} (\partial_{\mu_{i,f}})^{d_{i,f}} f(x) \quad , \quad \sum d_{i,f} = d \quad , \quad (2.13)$$

being the product of E_f fields of type f with a total number of d derivatives. Such a counterterm gives a (tree-level) contribution to the diagram with E_f external lines of type f which is a polynomial of order d in the momenta. Thus, to cancel the diverging $\mathcal{P}_D(p_i)$ one can always find a sum of such local counterterms involving up to D derivatives. Now, such a counterterm will also appear as a new vertex in loop diagrams and it is important to check that it does not render the theory less renormalizable. Indeed, with at most D derivatives, we have $\Delta_{\text{c.t.}} \geq 4 - \sum_f E_f \mathcal{D}_f - D$. But in a renormalizable theory the possible values of D are constrained by $D \leq 4 - \sum_f E_f \mathcal{D}_f$ so that $4 - \sum_f E_f \mathcal{D}_f - D \geq 0$ and $\Delta_{\text{c.t.}} \geq 0$, so the counterterm is part of the finitely many renormalizable interactions. Moreover, if one includes *all* (finitely many) renormalizable interactions in the bare Lagrangian, then all possible counterterms are necessarily of the same form as the terms already present, and they just renormalize the couplings, masses and wave-functions

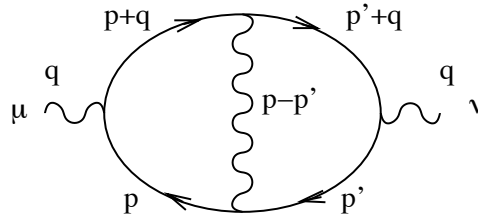


Figure 3: A two-loop diagram with overlapping divergences: two divergent subdiagrams share a common propagator.

There is a caveat in the argument that taking a derivative with respect to the external momenta lowers the superficial degree of divergence by one unit and which has to do with so-called *overlapping divergences*. An overlapping divergence occurs if two divergent subdiagrams share a common line (propagator), see e.g. the two-loop contribution to the vacuum polarization shown in Fig. 3. Then taking a derivative with respect to an external momentum typically results in a *sum* of terms with the large momentum behaviour in one of the sub-loops improved in one term but not in the other, and vice versa for the other sub-loop. To be specific, for the example of Fig. 3 one has (neglecting the electron mass for simplicity)

$$\Pi_{\mu\nu}^{\text{*overlap}}(q) \sim \int \frac{d^4 p d^4 p'}{(p-p')^2} \text{tr} \left[\frac{1}{\not{p}'} \gamma_\nu \frac{1}{\not{p}' + \not{q}} \gamma_\rho \frac{1}{\not{p} + \not{q}} \gamma_\mu \frac{1}{\not{p}} \gamma^\rho \right] . \quad (2.14)$$

Taking a derivative with respect to the external momentum q gives a sum of two integrals:

$$\frac{\partial}{\partial q^\sigma} \Pi_{\mu\nu}^{\text{*overlap}}(q) = \int \frac{d^4 p d^4 p'}{(p-p')^2} \text{tr} \left[-\frac{1}{\not{p}'} \gamma_\nu \frac{(\not{p}' + \not{q}) \gamma^\sigma (\not{p}' + \not{q})}{[(p' + q)^2]^2} \gamma_\rho \frac{1}{\not{p} + \not{q}} \gamma_\mu \frac{1}{\not{p}} \gamma^\rho \right] \quad (2.15)$$

$$- \frac{1}{\not{p}'} \gamma_\nu \frac{1}{\not{p}' + \not{q}} \gamma_\rho \frac{(\not{p} + \not{q}) \gamma^\sigma (\not{p} + \not{q})}{[(p + q)^2]^2} \gamma_\mu \frac{1}{\not{p}} \gamma^\rho \right] . \quad (2.16)$$

In the first term the large p' behavior is improved but *not* the large p behaviour, while in the second term things are reversed. Of course, the superficial degree of divergence is lowered by one unit, and taking two more derivatives would result in a sum of terms all with a $D' = -1$, but the criteria of the theorem cited above are not satisfied and one cannot conclude that up to a divergent subgraph the integral is convergent. Although this is true for each individual term in the sum, the trouble is that for each term the diverging subgraph is a different one. For example, in a given term in the sum, the divergence may be traced to a divergent p sub-integration corresponding to the divergent subgraph γ_1 in Fig. 4, while in another term the divergence would be due to the p' sub-integration corresponding to the subdiagram γ_2 in Fig. 4.

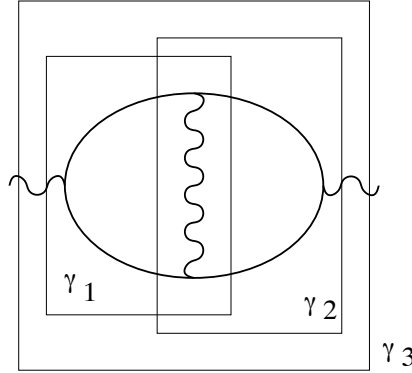


Figure 4: The nested sequences of subdiagrams used in the BPHZ construction: $\gamma_1 \subset \gamma_3$ and $\gamma_2 \subset \gamma_3$.

Although such overlapping divergences are more complicated, the previous discussion shows that one still has a sum of divergences associated with the various subdiagrams. To deal with such overlapping divergences in a systematic way and show that the usual counterterms are exactly sufficient to cancel these divergences, too, was the achievement of Bogoliubov, Parasiuk, Hepp and Zimmermann (BPHZ).

2.3 Bogoliubov-Parasiuk-Hepp-Zimmermann prescription and theorem

One defines a *forest* U as a family of nested (sub)graphs γ_i : if $\gamma_1 \in U$ and $\gamma_2 \in U$ then either $\gamma_1 \subset \gamma_2$ or $\gamma_2 \subset \gamma_1$ or $\gamma_1 \cap \gamma_2 = \emptyset$. For the diagram of Fig. 4 e.g., γ_1 and γ_2 cannot be in the same forest. More generally, overlapping loops cannot be in the same forest. On the other hand, in Fig. 2 the vacuum polarization subdiagram and the whole diagram can be in the same forest. Hence the notion of forest is exactly what one needs to distinguish overlapping subdiagrams (divergences) from non-overlapping ones. Again, in Fig. 4 the list of all forests is:

$$U_0 = \emptyset, \quad U_1 = \{\gamma_1\}, \quad U_2 = \{\gamma_2\}, \quad U_3 = \{\gamma_3\}, \quad U_4 = \{\gamma_1, \gamma_3\}, \quad U_5 = \{\gamma_2, \gamma_3\}. \quad (2.17)$$

For each diagram G one then considers such a family $\mathcal{F}(G)$ of all forests of G . Then we have the *BPHZ prescription*: For each forest $U_i \in \mathcal{F}(G)$ consider the nested sequence $\gamma_a \subset \gamma_b \subset \dots$ of the (sub)graphs in U_i . Starting with the innermost γ_a , one defines a *subtraction term* $S(U_i)$ by replacing the integrand I_b of each subdiagram γ_b (in the nested sequence) with superficial degree of divergence D_b by the first $D_b + 1$ terms in its Taylor expansion (e.g. around 0) in the momenta flowing into (or out of) this subdiagram. Since at each step one only keeps a polynomial in the corresponding momenta, the final subtraction term is still a polynomial in the external momenta.

To see how this works, consider the diagram of Fig. 3, resp. Fig. 4:

- $U_0 = \emptyset$: no subtraction, $S(U_0) = 0$.
- $U_1 = \{\gamma_1\}$: This subdiagram is just the vertex function $e\Gamma_{1\text{-loop}}^\mu(p' + q, p')$ and a one-loop computation shows that $e\Gamma_{1\text{-loop}}^\mu(p' + q, p') = eL\gamma^\mu + e\Gamma_{\text{finite}}^\mu(p' + q, p')$, with $\Gamma_{\text{finite}}^\mu(p' + q, p')$ being the finite (renormalized) vertex function and $L = -(Z_2 - 1)_{e^2}$. This subdiagram has $D = 0$ and hence the corresponding subtraction term is just the first term in the Taylor expansion: $S(U_1) = -e[(Z_2 - 1)_{e^2} + c]\gamma^\mu$ with $c = -\Gamma_{\text{finite}}^\mu(0, 0)$ and $(Z_2 - 1)_{e^2} = -\frac{e^2}{8\pi^2} \frac{1}{\epsilon} + \text{finite}$. Obviously, this corresponds to a counterterm $\sim A_\mu \bar{\psi} \gamma^\mu \psi$.
- $U_2 = \{\gamma_2\}$: This similarly gives the same subtraction term $S(U_2) = -e[(Z_2 - 1)_{e^2} + c]\gamma^\mu$.
- $U_3 = \{\gamma_3\}$: This is just the full diagram we want to study. It has $D = 2$ and thus the subtraction term must be a second-order polynomial in the q . By gauge invariance this must be of the form $(\eta_{\mu\nu}q^2 - q_\mu q_\nu)C$ with $C = \pi_{\text{overlap}}(q^2 = 0)$. We are not claiming that this is the only divergence of this diagram, but BPHZ tell us that this is the subtraction term for this forest: $S(U_3) = (\eta_{\mu\nu}q^2 - q_\mu q_\nu) \times \pi_{\text{overlap}}(q^2 = 0)$. This corresponds to an order e^4 contribution to Z_3 .
- $U_4 = \{\gamma_1, \gamma_3\}$: Here we must first replace the integrand of the γ_1 -part by $S(U_1)$ as defined above, so that the integrand for the γ_3 part is the same as for $\Pi_{1\text{-loop}}^{\mu\nu}$ but with $e\gamma^\mu$ in the left vertex replaced by $e[-(Z_2 - 1)_{e^2} - c]\gamma^\mu$. Then the subtraction term is $S(U_4) = (q^2\eta^{\mu\nu} - q^\mu q^\nu)[-(Z_2 - 1)_{e^2} - c](Z_3 - 1)_{e^2}$. This also corresponds to an order e^4 contribution to Z_3 .
- $U_5 = \{\gamma_2, \gamma_3\}$: same subtraction term as for U_4 .

BPHZ theorem: The original (regularized) diagram *minus* the sum of all subtraction terms is finite. The latter is exactly generated by the (usual) counterterms.

We will not give the proof of this theorem, but just mention that it proceeds by recursion. One assumes that all diagrams at order N are made finite by the counterterms. Then one finds a recursion relation for the large momentum behavior at the next order $N + 1$ (including on subspaces $p'_\mu || p_\mu$, etc, as relevant for overlapping divergences). The solution to this recurrence relation uses the sum over forests. It can then be shown that the original diagram minus the sum of all subtraction terms is indeed finite.

2.4 Summary of the renormalization program and proof

- Δ_i is the (mass) dimension of the coupling g_i (provided one only deals with scalars, spin- $\frac{1}{2}$, photons or gravitons). Renormalizable interactions have $\Delta_i \geq 0$, a renormalizable theory only has interactions with $\Delta_i \geq 0$. The superficial degree of divergence (divergence in the region where all $k_i \rightarrow \infty$) is $D = 4 - \sum_f E_f \mathcal{D}_f - \sum_i N_i \Delta_i$. In a renormalizable theory only a finite set of E_f gives $D \geq 0$. Diagrams with $D < 0$ are superficially convergent.
- A superficially convergent diagram can be divergent only due to divergent subdiagrams. Then the same counterterms that render finite these subdiagrams also make the whole diagram finite. Thus we only have to deal with superficially divergent diagrams, i.e. only with finitely many E_f if the theory is renormalizable.

- Ignoring overlapping divergences for the moment, a superficially divergent diagram with $D \geq 0$ is made superficially convergent by $D + 1$ derivatives with respect to the external momenta. This implies that its diverging part is a polynomial in the external momenta and this can be subtracted by a local counterterm. In a renormalizable theory, even though at each order in perturbation theory there are new diverging diagrams, for fixed E_f , the degree D cannot increase and the structure of the polynomial, i.e. of the counterterms remains the same, only the coefficients (the Z 's) get higher and higher order contributions. Since $D \geq 0$ only for finitely many E_f , only finitely many counterterms are needed. Adding these counterterms renders all these diagrams finite to all orders in perturbation theory. On the other hand, in a non-renormalizable theory, at every order in perturbation theory, diagrams with more and more external lines become divergent and one needs an infinite number of counterterms to render them all finite.
- To deal with overlapping divergences, BPHZ define the forests with nested boxes (subdiagrams) γ_a of degree $D_a = D(\gamma_a)$, each forest giving a subtraction term obtained by replacing each subdiagram in the nested sequence by its $D_a + 1$ first terms in a Taylor series in the “external” momenta. This corresponds to the same counterterms as defined above. BPHZ show that this procedure renders also finite these diagrams with overlapping divergences to all orders in perturbation theory. Again, in a renormalizable theory, finitely many counterterms make all diagrams finite.

2.5 The criterion of renormalizability

Renormalizability restricts interactions to those with $\Delta_i = 4 - d_i - \sum_f n_{if}(1 + s_f) = 4 - d_i - \sum_f n_{if}\mathcal{D}_f \geq 0$ only. For a finite number of fields these are only finitely many interactions. Of course, there are also other restrictions on the possible interactions like Lorentz invariance or other symmetries one might want to preserve.

If the regulator preserves a given symmetry then the finite and the diverging parts of any Green's function (not necessarily of individual diagrams) at any order in perturbation theory must also preserve this symmetry. Hence the counterterms also respect this symmetry.

What if one allowed to include *non*-renormalizable interactions, like e.g. adding a term $\bar{\psi}\gamma_{\mu\nu}\psi F^{\mu\nu}$ to the QED Lagrangian? This interaction indeed has dimension 5 so that $\Delta = -1$ and it is non-renormalizable. For dimensional reasons its coupling constant must be of the form $\frac{\rho}{M}$ with dimensionless ρ and some mass scale M . Such a term would change the magnetic moment of electrons by an additional amount $\sim \rho \frac{m_e}{M}$ which would thus become an adjustable parameter. From experiment we know that this quantity must be extremely small. Said differently, if we take ρ of the same order as e , the mass scale M must be very large compared to m_e .

More generally, a non-renormalizable interaction has $\Delta_i < 0$ and coupling constants of dimension $\Delta_i = -|\Delta_i|$. Write $g_i = \frac{\tilde{g}_i}{M^{|\Delta_i|}}$. Then non-renormalized diagrams have divergences that behave as $(\tilde{g}_i)^N \left(\frac{\Lambda}{M}\right)^{|\Delta_i|N}$ at order N in perturbation theory, Λ being some UV cutoff. Also, infinitely many Green's functions are divergent and we need infinitely many counterterms, possibly with an arbitrary number of derivatives. However, once these divergences are cancelled, we have renormalized Green's functions $G(p_i)$, and at order N in perturbation theory (in g_i) they will behave as $(\tilde{g}_i)^N \left(\frac{p_i}{M}\right)^{|\Delta_i|N}$. Thus, as long as $|p_j| \ll M$ one can neglect the effect of these non-renormalizable interactions.

Theories like QED are presently thought to be only effective theories, in the sense that they provide the effective description of electromagnetic interactions at energies that are low compared to some scale at which new physics could be expected, like e.g. the grand unification scale of 10^{15} GeV or even the Planck scale of 10^{19} GeV. Such an effective theory then has an effective Lagrangian obtained by “integrating out” the very heavy additional fields that are present in such theories. This necessarily results in the generation of (infinitely) many non-renormalizable interactions in this effective Lagrangian with couplings $\frac{\tilde{g}_i}{M^{|\Delta_i|}}$, M being e.g. the grand unification scale. From the previous argument it is then clear that at energies well below this scale these additional non-renormalizable interactions are completely irrelevant, and this is why we only “see” the renormalizable interactions. Our “low-energy” world is described by renormalizable theories like QED not because such theories are somehow better behaved, but because these are the only relevant ones at low energies:

Renormalizable interactions are those that are relevant at low energies, while non-renormalizable interactions are irrelevant at low energies.

This is also the terminology in statistical mechanics where one studies infrared physics at scales $|p| \ll \frac{1}{a} \equiv \Lambda$, where a is the lattice spacing.

2.6 Exercises

2.6.1 Yukawa theory

Consider a 4-dimensional quantum field theory of a massive Dirac fermion ψ coupled to a massive scalar field ϕ via a Yukawa coupling $g\bar{\psi}\psi\phi$.

a) Determine whether this theory is renormalizable. Let $\Gamma^{n,m}$ be the one-particle irreducible $n + m$ point function with n external fermions and m external scalars. Determine which $\Gamma^{n,m}$ are superficially divergent and which ones are superficially convergent. Draw the one-, two-, and three-loop contributions to $\Gamma^{0,2}$ and $\Gamma^{2,1}$ and explicitly (by looking at the propagators and integrations) determine their superficial degree of divergence.

b) What happens if one replaces the Yukawa coupling by a derivative coupling $g'\bar{\psi}\gamma^\mu\psi\partial_\mu\phi$? Study the superficial degree of divergences of $\Gamma^{4,0}$ at one, two and three loops.

2.6.2 Quantum gravity

The Einstein-Hilbert action for gravity in $d \geq 3$ dimensions is

$$S_{\text{EH}} = \frac{1}{\kappa^2} \int d^d x \sqrt{g} R[g] \quad (2.18)$$

where $R[g]$ is the scalar (Ricci) curvature computed from the metric $g_{\mu\nu}$ and $\sqrt{g}$ denotes the square-root of the determinant of the metric. In perturbative quantum gravity one lets $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with “small” $h_{\mu\nu}$ and does a perturbative expansion of the action. Show that $g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} + \mathcal{O}(h^2)$ where one raises and lowers indices on h with η . Also show that $\sqrt{g} = 1 + \frac{1}{2}h^\mu_\mu$. (Hint: use the formula $\det g = \exp(\text{tr} \log g)$. Recall the formula for the Christoffel symbols and show that

$$2\Gamma^\rho_{\mu\nu} = \partial_\mu h^\rho_\nu + \partial_\nu h^\rho_\mu - \partial^\rho h_{\mu\nu} - h^{\rho\sigma}(\partial_\mu h_{\nu\sigma} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}) + \mathcal{O}(h^2\partial h) . \quad (2.19)$$

Recall that

$$R^\rho_{\nu\lambda\mu} = \partial_\lambda \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\lambda\mu} + \Gamma^\rho_{\lambda\sigma} \Gamma^\sigma_{\mu\nu} - \Gamma^\rho_{\mu\sigma} \Gamma^\sigma_{\lambda\nu} \quad , \quad R_{\nu\mu} = R^\rho_{\nu\rho\mu} \quad , \quad R = g^{\mu\nu} R_{\nu\mu} , \quad (2.20)$$

and argue *without keeping track of the indices* that $\partial\Gamma \sim \partial^2 h + h\partial^2 h + (\partial h)^2 + h^2\partial^2 h + h(\partial h)^2 + \dots$. Give the corresponding expansions (still without keeping track of the indices) of $R^\rho_{\nu\lambda\mu}$, R and finally of $\sqrt{g}R$. Deduce the large momentum behaviour of the h -propagator as well as the dimensions of the various interactions in space-time dimension d . Discuss whether this theory is renormalizable in d dimensions? What happens in $d = 2$ dimensions?

2.6.3 Non-renormalizable interactions from integrating out a massive field

Consider a theory of two scalars φ and ϕ with masses m and M where $M \gg m$. Let their action be

$$S[\varphi, \phi] = \int d^4x \left(-\frac{1}{2} \right) [(\partial\varphi)^2 + m^2\varphi^2 + (\partial\phi)^2 + M^2\phi^2 + gm\varphi\phi^2] . \quad (2.21)$$

(We have defined the coupling as $g \times m$ to have a dimensionless g .) Define the effective action for the light field φ as

$$e^{iS_{\text{eff}}[\varphi]} = \int \mathcal{D}\phi e^{iS[\varphi, \phi]} . \quad (2.22)$$

Show that the resulting determinant can be interpreted as a sum of one-loop diagrams with N ϕ -propagators and N external φ attached at the vertices. Compute the determinant by neglecting the ϕ momenta with respect to M and show that this results in infinitely many new couplings for φ of the form

$$\sum_{N=1}^{\infty} \frac{(-)^N}{N} \Lambda^4 \left(\frac{m g}{M^2} \right)^N \varphi^N , \quad (2.23)$$

where Λ is some large UV-cutoff which one can take of the order of M . Discuss the role of these new interactions for $N \leq 4$ and for $N > 4$.

3 Non-abelian gauge theories: formulation and quantization

Gauge invariance, and in particular non-abelian gauge invariance, plays a most important role in the formulation of the quantum field theories that seem to describe (almost?) all of the particle physics as it is known today: quantum electrodynamics to begin with and its embedding into the electro-weak theory based on the gauge group $SU(2) \times U(1)$, as well as the theory of strong interactions based on the gauge group $SU(3)$. Here, we will briefly show how to construct (classical) actions that are invariant under non-abelian gauge symmetries. Then we will discuss how to quantize these theories using the functional integral approach. This will involve the issues of gauge-fixing, Faddeev-Popov procedure and the appearance of so-called ghost fields. The gauge-fixed action no longer is gauge invariant but instead has a new symmetry, the BRST-symmetry which we will identify. This BRST symmetry will play a crucial role when showing that these gauge theories are renormalizable.

3.1 Non-abelian gauge transformations and gauge invariant actions

Recall from quantum electrodynamics that the classical action is invariant under the following gauge transformation:

$$\psi_l(x) \rightarrow \psi'_l(x) = e^{i\alpha(x)q_l} \psi_l(x) \quad , \quad A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu \alpha(x) = A_\mu(x) + \frac{i}{q_l} e^{i\alpha(x)q_l} \partial_\mu e^{-i\alpha(x)q_l} . \quad (3.1)$$

This is often called a $U(1)$ gauge invariance because obviously $g(x) = e^{i\alpha(x)q_l} \in U(1)$. There seems to be a different $g(x)$ for every different charge q_l , but note that q_l is the eigenvalue of the charge operator Q when acting on ψ_l . Hence we can write $g(x) = e^{i\alpha(x)Q}$. Using this $g(x)$ we can rewrite the transformations (3.1) as

$$\psi(x) \rightarrow \psi'_l(x) = g(x)\psi(x) \quad , \quad QA_\mu(x) \rightarrow QA'_\mu(x) = Q A_\mu(x) + ig(x)\partial_\mu g(x)^{-1} . \quad (3.2)$$

Of course, gauge invariant terms then are built from covariant derivatives, i.e.

$$D_\mu \psi = (\partial_\mu - iQA_\mu)\psi , \quad (3.3)$$

since this transforms as

$$\begin{aligned} D_\mu \psi \rightarrow D'_\mu \psi' &= \left(\partial_\mu - iQA_\mu + (g\partial_\mu g^{-1}) \right) g\psi = g \left(\partial_\mu \psi_l + (g^{-1}\partial_\mu g) - iQA_\mu + (\partial_\mu g^{-1}g) \right) \psi \\ &= g \left(\partial_\mu - iQA_\mu \right) \psi = g D_\mu \psi . \end{aligned} \quad (3.4)$$

Indeed, since g is just a phase, $|D_\mu \psi|^2$ obviously is invariant.

We have rewritten the simple $U(1)$ gauge transformation of the electromagnetic (gauge) field A_μ and of the matter fields ψ , as well as the definition of the covariant derivative, in a way that makes their generalization to other gauge groups almost obvious. Let us now consider the case where the abelian $U(1)$ group is replaced by some, generally non-abelian Lie group G , called the gauge group. In general this can be a product of so-called simple groups and $U(1)$ factors, like e.g. $SU(3) \times SU(2) \times U(1)$ for the standard model of electro-weak and strong interactions, or even some

more exotic groups like e.g. E_8 . For simplicity, one might simply think of $SU(N)$, the group of unitary $N \times N$ matrices with unit determinant. The structure of a Lie group is almost entirely captured by the commutators of its generators: the group elements close to the identity are always of the form $e^{i\theta^a t_a}$ with $\dim G$ small parameters θ^a and $\dim G$ generators t_a . The latter must satisfy closed (Lie algebra) commutation relations

$$[t_\alpha, t_\beta] = iC_{\alpha\beta}^\gamma t_\gamma, \quad (3.5)$$

with real structure constants $C_{\alpha\beta}^\gamma$ which satisfy the Jacobi identity $C_{[\alpha\beta}^\delta C_{\gamma]\delta}^\epsilon = 0$. It may happen that several groups like e.g. $SU(2)$ and $SO(3)$ have the same Lie algebra. In this case the groups are locally identical but not globally. We will be not so much concerned with the group elements themselves but rather with their representations. If we use a specific representation $\mathcal{R}$ we write $t_\alpha^\mathcal{R}$ or $(t_\alpha^\mathcal{R})_k^l$ for the $\dim \mathcal{R} \times \dim \mathcal{R}$ matrices of this representation. For compact Lie algebras (i.e. if $\text{tr } t_\alpha t_\beta$ is positive-definite), all finite dimensional representations are hermitian. This is the case of most interest in gauge theories and, hence, $(t_\alpha^\mathcal{R})^\dagger = t_\alpha^\mathcal{R}$. Of course, the corresponding representations of the group then are unitary. In this case one can also find a basis for the generators for which the $C_{\alpha\beta}^\gamma$ are antisymmetric in all 3 indices and one may then drop the distinction between upper and lower indices. Note that the Jacobi identity implies that the

$$(t_\alpha^{\text{adj}})_{\beta\gamma} = -iC_{\alpha\beta\gamma} \quad (3.6)$$

satisfy the algebra (3.5). This special representation is called the adjoint representation.

Consider a collection of matter fields ψ_l , $l = 1, \dots, r$ transforming in some r -dimensional representation $\mathcal{R}$ of G :

$$\psi_l(x) \rightarrow \psi'_l(x) = U^\mathcal{R}(g(x))_l^k \psi_k(x), \quad (3.7)$$

where $U^\mathcal{R}(g(x))$ is the $r \times r$ matrix associated to the group element $g(x) \in G$ in the representation $\mathcal{R}$. For $G = SU(N)$, the simplest example is the so-called fundamental (or “vector”) representation for which $U(g(x))_i^j = g(x)_i^j$ with i and j running simply from 1 to N . In the following, we will not write the indices explicitly so that we simply write

$$\psi(x) \rightarrow \psi'(x) = U^\mathcal{R}(g(x)) \psi(x), \quad (3.8)$$

with the understanding that ψ is a $\dim \mathcal{R}$ -dimensional column vector of matter fields. We want to construct a covariant derivative D_μ involving ∂_μ and some gauge field A_μ . In the abelian case the relevant combination was $A_\mu Q$ with Q the generator of the $U(1)$ group. Let us try

$$A_\mu = A_\mu^\alpha t_\alpha^\mathcal{R} \equiv A_\mu^\mathcal{R}, \quad (3.9)$$

since this can act on the $\dim \mathcal{R}$ -dimensional column vector of matter fields ψ , and a definition of the covariant derivative as

$$D_\mu^\mathcal{R} \psi = (\partial_\mu - iA_\mu^\mathcal{R}) \psi. \quad (3.10)$$

One often drops the superscript $\mathcal{R}$ on D_μ or A_μ , but one should remember that these quantities take values in the representation $\mathcal{R}$ of the gauge group determined by the matter fields. We want to determine the gauge transformation properties of A_μ in such a way that the covariant derivative simply transforms as

$$D_\mu \psi \rightarrow D'_\mu \psi' = U^\mathcal{R}(g) D_\mu \psi. \quad (3.11)$$

First, we look at the transformation of $\partial_\mu \psi$. To further simplify the notation we will simply assume that the representation is the fundamental representation where $U^{\mathcal{R}}(g(x)) = g(x)$, but one could replace everywhere g by $U^{\mathcal{R}}(g)$. Also we will drop the arguments x , although g , ψ and A_μ all depend on x . One has

$$\partial_\mu \psi \rightarrow \partial_\mu \psi' = \partial_\mu (g\psi) = g\partial_\mu \psi + (\partial_\mu g)\psi = g\left(\partial_\mu + (g^{-1}\partial_\mu g)\right)\psi \quad (3.12)$$

On the other hand,

$$-iA_\mu \psi \rightarrow -iA'_\mu \psi' = -iA'_\mu g\psi = -ig\left(g^{-1}A'_\mu g\right)\psi. \quad (3.13)$$

It is then obvious that the covariant derivative will transform covariantly, i.e. as in (3.12) provided A_μ transforms as $-ig^{-1}A'_\mu g + g^{-1}\partial_\mu g = -iA_\mu$, i.e. $A'_\mu = gA_\mu g^{-1} - i\partial_\mu g g^{-1}$. Since $\partial_\mu g^{-1} = -g^{-1}\partial_\mu g g^{-1}$ one has $-\partial_\mu g g^{-1} = g\partial_\mu g^{-1}$ so that one gets

$$A'_\mu = gA_\mu g^{-1} + ig\partial_\mu g^{-1} = g\left(A_\mu + i\partial_\mu\right)g^{-1}. \quad (3.14)$$

Note that for $G = U(1)$ and $g = e^{i\alpha Q}$ and the replacement $A_\mu \rightarrow QA_\mu$ one gets back (3.1). For a general matter representation one simply has

$$\begin{aligned} \psi' &= U^{\mathcal{R}}(g) \psi & , & & A'_\mu{}^{\mathcal{R}} &= U^{\mathcal{R}}(g) \left(A_\mu{}^{\mathcal{R}} + i\partial_\mu \right) U^{\mathcal{R}}(g^{-1}) \\ D_\mu{}^{\mathcal{R}} &= \partial_\mu - iA_\mu{}^{\mathcal{R}} & , & & (D_\mu{}^{\mathcal{R}}\psi)' &= U^{\mathcal{R}}(g) D_\mu{}^{\mathcal{R}} \psi . \end{aligned}$$

(3.15)

For a unitary representation, $(D_\mu \psi)^\dagger D_\mu \psi$ then is obviously invariant. This will be the ingredient to write gauge invariant matter kinetic terms.

It is often enough and simpler to consider only infinitesimal gauge transformations with $g(x) = e^{i\epsilon^\alpha t_\alpha}$ or $U^{\mathcal{R}}(g) = e^{i\epsilon^\alpha (x)t_\alpha^{\mathcal{R}}}$. It is convenient to define $\epsilon^{\mathcal{R}} = \epsilon^\alpha t_\alpha^{\mathcal{R}}$. Then the previous equations yield

$$\begin{aligned} \delta\psi &= \psi' - \psi & = & i\epsilon^{\mathcal{R}}\psi \\ \delta A_\mu{}^{\mathcal{R}} &= A'_\mu{}^{\mathcal{R}} - A_\mu{}^{\mathcal{R}} & = & \partial_\mu \epsilon^{\mathcal{R}} - i[A_\mu{}^{\mathcal{R}}, \epsilon^{\mathcal{R}}] . \end{aligned}$$

(3.16)

The latter equation reads in components

$$\delta A_\mu^\alpha = \partial_\mu \epsilon^\alpha + C_{\beta\gamma}^\alpha A_\mu^\beta \epsilon^\gamma, \quad (3.17)$$

which shows that, of course, the transformation of the gauge field components does not depend on the representation $\mathcal{R}$ of the matter fields. Using the generators of the adjoint representation (3.6) one may also write this as

$$\delta A_\mu^\alpha = (D_\mu^{\text{adj}})^\alpha_\gamma \epsilon^\gamma. \quad (3.18)$$

We still need to define the field strength $F_{\mu\nu}$. Since $D_\nu \psi$ transforms exactly as ψ , taking a further covariant derivative D_μ is defined in exactly the same way:

$$D_\mu^{\mathcal{R}} D_\nu^{\mathcal{R}} \psi = (\partial_\mu - iA_\mu^{\mathcal{R}})(\partial_\nu - iA_\nu^{\mathcal{R}})\psi = \partial_\mu \partial_\nu \psi - i(\partial_\mu A_\nu^{\mathcal{R}})\psi - iA_\nu^{\mathcal{R}} \partial_\mu \psi - iA_\mu^{\mathcal{R}} \partial_\nu \psi - A_\mu^{\mathcal{R}} A_\nu^{\mathcal{R}} \psi. \quad (3.19)$$

Antisymmetrizing in μ and ν eliminates all terms with a derivative of ψ . Hence

$$[D_\mu^{\mathcal{R}}, D_\nu^{\mathcal{R}}]\psi = -iF_{\mu\nu}^{\mathcal{R}}\psi, \quad (3.20)$$

with

$$F_{\mu\nu}^{\mathcal{R}} = \partial_\mu A_\nu^{\mathcal{R}} - \partial_\nu A_\mu^{\mathcal{R}} - i[A_\mu^{\mathcal{R}}, A_\nu^{\mathcal{R}}], \quad (3.21)$$

or with $F_{\mu,\nu}^{\mathcal{R}} = F_{\mu\nu}^\alpha t_\alpha^{\mathcal{R}}$:

$$F_{\mu\nu}^\alpha = \partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha + C_{\beta\gamma}^\alpha A_\mu^\beta A_\nu^\gamma. \quad (3.22)$$

It follows from (3.20) that $F_{\mu\nu}$ transforms covariantly:

$$F_{\mu\nu}^{\mathcal{R}} \rightarrow F_{\mu\nu}'^{\mathcal{R}} = U^{\mathcal{R}}(g) F_{\mu\nu}^{\mathcal{R}} U^{\mathcal{R}}(g^{-1}) \quad \text{or} \quad \delta F_{\mu\nu}^{\mathcal{R}} = -i[F_{\mu\nu}^{\mathcal{R}}, \epsilon^{\mathcal{R}}] \quad (3.23)$$

Obviously then, a gauge invariant scalar density is $\text{tr } F_{\mu\nu}^{\mathcal{R}} F^{\mathcal{R}\mu\nu}$. For compact groups we can choose a basis of generators such that

$$\text{tr } t_\alpha^{\mathcal{R}} t_\beta^{\mathcal{R}} = C_{\mathcal{R}} \delta_{\alpha\beta} \Rightarrow \text{tr } F_{\mu\nu}^{\mathcal{R}} F^{\mathcal{R}\mu\nu} = C_{\mathcal{R}} F_{\mu\nu}^\alpha F^{\alpha\mu\nu}. \quad (3.24)$$

With the appropriate normalization, this will be the gauge kinetic term, generalizing $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$. Note that, contrary to the abelian gauge theory of QED, in the non-abelian case, $F_{\mu\nu}^\alpha F^{\alpha\mu\nu}$ contains cubic and quartic terms in the gauge fields. This will lead to cubic and quartic vertices involving only gauge boson lines.

In QED, the coupling constant is given by the electric charge, i.e. the eigenvalue of the charge operator Q . By analogy we may include the Yang-Mills coupling constant g_{YM} in the generators t_α and their representations $t_\alpha^{\mathcal{R}}$. In this case $C_{\mathcal{R}} \sim g_{\text{YM}}^2$ and $C_{\beta\gamma}^\alpha \sim g_{\text{YM}}$. This is what [7] does, and we will follow this convention. Alternatively one can use conventionally normalized Lie algebra generators (where the $C_{\mathcal{R}}$ do not include any factors of g_{YM}). Then a factor of g_{YM} does appear explicitly in front of the A_μ in the covariant derivative, and similarly in front of the quadratic term in $F_{\mu\nu}$. These explicit factors of g_{YM} can then be removed by rescaling $A_\mu^{\mathcal{R}} \rightarrow \frac{1}{g_{\text{YM}}} A_\mu^{\mathcal{R}}$. The only appearance of the coupling then is a factor $\frac{1}{g_{\text{YM}}}$ in front of every $F_{\mu\nu}$. If the gauge group is simple, there is only a single coupling constant g_{YM} . However, in the general case where the gauge group is a product of simple and $U(1)$ factors, every simple or $U(1)$ factor can have its own coupling constant.

The unique Lorentz and gauge invariant Lagrangian quadratic in the field strength then is $\mathcal{L}_{\text{gauge}}[F_{\mu\nu}] = -\frac{1}{4}F_{\mu\nu}^\alpha F^{\alpha\mu\nu}$. Of course, there is one more possibility, $\theta_{\alpha\beta}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}^\alpha F_{\rho\sigma}^\beta$ but this is a total derivative. Hence the Lagrangian for matter and gauge fields is

$$\mathcal{L}[A_\mu, \psi] = -\frac{1}{4}F_{\mu\nu}^\alpha F^{\alpha\mu\nu} + \mathcal{L}_{\text{matter}}[\psi, D_\mu\psi], \quad (3.25)$$

where we collectively denoted all matter fields by ψ . If temporarily ψ denotes spin $\frac{1}{2}$ fields and ϕ scalars, an example of $\mathcal{L}_{\text{matter}}$ would be

$$\mathcal{L}_{\text{matter}}[\psi, D_\mu\psi, \phi, D_\mu\phi] = -\bar{\psi}(\gamma^\mu D_\mu + m)\psi - (D_\mu\phi)^\dagger D_\mu\phi - \tilde{m}^2\phi^\dagger\phi - V(\phi^\dagger\phi). \quad (3.26)$$

We only need the fact that the matter Lagrangian is gauge invariant. This will be the case if $\mathcal{L}_{\text{matter}}[\psi, \partial_\mu\psi, \phi, \partial_\mu\phi]$ is invariant under global (rigid) transformation by elements of G .

3.2 Quantization

Just as with QED, direct canonical quantization does not work due to the presence of constraints, in particular one has $\Pi_\alpha^\mu = F_\alpha^{\mu 0}$ and hence $\Pi_\alpha^0 = 0$ which is a primary constraint. Together with a secondary constraint which does not involve A_0 these are first class constraints. As usual, they have to be eliminated by a gauge choice. In the non-abelian case, a convenient choice is the axial gauge

$$A_3^\alpha = 0. \quad (3.27)$$

Then A_1^α and A_2^α become canonical variables, while A_0^α is given by the solution of the secondary constraint. One can then go through the canonical formulation of the functional integral in Hamiltonian form, as we also outlined for QED. To begin with one only integrates over the canonical fields A_i^α , $i = 1, 2$ and their conjugate fields Π_i^α , and after a few Gaussian integrations and some rewritings,¹⁶ one ends up with a Lagrangian version of the functional integral as

$$\langle T(\mathcal{O}_a \dots \mathcal{O}_N) \rangle_{\text{vac}} = \int \mathcal{D}A_\mu^\alpha \mathcal{D}\psi_l \prod_{x,\alpha} \delta(A_3^\alpha) \mathcal{O}_a \dots \mathcal{O}_N e^{i \int d^4x \mathcal{L}[A_\mu, \psi]} \quad (3.28)$$

where one integrates over all 4 components with the manifestly gauge invariant action $\int d^4x \mathcal{L}[A_\mu, \psi]$. Of course, There is also some irrelevant (infinite) normalization constant which we have suppressed. Now, the integration over A_3^α is effectively suppressed by the functional $\delta(A_3^\alpha)$ which imposes our gauge choice. Note that the derivation ensures that this gives a manifestly unitary theory. On the other hand, Lorentz invariance is not manifest.

One is usually interested in computing the vacuum expectation values of (time-ordered) *gauge invariant* operators $\mathcal{O}_1, \dots, \mathcal{O}_N$, although, at intermediate stages one may also compute and use non-gauge invariant objects as e.g. the gauge field propagator $\langle T(A_\mu^\alpha(x) A_\nu^\beta(y)) \rangle_{\text{vac}}$. Note that the measure $\mathcal{D}A_\mu^\alpha$ is gauge invariant: It is easy to show that the Jacobian for the gauge transformation (3.17) equals one.¹⁷ In the absence of (massless) chiral fermions one can also show that the matter measure is gauge invariant. However, if chiral fermions are present this is no longer guaranteed. If the matter measure is not invariant, one has a so-called anomaly, and the entire following discussion breaks down. It is thus important that, if present, gauge anomalies cancel between the contributions of the different chiral matter fields.

3.2.1 Faddeev-Popov method

We want to show that one can rewrite the expectation values (3.28) in different equivalent ways, corresponding to different gauge choices. First it is useful to rewrite (3.28) as

$$\langle T(\mathcal{O}_a \dots \mathcal{O}_N) \rangle_{\text{vac}} = \int \mathcal{D}A_\mu^\alpha \prod_{x,\alpha} \delta(A_3^\alpha) G[A_\mu^\alpha] \quad \text{with} \quad G[A_\mu^\alpha] = \int \mathcal{D}\psi_l \prod_a \mathcal{O}_a[A_\mu, \psi] e^{i \int d^4x \mathcal{L}[A_\mu, \psi]}, \quad (3.29)$$

the important point being that $G[A_\mu^\alpha]$ is gauge invariant. Indeed, with a gauge invariant matter measure and gauge invariant $\mathcal{O}_i$ we have $G[A_\mu^\alpha] = \int \mathcal{D}\psi_l \prod_a \mathcal{O}_a[A'_\mu, \psi] e^{i \int d^4x \mathcal{L}[A'_\mu, \psi]}$. Upon changing

¹⁶As before, in this section, we will simply write $\langle (\dots) \rangle_{\text{vac}}$ instead of $\langle \text{vac, out} | (\dots) | \text{vac, in} \rangle$.

¹⁷ $J_{\alpha\mu x, \beta\nu y} = \frac{\delta A_\mu'^\alpha(x)}{\delta A_\nu^\beta(y)} = \delta^{(4)}(x-y) \delta_\nu^\mu \left[\delta_\beta^\alpha + C_{\beta\gamma}^\alpha \epsilon^\gamma(x) \right]$, and due to the antisymmetry of the structure constants $C_{\alpha\gamma}^\alpha = 0$ and $\text{Det} J = \exp \text{Tr} \log J = \exp(\mathcal{O}(\epsilon^2))$.

the integration variable from ψ_l to the gauge transformed ψ'_l and using the gauge invariance of the measure and of the $\mathcal{O}_a$ this is

$$G[A'^\alpha_\mu] = \int \mathcal{D}\psi'_l \prod_a \mathcal{O}_a[A'_\mu, \psi'] e^{i \int d^4x \mathcal{L}[A'_\mu, \psi']} = \int \mathcal{D}\psi_l \prod_a \mathcal{O}_a[A_\mu, \psi] e^{i \int d^4x \mathcal{L}[A_\mu, \psi]} = G[A^\alpha_\mu] . \quad (3.30)$$

The expression of eq. (3.29) is of the general form

$$I_G = \int \mathcal{D}A_\mu B[f[A_\mu]] \text{Det}\mathcal{F}[A_\mu] G[A_\mu] \quad , \quad \mathcal{F}^\alpha_\beta(x, y) = \left. \frac{\delta f^\alpha[A'(x)]}{\delta \epsilon^\beta(y)} \right|_{\epsilon=0} , \quad (3.31)$$

where A' denotes the gauge field A , gauge transformed with parameter ϵ . Indeed, if we choose $f^\alpha[A] = A^\alpha_3$ and $B[f] = \prod_{x,\alpha} \delta(f^\alpha(x))$, we have $A'^\alpha_3 = \partial_3 \epsilon^\alpha$ (since $A_3 = 0$) so that $\mathcal{F}^\alpha_\beta(x, y) = \delta^\alpha_\beta \frac{\partial}{\partial x^3} \delta^{(4)}(x - y)$ and then $\text{Det}\mathcal{F}$ is just an irrelevant constant.

We want to show that I_G does not depend on the choice of the *gauge-fixing function* f or on B . Intuitively, what happens is the following: due to the gauge invariance of $G[A_\mu]$ there are many gauge-equivalent configurations and integrating over all A_μ would result in an infinite factor equal to the “volume” of a “gauge slice”. The gauge-fixing condition restricts the functional integration to exactly one configuration among the gauge equivalent ones. There are many different ways to do this and the factor $\text{Det}\mathcal{F}$ ensures the independence of the specific choice of f or B . This is the functional analogue of the well known fact that $\int dx \delta(f(x)) f'(x) g(x) = g(x_0)$ does not depend on f , as long as f has a single root x_0 .

Faddeev-Popov theorem : For gauge invariant $G[A_\mu]$, the functional integral I_G is independent of the gauge-fixing function f and depends on B only through an irrelevant overall factor.

To prove this theorem we first change variables from A_μ to some $A'_\mu \equiv A^g_\mu$ which we identify with A_μ gauge transformed by some $g(x) \in G$. Using the invariance of the measure and of $G[A_\mu]$ we get

$$I_G = \int \mathcal{D}A_\mu B[f[A^g_\mu]] \text{Det}\mathcal{F}[A^g_\mu] G[A_\mu] . \quad (3.32)$$

Since we only introduced g by a change of variables, this cannot depend on g . We multiply both sides by some weight function $\rho[g]$ such that $\int \mathcal{D}g \rho[g] = C_0$ is finite: ¹⁸

$$C_0 I_G = \int \mathcal{D}A_\mu G[A_\mu] H[A_\mu] \quad , \quad H[A_\mu] = \int \mathcal{D}g \rho[g] B[f[A^g_\mu]] \text{Det}\mathcal{F}[A^g_\mu] . \quad (3.33)$$

where now

$$\mathcal{F}^\alpha_\beta[A^g_\mu](x, y) = \left. \frac{\delta f^\alpha[(A^g)'](x)]}{\delta \epsilon^\beta(y)} \right|_{\epsilon=0} . \quad (3.34)$$

Here $(A^g)'$ is the result of gauge transforming A_μ first with $g = e^{i\theta^\alpha t_\alpha}$ and then with $e^{i\epsilon^\alpha t_\alpha}$. Since the gauge transformations form a group this amounts to gauge transforming A_μ with $\tilde{g} = e^{i\tilde{\theta}^\alpha t_\alpha} = e^{i\epsilon^\alpha t_\alpha} e^{i\theta^\alpha t_\alpha}$. Thus

$$\mathcal{F}^\alpha_\beta[A^g_\mu](x, y) = \left. \frac{\delta f^\alpha[A^{\tilde{g}(g,\epsilon)}(x)]}{\delta \epsilon^\beta(y)} \right|_{\epsilon=0} = \int d^4z \left. \frac{\delta f^\alpha[A^{\tilde{g}}(x)]}{\delta \tilde{\theta}^\gamma(z)} \right|_{\tilde{\theta}=\theta} \times \left. \frac{\delta \tilde{\theta}^\gamma(z)}{\delta \epsilon^\beta(y)} \right|_{\epsilon=0} . \quad (3.35)$$

¹⁸This is a tricky point. The integration over the full gauge group, i.e. the functional integration over all $g(x)$ typically would produce an infinite result C_0 . Instead, one should regulate this by discretizing space-time and working in finite volume and with a finite time interval so that $\mathcal{D}g$ becomes a finite product of integrations over the group at finitely many space-time points $\prod_i dg(x_i)$.

We let $\mathcal{G}[g]^\gamma_\beta(z, y) = \frac{\delta\tilde{\theta}^\gamma(z)}{\delta\epsilon^\beta(y)}\Big|_{\epsilon=0}$. This can be computed from the Baker-Campbell-Hausdorff formula for $i\tilde{\theta}^\alpha t_\alpha = \log(e^{i\epsilon^\alpha t_\alpha} e^{i\theta^\alpha t_\alpha})$ in which one only needs to keep the terms linear in ϵ :

$$\mathcal{G}[g]^\alpha_\beta(z, y) = \delta^{(4)}(z - y) \left[\delta^\alpha_\beta - \frac{1}{2} C_{\alpha\beta\gamma} \theta^\gamma + \frac{1}{12} C_{\alpha\kappa\delta} C_{\delta\gamma\beta} \theta^\kappa \theta^\gamma + \mathcal{O}(\theta^3) \right]. \quad (3.36)$$

Obviously, this only depends on the $\theta^\alpha(x)$ i.e. on $g(x)$. One has

$$\begin{aligned} \log \text{Det } \mathcal{G} &= \text{Tr } \log \mathcal{G} = \delta^{(4)}(0) \int d^4x \left[\frac{1}{12} C_{\alpha\kappa\delta} C_{\alpha\delta\gamma} \theta^\kappa \theta^\gamma - \frac{1}{2} \left(\frac{1}{2}\right)^2 C_{\alpha\delta\gamma} \theta^\gamma C_{\delta\alpha\kappa} \theta^\kappa + \mathcal{O}(\theta^3) \right] \\ &= \delta^{(4)}(0) \int d^4x \left[\frac{C_{\text{adj}}}{24} \theta^\alpha \theta^\alpha + \mathcal{O}(\theta^3) \right] \simeq \frac{\Lambda^4}{(2\pi)^4} \int d^4x \left[\frac{C_{\text{adj}}}{24} \theta^\alpha \theta^\alpha + \mathcal{O}(\theta^3) \right], \end{aligned} \quad (3.37)$$

where we used $C_{\alpha\gamma\delta} C_{\beta\gamma\delta} = C_{\text{adj}} \delta_{\alpha\beta}$ and the fact that $\delta^{(4)}(0) = \delta^{(4)}(x - x) = \int \frac{d^4p}{(2\pi)^4}$. One can work out the form of the higher-order terms in θ and show that $\rho[g] = (\text{Det } \mathcal{G})^{-1}$ yields a reasonable weight function. With this choice, (3.33) yields $(\mathcal{D}g \simeq \mathcal{D}\theta)$

$$H[A_\mu] = \int \mathcal{D}g \rho[g] B[f[A_\mu^g]] \text{Det} \frac{\delta f[A_\mu^g]}{\delta \theta} \text{Det } \mathcal{G} = \int \mathcal{D}\theta B[f[A_\mu^g]] \text{Det} \frac{\delta f[A_\mu^g]}{\delta \theta} = \int \mathcal{D}f B[f] \equiv C_B. \quad (3.38)$$

Obviously, this is independent of the choice of f and depends on B only through the constant C_B . Substituting this result into the first equation (3.33), the same is also true for I_G , i.e. I_G/C_B is independent of f and B , which was to be shown. Actually, there is an obvious problem here. Since $H[A_\mu]$ doesn't depend on A_μ , one sees that $C_0 I_G = \int \mathcal{D}A_\mu G[A_\mu] H[A_\mu] = C_B \int \mathcal{D}A_\mu G[A_\mu]$ is actually infinite as it overcounts all the gauge-equivalent configurations. The reason is that, as already noted in the previous footnote, the C_0 is also infinite since our $\mathcal{G}[g]$ yields a finite integral as we integrate over the group at fixed x , but we still get an infinite factor from the infinitely many space-time points. To do things properly, we should define the gauge group on a finite space-time lattice. This makes the integral over the gauge group finite. In particular it also gives a natural momentum cutoff Λ as already used above.

Since for the special choice $f^\alpha[A_\mu] = A_3^\alpha$ and $B[f] = \prod_{x,\alpha} \delta(f)$ (in this case $C_B = 1$) the integral I_G equals (3.29), the theorem tells us that the vacuum expectation value of a time-ordered product of gauge invariant operators can be evaluated with any choice of gauge-fixing function and function B , provided $\int \mathcal{D}f B[f]$ converges and f indeed fixes the gauge:

$$\langle T(\mathcal{O}_a \dots \mathcal{O}_N) \rangle_{\text{vac}} = \int \mathcal{D}A_\mu B[f[A_\mu]] \text{Det } \mathcal{F}[A_\mu] \int \mathcal{D}\psi_l \prod_a \mathcal{O}_a[A_\mu, \psi] e^{i \int d^4x \mathcal{L}[A_\mu, \psi]}, \quad (3.39)$$

where, as usual, we dropped all irrelevant constant normalization factors. This also shows that one can proceed in two steps : first do the functional integral over the matter fields without worrying about the gauge-fixing procedure and only then in the second step when doing the functional integral over the gauge field one needs to insert the gauge-fixing factors and Faddeev-Popov determinant.¹⁹

¹⁹Of course, this is only true if, as we assumed, the gauge-fixing function f only depends on A_μ and not on the matter fields.

3.2.2 Gauge-fixed action, ghosts and Feynman rules

Let us now make a convenient choice of gauge-fixing function f and function B which, in particular, will be manifestly Lorentz invariant:

$$f_\alpha = \partial_\mu A_\alpha^\mu \quad , \quad B[f] = \exp \left[-\frac{i}{2\xi} \int d^4x f_\alpha(x) f_\alpha(x) \right] . \quad (3.40)$$

As one can see from (3.39), the factor $B[f]$ just contributes an extra term

$$\mathcal{L}_{\text{g.f.}} = -\frac{1}{2\xi} f_\alpha f_\alpha \quad (3.41)$$

to the classical Lagrangian. One can then read off the gauge boson propagator $-i\Delta_{\alpha\mu,\beta,n}(p)$ from the quadratic part of $\mathcal{L} + \mathcal{L}_{\text{g.f.}}$ and finds

$$\Delta_{\alpha\mu,\beta,\nu}(p) = \delta_{\alpha\beta} \left(\eta_{\mu\nu} + (\xi - 1) \frac{p_\mu p_\nu}{p^2} \right) \frac{1}{p^2 - i\epsilon} . \quad (3.42)$$

Except for the extra $\delta_{\alpha\beta}$ this is just like the photon propagator, which should not be surprising, given our choice of f_α .

Next, we must evaluate $\mathcal{F}$ and compute its determinant. Since a fermionic gaussian integral equals the determinant of the quadratic form, it is a convenient trick to rewrite $\text{Det}\mathcal{F}$ as

$$\text{Det}\mathcal{F} = \int \mathcal{D}\omega \mathcal{D}\omega^* \exp \left[i \int d^4x d^4y \omega_\alpha^*(x) \mathcal{F}_{\alpha,\beta}(x,y) \omega_\beta(y) \right] , \quad (3.43)$$

where ω_α and ω_α^* are anticommuting, i.e. fermionic fields. On the other hand, given the nature of $\mathcal{F}$ we will soon see that they cannot be spin $\frac{1}{2}$ fields but must be scalars. Moreover, ω_α and ω_α^* are independent *real* scalars. They do not obey the usual spin statistics relation and hence are named ghosts. However, this is not a problem since we do not want to obtain any Lorentz invariant S -matrix for scattering of these ghost fields. More precisely, ω_α is called a ghost and ω_α^* an anti-ghost. Note that the ghost and anti-ghost carry an index α just like the gauge field A_μ . One sometimes says that they are in the adjoint representation of G , just as A_μ . One has $f_\alpha[A'_\mu] = f_\alpha[A_\mu] + \partial_\mu (\partial^\mu \epsilon_\alpha + C_{\gamma\beta}^\alpha A_\mu^\gamma \epsilon^\beta)$ and hence

$$\mathcal{F}_{\alpha,\beta}(x,y) = \frac{\delta f_\alpha[A'_\mu](x)}{\delta \epsilon^\beta(y)} = \frac{\partial}{\partial x^\mu} \left(\frac{\partial}{\partial x_\mu} \delta_{\alpha\beta} + C_{\gamma\beta}^\alpha A_\mu^\gamma(x) \right) \delta^{(4)}(x-y) . \quad (3.44)$$

Using (3.43) we get

$$\text{Det}\mathcal{F} = \int \mathcal{D}\omega \mathcal{D}\omega^* \exp \left[i \int d^4x \mathcal{L}_{\text{ghost}}(x) \right] , \quad \mathcal{L}_{\text{ghost}} = -\partial_\mu \omega_\alpha^* \partial^\mu \omega_\alpha - C_{\gamma\beta}^\alpha \partial_\mu \omega_\alpha^* A_\mu^\gamma \omega_\beta . \quad (3.45)$$

Thus the effect of the so-called Faddeev-Popov determinant $\text{Det}\mathcal{F}$ is to add the ghost Lagrangian $\mathcal{L}_{\text{ghost}}$ and the gauge-fixing Lagrangian $\mathcal{L}_{\text{g.f.}}$ to the classical Lagrangian $\mathcal{L}$:

$$\mathcal{L}_{\text{mod}} = \mathcal{L} + \mathcal{L}_{\text{g.f.}} + \mathcal{L}_{\text{ghost}} . \quad (3.46)$$

Note that for an abelian theory like QED, the structure constants $C_{\beta\gamma}^\alpha$ vanish and the ghosts do not couple to any of the other fields. Equivalently then, their functional integral only leads to the

determinant of $\partial_\mu \partial^\mu$ which is an irrelevant constant. This is why we did not have to bother about the ghosts in QED. In the non-abelian gauge theory, with our choice of gauge-fixing function, the ghosts do couple to the gauge fields and do make important contributions to loop diagrams.²⁰ Their propagator $\frac{-i}{(2\pi)^4} \Delta_{\alpha\beta}(p)$ can be read off from (3.46) to be the same as for a massless spin-0 fermion:

$$\Delta_{\alpha\beta} = \frac{\delta_{\alpha\beta}}{p^2 - i\epsilon} , \quad (3.47)$$

while the ghost - anti-ghost - gauge field vertex involves the structure constant $C_{\beta\gamma}^\alpha$ as well as a factor of momentum. More precisely, taking all momenta as incoming (q for the ghost with label β , p for the anti-ghost with label α and k for the gauge boson with labels μ and γ), one reads from (3.45) that this vertex contributes

$$i(2\pi)^4 \delta^{(4)}(p + q + k) i p_\mu C_{\alpha\beta\gamma} . \quad (3.48)$$

We already gave the gauge boson propagator. The cubic and quartic gauge boson couplings simply follow from the cubic and quartic terms in $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$. Again, with all momenta taken as incoming, the vertex for the coupling of three gauge bosons with (p, μ, α) , (q, ν, β) and (k, ρ, γ) is

$$i(2\pi)^4 \delta^{(4)}(p + q + k) (-i C_{\alpha\beta\gamma}) \left[p_\nu \eta_{\mu\rho} - p_\rho \eta_{\mu\nu} + q_\rho \eta_{\nu\mu} - q_\mu \eta_{\nu\rho} + k_\mu \eta_{\rho\nu} - k_\nu \eta_{\rho\mu} \right] , \quad (3.49)$$

while the vertex for the coupling of four gauge bosons (the fourth one having (l, λ, δ)) is

$$i(2\pi)^4 \delta^{(4)}(p + q + k + l) \left[\begin{aligned} & -C_{\epsilon\alpha\beta} C_{\epsilon\gamma\delta} (\eta_{\mu\rho} \eta_{\nu\lambda} - \eta_{\mu\lambda} \eta_{\nu\rho}) \\ & -C_{\epsilon\alpha\gamma} C_{\epsilon\delta\beta} (\eta_{\mu\lambda} \eta_{\rho\nu} - \eta_{\mu\nu} \eta_{\lambda\rho}) \\ & -C_{\epsilon\alpha\delta} C_{\epsilon\beta\gamma} (\eta_{\mu\nu} \eta_{\rho\lambda} - \eta_{\mu\rho} \eta_{\lambda\nu}) \end{aligned} \right] . \quad (3.50)$$

Note that the Lagrangian preserves the ghost-number which translates into the fact that every vertex with one incoming ghost line also has exactly one outgoing ghost line.

3.2.3 BRST symmetry

One can now start computing Feynman diagrams and generating functionals using the Lagrangian $\mathcal{L}_{\text{mod}} = \mathcal{L} + \mathcal{L}_{\text{g.f.}} + \mathcal{L}_{\text{ghost}}$ of (3.46) and integrating over the matter fields, gauge and ghost fields. Note that $\mathcal{L}_{\text{mod}}$ no longer is gauge invariant: this was the whole point about gauge-fixing. All terms in $\mathcal{L}_{\text{mod}}$ have $\Delta = 0$ and hence this Lagrangian is *renormalizable by power-counting* as discussed in section 2: there are only finitely many divergent Green's functions and they are made finite by the addition of finitely many counterterms with their coefficients fixed at any given order in perturbation theory. Moreover, these counterterms themselves all have $\Delta_{\text{ct}} \geq 0$ and do not upset the renormalizability.

However, we want more than this: the counterterms should be of the form of the initial terms in the Lagrangian. In particular, one should still be able to interpret $\mathcal{L}_{\text{mod}} + \mathcal{L}_{\text{ct}}$ as arising from the gauge-fixing of some gauge invariant Lagrangian, but now with renormalized parameters. We must find out what is the remnant of the gauge invariance of the original classical Lagrangian. This turns out to be the BRST-symmetry.

²⁰Of course, in axial gauge, $A_3 = 0$, one finds $\mathcal{F}_{\alpha\beta}(x, y) = \frac{\delta}{\delta x^3} \delta^{(4)}(x - y) \delta_{\alpha\beta}$ and the ghost Lagrangian simply is $-\partial_3 \omega_\alpha^* \omega_\alpha$ so that the ghosts again decouple.

Let us first do one more rewriting of our Lagrangian. Introducing an auxiliary field h_α we can rewrite the gauge-fixing term as $\frac{\xi}{2}h_\alpha h_\alpha + h_\alpha f_\alpha$ since doing the gaussian integration over h_α reproduces $-\frac{1}{2\xi}f_\alpha f_\alpha$. Thus, our starting point is

$$\mathcal{L}_{\text{new}} = \mathcal{L} + \frac{\xi}{2}h_\alpha h_\alpha + h_\alpha f_\alpha + \omega_\alpha^* \rho_\alpha , \quad (3.51)$$

where

$$\rho_\alpha(x) = \partial_\mu \partial^\mu \omega_\alpha(x) + C_{\beta\gamma}{}^\alpha \partial_\mu (A_\beta^\mu(x) \omega_\gamma(x)) = \int d^4y \mathcal{F}_{\alpha\beta}(x, y) \omega_\beta(y) . \quad (3.52)$$

The BRST symmetry is defined to act on the “ordinary” fields, i.e. the matter and gauge fields, just like an ordinary infinitesimal gauge transformation but with the (real) parameter ϵ_α replaced by the (real) ghost field ω_α . It is thus a fermionic symmetry that increases the ghost-number by one unit. To begin with, we simply set $\epsilon_\alpha(x) = \theta \omega_\alpha(x)$ with θ an anticommuting parameter ($\theta\omega = -\omega\theta$, $\theta A_\mu = A_\mu\theta$, etc, and $\theta\theta = 0$). Then

$$\begin{aligned} \delta_\theta A_\mu^\alpha &= \theta (\partial_\mu \omega^\alpha + C_{\beta\gamma}{}^\alpha A_\mu^\beta \omega^\gamma) \\ \delta_\theta \psi &= i \theta \omega^\alpha t_\alpha^\mathcal{R} \psi \\ \delta_\theta \psi^\dagger &= -i \theta \omega^\alpha \psi^\dagger t_\alpha^\mathcal{R} = i \theta \psi^\dagger t_\alpha^\mathcal{R} \omega^\alpha . \end{aligned} \quad (3.53)$$

This is conveniently rewritten by defining a fermionic operator s such that for any functional F we simply let

$$\delta_\theta F = \theta s F . \quad (3.54)$$

Consistency with the fermionic character requires

$$s(FG) = (sF)G \pm F sG , \quad (3.55)$$

with a minus sign if F is fermionic (contains an odd number of anticommuting fields). We write furthermore $\omega^\mathcal{R} = \omega^\alpha t_\alpha^\mathcal{R}$, so that we can reformulate (3.53) as $s A_\mu^\mathcal{R} = D_\mu \omega^\mathcal{R}$ and $s \psi = i \omega^\mathcal{R} \psi$ or, dropping the superscript $\mathcal{R}$, simply

$$\begin{aligned} s A_\mu &= \partial_\mu \omega - i[A_\mu, \omega] \\ s \psi &= i \omega \psi \quad , \quad s \psi^\dagger = i \psi^\dagger \omega . \end{aligned}$$

(3.56)

Note that $s\psi^\dagger$ is not the hermitian conjugate of $s\psi$, but rather $(s\psi)^\dagger = (i\omega\psi)^\dagger = -i\psi^\dagger\omega = -s\psi^\dagger$. More generally, one sees that $(sF)^\dagger = \mp sF^\dagger$ with a minus sign if F is fermionic. Let us complete (3.56) by the rules how s acts on the ghosts, chosen in such a way that s is a nilpotent operation, i.e. $s^2 = 0$ on all fields. Choosing

$$s \omega = i \omega \omega \quad , \quad s \omega^* = -h \quad , \quad s h = 0 ,$$

(3.57)

where h is the auxiliary field introduced above, we have

$$\begin{aligned}
s^2 A_\mu &= s(\partial_\mu \omega - i[A_\mu, \omega]) = \partial_\mu(s\omega) - i(sA_\mu)\omega - iA_\mu s\omega + i(s\omega)A_\mu - i\omega sA_\mu \\
&= i\partial_\mu(\omega\omega) - i(\partial_\mu \omega - iA_\mu \omega + i\omega A_\mu)\omega - iA_\mu i\omega\omega - \omega\omega A_\mu - i\omega(\partial_\mu \omega - i\omega A_\mu \omega + i\omega A_\mu) = 0 , \\
s^2 \psi &= s(i\omega\psi) = i(s\omega)\psi - i\omega s\psi = -\omega\omega\psi + \omega\omega\psi = 0 , \\
s^2 \omega &= s(i\omega\omega) = i(s\omega)\omega - i\omega(s\omega) = -\omega\omega\omega + \omega\omega\omega = 0 , \\
s^2 \omega^* &= s(-h) = 0 ,
\end{aligned} \tag{3.58}$$

which shows that the BRST operator s is nilpotent:

$$s^2 = 0 . \tag{3.59}$$

It is now easy to show that $\mathcal{L}_{\text{new}}$ is BRST invariant. First note that on functionals of A_μ and the matter fields ψ only the BRST transformation is simply a gauge transformation with ϵ replaced by ω . Thus any gauge invariant functional of the gauge and matter fields only automatically also is BRST invariant. Hence

$$s\mathcal{L} = 0 . \tag{3.60}$$

Next note that $\mathcal{F}\epsilon$ was defined as the gauge variation of f (which only depends on A_μ) and hence

$$\rho_\alpha = sf_\alpha \Rightarrow s\rho = 0 . \tag{3.61}$$

It follows that

$$s\mathcal{L}_{\text{new}} = s\left(\frac{\xi}{2}h_\alpha h_\alpha + h_\alpha f_\alpha + \omega_\alpha^* \rho_\alpha\right) = h_\alpha sf_\alpha + (s\omega_\alpha^*)\rho_\alpha = h_\alpha \rho_\alpha - h_\alpha \rho_\alpha = 0 . \tag{3.62}$$

Actually one has

$$\mathcal{L}_{\text{new}} = \mathcal{L} + s\Psi , \quad \Psi = -\omega_\alpha^* f_\alpha - \frac{\xi}{2}\omega_\alpha^* h_\alpha , \tag{3.63}$$

with $\mathcal{L}$ being the gauge invariant Lagrangian of the gauge and matter fields only. In this form, BRST invariance is obvious. This also suggests how to obtain more general gauge-fixings: Any $\mathcal{L}_{\text{new}}$ of this form with an arbitrary functional Ψ of ghost-number -1 will provide a BRST invariant starting point for quantizing the gauge theory (provided the quadratic term in the gauge fields is non-degenerate so that one can define a propagator). The most general local Ψ of ghost number -1 , dimension 3 (and invariant under Lorentz and global “gauge” transformations²¹) necessarily is of the form

$$\Psi = -\omega_a^* \beta^{ab} \partial_\mu A_b^\mu - c_1 \omega_a^* h_a - c_2 \alpha_{abc} \omega_a^* \omega_b \omega_c . \tag{3.64}$$

with some globally G -invariant tensors β_{ab} and α_{abc} .

Below, we will use this BRST invariance to show that all counterterms also must respect this BRST invariance (but with a renormalized coupling constant which we had hidden in the normalization of the Lie algebra generators). Hence

$$\mathcal{L}_{\text{new}} + \mathcal{L}_{\text{ct}} = \tilde{\mathcal{L}} + s\tilde{\Psi} , \tag{3.65}$$

respects the same symmetries as the original $\mathcal{L}_{\text{new}}$ with a renormalized Yang-Mills coupling constant and yields finite Green’s functions.

²¹Global gauge transformations or, more precisely, global G -transformations correspond to gauge transformations with constant ϵ^a , and invariance under such global G -transformations just means that all gauge group indices $a, b, \dots$ are contracted with G -invariant tensors, in particular with δ_{ab} . Note that the Dirac matter fields cannot appear since the Lorentz invariant of lowest dimension is $\bar{\psi}\psi$ which already has dimension 3 but ghost number 0. With scalar matter fields ϕ , however, one could e.g. have a term $\omega_a^* \phi^\dagger t_a \phi$, corresponding to a gauge-fixing function $f_a = \phi^\dagger t_a \phi$ (“unitary gauge”).

3.3 BRST cohomology

We already noted that the BRST symmetry is nilpotent, i.e. $s^2 = 0$. Obviously also, it increases the ghost-number by one unit.

At an algebraic level, this is quite similar to the behavior of the exterior derivative $d = dx^\mu \partial_\mu$ acting on differential forms²² of degree p and yielding a differential form of degree $p + 1$. As is well known, one has $d^2 \equiv dd = 0$ (which just states that $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$). A p -form $\xi^{(p)}$ is called closed if $d\xi^{(p)} = 0$. If there exists a (globally well-defined) $(p - 1)$ -form $\zeta^{(p-1)}$ such that $\xi^{(p)} = d\zeta^{(p-1)}$ then $\xi^{(p)}$ is called exact. Since $d^2 = 0$, obviously every exact form is also closed. It is then interesting to find out which p -forms are closed without being exact. As an example consider a space-time with space being just the two-dimensional sphere with coordinates θ and ϕ defined in the usual way. Then $d\phi$ is well-defined everywhere except at $\theta = 0$ or $\theta = \pi$ (north and south pole), and the 2-form $\Omega = \sin \theta d\theta \wedge d\phi$ is well-defined everywhere. One has $d\Omega = 0$, so it is closed. Although one has $\Omega = d(-\cos \theta d\phi)$ or $\Omega = d(-\sin \theta d\theta \phi)$, neither $\cos \theta d\phi$ nor $\sin \theta d\theta \phi$ are well-defined everywhere on the sphere and one finds that $\Omega \neq d\zeta^{(1)}$. Hence Ω is closed but not exact. Up to multiplication by a constant²³ this is the only closed and non-exact 2-form on the sphere. Actually, Ω is the volume-form on the sphere, and the volume form on any compact manifold is always closed but not exact. For a given (compact) manifold $\mathcal{M}$, the vector space of closed p -forms that are not exact is called the p^{th} de Rham cohomology and is denoted $H^{(p)}(\mathcal{M})$.

Similarly, one defines the classes of BRST-closed functionals and BRST-exact functionals as follows:

The functional $F[A_\mu, \psi, \omega, \omega^*, h]$ is BRST closed if $sF = 0$,
it is BRST exact if there exists a functional G such that $F = sG$.

(3.66)

Of course, a BRST-closed functional is a BRST invariant functional. Moreover, the ghost-number of a given monomial of the fields is defined as the number of ghost-fields minus the number of anti-ghost fields. The Lagrangian $\mathcal{L}_{\text{new}}$ e.g. has ghost-number zero. The BRST operator s always increases the ghost-number by one unit. Since one cannot have cancellations between terms of different ghost-numbers it follows that one can define the space of BRST-closed functionals of a fixed ghost-number and, similarly the space of BRST-exact functionals of a fixed ghost-number. Hence, the ghost-number plays a role analogous to the degree p of the differential forms. In particular, one defines the BRST-cohomology classes at ghost-number n :

The BRST-cohomology at ghost-number n is given by
the BRST-closed functionals of ghost-number n , modulo the BRST-exact functionals.

(3.67)

We have seen above that the gauge invariant Lagrangian $\mathcal{L}[A, \psi]$ is BRST invariant, i.e. $s\mathcal{L}[A, \psi] = 0$, so it is BRST closed. Moreover, it is of ghost-number zero. The gauge-fixing and ghost terms are

²²In exterior calculus, one defines an antisymmetric “wedge product” so that $dx^\mu \wedge dx^\nu = -dx^\nu \wedge dx^\mu$. For every rank- p antisymmetric tensor field $\xi_{\mu_1 \mu_2 \dots \mu_p}(x)$ one can then define a p -form as $\xi^{(p)} = \frac{1}{p!} \xi_{\mu_1 \mu_2 \dots \mu_p} dx^{\mu_1} \wedge dx^{\mu_2} \wedge \dots \wedge dx^{\mu_p}$. In d dimensions, the maximal degree of a form is $p = d$.

²³For example, $\cos \theta \Omega = \cos \theta \sin \theta d\theta \wedge d\phi = da$ with $a = \frac{1}{2} \sin^2 \theta d\phi$ a well-defined one-form. Hence $\cos \theta \Omega$ is exact.

of the form $s\Psi$ with Ψ of ghost-number -1 , i.e. are BRST-exact terms. Since correlation functions of gauge invariant operators do not depend on Ψ (as long as it provides some gauge-fixing), the physics determined by $\mathcal{L}_{\text{new}} = \mathcal{L}[A, \psi] + s\Psi$ depends only on the BRST-closed $\mathcal{L}$ and is independent of the BRST-exact terms $s\Psi$: it only depends on the BRST cohomology class of $\mathcal{L}_{\text{new}}$. Later-on we will show that the BRST-cohomology at ghost-number zero is precisely given by the gauge invariant functionals of A_μ and ψ (and $\psi^\dagger$) only, i.e. independent of the ghost, anti-ghost and h -fields.

BRST-charge :

One can introduce a fermionic BRST charge operator (acting on a “Hilbert space”) by demanding that for any (Heisenberg picture) field operator Φ one has

$$\delta_\theta \Phi(x) = i[\theta Q, \Phi(x)] . \quad (3.68)$$

The right-hand-side equals $i\theta[Q, \Phi(x)]_\mp$, i.e. a commutator if Φ is bosonic and an anti-commutator if Φ is fermionic, while the left-hand-side equals $\theta s\Phi(x)$. Hence the BRST-charge operator must have the following (anti)commutation relations with the field operators:

$$s\Phi(x) = i[Q, \Phi(x)]_\mp . \quad (3.69)$$

By taking the hermitian conjugate of either (3.68) (θ is real) or (3.69) and comparing with the corresponding relations for $\Phi^\dagger$, one sees that $Q^\dagger = -Q$. This then implies $(\theta Q)^\dagger = Q^\dagger \theta = -Q\theta = \theta Q$ as expected for a symmetry generator. Using the Jacobi identity, one finds that

$$0 = -s s \Phi = -is[Q, \Phi]_\mp = [Q, [Q, \Phi]_\mp]_\pm = \frac{1}{2}[[Q, Q]_+, \Phi]_- = [Q^2, \Phi]_- \quad \forall \Phi . \quad (3.70)$$

An operator that commutes with all fields is either the identity or has to vanish. Since Q increases the ghost-number by one, the first option is excluded and one concludes

$Q^2 = 0 .$

(3.71)

The “Hilbert space” on which Q acts is some “big” space of states which must include the Hilbert space of physical states, but also states including an arbitrary number of ghost and anti-ghost excitations, as well as non-physical polarization states of the gauge bosons. This is the space which naturally arises upon the Faddeev-Popov quantization of the gauge theory. In particular, this is not a Hilbert space in the strict mathematical sense since the inner product cannot be positive definite: for any state $|\gamma\rangle \neq 0$ the state $Q|\gamma\rangle$ has zero norm. One must then characterize the Hilbert space of physical states. This can be conveniently done using this BRST charge as follows. Matrix elements between physical states must be gauge invariant, i.e. independent of the “gauge-fixing functional” Ψ . Hence, under an infinitesimal variation $\Psi(x) \rightarrow \Psi(x) + \delta\Psi(x)$, the gauge-fixed Lagrangian changes by $\delta\mathcal{L}_{\text{new}} = s\delta\Psi(x)$ and any matrix element changes by

$$\delta\langle\alpha|\beta\rangle = i\langle\alpha|s\delta\Psi|\beta\rangle = -\langle\alpha|[Q, \delta\Psi]_+|\beta\rangle = -\langle\alpha|Q\delta\Psi|\beta\rangle - \langle\alpha|\delta\Psi Q|\beta\rangle . \quad (3.72)$$

If $\langle\alpha|$ and $|\beta\rangle$ are physical states this should vanish. Since $\delta\Psi(x)$ is arbitrary one concludes $Q|\beta\rangle = 0$ and $\langle\alpha|Q = 0$. Hence, physical states $|\text{phys}\rangle$ must obey

$$Q|\text{phys}\rangle = 0 . \quad (3.73)$$

Also, changing $|\beta\rangle \rightarrow |\beta\rangle + Q|\gamma\rangle$ does not change $\langle\alpha|\beta\rangle$ if $\langle\alpha|$ is physical since $\langle\alpha|Q|\gamma\rangle = 0$. Moreover, a physical state should have ghost-number zero. In conclusion, physical states are determined by $Q|\text{phys}\rangle = 0$ subject to the equivalence relation $|\text{phys}\rangle \simeq |\text{phys}\rangle + Q|\gamma\rangle$ and the condition of having ghost-number zero:

Physical states are given by the ghost-number zero cohomology class of the BRST-operator Q .

(3.74)

Of course, we have only shown that this cohomology class contains the physical Hilbert space, i.e. that the above conditions are necessary. However, one can also show that they are sufficient and thus (3.74) exactly defines the physical Hilbert space. Note that in any covariant quantization, the modes of the fields A_0^a , when acting on the vacuum, generate states of negative norm (since $\eta_{00} = -1$). One can then also show that the space of physical state as defined by (3.74) has a positive definite norm.

Remarks :

- The Faddeev-Popov procedure always leads to a Lagrangian $\mathcal{L}_{\text{new}}$ that is bilinear in the ghost and anti-ghost fields. (Recall that the ghosts appeared from expressing the Faddeev-Popov determinant $\text{Det } \mathcal{F}$ as an integral over ω and ω^* of $e^{i \int \omega^* \mathcal{F} \omega}$.) Then, calling these fields the bare fields and rewriting $\mathcal{L}_{\text{new}}$ in terms of renormalized fields will generate various counterterms of the same form as the original terms contained in $\mathcal{L}_{\text{new}}$. In particular, one will only get counterterms that are at most bilinear in the ghost and anti-ghost fields. This turns out to be sufficient with the gauge choice $f_\alpha = \partial_\mu A_\alpha^\mu$, but for more general choices of f_α one might need counterterms that involve two ghost and two anti-ghost fields. Although one does not have to worry about diagrams with 4 external (anti)ghost lines, such diagrams can well appear as divergent subgraphs, requiring a corresponding counterterm.
- One can consider more general gauge-fixing functionals that do not rely on the Faddeev-Popov procedure. As we have seen, all that is needed for BRST invariance is to define some BRST-operation s such that s acts on A_μ and ψ as a gauge transformation with the parameter replaced by a ghost field, and with the action on the ghost and other fields (anti-ghost, h -field, and possibly others) defined such that $s^2 = 0$. Then for any $\tilde{\Psi}$ of ghost-number -1 and gauge invariant $\mathcal{L}$, a BRST-invariant Lagrangian is $\tilde{\mathcal{L}}_{\text{new}} = \mathcal{L} + s\tilde{\Psi}$. Then just as before, there is a corresponding BRST charge Q that defines the physical states as the zero ghost-number cohomology class and matrix elements between physical states (in particular also vacuum expectation values of time-ordered products of gauge invariant operators) do not depend on the choice of $\tilde{\Psi}$. In particular, they are the same as with the Ψ from the Faddeev-Popov procedure and thus the same as in axial gauge.
- One can prove independently of comparison with any particular gauge, that the space of physical states, defined as the zero ghost-number cohomology of Q , has a positive definite norm, contains no ghosts or anti-ghosts and has a unitary S -matrix.
- Finally let us just mention that this BRST formalism can be rather straightforwardly extended to other local symmetries as appear e.g. in general relativity or in string theory. If the natural

formulation of these symmetries is “too large” in the sense that one has introduced too many “gauge” parameters and actually not all symmetries are independent, one has to introduce “ghosts of ghosts”. In all these setting, the BRST operator always increases the ghost-number by one unit.

We have already mentioned that the BRST invariant functionals of ghost-number zero are just the gauge invariant functionals of A_μ and the matter fields, up to adding BRST-exact terms. Let us now prove this result.

Cohomology theorem :

The zero ghost-number cohomology consists of gauge invariant functionals of A_μ and ψ (and $\psi^\dagger$) only, i.e. the most general BRST invariant functional of ghost-number zero is of the form $I = I_0[A_\mu, \psi] + s\Psi[A_\mu, \psi, \omega, \omega^*, h]$.

The *proof* is relatively simple: suppose $sI = 0$. Write $I = \sum_{N=0}^{\infty} I_N$ where I_N contains all terms that have a total number of fields ω^* and h equal to N . (Of course, we do not allow negative powers of the fields.) Since $s\omega^* = -h$ and $sh = 0$, s does not change this total number N and one cannot have any cancellations between the different sI_N . Hence $sI_N = 0$ for all N separately. Introduce $t = \omega_\alpha^* \frac{\delta}{\delta h_\alpha}$. One may similarly write $s = -h_\beta \frac{\delta}{\delta \omega_\beta^*} + \dots$ where the unwritten terms do not involve ω^* or h . It follows that $st + ts = -\hat{N}$ where $\hat{N} = \omega_\alpha^* \frac{\delta}{\delta \omega_\alpha^*} + h_\alpha \frac{\delta}{\delta h_\alpha}$ is such that $\hat{N}I_N = NI_N$. Thus

$$-NI_N = -\hat{N}I_N = (st + ts)I_N = stI_N, \quad (3.75)$$

and we conclude that for every $N \neq 0$ one has $I_N = s(-\frac{1}{N}tI_N)$, i.e. I_N is BRST-exact. Hence, $I = I_0 + s\Psi$ with $\Psi = \sum_{N=1}^{\infty} (-\frac{1}{N}tI_N)$. Now, I_0 contains no h and no ω^* and, having ghost-number zero, it cannot contain any ω either. Thus $I_0 = I_0[A_\mu, \psi]$, as was to be proven. Finally, I_0 cannot be BRST exact. Indeed, it is easy to see that a BRST exact functional of ghost number zero must contain at least either h or $\omega\omega^*$.

3.4 Exercices

3.4.1 Some one-loop vertex functions

Use the action and Feynman rules as given above.

- Draw all Feynman diagrams contributing to the 1PI gauge-boson-ghost-antighost vertex function $\Gamma_{\alpha\beta\gamma}^\mu(p_1, p_2, p_3)$ at one loop (2 diagrams). For each diagram determine the superficial degree of divergence.
- Same question for the 1PI three gauge-boson vertex function $\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(p_1, p_2, p_3)$ at one loop. (There are 8 diagrams if one counts as different diagrams two fermion loops with opposite orientations of the arrows.)
- Same question for the 1PI vertex function for two ghosts and two antighosts $\Gamma_{\alpha\beta\gamma\delta}(p_1, p_2, p_3, p_4)$. (There are 4 diagrams. Pay attention to the relative signs between these diagrams.)

3.4.2 More general gauge-fixing functions

Instead of the gauge-fixing function $f^\alpha = \partial_\mu A^{\alpha\mu}$ now choose $f^\alpha = (D_\mu A^\mu)^\alpha$.

- a) Give the explicit form of the corresponding $\mathcal{L}_{\text{gf}} + \mathcal{L}_{\text{gh}}$ and deduce the corresponding vertices involving the ghost and antighost fields. In particular give the vertex between the ghost, antighost and two gauge bosons.
- b) Show that, in addition to the diagrams present in c) of the previous exercise, there are new types of Feynman diagrams contributing to the one-loop 4-point 2 ghost- 2 antighost vertex function $\Gamma_{\alpha\beta\gamma\delta}(p_1, p_2, p_3, p_4)$. Determine their superficial degrees of divergence. Does it come as a surprise if you find the same answer as in c) of the previous exercise?

3.4.3 Two ghost - two antighost counterterms

In the Lagrangian used in exercise 3.4.1 there are no terms $\sim \omega^* \omega^* \omega \omega$, and thus we expect that one should not need such counterterms either. This would mean that the vertex function $\Gamma_{\alpha\beta\gamma\delta}(p_1, p_2, p_3, p_4)$ given in c) of this exercise should not be divergent, although it had a superficial degree of divergence equal to 0. Write the expressions for the 4 corresponding Feynman diagrams and try to argue that there are cancellations between the diagrams.

4 Renormalization of non-abelian gauge theories

4.1 Slavnov-Taylor identities and Zinn-Justin equation

4.1.1 Slavnov-Taylor identities

Recall that the vacuum expectation values of time-ordered products of field operators can be obtained from the generating functional

$$Z[J] = \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}\omega \mathcal{D}\omega^* \mathcal{D}h \exp \left[i \int d^4x (\mathcal{L}_{\text{new}} + \tilde{\chi}^n J_n) \right], \quad (4.1)$$

where, to simplify the notation, $\tilde{\chi}^n$ stands collectively for any of the fields A_μ , ψ , $\bar{\psi}$, ω , ω^* or h . Of course, we do not really need to compute ghost correlation functions to get S -matrix elements between physical states, but they certainly can and do appear in subdiagrams. Also, the form of $\mathcal{L}_{\text{new}}$ originally was derived only for time-ordered products of gauge invariant operators, but we can take (4.1) as a definition for the gauge-dependent vacuum expectation values of time-ordered products of the gauge non-invariant field operators. We have seen in the last subsection that the BRST invariance implies that in the end matrix elements between gauge invariant states do not depend on the gauge-fixing functional. We will now derive the implications of BRST invariance of the action (and measure) for the generating functional $Z[J]$ as defined in (4.1). These are the Slavnov-Taylor identities.

As usual, we define $Z[J] = e^{iW[J]}$, as well as $\chi^n = \frac{\delta_R W[J]}{\delta J_n} = \langle \tilde{\chi}_{\text{operator}}^n \rangle$ with $J_{n,\chi}$ being the solution of this equation for given χ^n . Then the Legendre transform is $\Gamma[\chi] = W[J_\chi] - \int d^4x \chi^n J_{n,\chi}$. Note that, since χ^n can be either bosonic or fermionic (i.e. anticommuting), one must specify whether functional derivatives should act from the left (L) or from the right (R). Indeed, for fermionic χ^n one has $\frac{\delta_L \Gamma}{\delta \chi^n} = -\frac{\delta_R \Gamma}{\delta \chi^n}$.

The invariance of $\int \mathcal{L}_{\text{new}}$ and of the functional integral measure under

$$\delta_\theta \chi^n = \theta \Delta^n, \quad \Delta^n = s \chi^n, \quad (4.2)$$

(dropping the tilde on the χ^n) implies the Slavnov-Taylor identity

$$\int d^4x \langle \Delta^n(x) \rangle_{J_\chi} \frac{\delta_L \Gamma}{\delta \chi^n(x)} = 0. \quad (4.3)$$

Of course, invariance of the functional integral measure under BRST transformations is not guaranteed and needs to be verified. This will be discussed at the end of this subsection. The conclusion will be that the measure is indeed invariant under BRST transformations provided the measure for the matter fields is gauge invariant, i.e. there are no gauge anomalies. For now, we assume that this is the case and the Slavnov-Taylor identity (4.3) does hold. For a *linear* symmetry ($\Delta^n = c_m^n \chi^m$) one would have $\langle \Delta^n(x) \rangle_{J_\chi} = c_m^n \langle \chi^m(x) \rangle_{J_\chi} = c_m^n \chi^m$, and then the Slavnov-Taylor identity just states that Γ is invariant under this symmetry. At present, however, the *BRST symmetry is non-linear*, e.g. $s\psi = i\omega\omega$, etc, and the Slavnov-Taylor identity (4.3) does *not* tell us that the effective action Γ is also BRST invariant.

4.1.2 Zinn-Justin equation

In order to nevertheless exploit the content of the Slavnov-Taylor identity, one uses the trick to also introduce sources for the “composite” fields Δ^n . (This is somewhat similar to what we did when discussing the renormalization of composite operators and the Callan-Symanzik equations for vertex functions $\Gamma^{(n)}$ with l additional insertions of some composite operator like e.g. ϕ^2 .) To begin with, one defines

$$Z[J, K] \equiv e^{iW[J, K]} = \int \mathcal{D}\chi^n \exp \left[i \int d^4x (\mathcal{L}_{\text{new}} + \chi^n J_n + \Delta^n K_n) \right] . \quad (4.4)$$

The K_n are like additional (position-dependent) coupling constants. In particular,

$$\frac{\delta_R W[J, K]}{\delta K_n(x)} = \langle \Delta^n(x) \rangle_{J, K} . \quad (4.5)$$

Note that the addition of $\Delta^n K_n = s(\chi^n K_n)$ amounts just to a modification of the gauge fixing functional Ψ . In particular, the exponent in (4.4) is still BRST invariant. The effective action $\Gamma[\chi, K]$ is obtained by performing the Legendre transformation with respect to the sources J_n only, while keeping these extra couplings K_n :

$$\Gamma[\chi, K] = W[J_{\chi, K}, K] - \int d^4x \chi^n(x) J_{\chi, K}^n(x) , \quad (4.6)$$

where $J_{\chi, K}^n$ is the solution of

$$\frac{\delta_R W[J, K]}{\delta J_n} = \chi^n \quad (4.7)$$

for given χ^n . (For obvious notational reasons, we write $J_{\chi, K}^n$ rather than $(J_n)_{\chi, K}$.) Note that these definitions imply on the one hand, that χ^n and J_n have the same statistics (both bosonic or both fermionic) and Δ^n and K_n also have the same statistics. On the other hand, any field from the first group (χ^n or J_n) has opposite statistics from the corresponding field of the second group (Δ^n or K_n). Through the usual manipulations one finds from (4.6)

$$\frac{\delta_R \Gamma[\chi, K]}{\delta K_n(x)} = \frac{\delta_R W[J, K]}{\delta K_n(x)} \Big|_{J=J_{\chi, K}} + \int d^4y \frac{\delta_R W[J, K]}{\delta J_m(y)} \Big|_{J=J_{\chi, K}} \frac{\delta_R J_{\chi, K}^m(y)}{\delta K_n(x)} - \int d^4y \chi^m(y) \frac{\delta_R J_{\chi, K}^m(y)}{\delta K_n(x)} . \quad (4.8)$$

The last two terms cancel by (4.7) while, by (4.5), the first term is just $\langle \Delta^n(x) \rangle_{J_{\chi, K}, K}$. Hence

$$\frac{\delta_R \Gamma[\chi, K]}{\delta K_n(x)} = \langle \Delta^n(x) \rangle_{J_{\chi, K}, K} . \quad (4.9)$$

It is completely straightforward to generalize the Slavnov-Taylor identity to the case where the additional couplings $\sim \Delta^n K_n$ are present:

$$\int d^4x \langle \Delta^n(x) \rangle_{J_{\chi, K}, K} \frac{\delta_L \Gamma[\chi, K]}{\delta \chi^n(x)} = 0 . \quad (4.10)$$

Using (4.9), we get the Zinn-Justin equation:

$$\int d^4x \frac{\delta_R \Gamma[\chi, K]}{\delta K_n(x)} \frac{\delta_L \Gamma[\chi, K]}{\delta \chi^n(x)} = 0 . \quad (4.11)$$

4.1.3 Antibracket

The Zinn-Justin equation displays a nice symmetry between the roles of K_n and χ^n . This is further emphasized by introducing the notion of antibracket. The antibracket of two (bosonic) functionals²⁴ $F[\chi, K]$ and $G[\chi, K]$ depending on two sets of arguments χ^n and K_n having *opposite* statistics is defined as

$$(F, G) = \int d^4x \left(\frac{\delta_R F[\chi, K]}{\delta \chi^n(x)} \frac{\delta_L G[\chi, K]}{\delta K_n(x)} - \frac{\delta_R F[\chi, K]}{\delta K_n(x)} \frac{\delta_L G[\chi, K]}{\delta \chi^n(x)} \right) . \quad (4.12)$$

This is somewhat similar to the definition of the Poisson bracket but, as we will see, it is symmetric rather than antisymmetric under the exchange of F and G . Since the χ^n always have the opposite statistics from the K_n , one of the two functional derivatives in the definition of the antibracket always yields an anticommuting expression while the other is commuting. Since exchanging a left with a right derivative yields a minus sign for an anticommuting expression, we can flip left and right derivatives in either of the two terms in (4.12) provided we include one extra minus sign. Thus

$$\begin{aligned} (F, G) &= \int d^4x \left(\frac{\delta_R F[\chi, K]}{\delta \chi^n(x)} \frac{\delta_L G[\chi, K]}{\delta K_n(x)} + \frac{\delta_L F[\chi, K]}{\delta K_n(x)} \frac{\delta_R G[\chi, K]}{\delta \chi^n(x)} \right) \\ &= \int d^4x \left(\frac{\delta_R F[\chi, K]}{\delta \chi^n(x)} \frac{\delta_L G[\chi, K]}{\delta K_n(x)} + \frac{\delta_R G[\chi, K]}{\delta \chi^n(x)} \frac{\delta_L F[\chi, K]}{\delta K_n(x)} \right) \\ &= - \int d^4x \left(\frac{\delta_R G[\chi, K]}{\delta K_n(x)} \frac{\delta_L F[\chi, K]}{\delta \chi^n(x)} + \frac{\delta_R F[\chi, K]}{\delta K_n(x)} \frac{\delta_L G[\chi, K]}{\delta \chi^n(x)} \right) . \end{aligned} \quad (4.13)$$

This shows that the antibracket is *symmetric* under the interchange of F and G :

$$(F, G) = (G, F) . \quad (4.14)$$

As is obvious from the last line in (4.13), the Zinn-Justin equation (4.11) can then be written as

$$\boxed{(\Gamma, \Gamma) = 0} . \quad (4.15)$$

4.1.4 Invariance of the measure under the BRST transformation

Let us now come back to the question of whether the functional integral measure is invariant under the BRST transformation or not. Clearly, even if the measure is invariant under gauge transformations, this does not immediately imply invariance under BRST transformations. The reason is that the latter are non-linear transformations of the fields $A_\mu, \psi, \bar{\psi}, \omega, \omega^*, h$ while the former are linear transformations of A_μ, ψ and $\bar{\psi}$ only. A further technical complication appears since the BRST transformation mixes commuting and anticommuting field variables. Nevertheless, we will now show that the relevant Jacobian equals unity provided the measure for the matter fields alone is gauge invariant.

²⁴A bosonic functional is a sum of terms each of which contains an *even* number of anticommuting fields. Hence it is commuting. Similarly, a fermionic functional is made from terms containing an *odd* number of anticommuting fields and hence is anticommuting.

We want to compute the Jacobian for a change of integration variables of the form $\chi^n \rightarrow \chi'^n = \chi^n + \theta \Delta^n \equiv \chi^n + \theta s \chi^n$ where χ^n stands for the various commuting and anticommuting fields and θ is a single anticommuting parameter satisfying $\theta^2 = 0$. Although the Δ^n depend non-linearly on the χ^r , the fact that $\theta^2 = 0$ will simplify the Jacobian enormously. We let

$$J_r^n(x, y) = \frac{\delta_L \chi'^n(x)}{\delta \chi^r(y)} = \delta_r^n \delta^{(4)}(x - y) + \frac{\delta_L(\theta \Delta^n(x))}{\delta \chi^r(y)} \equiv \delta_r^n \delta^{(4)}(x - y) + R_r^n(x, y) . \quad (4.16)$$

Note that R is linear in θ , i.e. $R \equiv \theta \widehat{R}$, and hence $R^2 = 0$, so that $\log(1 + R) = R$. We can then use the standard relation between the determinant and the trace of the logarithm to get

$$\text{Det } J = \exp(\text{Tr} \log(\mathbf{1} + R)) = \exp(\text{Tr } R) = 1 + \text{Tr } R , \quad (4.17)$$

where Tr includes a sum over the different field types of functional traces for every field. Now, for anticommuting χ^r (i.e. the ghost, antighost or matter fermions), one has $\frac{\delta_L \theta \Delta^n}{\delta \chi^r} = -\theta \frac{\delta_L \Delta^n}{\delta \chi^r}$. It follows that one has

$$\text{Tr } R = \theta \text{Str } \widehat{R} , \quad (4.18)$$

where Str is a functional “supertrace” which is just an ordinary functional trace but with minus signs inserted for the anticommuting fields. With obvious notations we have

$$\text{Str } \widehat{R} = \text{Tr } \widehat{R}_A - \text{Tr } \widehat{R}_\psi - \text{Tr } \widehat{R}_{\bar{\psi}} + \text{Tr } \widehat{R}_h - \text{Tr } \widehat{R}_\omega - \text{Tr } \widehat{R}_{\omega^*} . \quad (4.19)$$

Recall that $A'_\mu = A_\mu + \theta s A_\mu = A_\mu + \theta(\partial_\mu \omega - i[A_\mu, \omega])$ so that $\Delta^{A_\mu} = (\partial_\mu \omega - i[A_\mu, \omega])$. Similarly, $\Delta^\omega = i\omega\omega$ and $\Delta^{\omega^*} = -h$, as well as $\Delta^\psi = i\omega\psi$ and $\Delta^h = 0$. Obviously then, $\text{Tr } \widehat{R}_{\omega^*} = \text{Tr } \widehat{R}_h = 0$. Also $(\widehat{R}_\omega)_{\alpha\beta}(x, y) = C_{\alpha\beta\gamma} \omega^\gamma(x) \delta^{(4)}(x - y)$ so that its trace vanishes (after an appropriate gauge invariant regularization). Similarly,

$$(\widehat{R}_A)_{\alpha\beta\nu}^\mu(x, y) = \frac{\delta \Delta_\alpha^\mu(x)}{\delta A_\beta^\nu(y)} = \delta_\nu^\mu C_{\alpha\beta\gamma} \omega^\gamma \delta^{(4)}(x - y) . \quad (4.20)$$

This is again antisymmetric and its trace vanishes (after an appropriate gauge invariant regularization). Actually, this is the same operator as the one we encountered when we discussed the invariance of the gauge field measure under gauge transformations (except that now $\epsilon \rightarrow \omega$). Obviously, $\text{Tr } \widehat{R}_A = 0$ is equivalent to the statement that the gauge field measure is gauge invariant. Finally, one has to study $\text{Tr } R_\psi$ and $\text{Tr } R_{\bar{\psi}}$. Once more, the vanishing of the trace is equivalent to the gauge invariance of the matter measure. Although formally this trace vanishes, just as for the other fields, here we may encounter a problem if it is not possible to regularize the trace in a gauge invariant way. This is precisely the case for chiral fermions and one has a so-called anomaly as we will extensively discuss in the next section.

We conclude that in the absence of anomalies (no chiral matter fermions), the full functional integral measure is indeed invariant under BRST transformations. Moreover, the BRST invariance of the measure is equivalent to the gauge invariance of the matter measure alone.

4.2 Renormalization of gauge theories

4.2.1 The general structure and strategy

As already emphasized, non-abelian gauge theories are renormalizable by power counting. As discussed in sect. 2, it follows from power counting that there only are finitely many Green-functions (or equivalently only finitely many 1PI n -point vertex functions $\Gamma^{(n)}$) that are divergent with their superficial degree of divergence being some fixed finite number (independent of the order of perturbation theory). Relying on the BPHZ theorem, we then know that all divergences can be removed by finitely many local counterterms whose coefficients can be determined order by order in perturbation theory. However, what this does not tell us is whether the divergent parts of the Green functions or of the $\Gamma^{(n)}$ – and hence the corresponding counterterms – share the same symmetries as the original action. If one needed to add non gauge invariant (or actually non BRST invariant) counterterms at some order N of perturbation theory to cancel some non-invariant divergence, these non-invariant counterterms would almost certainly lead to non-invariant contributions at some higher order $N' > N$ of perturbation theory. Thus, when one asks whether a (non-abelian) gauge theory is renormalizable the question is whether all counterterms share the same symmetries as the original (gauge-fixed) action, namely the BRST symmetry.

The proof of the renormalizability of non-abelian gauge theories thus amounts to showing that the divergent parts Γ_∞ of the quantum effective action Γ (which generates the 1PI vertex functions $\Gamma^{(n)}$) due to the loops still has the BRST symmetry, or some deformation thereof obtained after changing the coupling constant and field normalizations. Instead of actually computing the loop diagrams, we will exploit the algebraic structure coded in the Zinn-Justin equation to show that this is indeed the case.

As usual, we rewrite the original bare action (including now the extra couplings $\sim \int \Delta^n K_n$) in terms of renormalized fields and coupling constants and masses:

$$S_B[\chi_B, K_B] = S[\chi, K] = S_R^{(0)}[\chi, K] + S_{\text{c.t.}}[\chi, K] , \quad (4.21)$$

where $S_R^{(0)}$ has the same form as the original (bare) action but with all masses and couplings equal to their renormalized values, and $S_{\text{c.t.}}$ are the counterterms, c.f. the general discussion and examples studied in sections ?? and ??. There are now two ways to rephrase the question of renormalizability of the non-abelian gauge theories. If one insists that S_B is BRST invariant, as well as the “tree-level” renormalized action $S_R^{(0)}$, then the counterterms will also be BRST invariant. The non-trivial question then is whether these counterterms are enough to cancel all the divergent parts that arise in any N -loop diagram. Equivalently, we may just begin with the usual BRST invariant $S_R^{(0)}$ and compute the divergent parts Γ_∞ of Γ and adjust the counterterms $S_{\text{c.t.}}$ order by order in perturbation theory to cancel these divergent parts. The question then is whether these counterterms are BRST invariant, i.e. whether the bare action can be BRST invariant.

Actually, there is a slight subtlety here: in the non-abelian case, the gauge and BRST transformations explicitly depend on the gauge coupling constant (which we have hidden in the generators t_α , respectively in the structure constants). Thus, when one states that some quantity is gauge or

BRST invariant, one has to specify what is the coupling constant. Now the bare action S_B is BRST invariant using the bare coupling constant, while the renormalized action $S_R^{(0)}$ should be BRST invariant using the renormalized coupling constant. The precise formulation of renormalizability then is whether we can find counterterms that cancel the divergent part Γ_∞ and which equal the difference of the g_B -BRST invariant bare action S_B and the g_R -BRST invariant action $S_R^{(0)}$. Equivalently, starting from a g_R -BRST invariant $S_R^{(0)}$, we must show that $S_R^{(0)} - \Gamma_\infty$ is g_B -BRST invariant. In the sequel, we will call a conterterm or Γ_∞ simply “BRST-invariant” if it fulfills this requirement.

In perturbation theory, Γ has an expansion

$$\Gamma[\chi, K] = \sum_{N=0}^{\infty} \Gamma_N[\chi, K] \quad , \quad \Gamma_0 = S_R^{(0)} \quad , \quad \Gamma_N = \Gamma_{N\text{-loop}} + \sum_{M=1}^N \Gamma_{(N-M)\text{-loop}}^{\text{c.t.}, M} \quad , \quad (4.22)$$

where $\Gamma_{(N-M)\text{-loop}}^{\text{c.t.}, M}$ is an $(N-M)$ -loop contribution involving “vertices” from lower-order counterterms of total order M , resulting in a contribution at the same order in perturbation theory as $\Gamma_{N\text{-loop}}$. (In QED e.g., the order α^2 contributions to the vacuum polarisation correspond to $\Gamma_2^{\mu\nu}$ and are given by the genuine two-loop diagrams ($\Gamma_{2\text{-loop}}^{\mu\nu}$), one-loop diagrams with one-loop (order α) counterterms inserted ($\Gamma_{1\text{-loop}}^{\mu\nu \text{ c.t.}, 1}$), as well as a new order- α^2 counterterm ($\Gamma_{0\text{-loop}}^{\mu\nu \text{ c.t.}, 2}$).

One can again introduce a loop-counting parameter λ and formally replace $S_R^{(0)}$ by $\frac{1}{\lambda} S_R^{(0)}$. Then an N -loop term will have a coefficient λ^{N-1} , and $\Gamma_{N\text{-loop}}$ will be multiplied by λ^{N-1} . We can similarly assign a factor λ^M to every counterterm that arises from an M -loop diagram, in addition to an overall $\frac{1}{\lambda}$ in front of $S_{\text{c.t.}}$: $\frac{1}{\lambda} S_{\text{c.t.}} = \frac{1}{\lambda} \sum_{M=1}^{\infty} \lambda^M S_{\text{c.t.}}^M$. Then $\Gamma_{(N-M)\text{-loop}}^{\text{c.t.}, M}$ will be accompanied by a factor $\lambda^{N-M-1} \lambda^M = \lambda^{N-1}$. As a result, we see that $\Gamma = \sum_{N=0}^{\infty} \lambda^{N-1} \Gamma_N$. Inserting this into the Zinn-Justin equation $(\Gamma, \Gamma) = 0$ and collecting the coefficients of the λ^{N-2} -terms (recall that the antibracket is symmetric) yields

$$\sum_{N'=0}^N (\Gamma_{N'}, \Gamma_{N-N'}) = 0 \quad , \quad N \geq 0 \quad . \quad (4.23)$$

We want to show that Γ can be made finite by choosing “BRST-invariant” local counterterms order by order in perturbation theory. We will show this by induction in N . First, for $N = 0$ one has $\Gamma_0 = S_R^{(0)}$. Now, $S_R^{(0)}$ is expressed in terms of the renormalized (finite) parameters (couplings and masses) and the renormalized fields. Obviously, no counterterms are needed at this order. Suppose then that all $\Gamma_{N'}$ with $N' \leq N-1$ are finite, i.e. we assume that all divergences in $\Gamma_{N'\text{-loop}}$ have been cancelled by the contributions $\sum_{M=1}^{N'} \Gamma_{(N'-M)\text{-loop}}^{\text{c.t.}, M}$, induced by appropriately chosen (“BRST-invariant”) counterterms $S_{\text{c.t.}}^M$ with $M \leq N'$. This shows that in the sum (4.23) the only terms that can involve divergences are the terms $N' = 0$ and $N' = N$. Thus, isolating the possibly divergent parts, eq. (4.23) yields

$$(S_R^{(0)}, \Gamma_{N,\infty}) = 0 \quad . \quad (4.24)$$

In agreement with our above remark, we can interpret this constraint in two different ways.

- First, we may consider that Γ_N contains the N -loop contributions $\Gamma_{N\text{-loop}}$, the contributions from lower-order counterterms in loops $\sum_{M=1}^{N-1} \Gamma_{(N-M)\text{-loop}}^{\text{c.t.}, M}$ (the corresponding counterterms are

“BRST-invariant” by the induction hypothesis) and an order- N counterterm $\Gamma_{\text{tree}}^{\text{c.t.},N}$. The latter is taken “BRST invariant” and we must then show that it can indeed cancel the divergent part from the other contributions. This will be the case if the latter are “BRST invariant”. Since the counterterm is already supposed to be “BRST invariant”, it is equivalent to show that the divergent part $\Gamma_{N,\infty}$ of the full Γ_N is “BRST invariant”. (Of course, if this is the case, the divergent parts just cancel and $\Gamma_{N,\infty} = 0$.)

- Second, we may consider that we have not yet included an order- N counterterm $\Gamma_{\text{tree}}^{\text{c.t.},N}$ in the computation of Γ_N which now only contains $\Gamma_{N-\text{loop}}$ and the contributions from lower-order counterterms in loops $\sum_{M=1}^{N-1} \Gamma_{(N-M)-\text{loop}}^{\text{c.t.},M}$. Then, to be able to cancel the divergent part $\Gamma_{N,\infty}$ of Γ_N by an appropriate “BRST invariant” counterterm, we must show that $\Gamma_{N,\infty}$ is “BRST invariant”.

We see that, whatever interpretation we adopt, it is necessary and sufficient to show that $\Gamma_{N,\infty}$ is “BRST invariant”. This will be done by :

- exploiting the content of (4.24) which we derived from the Zinn-Justin equation,
- using the various linear symmetries²⁵ of the tree-level action $S_R^{(0)}$ which must be also be symmetries of the effective action Γ at every order, i.e. of Γ_N , and in particular also of its diverging part,
- using the fact that $\Gamma_{N,\infty}$ is a local functional of the fields of dimension less or equal to 4 (here we rely of course on BPHZ to exclude any trouble with overlapping divergences).

Let us recall that all computations are done in the presence of the extra couplings $\sim \Delta^n K_n$, i.e. with $S_R^{(0)}[\chi^n, K_n] = \int (\mathcal{L}_{\text{new},R}[\chi^n] + \Delta^n K_n)$. Here $\Delta^n = s\chi^n$ with the BRST-transformation s being the one involving the same renormalized coupling as appears in $S_R^{(0)}$. Let us also insist, that eq. (4.24) will only tell us something about the diverging part of Γ_N , not its finite part.

4.2.2 Constraining the divergent part of Γ

Let us now constrain the divergent part $\Gamma_{N,\infty}$ of Γ_N using dimensional arguments, the linear symmetries and equation (4.24). We begin by showing that $\Gamma_{N,\infty}$ can depend at most linearly on the various K_n .

Dimensional arguments

The gauge field A_μ has dimension 1, just as scalar matter fields. The fermionic matter fields, ψ and $\bar{\psi}$ have dimension $\frac{3}{2}$. The dimensions of the ghost and antighost fields can be read from their kinetic term (or their propagator). This depends on the choice of gauge-fixing function f_α . If $f_\alpha = \partial_\mu A^\mu + \dots$, the ghost Lagrangian is $\sim \omega^* \partial_\mu D^\mu \omega + \dots$ and $\dim \omega^* + \dim \omega = 2$. Since in any ghost-number zero

²⁵These are Lorentz transformations, global gauge transformations, ghost phase transformations related to ghost number conservation and possibly antighost translations for certain choices of the gauge-fixing function like e.g. $f_\alpha = \partial_\mu A^\mu_\alpha$.

functional one always has as many ghosts as antighosts it does not matter how we distribute the 2 between the ghost and antighost and we can choose $\dim \omega = \dim \omega^* = 1$. As one sees from $s A_\mu = D_\mu \omega$, $s\psi = i\omega\psi$, etc, s changes dimensions by $\dim \omega = 1$. It follows that for any field χ^n one has $\dim \Delta^n = \dim \chi^n + 1$. Finally, $\Delta^n K_n$ must have dimension 4 so that $\dim K_n = 4 - \dim \Delta^n$. To summarize:

$$\begin{aligned} \dim A_\mu &= \dim \omega = \dim \omega^* = 1 \quad , \quad \dim \psi = \dim \bar{\psi} = \frac{3}{2} \quad , \\ \dim \Delta^A &= \dim \Delta^\omega = \dim \Delta^{\omega^*} = 2 \quad , \quad \dim \Delta^\psi = \dim \Delta^{\bar{\psi}} = \frac{5}{2} \quad , \\ \dim K_A &= \dim K_\omega = \dim K_{\omega^*} = 2 \quad , \quad \dim K_\psi = \dim K_{\bar{\psi}} = \frac{3}{2} \quad . \end{aligned} \quad (4.25)$$

Since the divergent part of Γ must be a local functional of all the fields of at most dimension 4, we see that it can be at most quadratic in the K_n .

Ghost number conservation

Since $\mathcal{L}_{\text{new}}$ is invariant under global ghost/antighost phase rotations $\omega \rightarrow e^{i\alpha}\omega$, $\omega^* \rightarrow e^{-i\alpha}\omega^*$ it follows that the ghost number is conserved (ω has ghost number +1 and ω^* ghost number -1). This will remain true in the presence of the extra couplings $\sim \Delta^n K_n$ if we assign ghost numbers to the K_n which are the opposite of the ghost numbers of Δ^n . Since $\Delta^n = s\chi^n$ and s increases the ghost numbers by one unit, one has $n_{\text{gh}}\Delta^n = n_{\text{gh}}\chi^n + 1$. Thus

$$\begin{aligned} n_{\text{gh}}(A_\mu) &= n_{\text{gh}}(\psi) = n_{\text{gh}}(\bar{\psi}) = 0 \quad , \quad n_{\text{gh}}(\omega) = 1 \quad , \quad n_{\text{gh}}(\omega^*) = -1 \quad , \\ n_{\text{gh}}(\Delta^A) &= n_{\text{gh}}(\Delta^\psi) = n_{\text{gh}}(\Delta^{\bar{\psi}}) = 1 \quad , \quad n_{\text{gh}}(\Delta^\omega) = 2 \quad , \quad n_{\text{gh}}(\Delta^{\omega^*}) = 0 \quad , \\ n_{\text{gh}}(K_A) &= n_{\text{gh}}(K_\psi) = n_{\text{gh}}(K_{\bar{\psi}}) = -1 \quad , \quad n_{\text{gh}}(K_\omega) = -2 \quad , \quad n_{\text{gh}}(K_{\omega^*}) = 0 \quad , \quad . \end{aligned} \quad (4.26)$$

Hence, all K_n have negative or zero ghost number.

Now, $\Gamma_{N,\infty}$ must have ghost number zero and we have seen above that it can be at most quadratic in the K_n . Let's see which quadratic terms could appear. Terms involving two K_A, K_ω or K_{ω^*} have dimension 4 and hence cannot involve any other field. Such terms have negative ghost numbers and are excluded except for $K_{\omega^*}K_{\omega^*}$. The latter term, however, is excluded since $\Delta^{\omega^*} \equiv s\omega^* = -h$ is a linear transformation so that $\frac{\delta_R \Gamma}{\delta K_{\omega^*}} = \langle \Delta^{\omega^*} \rangle = \Delta^{\omega^*} \equiv -h$ tells us that $\Gamma[\chi^n, K_n]$ must be linear in K_{ω^*} and cannot have a term $\sim K_{\omega^*}K_{\omega^*}$. Actually, this argument tells us even a bit more: the K_{ω^*} dependence of $\Gamma[\chi, K]$ must be precisely a term $\int d^4x (-h)K_{\omega^*}$. This is just the (finite) term already present in the tree-level action $S_R^{(0)}[\chi, K]$ and, in particular, the diverging part Γ_∞ cannot contain any term linear in K_{ω^*} . Similarly, $\Delta^h = sh = 0$ and we conclude that Γ_∞ cannot contain a term linear in K_h either.

Next, since the ghost is a Lorentz scalar, all Δ^n transform under the same representation of the Lorentz group as the fields χ^n and the K_n must transform in such a way that $\Delta^n K_n$ is a Lorentz scalar. It follows that K_ω and K_{ω^*} are Lorentz scalars, K_{A^μ} is a four-vector and K_ψ transforms as the spinor $\bar{\psi}$, while $K_{\bar{\psi}}$ transforms as ψ . Thus a term quadratic in the K_n and involving at least one K_ψ or $K_{\bar{\psi}}$ must necessarily involve both of them, i.e. be $\sim K_\psi K_{\bar{\psi}}$. This expression has ghost

number -2 and dimension 3 . There is no dimension 1 field of ghost number $+2$ available to make a dimension 4 , ghost number 0 term. Thus we conclude that $\Gamma_{N,\infty}$ is at most linear in any of the K_n :

$$\Gamma_{N,\infty}[\chi, K] = \Gamma_{N,\infty}[\chi] + \int d^4x \tilde{\Delta}_N^n(x) K_n(x) , \quad (4.27)$$

with $\Gamma_{N,\infty}[\chi] \equiv \Gamma_{N,\infty}[\chi, 0] = \int d^4x \gamma_{N,\infty}(x)$. Similarly, we had for the (tree-level) action

$$S_R^{(0)}[\chi, K] = S_R^{(0)}[\chi] + \int d^4x \Delta^n(x) K_n(x) . \quad (4.28)$$

Just as Δ^n is a local expression in the fields χ^n , the same must be true for the $\tilde{\Delta}_N^n$. Note also that the $\tilde{\Delta}_N^n$ must have the same ghost numbers and dimensions as the Δ^n . Finally, we have seen above that $\Gamma_{N,\infty}$ cannot contain any terms linear in K_{ω^*} or K_h and we conclude that

$$\tilde{\Delta}_N^{\omega^*} = \tilde{\Delta}_N^h = 0 . \quad (4.29)$$

Exploiting $(S_R^{(0)}, \Gamma_{N,\infty}) = 0$

In order to get further information on the form of $\gamma_{N,\infty}(x)$ and $\tilde{\Delta}_N^n(x)$, we insert (4.27) and (4.28) into $(S_R^{(0)}[\chi, K], \Gamma_{N,\infty}[\chi, K]) = 0$:

$$\begin{aligned} 0 &= \int d^4x \left(\frac{\delta_R S_R^{(0)}[\chi, K]}{\delta K_n(x)} \frac{\delta_L \Gamma_{N,\infty}[\chi, K]}{\delta \chi^n(x)} + \frac{\delta_R \Gamma_{N,\infty}[\chi, K]}{\delta K_n(x)} \frac{\delta_L S_R^{(0)}[\chi, K]}{\delta \chi^n(x)} \right) \\ &= \int d^4x \left\{ \Delta^n(x) \left(\frac{\delta_L \Gamma_{N,\infty}[\chi]}{\delta \chi^n(x)} + \int d^4y \frac{\delta_L \tilde{\Delta}_N^m(y)}{\delta \chi^n(x)} K_m(y) \right) \right. \\ &\quad \left. + \tilde{\Delta}_N^n(x) \left(\frac{\delta_L S_R^{(0)}[\chi]}{\delta \chi^n(x)} + \int d^4y \frac{\delta_L \Delta^m(y)}{\delta \chi^n(x)} K_m(y) \right) \right\} . \end{aligned} \quad (4.30)$$

The terms without K_m and the coefficients of $K_m(y)$ must vanish separately, giving two equations:

$$\int d^4x \left(\Delta^n(x) \frac{\delta_L \Gamma_{N,\infty}[\chi]}{\delta \chi^n(x)} + \tilde{\Delta}_N^n(x) \frac{\delta_L S_R^{(0)}[\chi]}{\delta \chi^n(x)} \right) = 0 , \quad (4.31)$$

and

$$\int d^4x \left(\Delta^n(x) \frac{\delta_L \tilde{\Delta}_N^m(y)}{\delta \chi^n(x)} + \tilde{\Delta}_N^n(x) \frac{\delta_L \Delta^m(y)}{\delta \chi^n(x)} \right) = 0 . \quad (4.32)$$

The second equation (4.32) is a set of functional first-order partial linear differential equations for the $\tilde{\Delta}_N^n$. Note that these equations do constrain the functional form of the $\tilde{\Delta}_N^n$ but not their overall normalization. Inserting the solutions $\tilde{\Delta}_N^n$ into the first equation then should determine the functional form of $\Gamma_{N,\infty}[\chi]$. Note that these two equations (4.31) and (4.32) do not explicitly involve the K_n anymore. Indeed, all we really want to determine in the end is the form of $\Gamma_{N,\infty}[\chi] = \Gamma_{N,\infty}[\chi, 0]$ although, once we know $\tilde{\Delta}_N^n(x)$, we also know $\Gamma_{N,\infty}[\chi, K] = \Gamma_{N,\infty}[\chi, 0] + \int \tilde{\Delta}_N^n(x) K_n(x)$.

To solve these equations in practice, it is useful to give them a more physical interpretation. Define

$$\Gamma_N^{(\epsilon)}[\chi] = S_R^{(0)}[\chi] + \epsilon \Gamma_{N,\infty}[\chi] \quad \text{and} \quad \Delta_N^{(\epsilon)n}(x) = \Delta^n(x) + \epsilon \tilde{\Delta}_N^n(x) . \quad (4.33)$$

Consider the “deformed BRST transformation”

$$s^{(\epsilon)}\chi^n(x) = \Delta_N^{(\epsilon)n}(x) . \quad (4.34)$$

Obviously, for $\epsilon = 0$ this reduces to the ordinary BRST transformation. It is then not difficult to see that, *up to first order in ϵ* , $\Gamma_N^{(\epsilon)}[\chi]$ is invariant under this transformation and that this transformation is nilpotent:

$$s^{(\epsilon)}\Gamma_N^{(\epsilon)}[\chi] = 0 + \mathcal{O}(\epsilon^2) \quad \text{by (4.31)} \quad \text{and} \quad s^{(\epsilon)}s^{(\epsilon)}\chi^n = 0 + \mathcal{O}(\epsilon^2) \quad \text{by (4.32)} . \quad (4.35)$$

Deformed BRST symmetry

First note (again) that, since $\tilde{\Delta}_N^n$ couples to K_n in $\Gamma_{N,\infty}[\chi, K]$, in the same way as Δ^n couples to K_n in $S_R^{(0)}$, both $\tilde{\Delta}_N^n$ and Δ^n must have the same ghost numbers, dimensions and Lorentz transformation properties.

The strategy now is to write the most general nilpotent transformation of the fields that increases the ghost numbers of the fields by one unit, increases their dimensions by $\dim \omega$ and that is a deformation of the usual BRST transformation in the sense that it depends on some small parameter ϵ and reduces to the ordinary BRST transformation in the $\epsilon \rightarrow 0$ limit. (Of course, nilpotency is only required up to terms of order ϵ^2 .) These requirements will turn out to be stringent enough to show that any such deformed BRST transformation must be identical to the ordinary BRST transformation, up to changes in the normalization of the fields and a change of the gauge coupling constant.

It is easy to see that the deformed BRST symmetry must act as

$$\begin{aligned} \delta_\theta \psi &= i\theta \omega^\alpha T_\alpha \psi , \\ \delta_\theta A_{\alpha\mu} &= \theta (B_{\alpha\beta} \partial_\mu \omega_\beta + D_{\alpha\beta\gamma} A_{\beta\mu} \omega_\gamma) , \\ \delta_\theta \omega_\alpha &= -\frac{1}{2}\theta E_{\alpha\beta\gamma} \omega_\beta \omega_\gamma . \end{aligned} \quad (4.36)$$

(Having the correct ghost number would also allow e.g. a term of the form $\delta\psi \sim i\omega_\alpha \omega_\beta \omega_\gamma^* F_{\alpha\beta\gamma}^\delta T_\delta \psi$, but this is excluded by the argument about the dimensions.) Recall from (4.29) that $\tilde{\Delta}_N^{\omega^*} = \tilde{\Delta}_N^h = 0$ and, hence, the deformed BRST transformation of ω^* and h just equals the undeformed one:

$$\delta\omega_\alpha^* = -h_\alpha \quad , \quad \delta h_\alpha = 0 . \quad (4.37)$$

Let us now exploit the nilpotency of these transformations (4.36) and (4.37). First, $\delta_{\theta_1}\delta_{\theta_2}\omega = 0$ yields $E_{\alpha\beta\gamma}E_{\beta\delta\epsilon}\omega_\delta\omega_\epsilon\omega_\gamma = 0$ so that $E_{\alpha\beta\gamma}E_{\beta\delta\epsilon}$ must vanish after antisymmetrizing in δ, ϵ, γ . i.e. the $E_{\alpha\beta\gamma}$ satisfy the Jacobi identity and must be the structure constants of some Lie algebra. As $\epsilon \rightarrow 0$,

these are just the $C_{\alpha\beta\gamma}$ and we conclude that this Lie algebra is the same as the original Lie algebra $\mathcal{G}$ and the $E_{\alpha\beta\gamma}$ differ from the $C_{\alpha\beta\gamma}$ just by the normalization:

$$E_{\alpha\beta\gamma} = \mathcal{Z} C_{\alpha\beta\gamma} . \quad (4.38)$$

(Note that if the Lie algebra is semi-simple one may have a different $\mathcal{Z}_i$ for every simple factor.) Next, looking at $\delta_{\theta_1} \delta_{\theta_2} A_\mu$ gives $D_{\alpha\beta\gamma} D_{\beta\delta\epsilon} - D_{\alpha\beta\epsilon} D_{\beta\delta\gamma} = E_{\beta\epsilon\gamma} D_{\alpha\delta\beta} = \mathcal{Z} C_{\beta\epsilon\gamma} D_{\alpha\delta\beta}$ and $B_{\alpha\beta} E_{\beta\gamma\delta} = D_{\alpha\beta\delta} B_{\beta\gamma}$. The first equation implies that the matrices $\hat{D}_\gamma$ with $(\hat{D}_\gamma)_{\alpha\beta} = \frac{1}{\mathcal{Z}} D_{\gamma\alpha\beta}$ are the generators of the adjoint representation of the Lie algebra $\mathcal{G}$, so that

$$D_{\alpha\beta\gamma} = \mathcal{Z} C_{\alpha\beta\gamma} . \quad (4.39)$$

Hence, $D_{\alpha\beta\gamma} = E_{\alpha\beta\gamma}$ and the second condition involving $B_{\alpha\beta}$ just states that the matrix B commutes with all matrices $\hat{D}$ and thus is proportional to the unit matrix:

$$B_{\alpha\beta} = \mathcal{Z} \mathcal{N} \delta_{\alpha\beta} . \quad (4.40)$$

(Again, for a semi-simple $\mathcal{G}$ we can have a different $\mathcal{N}_i$ for every simple factor.) Finally, looking at $\delta_{\theta_1} \delta_{\theta_2} \psi$ implies $[T_\alpha, T_\beta] = i E_{\alpha\beta\gamma} T_\gamma = i \mathcal{Z} C_{\alpha\beta\gamma} T_\gamma$ and we identify

$$T_\alpha = \mathcal{Z} t_\alpha . \quad (4.41)$$

We conclude that, apart from the new constants $\mathcal{Z}$ and $\mathcal{N}$, the deformed BRST symmetry $\tilde{s}$ must act exactly as the ordinary BRST symmetry s we started with. Let us summarize:

$$\begin{aligned} \tilde{s} \psi &= i \mathcal{Z} \omega \psi , \\ \tilde{s} A_\mu &= \mathcal{Z} (\mathcal{N} \partial_\mu \omega - i [A_\mu, \omega]) , \\ \tilde{s} \omega &= i \mathcal{Z} \omega \omega \\ \tilde{s} \omega^* &= -h \\ \tilde{s} h &= 0 . \end{aligned} \quad (4.42)$$

We can absorb these new constants by redefining the normalizations of the ghost fields and redefining the coupling constant (hidden in the $C_{\alpha\beta\gamma}$) as follows:

$$\tilde{\omega}^\alpha = \mathcal{Z} \mathcal{N} \omega^\alpha , \quad \tilde{A}_\mu^\alpha = A_\mu^\alpha , \quad \tilde{C}_{\alpha\beta\gamma} = \frac{1}{\mathcal{N}} C_{\alpha\beta\gamma} . \quad (4.43)$$

Then the generators are redefined accordingly by $\tilde{t}_\alpha = \frac{1}{\mathcal{N}} t_\alpha$ and thus $\tilde{\omega} = \tilde{\omega}^\alpha \tilde{t}_\alpha = \mathcal{Z} \omega^\alpha t_\alpha = \mathcal{Z} \omega$ and $\tilde{A}_\mu = \frac{1}{\mathcal{N}} A_\mu$. Furthermore, one can redefine the normalization of ψ by some factor and the one of ω^* and h by another (common) factor. These latter factors are not yet determined by (4.42). In terms of the redefined fields and couplings, the deformed BRST transformations (4.42) take exactly the form of the undeformed ones. In particular, we have

$$\tilde{s}^2 = 0 \quad (4.44)$$

and any gauge invariant functional $F[\tilde{A}_\mu, \psi]$ (with structure constants $\tilde{C}_{\alpha\beta\gamma}$) automatically also is $\tilde{s}$ invariant.

Let us now construct $\Gamma_N^{(\epsilon)}[\chi] = \int \mathcal{L}_N^{(\epsilon)}[\chi]$ with a Lagrangian $\mathcal{L}_N^{(\epsilon)}$ that is invariant under the deformed BRST transformation $\tilde{s}$. As we have just seen, the latter is the same as the original BRST transformation – up to changes of normalizations. Hence, our task amounts to constructing the most general BRST invariant Lagrangian in the redefined fields and couplings. Strictly speaking, we should only require invariance to first order in the deformation parameter ϵ , but because of the rather rigid algebraic structure, this actually results in invariance to all orders in ϵ . Of course, as explained before, we also require that it is of dimension less or equal to four and invariant under all linear symmetries of S_R , namely Lorentz symmetry, global gauge symmetry, ghost phase rotations (implying total ghost number zero), and (in most gauges) antighost translations, in which case ω^* must appear as $\partial_\mu \omega^*$. Obviously, once we have constrained the form of $\mathcal{L}_N^{(\epsilon)} = \mathcal{L}_{\text{new}} + \epsilon \mathcal{L}_{N,\infty}$ we also have constrained the form of the diverging part $\mathcal{L}_{N,\infty}$ of $\mathcal{L}_N$.

It now follows from the above cohomology theorem applied to $\tilde{s}$ that the most general $\tilde{s}$ invariant local function is of the form $\mathcal{L}_N^{(\epsilon)} = \mathcal{L}'_N[A, \psi] + \tilde{s}\tilde{\Psi}$ with $\mathcal{L}'_N[A, \psi]$ invariant under the gauge symmetry with the redefined coupling and $\tilde{\Psi}$ of ghost number -1 . The additional requirements cited above imply that $\mathcal{L}'_N[A, \psi]$ is of dimension less or equal four, while $\tilde{\Psi}$ must be of dimension less or equal 3 and contain ω^* only as $\partial_\mu \omega^*$. Taking also into account Lorentz and global gauge invariance, we arrive at

$$\begin{aligned} \mathcal{L}_N^{(\epsilon)} &= -\frac{1}{4} \mathcal{Z}_A \tilde{F}^{\mu\nu}_\alpha \tilde{F}^{\alpha\mu\nu} + \tilde{\mathcal{L}}'_N[\psi, \tilde{D}_\mu \psi] + \tilde{s} \left(-\frac{\xi'}{2} \omega^*_\alpha h_\alpha + \frac{\mathcal{Z}_\omega}{\mathcal{Z}\mathcal{N}} \partial_\mu \omega^*_\alpha A^\mu_\alpha \right) \\ &= -\frac{1}{4} \mathcal{Z}_A \tilde{F}^{\mu\nu}_\alpha \tilde{F}^{\alpha\mu\nu} + \tilde{\mathcal{L}}'_N[\psi, \tilde{D}_\mu \psi] + \frac{\xi'}{2} h_\alpha h_\alpha + \frac{\mathcal{Z}_\omega}{\mathcal{Z}\mathcal{N}} h_\alpha \partial_\mu A^\mu_\alpha - \mathcal{Z}_\omega \partial_\mu \omega^*_\alpha (\tilde{D}^\mu \omega)_\alpha, \end{aligned} \quad (4.45)$$

where

$$\tilde{F}^\alpha_{\mu\nu} = \partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu + \tilde{C}_{\alpha\beta\gamma} A^\beta_\mu A^\gamma_\nu, \quad \tilde{D}_\mu \psi = \partial_\mu \psi - i A^a_\mu \tilde{t}^{\mathcal{R}}_\alpha \psi, \quad (\tilde{D}_\mu \omega)_\alpha = \partial_\mu \omega^\alpha + \tilde{C}_{\alpha\beta\gamma} A^\beta_\mu \omega^\gamma. \quad (4.46)$$

This is of the same form as the initial $\mathcal{L}_{\text{new}}$ except for the appearance of the “renormalization constants” $\mathcal{Z}_A$, $\mathcal{Z}_\omega$, a changed coupling via the $\frac{1}{\mathcal{N}}$ in $\tilde{C}_{\alpha\beta\gamma}$, as well as further “renormalization constants” hidden in $\tilde{\mathcal{L}}'_N[\psi, \tilde{D}_\mu \psi]$, and the replacement $\xi \rightarrow \xi'$. Let us insist that, on the redefined fields, the deformed BRST symmetry is just a BRST symmetry with a redefined coupling constant $\tilde{g} = g/\mathcal{N}$, and $\mathcal{L}_N^{(\epsilon)}$ is invariant under this $\tilde{g}$ -BRST symmetry for all ϵ .

4.2.3 Conclusion and remarks

It is enough to recall that $\mathcal{L}_N^{(\epsilon)} = \mathcal{L}_{\text{new}} + \epsilon \mathcal{L}_{N,\infty}$ to see that the diverging part $\mathcal{L}_{N,\infty}$ of $\mathcal{L}_N$ is “BRST-invariant” in the sense explained above, i.e. it is the difference of a $\tilde{g}$ -BRST and a g -BRST invariant local term. Since ϵ was arbitrary, the same is true for $-\mathcal{L}_{N,\infty}$ which is the required counterterm. One can now redo the argument order by order in the loop-expansion to see that the sum of all counterterms up to a given order N is “BRST-invariant” in this sense. Hence we conclude that to any order in perturbation theory the renormalized action is g -BRST invariant precisely if the bare action is g_B -BRST invariant. Said differently, by appropriately choosing the field renormalization

constants and coupling constant renormalization we get a finite quantum effective action Γ_N at every order N of perturbation theory.

Let us look more explicitly at the renormalization of the gauge field and of the gauge coupling constant g . Recall that the latter was included in the Lie algebra generators as $t_\alpha = g \hat{t}_\alpha$ and accordingly in the structure constants as $C_{\alpha\beta\gamma} = g \hat{C}_{\alpha\beta\gamma}$ with $\hat{t}_\alpha$ and $\hat{C}_{\alpha\beta\gamma}$ the more conventionally normalized, coupling-independent generators and structure constants. Thus $F_{\mu\nu}^\alpha = \partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha + g \hat{C}_{\beta\gamma}^\alpha A_\mu^\beta A_\nu^\gamma$, and similarly for $F_{B\mu\nu}^\alpha$ with A_B and g_B . If we let

$$A_{B\mu}^\alpha = \sqrt{Z_A} A_\mu^\alpha \quad , \quad g_B = \sqrt{Z_g} g \quad , \quad (4.47)$$

the relation between the bare and renormalized F^2 term is

$$\begin{aligned} -\frac{1}{4} F_{B\mu\nu}^\alpha F_B^{\alpha\mu\nu} &= -\frac{1}{4} F_{\mu\nu}^\alpha F^{\alpha\mu\nu} - \frac{1}{4} (Z_A - 1) (\partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha) (\partial^\mu A^{\alpha\nu} - \partial^\nu A^{\alpha\mu}) \\ &\quad - \frac{1}{2} g (\sqrt{Z_g} Z_A^{3/2} - 1) (\partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha) \hat{C}_{\beta\gamma}^\alpha A^{\beta\mu} A^{\gamma\nu} \\ &\quad - \frac{1}{4} g^2 (Z_g Z_A^2 - 1) \hat{C}_{\beta\gamma}^\alpha A_\mu^\beta A_\nu^\gamma \hat{C}_{\delta\epsilon}^\alpha A^{\delta\mu} A^{\epsilon\nu} . \end{aligned} \quad (4.48)$$

To determine only Z_A , at a given order, it is enough to compute the two-gauge boson function, i.e. the vacuum-polarization. However, to determine Z_g and hence the renormalization of the coupling constant, one has to determine two of the three counterterms on the rhs of (4.48) and hence compute e.g. the vacuum-polarization and the 3-gauge boson vertex function. This is related to the fact that the bare and renormalized actions have their BRST invariance defined with different couplings. In an abelian gauge theory like QED, there is no coupling dependence of the BRST symmetry and the Ward identity implies $(Z_g Z_A)_{\text{abelian}} = 1$ so that we could read the coupling constant renormalization from the vacuum-polarization. Below, we will discuss a special type of gauge fixing, the so-called background gauge, where an analogous relation holds also in the non-abelian theory and, hence, it is enough to compute a single Green's function to obtain the coupling constant renormalization and the β -function.

We have used the anti-ghost translation invariance of the so-called ξ -gauges $f_\alpha = \partial_\mu A_\alpha^\mu$, and this is why we get back this form. However, if one uses a gauge-fixing with $f_\alpha = \partial_\mu A_\alpha^\mu + a_{\alpha\beta\gamma} A_\beta^\mu A_{\mu\gamma}$ with some global G -tensor $a_{\alpha\beta\gamma}$, there is no ω^* -translation invariance to enforce the absence of BRST invariant terms with two ghosts and two anti-ghosts, like e.g. $\tilde{s} (b_{\alpha\beta\gamma} \omega_\alpha^* \omega_\beta^* \omega_\gamma) = -b_{\alpha\beta\gamma} (2h_\alpha \omega_\beta^* \omega_\gamma + \frac{1}{2} \mathcal{Z} C_{\gamma\delta\epsilon} \omega_\alpha^* \omega_\beta^* \omega_\delta \omega_\epsilon)$. Hence, such terms must be allowed in the action (and counterterms) from the beginning. However, such terms cannot be obtained from the Faddeev-Popov procedure (which always gives a ghost term linear in the ghost and linear in the anti-ghost). To get these terms necessitates a more general Lagrangian $\mathcal{L}_{\text{new}} = \mathcal{L}[A, \psi] + s\Psi[A, \psi, \omega, \omega^*, h]$, as already discussed above.

We can conclude that, as long as we start with a general Lagrangian of this form, with gauge invariant $\mathcal{L}$ of dimension less or equal 4, there is a counterterm available to cancel any divergence at any loop order, and the theory is renormalizable in the usual sense.

Heavy use was made of the Slavnov-Taylor identity $\int d^4x \langle s \chi^n \rangle \frac{\delta \Gamma}{\delta \chi^n} = 0$ from which the Zinn-Justin equation was obtained. The derivation of the Slavnov-Taylor identity assumed that

$\mathcal{D}\chi_n e^{i\int \mathcal{L}_{\text{new}}[\chi^n]}$ is BRST invariant and in particular that $\mathcal{D}\psi \mathcal{D}A_\mu e^{i\int \mathcal{L}[A,\psi]}$ is gauge invariant. Now $\int \mathcal{L}[A,\psi]$ is gauge invariant by construction but it may happen that $\mathcal{D}\psi$ is not gauge invariant. This typically shows up at one loop and then $(S_R, \Gamma_1) \neq 0$. Equivalently, certain one-loop diagrams are then seen to violate the corresponding Ward identities of covariant current conservation. In this case the theory is said to be anomalous. Anomalies are the subject of the next chapter. Fortunately, as we will see, anomalies can only appear for certain gauge groups and only in the presence of chiral fermions. Moreover, in potentially anomalous theories, by carefully arranging the content of the chiral matter fermions, the anomalies can be made to cancel. This is the case, in particular, for the standard model based on the gauge group $SU(3) \times SU(2) \times U(1)$ with the content of chiral matter fermions observed in nature.

4.3 Background field gauge

Suppose we can find a gauge-fixing procedure such that the one-loop quantum effective action²⁶ $\Gamma_{1\text{-loop}}[a, \psi_0, \omega_0, \omega_0^*]$ must be gauge invariant under the *same* transformation as was the tree-level renormalized action $S_R^{(0)}$, i.e. $\delta a_\mu^\alpha = \partial_\mu \epsilon^\alpha + g \hat{C}_{\beta\gamma}^\alpha a_\mu^\beta \epsilon^\gamma$ and $\delta \psi_0 = ig \epsilon^\alpha \hat{t}_\alpha^R \psi_0$, as well as some appropriate transformations for ω_0 and ω_0^* . Then necessarily we have

$$S_R^{(0)}[A, \dots] + S_{\text{c.t.}}^{1\text{-loop}}[A, \dots] + \dots = -\frac{1}{4} F_{\mu\nu}^\alpha F^{\alpha\mu\nu} - \frac{1}{4} (Z_A - 1) F_{\mu\nu}^\alpha F^{\alpha\mu\nu} + \dots = -\frac{1}{4} Z_A F_{\mu\nu}^\alpha F^{\alpha\mu\nu} + \dots \quad (4.49)$$

On the other hand, this must be the bare action. Comparing with (4.48), we see that then $Z_g Z_A = 1$ and, hence, $g_B = (Z_A)^{-1/2} g$. Since the one-loop β -function is given by the coefficient of $\frac{1}{\epsilon}$ in the expression of g_B in terms of g , we then find that²⁷

$$\beta_{1\text{-loop}} = -\frac{1}{2} g \times \left(\text{coeff of } \frac{1}{\epsilon} \text{ in } Z_A \right). \quad (4.50)$$

This will allow us to extract the one-loop β -function from the computation of the single coefficient of the F^2 -term in $\Gamma_{1\text{-loop}}$.

The quantum effective action $\Gamma[a, \psi_0, \bar{\psi}_0, \omega_0, \omega_0^*]$ was defined before in terms of the Legendre transformation of the generating functional W of connected Green's functions which is obtained from the functional integral of $e^{iS[A,\psi,\omega,\omega^*]}$ supplemented with the sources. One can convince oneself that an alternative definition is provided by

$$e^{i\Gamma[a,\psi_0,\bar{\psi}_0,\omega_0,\omega_0^*]} = \int_{1\text{PI}} \mathcal{D}A \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}\omega \mathcal{D}\omega^* e^{S[a+A,\psi_0+\psi,\bar{\psi}_0+\bar{\psi},\omega_0+\omega,\omega_0^*+\omega^*]}, \quad (4.51)$$

where, of course, S includes the gauge-fixing and ghost terms, and the subscript “1PI” instruct us to take into account only 1PI diagrams when diagrammatically evaluating this functional integral. In practice, this is not too convenient, and for gauge theories comes with a few subtleties. However,

²⁶Although we do not want to compute scattering amplitudes or Green's functions for external ghosts, nothing prevents us from computing Γ also with ghost “background fields” ω_0 and ω_0^* .

²⁷To lowest order in perturbation theory $g_B = (Z_A)^{-1/2} g = (1 + (Z_A - 1))^{-1/2} g \simeq (1 - \frac{1}{2}(Z_A - 1))g$.

as long as we are only interested in 1-loop diagrams – which are always 1PI – this is an effective prescription to get $\Gamma_{1\text{-loop}}$.

With this motivation in mind, let us now introduce the background field gauge. We want to find a gauge-fixing for A_μ such that $\Gamma[a]$ still is invariant under gauge transformations of the “background” field a_μ .

We introduce two types of gauge transformations: the “background field gauge transformations” δ_B and the “quantum field gauge transformations” δ_g . The background field gauge transformation is defined to act on the background *and* quantum fields as follows:

$$\begin{aligned} \delta_B a_\mu &= \partial_\mu \epsilon - i[a_\mu, \epsilon] \equiv D_\mu^B \epsilon & \Leftrightarrow & \delta_B a_\mu^\alpha = \partial_\mu \epsilon^\alpha + C_{\beta\gamma}^\alpha a_\mu^\beta \epsilon^\gamma , \\ \delta_B A_\mu &= -i[A_\mu, \epsilon] & \Leftrightarrow & \delta_B A_\mu^\alpha = C_{\beta\gamma}^\alpha A_\mu^\beta \epsilon^\gamma . \end{aligned} \quad (4.52)$$

Thus under this background gauge transformation a_μ transforms as a gauge field and A_μ as an adjoint matter field. Obviously, on the sum of both fields δ_B acts as an ordinary gauge transformation:

$$\delta_B(a_\mu + A_\mu) = \partial_\mu \epsilon - i[a_\mu + A_\mu, \epsilon] . \quad (4.53)$$

Similarly, the action on the matter quantum and background fields is defined as

$$\delta_B \psi_0 = i\epsilon^\alpha t_\alpha^{\mathcal{R}} \psi_0 \quad , \quad \delta_B \psi = i\epsilon^\alpha t_\alpha^{\mathcal{R}} \psi \quad \Rightarrow \quad \delta_B(\psi_0 + \psi) = i\epsilon^\alpha t_\alpha^{\mathcal{R}}(\psi_0 + \psi) , \quad (4.54)$$

and for the ghost fields

$$\delta_B \omega_0^\alpha = C_{\beta\gamma}^\alpha \omega_0^\beta \epsilon^\gamma \quad , \quad \delta_B \omega^\alpha = C_{\beta\gamma}^\alpha \omega^\beta \epsilon^\gamma \quad \Rightarrow \quad \delta_B(\omega_0^\alpha + \omega^\alpha) = C_{\beta\gamma}^\alpha (\omega_0^\beta + \omega^\beta) \epsilon^\gamma , \quad (4.55)$$

and idem for the antighost fields. On the other hand, the quantum field gauge transformations δ_g should not act on the background fields and will be defined such that they do act as standard gauge transformations on the sums $a_\mu + A_\mu$ and $\psi_0 + \psi$. Thus

$$\delta_g a_\mu = \delta_g \psi_0 = \delta_g \omega_0 = \delta_g \omega_0^* = 0 , \quad (4.56)$$

and

$$\begin{aligned} \delta_g A_\mu &= \partial_\mu \epsilon - i[a_\mu + A_\mu, \epsilon] & \Rightarrow & \delta_g(a_\mu + A_\mu) = \partial_\mu \epsilon - i[a_\mu + A_\mu, \epsilon] \\ \delta_g \psi &= i\epsilon^\alpha t_\alpha^{\mathcal{R}}(\psi_0 + \psi) & \Rightarrow & \delta_g(\psi_0 + \psi) = i\epsilon^\alpha t_\alpha^{\mathcal{R}}(\psi_0 + \psi) . \end{aligned} \quad (4.57)$$

In order to compute the one-loop quantum effective action as the functional integral (??) one has to choose a gauge-fixing function f_α in the action S . The choice for the background field gauge is

$$f_\alpha = (D_\mu^B A^\mu)_\alpha \equiv \partial_\mu A_\alpha^\mu + C_{\alpha\beta\gamma} A_\mu^\beta A^{\gamma\mu} . \quad (4.58)$$

Note that this is a generalization of the standard gauge fixing function to a non-vanishing background field a_μ . It is an appropriate choice for the present purpose, since D_μ^B is covariant under δ_B and thus $D_\mu^B A^\mu$ transforms as A^μ under δ_B :

$$\delta_B f_\alpha = C_{\alpha\beta\gamma} f^\beta \epsilon^\gamma \quad \Rightarrow \quad \delta_B(f_\alpha f_\alpha) = 0 , \quad (4.59)$$

and we see that the gauge-fixing term is invariant under the background gauge transformations δ_B . On the other hand, if f_α is to fix the gauge in the functional integral, it better be not invariant under the quantum field gauge transformations δ_g . Indeed, we have

$$\delta_g f_\alpha = \delta_g (D_\mu^B A^\mu)_\alpha = (D_\mu^B \delta_g A^\mu)_\alpha = (D_\mu^B D_B^\mu \epsilon)_\alpha - i(D_\mu^B [A^\mu, \epsilon])_\alpha , \quad (4.60)$$

from which we read the corresponding ghost Lagrangian

$$\mathcal{L}_{\text{gh}}[a_\mu, A_\mu, \omega, \omega^*] = -(D_\mu^B \omega^*)_\alpha (D_B^\mu \omega - [A^\mu, \omega])_\alpha . \quad (4.61)$$

Since ω , ω^* and A_μ transform as adjoint matter fields under δ_B , and since D_μ^B is covariant with respect to δ_B , we conclude that this ghost Lagrangian is invariant under the background field gauge transformations, $\delta_B \mathcal{L}_{\text{gh}} = 0$. Since also ω_0 and ω_0^* transform as adjoint matter fields under δ_B , the same remains true if one replaces ω by $\omega_0 + \omega$ and ω^* by $\omega_0^* + \omega^*$:

$$\delta_B \mathcal{L}_{\text{gh}}[a_\mu, A_\mu, \omega_0 + \omega, \omega_0^* + \omega^*] = 0 , \quad (4.62)$$

so that finally

$$\delta_B \mathcal{L}_{\text{mod}}[a_\mu, A_\mu, \psi_0 + \psi, \omega_0 + \omega, \omega_0^* + \omega^*] = 0 \quad , \quad \mathcal{L}_{\text{mod}} = \mathcal{L} - \frac{1}{2\xi} f_\alpha f_\alpha + \mathcal{L}_{\text{gh}} . \quad (4.63)$$

Note that $\mathcal{L}[a + A, \psi_0 + \psi] = -\frac{1}{4} F_{\mu\nu}[a + A] F^{\mu\nu}[a + A] + \mathcal{L}_{\text{matter}}[a + A, \psi_0 + \psi]$ with $F_{\mu\nu}[a + A] = F_{\mu\nu}[a] + D_\mu^B A_\nu - D_\nu^B A_\mu - i[A_\mu, A_\nu]$.

Our goal was to determine $\Gamma[a_\mu, \dots]$ as $\int_{\text{1PI}} \mathcal{D}A \dots \exp(i \int \mathcal{L}_{\text{mod}}(a + A, \dots))$. We see that due to the gauge fixing and ghost terms, $\mathcal{L}_{\text{new}}$ does not only depend on the sum $a + A$ but on a and A separately. It is then not obvious any more that this does generate *all* 1PI diagrams. However, if we restrict ourselves to the one-loop quantum effective action, things are simpler and the separate dependence on a and A of the gauge-fixing and ghost terms does not cause any trouble. Since $\mathcal{L}_{\text{mod}}$ is invariant under the background gauge transformations δ_B , as is also the measure $\mathcal{D}A \mathcal{D}\psi \mathcal{D}\omega \mathcal{D}\omega^*$ (excluding anomalies), it is then obvious that $\Gamma_{1\text{-loop}}[a, \psi_0, \dots]$ must also be invariant under these transformations:

$$\delta_B \Gamma_{1\text{-loop}}[a, \psi_0, \dots] = 0 . \quad (4.64)$$

But the background gauge transformations act on the background fields just as ordinary gauge transformations (in particular with the same normalisation of $g a_\mu^\alpha \hat{t}_\alpha$ as explained at the beginning of this subsection). Then eq. (4.49) must hold, and hence $Z_g Z_A = 1$ so that we can use (4.50) to compute the one-loop β -function.

4.4 One-loop β -functions for Yang-Mills and supersymmetric Yang-Mills theories

To extract the one-loop β -function we need to compute the coefficient of $\frac{1}{\epsilon}$ in Z_A in the background field gauge. $(Z_A - 1)$ is the coefficient of $-\frac{1}{4}F^2$ in the counterterm. Hence we need the coefficient of $\frac{1}{\epsilon}$ in front of the $-\frac{1}{4}F^2$ piece of the counterterm. To extract this term, we may simply consider a constant background gauge field a_μ and all other background fields vanishing. In particular then, $F_{\mu\nu}^\alpha(a) = C_{\beta\gamma}^\alpha a_\mu^\beta a_\nu^\gamma$.

4.4.1 β -function for Yang-Mills theory

As explained above, to compute the one-loop contribution to $\Gamma[a]$ we have to identify the part in $\int \mathcal{L}_{\text{mod}}[a, A, \psi, \omega, \omega^*]$ that is quadratic in the quantum fields $A, \psi, \omega, \omega^*$. In particular, $F_{\mu\nu}[a + A] F^{\mu\nu}[a + A]|_{A^2\text{-part}} = (D_\mu^B A_\nu - D_\nu^B A_\mu)(D_B^\mu A^\nu - D_B^\nu A^\mu) - 2F_{\mu\nu}[a][A^\mu, A^\nu]$. The expansion of $\mathcal{L}_{\text{mod}}$ then yields

$$\int \mathcal{L}_{\text{mod}}[a, A, \psi, \omega, \omega^*]|_{\text{quadratic part}} = \int d^4x d^4y \left(-\frac{1}{2} A^{\alpha\mu}(x) \mathcal{D}_{x\alpha\mu, y\beta\nu}^A[a] A^{\beta\nu}(y) - \bar{\psi}_k(x) \mathcal{D}_{xk, yl}^\psi[a] \psi_l(y) - \omega^{*\alpha}(x) \mathcal{D}_{x\alpha, y\beta}^\omega[a] \omega^\beta(y) \right), \quad (4.65)$$

where

$$\mathcal{D}_{x\alpha, y\beta}^\omega[a] = \left(-\delta_{\alpha\gamma} \frac{\partial}{\partial x^\mu} - C_{\alpha\delta\gamma} a_\mu^\delta(x) \right) \left(-\delta_{\gamma\beta} \frac{\partial}{\partial y_\mu} + C_{\gamma\epsilon\beta} a_\epsilon^\mu(y) \right) \delta^{(4)}(x - y), \quad (4.66)$$

or, using a matrix notation with $a_\mu^{\text{adj}} = a_\mu^\delta t_\delta^{\text{adj}}$ and the generators of the adjoint representation given in (3.6),

$$\mathcal{D}_{x, y}^\omega[a] = \left(-\frac{\partial}{\partial x^\mu} + i a_\mu^{\text{adj}}(x) \right) \left(-\frac{\partial}{\partial y_\mu} - i a_{\text{adj}}^\mu(y) \right) \delta^{(4)}(x - y). \quad (4.67)$$

Similarly, one has

$$\mathcal{D}_{x, y}^\psi[a] = \left(-\gamma^\mu \frac{\partial}{\partial y_\mu} - i \gamma^\mu a_\mu^{\mathcal{R}}(y) + m \right) \delta^{(4)}(x - y), \quad (4.68)$$

and

$$\begin{aligned} \mathcal{D}_{x\mu, y\nu}^A[a] = & \left\{ \eta_{\mu\nu} \left(-\frac{\partial}{\partial x^\rho} + i a_\rho^{\text{adj}}(x) \right) \left(-\frac{\partial}{\partial y_\rho} - i a_{\text{adj}}^\rho(y) \right) - \left(-\frac{\partial}{\partial x^\nu} + i a_\nu^{\text{adj}}(x) \right) \left(-\frac{\partial}{\partial y^\mu} - i a_\mu^{\text{adj}}(y) \right) \right. \\ & \left. - i F_{\mu\nu}^{\text{adj}}[a] + \frac{1}{\xi} \left(-\frac{\partial}{\partial x^\mu} + i a_\mu^{\text{adj}}(x) \right) \left(-\frac{\partial}{\partial y^\nu} - i a_\nu^{\text{adj}}(y) \right) \right\} \delta^{(4)}(x - y). \end{aligned} \quad (4.69)$$

Then

$$i\Gamma_{1\text{-loop}}[a] = -\frac{1}{2} \text{Tr} \log \mathcal{D}^A[a] + \text{Tr} \log \mathcal{D}^\psi[a] + \text{Tr} \log \mathcal{D}^\omega[a]. \quad (4.70)$$

These traces are evaluated as usual. We will sketch the computation for $\text{Tr} \log \mathcal{D}^\omega[a]$, the others being similar.

Since we will only consider constant fields a_μ it is most convenient to Fourier transform. One has in general $\mathcal{D}_{x, y} \equiv \langle x | \mathcal{D} | y \rangle$ and hence

$$\langle p | \mathcal{D} | q \rangle = \int d^4x d^4y \langle p | x \rangle \langle x | \mathcal{D} | y \rangle \langle y | q \rangle = \frac{1}{(2\pi)^4} \int d^4x d^4y e^{-ipx} \langle x | \mathcal{D} | y \rangle e^{iqy}. \quad (4.71)$$

With a $\mathcal{D}_{x, y}$ of the form $\mathcal{D}_{x, y} = f(-\frac{\partial}{\partial x^\mu}) g(-\frac{\partial}{\partial y_\mu}) \delta^{(4)}(x - y)$ this yields

$$\begin{aligned} \langle p | \mathcal{D} | q \rangle &= \frac{1}{(2\pi)^4} \int d^4x d^4y e^{-ipx} \left[f\left(-\frac{\partial}{\partial x^\mu}\right) g\left(-\frac{\partial}{\partial y_\mu}\right) \delta^{(4)}(x - y) \right] e^{iqy} = \frac{1}{(2\pi)^4} \int d^4x e^{i(q-p)x} f(-ip) g(iq) \\ &= f(-ip) g(iq) \delta^{(4)}(p - q) \equiv M(p) \delta^{(4)}(p - q). \end{aligned} \quad (4.72)$$

Hence $\langle p | \mathcal{D}^2 | q \rangle = \int d^4k \langle p | \mathcal{D} | k \rangle \langle k | \mathcal{D} | q \rangle = M^2(p) \delta^{(4)}(p - q)$, and similarly for any power n and any function of $\mathcal{D}$, and thus

$$\text{Tr} \log \mathcal{D} = \int d^4p \langle p | \text{tr} \log \mathcal{D} | p \rangle = \int d^4p \delta^{(4)}(p - p) \text{tr} \log M(p). \quad (4.73)$$

Of course, $\delta^{(4)}(p-p)$ arises because we work with constant fields and it has to be interpreted as $\int \frac{d^4x}{(2\pi)^4} e^{ix(p-p)} = \frac{1}{(2\pi)^4} \int d^4x$. Our $M(p)$ are all of the form $M(p) = M_0(p) + M_1(p) + M_2(p)$ with $iM_0(p)$ being the inverse propagator for the given field and M_1 and M_2 are linear and bilinear in the background fields a_μ . One has

$$\text{tr} \log M = \text{tr} \log M_0 + \text{tr} \log (1 + M_0^{-1}(M_1 + M_2)) = \text{tr} \log M_0 + \text{tr} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (M_0^{-1}(M_1 + M_2))^n. \quad (4.74)$$

We need to pick out the quartic term in a_μ :

$$\text{tr} \log M|_{a^4} = -\frac{1}{2} \text{tr} (M_0^{-1} M_2)^2 + \text{tr} (M_0^{-1} M_1)^2 M_0^{-1} M_2 - \frac{1}{4} \text{tr} (M_0^{-1} M_1)^4. \quad (4.75)$$

Since $-iM_0^{-1}$ is just the propagator this expansion generates the relevant one-loop Feynman diagrams with 4 external (background) gauge fields a_μ attached. There are vertices with one and vertices with two gauge fields, corresponding to M_1 and M_2 . Obviously, one could have evaluated these diagrams directly using the appropriate Feynman rules. The present alternative computation has the advantage of giving all combinatorial factors in a straightforward way.

Obviously, the $\mathcal{D}^\omega[a]$ corresponds to the ghost loop and the above expression gives

$$\begin{aligned} \text{Tr} \log \mathcal{D}^\omega[a] &= \int d^4x \int \frac{d^4p}{(2\pi)^4} \text{tr} \log M^\omega(p), \\ M^\omega(p) &= (-ip_\mu + i a_\mu^{\text{adj}})(ip^\mu - i a_\mu^{\text{adj}}) = p_\mu p^\mu - 2p^\mu a_\mu^{\text{adj}} + a_\mu^{\text{adj}} a_\mu^{\text{adj}}. \end{aligned} \quad (4.76)$$

We read off $M_0^\omega(p) = p_\mu p^\mu$, $M_1^\omega(p) = -2p^\mu a_\mu^{\text{adj}}$ and $M_2^\omega(p) = a_\mu^{\text{adj}} a_\mu^{\text{adj}}$ and (4.75) then gives

$$\int \frac{d^4p}{(2\pi)^4} \text{tr} \log M^\omega|_{a^4} = \int \frac{d^4p}{(2\pi)^4} \left(-\frac{1}{2} \frac{\text{tr} a_\lambda a^\lambda a_\sigma a^\sigma}{(p^2 - i\epsilon)^2} + 4 \frac{\text{tr} p_\mu a^\mu p_\nu a^\nu a_\lambda a^\lambda}{(p^2 - i\epsilon)^3} - 4 \frac{\text{tr} (p_\mu a^\mu)^4}{(p^2 - i\epsilon)^4} \right), \quad (4.77)$$

where all the a_μ are in the adjoint representation although we did not write it explicitly. Since all a_μ are constant, corresponding to vanishing external momenta, the loop integrals are particularly simple. Using dimensional regularization ($d = 4 - \epsilon$) and introducing an IR regulator μ , we have²⁸

$$\begin{aligned} I &\equiv \int \frac{d^d p}{(2\pi)^4} \frac{1}{(p^2 + \mu^2 - i\epsilon)^2} = \frac{i}{16\pi^2} (\pi\mu^2)^{-\epsilon/2} \left(\frac{2}{\epsilon} - \gamma + \mathcal{O}(\epsilon) \right) \\ I_{\mu\nu} &\equiv \int \frac{d^d p}{(2\pi)^4} \frac{p_\mu p_\nu}{(p^2 + \mu^2 - i\epsilon)^3} = \frac{1}{4} \eta_{\mu\nu} I \\ I_{\mu\nu\rho\sigma} &\equiv \int \frac{d^d p}{(2\pi)^4} \frac{p_\mu p_\nu p_\rho p_\sigma}{(p^2 + \mu^2 - i\epsilon)^4} = \frac{1}{24} (\eta_{\mu\nu} \eta_{\rho\sigma} + \eta_{\mu\rho} \eta_{\nu\sigma} + \eta_{\mu\sigma} \eta_{\nu\rho}) I. \end{aligned} \quad (4.78)$$

Thus

$$\begin{aligned} \int \frac{d^4p}{(2\pi)^4} \text{tr} \log M^\omega|_{a^4} &= I \left(-\frac{1}{2} \text{tr} a_\lambda a^\lambda a_\sigma a^\sigma + \eta_{\mu\nu} \text{tr} a^\mu a^\nu a_\lambda a^\lambda - \frac{1}{6} (\eta_{\mu\nu} \eta_{\rho\sigma} + \eta_{\mu\rho} \eta_{\nu\sigma} + \eta_{\mu\sigma} \eta_{\nu\rho}) \text{tr} a^\mu a^\nu a^\rho a^\sigma \right) \\ &= \frac{1}{6} I \left(\text{tr} a_\lambda a^\lambda a_\sigma a^\sigma - \text{tr} a_\lambda a_\sigma a^\lambda a^\sigma \right) = -\frac{1}{6} I \left(\text{tr} a_\lambda a_\sigma a^\lambda a^\sigma - \text{tr} a_\sigma a_\lambda a^\lambda a^\sigma \right). \end{aligned} \quad (4.79)$$

Recall that for constant a_μ in the adjoint representation one has $\text{tr}_{\text{adj}} F_{\mu\nu} F^{\mu\nu} = \text{tr}_{\text{adj}} (-i)[a_\mu, a_\nu](-i)[a^\mu, a^\nu] = -2 \text{tr}_{\text{adj}} (a_\mu a_\nu a^\mu a^\nu - a_\mu a_\nu a^\nu a^\mu)$, so that finally

$$\text{Tr} \log \mathcal{D}^\omega[a] = \int d^4x \frac{1}{12} I \text{tr}_{\text{adj}} F_{\mu\nu} F^{\mu\nu}. \quad (4.80)$$

²⁸Note that when computing $I_{\mu\nu}$ one replaces $p_\mu p_\nu$ by $\frac{1}{d} \eta_{\mu\nu} p^2$. The remaining scalar integral yields $\frac{d}{4} I$ so that $I_{\mu\nu} = \frac{1}{4} \eta_{\mu\nu} I$ and not $\frac{1}{d} \eta_{\mu\nu} I$ as one might have naively expected. Similarly for $I_{\mu\nu\rho\sigma}$.

The evaluation of $\text{Tr log } \mathcal{D}^A[a]$ is very similar, except that the tensorial structure is a bit more complicated. In the final result, the $\frac{1}{12}$ is replaced by a $-\frac{5}{3}$. The matter determinant $\text{Tr log } \mathcal{D}^\psi[a]$ depends on the mass of the fermions. However, the structure of the divergence is mass independent, i.e. the term $\sim \frac{1}{\epsilon}$ does not depend on the mass, and if our only purpose is to compute the β -function we can just as well neglect the fermion mass.

In the limit where the fermion masses can be neglected, the three traces are:

$$\begin{aligned}\text{Tr log } \mathcal{D}^A|_{a^4\text{-piece}} &= I \int d^4x \left(-\frac{5}{3} \text{tr}_{\text{adj}} F_{\mu\nu} F^{\mu\nu} \right) \\ \text{Tr log } \mathcal{D}^\omega|_{a^4\text{-piece}} &= I \int d^4x \left(\frac{1}{12} \text{tr}_{\text{adj}} F_{\mu\nu} F^{\mu\nu} \right) \\ \text{Tr log } \mathcal{D}^\psi|_{a^4\text{-piece}} &= I \int d^4x \left(-\frac{1}{3} \text{tr}_{\mathcal{R}} F_{\mu\nu} F^{\mu\nu} \right),\end{aligned}\tag{4.81}$$

where (μ is an IR regulator) $I = \int \frac{d^d p}{(2\pi)^4} \frac{1}{(p^2 + \mu^2 - i\epsilon)^2} = \frac{i}{8\pi^2} \left(\frac{1}{\epsilon} + \text{finite} \right)$. If we let

$$\text{tr}_{\text{adj}} t_\alpha t_\beta = g^2 C_1 \delta_{\alpha\beta}, \quad \text{tr}_{\mathcal{R}} t_\alpha t_\beta = g^2 C_2 \delta_{\alpha\beta},\tag{4.82}$$

we get

$$\Gamma_{1\text{-loop}}^{\text{div}}[a] = -\frac{g^2}{2\pi^2} \frac{1}{\epsilon} \left(\frac{5}{6} C_1 + \frac{1}{12} C_1 - \frac{1}{3} C_2 \right) \int d^4x \left(-\frac{1}{4} F_{\mu\nu}^\alpha F^{\alpha\mu\nu} \right),\tag{4.83}$$

and, hence

$$Z_A|_{1\text{-loop}}^{\text{div}} = +\frac{g^2}{2\pi^2} \frac{1}{\epsilon} \left(\frac{11}{12} C_1 - \frac{1}{3} C_2 \right)\tag{4.84}$$

Finally, according to (4.50), we read off the one-loop β -function as

$$\boxed{\beta_{1\text{-loop}} = -\frac{g^3}{4\pi^2} \left(\frac{11}{12} C_1 - \frac{1}{3} C_2 \right)},\tag{4.85}$$

which is negative as long as $C_2 < \frac{11}{4} C_1$, i.e. as long as there are not “too many” matter fields. Of course, this is a famous result. For example, if the gauge group is $SU(N)$ and the matter representation $\mathcal{R}$ is the defining “vector” representation N , one has $C_1 = N$ and $C_2 = \frac{1}{2}$. For n_f flavours of quarks in the N -representation one has $C_2 = \frac{n_f}{2}$. Thus

$$\beta_{1\text{-loop}}^{SU(N)\text{-QCD with } n_f \text{ flavours of quarks}} = -\frac{g^3}{4\pi^2} \left(\frac{11}{12} N - \frac{n_f}{6} \right),\tag{4.86}$$

Standard QCD with gauge group $SU(3)$ could accommodate as many as $n_f = 16$ flavours of quarks and still remain asymptotically free. Recall from our earlier discussion that the relevant β -function at a given scale μ is the one that corresponds to the number of flavours of quarks having masses (well) below μ . Any quarks having masses (well) above μ do not contribute at this scale. Let us write $\beta_{n_f}(g) = -\beta_0^{n_f} \frac{g^3}{16\pi^2}$ and also use $\alpha(\mu) = \frac{g^2(\mu)}{4\pi}$. As shown before, the running of $g(\mu)$ is such that

$$\Lambda_{n_f} = \mu \exp \left(-\frac{8\pi^2}{\beta_0^{n_f} g^2(\mu)} \right) = \mu \exp \left(-\frac{2\pi}{\beta_0^{n_f} \alpha(\mu)} \right)\tag{4.87}$$

does not depend on μ . It is the RG-invariant mass scale characterizing the strength of the interaction. More precisely, it is independent of μ as long as μ does not come close to any of the quark masses and the effective number of flavours does not change. Since the running of the coupling below m_{n_f+1} is well approximated by β_{n_f} and above m_{n_f+1} by β_{n_f+1} , matching the coupling at $\mu = m_{n_f+1}$ leads to the RG scale matching condition

$$\left(\frac{\Lambda_{n_f}}{m_{n_f+1}}\right)^{\beta_0^{n_f}} = \left(\frac{\Lambda_{n_f+1}}{m_{n_f+1}}\right)^{\beta_0^{n_f+1}} . \quad (4.88)$$

It is useful to isolate the contributions of the various fields to the one-loop β -function. There are 3 basic contributions: the gauge and ghost fields, spin- $\frac{1}{2}$ matter fields and scalar matter fields. We already have computed the first two types of contribution, while the contribution of a complex scalar matter field can be obtained from the one of the ghost fields. Indeed, the ghosts are (anticommuting) scalars in the adjoint representation and with the obvious changes we get ($C_{\text{adj}} \equiv C_1$ and $C_{\mathcal{R}} \equiv C_2$)

$$\begin{aligned} \beta_{1\text{-loop}}^{\text{gauge and ghost}} &= -\frac{g^3}{4\pi^2} \frac{11}{12} C_{\text{adj}} \\ \beta_{1\text{-loop}}^{\text{Dirac fermion in } \mathcal{R}} &= -\frac{g^3}{4\pi^2} \left(-\frac{1}{3} C_{\mathcal{R}}\right) \\ \beta_{1\text{-loop}}^{\text{complex scalar in } \mathcal{R}} &= -\frac{g^3}{4\pi^2} \left(-\frac{1}{12} C_{\mathcal{R}}\right) . \end{aligned} \quad (4.89)$$

Note that we considered standard Dirac fermions. If one considers instead chiral (Weyl) fermions or Majorana fermions, one has to divide the corresponding matter contribution by a factor of 2.

4.4.2 β -functions in supersymmetric gauge theories

In supersymmetric theories the fields are part of supersymmetry multiplets and thus the field content is subject to important constraints. We will now show that this leads to simpler expressions for the one-loop β -functions.

$\mathcal{N} = 1$:

In gauge theories with the minimal amount of supersymmetry, so called unextended supersymmetry, often referred to as $\mathcal{N} = 1$ supersymmetry, the gauge fields have a corresponding partner, called the gaugino field, which is a spin- $\frac{1}{2}$ Majorana fermion $\lambda^\alpha(x)$, also in the adjoint representation and also massless. More precisely, the boson-fermion correspondence holds for *physical* states, so that the Majorana fermion field is the partner of the gauge and ghost fields. Together they form the gauge or vector multiplet. The matter fields are organized in so-called chiral multiplets, each multiplet consisting of a Majorana fermion $\psi(x)$ and a complex scalar $\phi(x)$, both fields in the same representation $\mathcal{R}$ of the gauge group. Since a Dirac field is equivalent to two Majorana fields, the β -function of a Majorana fermion is half that of a Dirac fermion. Hence, $\beta_{1\text{-loop}}^{\text{gaugino}} = \frac{g^3}{4\pi^2} \frac{1}{6} C_{\text{adj}}$, so that

$$\beta_{1\text{-loop}}^{\text{vector multiplet}} = \beta_{1\text{-loop}}^{\text{gauge and ghost and gaugino}} = -\frac{g^3}{4\pi^2} \frac{3}{4} C_{\text{adj}} . \quad (4.90)$$

Similarly, for a chiral multiplet

$$\beta_{1\text{-loop}}^{\text{chiral multiplet}} = \frac{1}{2}\beta_{1\text{-loop}}^{\text{Dirac fermion in } \mathcal{R}} + \beta_{1\text{-loop}}^{\text{complex scalar in } \mathcal{R}} = -\frac{g^3}{4\pi^2} \left(-\frac{1}{4} C_{\mathcal{R}} \right) . \quad (4.91)$$

Obviously then, in an $\mathcal{N} = 1$ supersymmetric gauge theory with $n_{C,i}$ chiral multiplets in the representation $\mathcal{R}_i$:

$$\beta_{1\text{-loop}}^{\mathcal{N}=1 \text{ with } n_{C,i} \text{ in } \mathcal{R}_i} = -\frac{g^3}{4\pi^2} \left(\frac{3}{4} C_{\text{adj}} - \frac{1}{4} \sum_i n_{C,i} C_{\mathcal{R}_i} \right) . \quad (4.92)$$

Note that such a theory with 3 chiral multiplets in the adjoint representation has a vanishing one-loop β -function. With appropriate masses and couplings, this theory actually has an extended supersymmetry.

$\mathcal{N} = 2$:

In $\mathcal{N} = 2$ extended supersymmetric gauge theory the $\mathcal{N} = 1$ multiplets are grouped into larger $\mathcal{N} = 2$ multiplets. Thus the $\mathcal{N} = 2$ vector or gauge multiplet consists of the $\mathcal{N} = 1$ vector multiplet together with a massless $\mathcal{N} = 1$ chiral multiplet in the adjoint representation. Hence

$$\beta_{1\text{-loop}}^{\mathcal{N}=2 \text{ vector multiplet}} = -\frac{g^3}{4\pi^2} \frac{1}{2} C_{\text{adj}} . \quad (4.93)$$

$\mathcal{N} = 2$ matter multiplets are so-called hypermultiplets which consist of two $\mathcal{N} = 1$ chiral multiplets:

$$\beta_{1\text{-loop}}^{\mathcal{N}=2 \text{ hypermultiplet}} = -\frac{g^3}{4\pi^2} \left(-\frac{1}{2} C_{\mathcal{R}} \right) , \quad (4.94)$$

so that an $\mathcal{N} = 2$ theory with $n_{H,i}$ hypermultiplets in representations $\mathcal{R}_i$ has its one-loop β -function given by

$$\beta_{1\text{-loop}}^{\mathcal{N}=2 \text{ with } n_{H,i} \text{ in } \mathcal{R}_i} = -\frac{g^3}{4\pi^2} \left(\frac{1}{2} C_{\text{adj}} - \frac{1}{2} \sum_i n_{C,i} C_{\mathcal{R}_i} \right) . \quad (4.95)$$

Again, we see that the special theory with a single adjoint hypermultiplet has a vanishing one-loop β -function.

$\mathcal{N} = 4$:

The maximally extended (global) supersymmetry is $\mathcal{N} = 4$ extended supersymmetric gauge theory. It has a single $\mathcal{N} = 4$ multiplet consisting of an $\mathcal{N} = 2$ vector and a massless $\mathcal{N} = 2$ adjoint hypermultiplet. As just noted, this theory has a vanishing one-loop β -function:

$$\beta_{1\text{-loop}}^{\mathcal{N}=4} = 0 . \quad (4.96)$$

Part II : Anomalies in QFT

This part, as well as most of Sect. 1 above, are reprinted from Ada Cubukcu and Adel Bilal “A pedagogical Introduction to Anomalies in Quantum Field Theory”, arXiv:2609.28835 [hep-th].

5 Overview

An anomaly is the breaking of a classical symmetry by quantum effects. This was first observed in 1969 as a non-conservation of the axial-vector current by Adler [1] and Jackiw and Johnson [2]. Since (global) symmetries in quantum mechanics impose selection rules, the breaking of such a symmetry implies that forbidden transitions actually become allowed. It was quickly realized that this current non-conservation provided an explanation for the observed disintegration rate of the neutral pion into two photons [3]. While tree-level Feynman diagrams essentially only reflect classical physics, typically quantum effects show up in loop diagrams. In particular, the current non-conservation is seen by explicitly computing a certain one-loop diagram with three vertices (triangle diagram). It has been shown by Adler and Bardeen that there are no higher-loop corrections to this anomaly [4, 5]. This is quite amazing, as it means that one can get an exact expression for this anomaly by solely computing the one-loop diagram - a rare situation in quantum field theory ! This anomaly is now referred to as the abelian anomaly or the Adler-Bardeen-Jackiw (ABJ) anomaly.

But anomalies also occur for local symmetries such as gauge symmetries in (abelian or non-abelian) gauge theories, as well as in quantum field theories on curved space-times, where they manifest themselves as a lack of general covariance or of local Lorentz symmetry (gravitational anomalies). A very nice overview is given in the first 5 chapters of a classical paper by Alvarez-Gaumé and Witten [6]. These anomalies again show up in one-loop diagrams in the presence of chiral fermions, and are often referred to as chiral anomalies (although “gauge anomalies” would be a more fitting name). A rather detailed discussion of these anomalies can be found in many places, including [7]-[13], as well as [14].²⁹ Again, only one-loop diagrams contribute to the anomaly, although there are different such diagrams that are anomalous (triangle, box).³⁰ However, the form of the anomaly is constrained by the so-called Wess-Zumino consistency condition [15] which relates the contributions of the different diagrams, so that only an explicit computation of the triangle is necessary.

The Wess-Zumino consistency condition is very conveniently reformulated as the anomaly being the BRST variation of the effective action and, hence, it is BRST closed. The freedom to add local counterterms to the classical action appears as the possibility to add BRST-exact terms to the anomaly. Relevant anomalies then appear as BRST cohomology classes at ghost number one. These cohomology classes are related by the descent equations [16] to certain characteristic classes in two more dimensions. The latter uniquely characterize the anomaly. This provides an intriguing link with index theorems for the Dirac operator [17], as spelled out in [18], which is particularly useful when discussing gravitational anomalies in higher dimensions.

²⁹Although reference [14] by one of the authors has inspired several parts of these notes, the present introduction to anomalies is written in a different spirit : it is meant to be at the same time more concise and more self-contained, omitting certain aspects and emphasizing others.

³⁰The same is true for the abelian anomaly.

Contrary to anomalies in global symmetries, an anomaly of a local (gauge) symmetry is a real problem, as it would eventually ruin the renormalizability [19], unitarity and/or Lorentz invariance of the theory. (Similar remarks apply to gravitational anomalies.) Luckily, not all gauge groups lead to anomalies [20], and when they do, different chiral fermions contribute additively to the anomaly. Thus with the right combination of chiral fermions it may happen that the individual anomalies cancel and the total anomaly vanishes. This is precisely the case in the standard model of particle physics as was first shown by Bouchiat, Iliopoulos and Meyer [21], as well as in several prominent higher-dimensional theories that appear as low-energy limits of certain superstring theories [22], such as the ten-dimensional type I supergravity coupled to super Yang-Mills theory for specific gauge groups, or type IIB supergravity in 10 dimensions.

6 Various manifestations of an anomaly

In this section, we will give a brief overview of how an anomaly can arise and how it can manifest itself. We begin by recalling Noether's theorem that asserts the existence of a classically conserved current for any symmetry of the classical action, and derive the conservation of the corresponding quantum current in the functional integral formalism. This conservation is seen to hold provided the functional integration measure also is invariant. Said differently, this shows that an anomaly can arise from a non-trivial transformation of the integration measure. We will explicitly compute this anomaly for chiral transformations (involving γ_5) of Dirac fermions, and more specifically see how it arises from the regularization process of the determinant of the Jacobian. Secondly, we will derive the anomalous Ward identities (also called Slavnov-Taylor identities) that are identities between n -point functions and the anomaly. This will allow us below, in Sect. 4, to do an explicit Feynman diagram computation of different anomalies. We will also look at anomalies under gauge transformations and how the (covariant) non-conservation of the gauge current is related to the non-invariance of an appropriately defined effective action, with this non-invariance again related to the anomalous transformation of the functional integration measure. We will mostly follow [14, 7]. In the next section, we will also explicitly see how the anomaly arises in the computation of Feynman diagrams when one cannot regularize a divergent Feynman diagram in a way that respects all the symmetries.

6.1 Noether's theorem and a Ward identity

Noether's theorem states that to any global symmetry of a classical field theory corresponds a classically conserved current. Let us give the proof for a relativistically invariant field theory described by an action $S = \int d^4x \mathcal{L}(x)$ with $\mathcal{L}(x)$ a scalar density. Recall that the classical field equations are the Euler-Lagrange equations obtained from requiring stationarity of the action : under arbitrary variations $\delta\psi$ of the fields ψ (subject to the boundary condition of vanishing at infinity) one has $\delta S = 0$ if the fields satisfy the Euler-Lagrange equations. Assume now that the action is invariant, i.e. $\delta S = 0$, under some infinitesimal variation $\psi(x) \rightarrow \psi(x) + \Delta\psi(x)$ where $\Delta\psi(x) = \epsilon \mathcal{F}(\psi(x))$, for constant infinitesimal ϵ . Then, for *non-constant* $\epsilon(x)$ there will be a non-vanishing variation of S that must involve the derivatives of ϵ and thus be of the form $\delta S = - \int d^4x j^\mu(x) \partial_\mu \epsilon(x)$ with some

four-vector $j^\mu(x)$. Upon integrating by parts, this is

$$\delta S = \int d^4x (\partial_\mu j^\mu(x)) \epsilon(x) . \quad (6.1)$$

In general, this has no reason to vanish. But if the fields satisfy the Euler-Lagrange equations then, as just explained, one must have $\delta S = 0$. Since $\epsilon(x)$ is arbitrary, this implies that j^μ is a conserved current :

$$\partial_\mu j^\mu(x) \Big|_{\text{field equations}} = 0 . \quad (6.2)$$

In the quantum theory, there will be a corresponding (quantum) current operator $J^\mu(x)$. (This is not to be confused with the source fields of the generating functionals $Z[J]$.) The question then is whether this quantum current is still conserved. More precisely, one needs to check whether $\partial_\mu J^\mu(x)$ vanishes as an operator, which means that all time-ordered vacuum expectation values of products of operators $O_j(\Psi(y_j))$ and a factor $\partial_\mu J^\mu(x)$ should vanish :

$$\langle \text{vac, out} | T \{ \partial_\mu J^\mu(x) \prod_j O_j(\Psi(y_j)) \} | \text{vac, in} \rangle = 0 . \quad (6.3)$$

If this is not the case, one has an anomaly. Note that (6.3) is not the same as the statement $\frac{\partial}{\partial x^\mu} \langle \text{vac, out} | T \{ J^\mu(x) \prod_j O_j(y_j) \} | \text{vac, in} \rangle = 0$, since the time-ordering involves various $\theta(x^0 - y_j^0)$ and $\theta(y_j^0 - x^0)$ and their time derivatives generate $\delta(x^0 - y_j^0)$ times (equal-time) commutators of J^0 and the corresponding O_j . Thus generically, if (6.3) holds, one has

$$\begin{aligned} \frac{\partial}{\partial x^\mu} \langle \text{vac, out} | T \{ J^\mu(x) \prod_j O_j(\Psi(y_j)) \} | \text{vac, in} \rangle \\ = \sum_j \delta(x^0 - y_j^0) \langle \text{vac, out} | T \{ [J^0(x), O_j(\Psi(y_j))] \prod_{k \neq j} O_k(\Psi(y_k)) \} | \text{vac, in} \rangle , \end{aligned} \quad (6.4)$$

which is called a Ward identity. If there is an additional anomalous term on the right-hand side of this equation, i.e. if $\partial_\mu J^\mu \neq 0$ and (6.3) does not hold, then the analogue of (6.4) is called an anomalous Ward identity.

Let us show how the non-conservation of the quantum current is related in the functional integral formalism to the non-invariance of the functional integral measure. Consider the same transformation of the fields as before, such that δS is given by (6.1), and assume furthermore that the operators $O_j(\Psi(y_j))$, or rather their classical counterparts $O_j(\psi(y_j))$, are invariant. We then assume that the measure transforms as

$$\mathcal{D}\psi' = \mathcal{D}\psi e^{i\hat{\mathcal{A}}} , \quad (6.5)$$

with some possibly non-trivial Jacobian factor $e^{i\hat{\mathcal{A}}}$. Since $\hat{\mathcal{A}}$ vanishes for $\epsilon = 0$, and since ϵ is infinitesimal, $\hat{\mathcal{A}}$ must be of the form $\hat{\mathcal{A}} = \int d^4x \epsilon(x) \mathcal{A}(x)$. We then proceed by first renaming the functional integration variables from ψ to ψ' , and then identifying the ψ' with the transformed fields $\psi + \delta\psi$:

$$\begin{aligned} \int \mathcal{D}\psi \prod_j O_j(\psi(y_j)) e^{iS(\psi)} &= \int \mathcal{D}\psi' \prod_j O_j(\psi'(y_j)) e^{iS(\psi')} = \int \mathcal{D}\psi e^{i \int \epsilon \mathcal{A}} \prod_j O_j(\psi(y_j)) e^{iS(\psi) + i \int \partial_\mu j^\mu \epsilon} \\ &= \int \mathcal{D}\psi \left(1 + i \int (\mathcal{A} + \partial_\mu j^\mu) \epsilon \right) \prod_j O_j(\psi(y_j)) e^{iS(\psi)} . \end{aligned} \quad (6.6)$$

In the last expression, the 1 in the parenthesis reproduces the initial expression and, hence, the terms linear in $\epsilon(x)$ must vanish :

$$\int \mathcal{D}\psi \left(\mathcal{A}(x) + \partial_\mu j^\mu(x) \right) \prod_j O_j(\psi(y_j)) e^{iS(\psi)} = 0 . \quad (6.7)$$

But as we have extensively discussed, the functional integral computes time-ordered vacuum expectation values of the corresponding products of operators, and this translates into the operator statement

$$\langle \text{vac, out} | T \{ \partial_\mu J^\mu(x) \prod_j O_j(\Psi(y_j)) \} | \text{vac, in} \rangle = - \langle \text{vac, out} | T \{ \mathcal{A}(x) \prod_j O_j(\Psi(y_j)) \} | \text{vac, in} \rangle . \quad (6.8)$$

The non-vanishing term on the right-hand side will appear as the anomalous contribution on the right-hand side of the Ward identities (6.4). In any case, it relates the non-conservation of the quantum current to the anomalous transformation of the functional integral measure.

There is an important subtlety when translating between the functional integral relation (6.7) and the operator statement (6.8). Indeed, we could have written (6.7) as

$$\frac{\partial}{\partial x^\mu} \int \mathcal{D}\psi j^\mu(x) \prod_j O_j(\psi(y_j)) e^{iS(\psi)} = - \int \mathcal{D}\psi \mathcal{A}(x) \prod_j O_j(\psi(y_j)) e^{iS(\psi)} , \quad (6.9)$$

with the space-time derivative pulled outside the functional integral. Then the corresponding operator statement would be

$$\frac{\partial}{\partial x^\mu} \langle \text{vac, out} | T \{ J^\mu(x) \prod_j O_j(\Psi(y_j)) \} | \text{vac, in} \rangle = - \langle \text{vac, out} | T \{ \mathcal{A}(x) \prod_j O_j(\Psi(y_j)) \} | \text{vac, in} \rangle . \quad (6.10)$$

But as discussed above, the difference between (6.8) and (6.10) are the equal-time commutator terms $\sim [J^0(x), O_j(\Psi(y_j))]$ which are $\sim \delta^{(4)}(x - y_j)$ and are sometimes referred to as contact terms. But the space-integral of J^0 is the corresponding charge operator which generates the transformations of the $O_j(\Psi(y_j))$ under the symmetry, and $[J^0(x), O_j(\Psi(y_j))]$ is a local version of this. Our assumption that the $O_j(\Psi(y_j))$ are invariant under the symmetry amounts to the vanishing of these commutators and, at present, (6.8) and (6.10) are equivalent. However, if the O_j transform non-trivially, additional terms will be present and one has to be more careful.³¹

6.2 Non-invariance of the fermionic measure under chiral transformations (abelian anomaly)

The first example of anomalies we will study is the so-called abelian anomaly. This anomaly arises from chiral transformations (i.e. involving γ_5) of massless Dirac fermions. More specifically, we

³¹A related subtlety is that the functional integral representation was derived in Hamiltonian form, and only for operators not involving the momentum fields could one easily pass to the Lagrangian version of it. One can show that the Lagrangian version is still valid without modification if there is only one momentum field, but it receives additional contact term contributions if there are more momentum fields involved. Typically, J^0 involves a momentum field, so one needs to be careful.

consider the usual matter Lagrangian for a Dirac field ψ coupled to a gauge field $A_\mu = A_\mu^\alpha t_\alpha$ through the covariant derivative ³²

$$\mathcal{L}_{\text{matter}}[\psi, \bar{\psi}, A] = -\bar{\psi}(\not{D} + m)\psi = -\bar{\psi}\left(\gamma^\mu(\partial_\mu - iA_\mu^\alpha t_\alpha) + m\right)\psi. \quad (6.11)$$

Recall that our signature is $(-+++)$ and the γ^μ matrices obey $(\gamma^0)^\dagger = -\gamma^0$, $(\gamma^i)^\dagger = \gamma^i$ and $\gamma_5^\dagger = \gamma_5$. The Dirac field ψ is taken to be in some representation $\mathcal{R}$ of some Lie group with (hermitian) generators t_α . We want to study quite generally the chiral transformations $\psi(x) \rightarrow \psi'(x) = e^{i\epsilon^\alpha(x)t_\alpha\gamma_5}\psi(x)$. Since $\bar{\psi} = \psi^\dagger i\gamma^0$, and since $\{\gamma_5, \gamma^\mu\} = 0$, one also has $\bar{\psi}(x) \rightarrow \bar{\psi}'(x) = \bar{\psi}^\dagger(x)e^{-i\epsilon^\alpha(x)t_\alpha\gamma_5}i\gamma^0 = \bar{\psi}(x)e^{i\epsilon^\alpha(x)t_\alpha\gamma_5}$. These transformations are *not* symmetries of the matter Lagrangian unless (i) ϵ is constant, (ii) we replace $\epsilon^\alpha t_\alpha$ by ϵt where $[t, t_\alpha] = 0$, and (iii) take $m = 0$. Then the chiral transformation $\psi \rightarrow e^{i\epsilon t\gamma_5}\psi$ is a symmetry of the *massless* matter Lagrangian ³³ with corresponding chiral Noether current

$$j_5^\mu = i\bar{\psi}\gamma^\mu\gamma_5\psi. \quad (6.12)$$

The anomaly we will soon compute means that this chiral current is not conserved in the quantum theory : $\langle \partial_\mu J_5^\mu \rangle_A \neq 0$, where the notation $\langle \dots \rangle_A$ is shorthand for $\langle \text{vac}, \text{out} | (\dots) | \text{vac}, \text{out} \rangle$ computed in the presence of the interaction Lagrangian density $i\bar{\psi}A_\mu^\alpha t_\alpha\psi$. We will be more precise shortly.

Let us now use the functional integral approach to compute vacuum expectation values of time-ordered products of operators $\mathcal{O}_i$ that involve the matter fields. The computation can then be divided into two parts: the integration that interests us most is the integration over matter fields. In fact, this is enough to study anomalies. This is due to the fact that the only matter field - gauge field coupling we consider here is through the covariant derivative, thus the currents $j_\alpha^\mu = \partial\mathcal{L}/\partial A_\mu^\alpha$ do not involve the gauge fields A_μ^α . As such, all current correlation functions only involve the matter fields and one can do the functional integral over these matter fields, keeping the gauge fields fixed. This is indicated by writing $\langle \dots \rangle_A$, as we have already done above. The functional integration over the gauge fields can be done in a second step, and only then one has to deal with all the subtleties of gauge-fixing and adding the Faddeev-Popov ghost fields - but this will not be our focus here.

In the rest of this subsection, we will solely be concerned with the transformation of the functional integral measure for the (Dirac) matter fields ψ and $\bar{\psi}$. Quite generally, we consider a local transformation:

$$\psi(x) \rightarrow \psi'(x) = U(x)\psi(x) \quad , \quad \bar{\psi}(x) \rightarrow \bar{\psi}'(x) = \bar{\psi}(x)\bar{U}(x) \quad , \quad \bar{U}(x) = i\gamma^0 U^\dagger(x) i\gamma^0, \quad (6.13)$$

where $U(x)$ is a unitary matrix acting on the indices of the representation of the gauge group and on the Dirac indices.³⁴ Of course, in order for the matter action to also be invariant, U has to be constant.

³²Note that no gauge coupling constant g explicitly appears. Our convention is that it is included in the gauge-group generators t_α , so that in particular $\text{tr } t_\alpha t_\beta \sim g^2$, and the structure constants are $C_{\alpha\beta\gamma} \sim g$.

³³As a matter of fact, in order for this transformation to be a classical symmetry, the mass term has to vanish. Indeed, the invariance comes from the fact that $\{\gamma_5, \gamma^\mu\} = 0$ so that the transformation of ψ and $\bar{\psi}$ compensate each other in the $\bar{\psi}\not{D}\psi$ term. Since the mass term doesn't involve the γ^μ , the transformation of this term doesn't cancel out.

³⁴Other complications arise if ψ and $\bar{\psi}$ transform under additional representations like flavor symmetry, but we will not tackle these additional complexities here.

Let us examine how this transformation affects the fermionic measures $\mathcal{D}\psi$ and $\mathcal{D}\bar{\psi}$. Since these are measures for anticommuting fields, they transform with the inverse Jacobian:

$$\mathcal{D}\psi \rightarrow \mathcal{D}\psi' = (\text{Det } \mathcal{U})^{-1} \mathcal{D}\psi \quad , \quad \mathcal{D}\bar{\psi} \rightarrow \mathcal{D}\bar{\psi}' = (\text{Det } \bar{\mathcal{U}})^{-1} \mathcal{D}\bar{\psi} \quad , \quad (6.14)$$

where the operators $\mathcal{U}$ and $\bar{\mathcal{U}}$ are defined by

$$\langle x | \mathcal{U} | y \rangle = U(x) \delta^{(4)}(x - y) \quad , \quad \langle x | \bar{\mathcal{U}} | y \rangle = \bar{U}(x) \delta^{(4)}(x - y) \quad . \quad (6.15)$$

Thus, the overall Jacobian is

$$\mathcal{D}\psi' \mathcal{D}\bar{\psi}' = e^{i\hat{\mathcal{A}}} \mathcal{D}\psi \mathcal{D}\bar{\psi} \quad , \quad e^{i\hat{\mathcal{A}}} = (\text{Det } \mathcal{U})^{-1} (\text{Det } \bar{\mathcal{U}})^{-1} \quad . \quad (6.16)$$

6.2.1 Transformation of the measure

a) Non-chiral transformation

Let us first study what happens if the transformation we consider is not chiral, i.e. if U does not involve γ_5 . It is easy to see that in this case the transformations of $\mathcal{D}\psi$ and $\mathcal{D}\bar{\psi}$ cancel each other out. Indeed, if U is a unitary non-chiral transformation of the form:

$$U(x) = e^{i\epsilon^\alpha(x)t_\alpha} \quad , \quad \text{with} \quad t_\alpha^\dagger = t_\alpha \quad \text{and} \quad [\gamma^\mu, t_\alpha] = 0 \quad , \quad (6.17)$$

one has

$$\bar{U}(x) = i\gamma^0 e^{-i\epsilon^\alpha(x)t_\alpha} i\gamma^0 = e^{-i\epsilon^\alpha(x)t_\alpha} (i\gamma^0)^2 = e^{-i\epsilon^\alpha(x)t_\alpha} = U^{-1}(x) \quad (6.18)$$

Thus, $\bar{\mathcal{U}} = \mathcal{U}^{-1}$ which means that the overall measure is invariant:

$$e^{i\hat{\mathcal{A}}} = (\text{Det } \mathcal{U})^{-1} (\text{Det } \bar{\mathcal{U}})^{-1} = (\text{Det } \mathcal{U})^{-1} (\text{Det } \mathcal{U}^{-1})^{-1} = 1 \quad . \quad (6.19)$$

This remains true even after regularization; since both determinants are regularized in the same way, they still cancel each other. Hence, there is no anomaly.

b) Chiral transformations

On the other hand, if the transformation we consider is chiral, i.e. involves a γ_5 matrix, the transformations of $\mathcal{D}\psi$ and $\mathcal{D}\bar{\psi}$ do not compensate each other, as we now proceed to show. Consider a unitary chiral transformation of the form:

$$U(x) = e^{i\epsilon^\alpha(x)t_\alpha\gamma_5} \quad , \quad \text{with} \quad t_\alpha^\dagger = t_\alpha \quad \text{and} \quad [\gamma^\mu, t_\alpha] = 0 \quad . \quad (6.20)$$

Then, due to the anti-commutativity of γ_5 with γ^0 one has:

$$\bar{U}(x) = i\gamma^0 e^{-i\epsilon^\alpha(x)t_\alpha\gamma_5} i\gamma^0 = e^{+i\epsilon^\alpha(x)t_\alpha\gamma_5} (i\gamma^0)^2 = e^{+i\epsilon^\alpha(x)t_\alpha\gamma_5} = U(x) \quad , \quad (6.21)$$

Then also $\text{Det } \bar{\mathcal{U}} = \text{Det } \mathcal{U}$, and the overall fermionic measure transforms as follows:

$$e^{i\hat{\mathcal{A}}} = (\text{Det } \mathcal{U})^{-1} (\text{Det } \bar{\mathcal{U}})^{-1} = (\text{Det } \mathcal{U})^{-2} = e^{-2 \text{Tr } \log \mathcal{U}} \quad , \quad (6.22)$$

where we used the operator analogue of the relation $\det M = e^{\text{tr } \log M}$ which is easily proven for any diagonalizable matrix M . Since the determinant of $\mathcal{U}$ no longer cancels out, one now has to compute it. This is where complexities leading to anomalies arise.

6.2.2 Defining and computing the determinant

Defining the determinant of an operator in an infinite dimensional space comes with a set of challenges. Let us begin by identifying what the matrix elements of the powers of $\mathcal{U}$ are, which will allow us to generalize this relation to all Taylor expandable functions. For a local kernel, we have $\langle x | \mathcal{U}^2 | y \rangle = \int d^4 z \langle x | \mathcal{U} | z \rangle \langle z | \mathcal{U} | y \rangle = \int d^4 z U(x) \delta^{(4)}(x - z) U(z) \delta^{(4)}(z - y) = U^2(x) \delta^{(4)}(x - y) = U^2(x) \langle x | y \rangle$, and similarly for all powers of $\mathcal{U}$. As such, we obtain the generic relation $\langle x | f(\mathcal{U}) | y \rangle = f(U(x)) \langle x | y \rangle$, which we can now apply to $\text{Tr} \log \mathcal{U}$:

$$\text{Tr} \log \mathcal{U} = \int d^4 x \langle x | \text{tr} \log(\mathcal{U}) | x \rangle = \int d^4 x \delta^{(4)}(x - x) \text{tr} \log(U(x)) = \int d^4 x \delta^{(4)}(0) i\epsilon^\alpha(x) \text{tr} t_\alpha \gamma_5, \quad (6.23)$$

where Tr is a functional and matrix trace, while tr is only a matrix trace (with respect to the γ matrices and the gauge representation matrices). This allows us, in turn to obtain the following formula for the determinant:

$$(\text{Det} \mathcal{U})^{-2} = e^{-2 \text{Tr} \log \mathcal{U}} = e^{i \int d^4 x \epsilon^\alpha(x) \mathcal{A}_\alpha(x)} \quad \text{with} \quad \mathcal{A}_\alpha(x) = -2 \delta^{(4)}(0) \text{tr} t_\alpha \gamma_5, \quad (6.24)$$

where $\mathcal{A}_\alpha$ is called the anomaly function or simply the anomaly.

However, this anomaly function needs to be regularized as it is ill-defined at this stage given that $\delta^{(4)}(0)$ is infinite and that $\text{tr} \gamma_5 t_\alpha$ vanishes.³⁵ The $\delta^{(4)}(0) = \delta^{(4)}(x - x)$ factor is to be interpreted through its Fourier transform $\int \frac{d^4 p}{(2\pi)^4} e^{ip(x-y)} \Big|_{x=y}$ as a large momentum (UV) divergence. We will then proceed by using the Fourier transform and regularize the anomaly through a cut-off function which limits the contribution of larger momenta. To this end, we introduce a smooth cut-off function f such that $f(0) = 1$, $f(\infty) = 0$, as well as $xf'(x) = 0$ at $x = 0$ and at $x = \infty$, like e.g. $f(x) = e^{-x}$. The precise reason for these conditions will become clear in the computation of eq. (6.29). We then rewrite the integral of the anomaly as the limit of a quantity dependent on the regulator Λ :

$$\int d^4 x \epsilon^\alpha(x) \mathcal{A}_\alpha(x) = -2 \lim_{\Lambda \rightarrow \infty} \text{Tr} \mathcal{E}_\Lambda, \quad \text{where} \quad \mathcal{E}_\Lambda = \epsilon^\alpha(\hat{X}) \gamma_5 t_\alpha f\left(-\frac{\hat{\mathcal{D}}^2}{\Lambda^2}\right), \quad (6.25)$$

where $\hat{X}$ is the quantum mechanical position operator and $\hat{\mathcal{D}}$ denotes the quantum mechanical operator that acts as: $\langle x | \hat{\mathcal{D}} | \chi \rangle = \mathcal{D} \langle x | \chi \rangle$. We choose to regulate the integral through $\hat{\mathcal{D}}^2$ which is both gauge and Lorentz invariant. Below, we will further discuss the importance of this choice. Let us now compute this quantity:

³⁵This is due to the fact that actually the trace over the matrices of the representation and over the Dirac matrices are taken separately, i.e. $\text{tr} \gamma_5 t_\alpha \equiv (\text{tr}_{\mathcal{R}} t_\alpha^{\mathcal{R}}) (\text{tr}_D \gamma_5)$ and $\text{tr}_D \gamma_5 = 0$.

$$\begin{aligned}
\text{Tr } \mathcal{E}_\Lambda &\equiv \int d^4x \text{ tr } \langle x | \epsilon^\alpha(\hat{X}) \gamma_5 t_\alpha f\left(-\frac{\not{D}^2}{\Lambda^2}\right) | x \rangle = \int d^4x \epsilon^\alpha(x) \int d^4p \text{ tr } \gamma_5 t_\alpha \langle x | f\left(-\frac{\not{D}^2}{\Lambda^2}\right) | p \rangle \langle p | x \rangle \\
&= \int d^4x \epsilon^\alpha(x) \int \frac{d^4p}{(2\pi)^4} e^{-ipx} \text{ tr } \gamma_5 t_\alpha f\left(-\frac{1}{\Lambda^2} \left[\gamma^\mu \left(\frac{\partial}{\partial x^\mu} - iA_\mu^\mathcal{R}(x) \right) \right]^2 \right) e^{ipx} \\
&= \int d^4x \epsilon^\alpha(x) \int \frac{d^4p}{(2\pi)^4} e^{-ipx} \text{ tr } \gamma_5 t_\alpha \epsilon^{ipx} f\left(-\frac{1}{\Lambda^2} \left[\gamma^\mu \left(\frac{\partial}{\partial x^\mu} + ip_\mu - iA_\mu^\mathcal{R}(x) \right) \right]^2 \right) \\
&= \int d^4x \epsilon^\alpha(x) \int \frac{d^4p}{(2\pi)^4} \text{ tr } \gamma_5 t_\alpha f\left(-\frac{1}{\Lambda^2} [i\not{p} + \not{D}]^2\right) \\
&= \int d^4x \epsilon^\alpha(x) \Lambda^4 \int \frac{d^4q}{(2\pi)^4} \text{ tr } \gamma_5 t_\alpha f\left(-\left[i\not{q} + \frac{\not{D}}{\Lambda}\right]^2\right) \\
&= \int d^4x \epsilon^\alpha(x) \Lambda^4 \int \frac{d^4q}{(2\pi)^4} \text{ tr } \gamma_5 t_\alpha f\left(q^2 - 2i\frac{q^\mu D_\mu}{\Lambda} - \frac{\not{D}^2}{\Lambda^2}\right). \tag{6.26}
\end{aligned}$$

When going from the second to the third line, we have moved the e^{ipx} past the differential operator $f(-\not{D}^2/\Lambda^2)$ which should be thought of as a power series in $\not{D}^2/\Lambda^2$. In doing so, each $\partial_\mu - iA_\mu(x)$ gets replaced by a $\partial_\mu + ip_\mu - iA_\mu(x)$, with the right-most differential operator acting on 1 (which we did not write explicitly). Also, in the last steps, we set $p^\mu = \Lambda q^\mu$ and use that $-\left[i\not{q} + \frac{\not{D}}{\Lambda}\right]^2 = \not{q}^2 - \frac{i}{\Lambda}(\not{q}\not{D} + \not{D}\not{q}) - \frac{1}{\Lambda^2}\not{D}^2 = q^2 - \frac{2i}{\Lambda}q^\mu D_\mu - \frac{1}{\Lambda^2}\not{D}^2$. We then Taylor expand $\Lambda^4 f$ around q^2 . This yields the following terms

$$\Lambda^4 \left(f(q^2) + f'(q^2) \left(-\frac{2i}{\Lambda} q^\mu D_\mu - \frac{1}{\Lambda^2} \not{D}^2 \right) + \frac{1}{2} f''(q^2) \left(-\frac{2i}{\Lambda} q^\mu D_\mu - \frac{1}{\Lambda^2} \not{D}^2 \right)^2 + \dots \right), \tag{6.27}$$

acting on 1. Note that the only term involving Dirac matrices are the powers of $\not{D}^2$. Two elements then come into play to select the non-vanishing term. First, in order for the Dirac trace involving γ_5 to be non vanishing, the term has to involve at least four γ^μ , i.e. two factors of $\not{D}^2$, which means expanding at least to second order. On the other hand, for the term to be non vanishing as we take the limit $\Lambda \rightarrow +\infty$, it has to contain no power of the inverse of Λ greater than four. Thus, the only term that contributes to $\lim_{\Lambda \rightarrow \infty} \text{Tr } \mathcal{E}_\Lambda$ is the term of second order: $\frac{1}{2} f''(q^2) (-\not{D}^2/\Lambda^2)^2$. i.e.

$$\lim_{\Lambda \rightarrow \infty} \text{Tr } \mathcal{E}_\Lambda = \int d^4x \epsilon^\alpha(x) \int \frac{d^4q}{(2\pi)^4} \frac{1}{2} f''(q^2) \text{ tr } \gamma_5 t_\alpha (-\not{D}^2)^2. \tag{6.28}$$

We can easily evaluate this integral after performing a Wick rotation:

$$\begin{aligned}
\int \frac{d^4q}{(2\pi)^4} \frac{1}{2} f''(q^2) &= \frac{i}{2(2\pi)^4} \text{vol}(S^3) \int_0^\infty dq q^3 f''(q^2) = \frac{i}{2(2\pi)^4} 2\pi^2 \frac{1}{2} \int_0^\infty d\xi \xi f''(\xi) \\
&= \frac{i}{32\pi^2} \left(\xi f'(\xi) \Big|_0^\infty - \int_0^\infty d\xi f'(\xi) \right) = \frac{i}{32\pi^2}, \tag{6.29}
\end{aligned}$$

where we used the special properties of the function f . Furthermore,

$$\begin{aligned}
\not{D}^2 &= \gamma^\mu \gamma^\nu D_\mu D_\nu = \frac{1}{2} \{ \gamma^\mu, \gamma^\nu \} D_\mu D_\nu + \frac{1}{2} [\gamma^\mu, \gamma^\nu] D_\mu D_\nu = \eta^{\mu\nu} D_\mu D_\nu + \frac{1}{4} [\gamma^\mu, \gamma^\nu] [D_\mu, D_\nu] \\
&= D^\mu D_\mu - \frac{i}{4} [\gamma^\mu, \gamma^\nu] F_{\mu\nu}, \tag{6.30}
\end{aligned}$$

where we used that the gauge field strength is defined as $F_{\mu\nu} = i[D_\mu, D_\nu]$. Using furthermore

$$\text{tr}_D \gamma_5 \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma = 4i\epsilon^{\mu\nu\rho\sigma} \quad , \quad \text{with } \epsilon^{0123} = +1 \quad , \quad (6.31)$$

we get

$$\text{tr} \gamma_5 t_\alpha (-\not{D})^2 = \left(\frac{i}{4}\right)^2 \text{tr}_D \gamma_5 [\gamma^\mu, \gamma^\nu] [\gamma^\rho, \gamma^\sigma] \text{tr}_{\mathcal{R}} t_\alpha F_{\mu\nu} F_{\rho\sigma} = -i\epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t_\alpha F_{\mu\nu} F_{\rho\sigma} \quad . \quad (6.32)$$

Putting everything together we finally obtain $\lim_{\Lambda \rightarrow \infty} \text{Tr} \mathcal{E}_\Lambda = \frac{1}{32\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t_\alpha F_{\mu\nu} F_{\rho\sigma}$ and hence for the anomaly function

$$\mathcal{A}_\alpha(x) = -\frac{1}{16\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t_\alpha F_{\mu\nu}(x) F_{\rho\sigma}(x) \quad . \quad (6.33)$$

A remark is in order : The trace over the Hilbert space should be taken with respect to some basis of square-normalizable, gauge-covariant functions which one could chose to be eigenfunctions of $i\not{D}$. Of course, we don't know this basis explicitly and, instead, we have taken the trace with respect to the improper, non-normalisable basis of position eigenstates $|x\rangle$. This is not a problem though, as long as we are only interested in the leading terms for large Λ . However, this might not be correct for the sub-leading terms.

Let us now specify to the case where t_α is replaced by an abelian generator t so that we consider abelian chiral transformations, in which case the chiral current was classically conserved. Then the anomaly in the quantum theory is immediately obtained from (6.33) by this replacement :

$$\mathcal{A}(x) = -\frac{1}{16\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t F_{\mu\nu}(x) F_{\rho\sigma}(x) \quad . \quad (6.34)$$

In view of the Ward identity (6.10), we can conclude that the chiral current non-conservation is given exactly by (minus) the anomaly function :

$$\partial_\mu \langle J_5^\mu(x) \rangle_A = -\mathcal{A}(x) \quad . \quad (6.35)$$

At first sight, it may seem a bit strange that the divergence of the chiral current J_5^μ - which does not depend on the gauge fields - involves the gauge field strength. Of course, this gauge-field dependence is a quantum effect. Below, we will relate it to a one-loop diagram, and the gauge-field dependence appears naturally through the matter-gauge field vertices. In our previous computation the gauge-field dependence is due to our regulator $f(-\not{D}^2/\Lambda^2)$ which is made from the gauge-covariant derivative. Had we used a regulator $f(-\partial_\mu \partial^\mu/\Lambda^2)$ there would be no chiral anomaly, but this regulator would break the gauge invariance of the theory. The lesson is that an anomaly arises if the regulator cannot maintain *all* the classical symmetries, and then one needs to make a choice which symmetries are more important. Since the gauge symmetry is fundamental for the consistency of the theory the regulator function must be taken gauge-invariant (if possible).

The computation of the abelian anomaly and the result (6.34) bear a strong resemblance with the index of the Dirac operator. It is instructive to spell out this relation in more detail, as we will do in the next subsection. Later-on we will come upon another occurrence of index theorems which will relate the chiral anomaly in d dimensions to the index of a Dirac operator in $d + 2$ dimensions.

6.2.3 Relation of the abelian anomaly with the index theorem

Let us then briefly explore the relation between the abelian anomaly and the index theorem. Let us begin by recalling the definition of the abelian anomaly as the limit of the regularized quantity (6.25) but now with $\epsilon^\alpha \rightarrow \epsilon$ and $t_\alpha \rightarrow t$:

$$\int d^4x \epsilon(x) \mathcal{A}(x) = -2 \lim_{\Lambda \rightarrow \infty} \text{Tr} \mathcal{E}_\Lambda \quad , \quad \text{where} \quad \mathcal{E}_\Lambda = \epsilon(\hat{X}) \gamma_5 t f\left(-\frac{\hat{\mathcal{D}}^2}{\Lambda^2}\right) . \quad (6.36)$$

Note that if we set $\epsilon = t = 1$, this relation becomes

$$\int d^4x \mathcal{A}(x) = -2 \lim_{\Lambda \rightarrow \infty} \text{Tr} \gamma_5 f(-\hat{\mathcal{D}}^2/\Lambda^2) . \quad (6.37)$$

As a matter of fact, we will show that the trace $\text{Tr} \gamma_5 f(-\hat{\mathcal{D}}^2/\Lambda^2)$ counts the number of positive chirality zero-eigenmodes of $i\mathcal{D}$ minus the number of its negative chirality zero-eigenmodes. This is called the index of $i\mathcal{D}$.

This argument is best done in Euclidean signature where $i\mathcal{D}$ is hermitian. Indeed, in Euclidean signature, all γ^μ can be chosen to be hermitian and then, with appropriate boundary conditions, $\int d^4x_E (i\mathcal{D}\chi)^\dagger \psi = \int d^4x_E \chi^\dagger i\mathcal{D}\psi$. Let us now compute the trace in the eigenbasis of $i\mathcal{D}$. Let then ψ_n be an eigenmode of $i\mathcal{D}$ with eigenvalue λ_n :

$$i\mathcal{D}\psi_n = \lambda_n \psi_n . \quad (6.38)$$

Since $\mathcal{D}\gamma_5 = -\gamma_5\mathcal{D}$, we have

$$i\mathcal{D}(\gamma_5\psi_n) = -\gamma_5 i\mathcal{D}\psi_n = -\gamma_5 \lambda_n \psi_n = (-\lambda_n)(\gamma_5\psi_n) , \quad (6.39)$$

so that $\gamma_5\psi_n$ is an eigenmode of $i\mathcal{D}$ with eigenvalue $-\lambda_n$.

Then, if $\lambda_n \neq 0$, ψ_n and $\gamma_5\psi_n$ are both eigenmodes of the hermitian $i\mathcal{D}$ with different eigenvalues and, as such they are orthogonal, and in particular linearly independent. It follows that the two combinations $\psi_n^\pm = \psi_n \pm \gamma_5\psi_n$ are both non-vanishing. Clearly, $\psi_n^\pm$ are eigenstates of γ_5 and have chirality ± 1 . However, they are no longer eigenmodes of $i\mathcal{D}$. Nevertheless, they are still eigenmodes of $(i\mathcal{D})^2 = -\mathcal{D}^2$ with eigenvalues λ_n^2 . One concludes that ψ_n^+ contributes $+f(\lambda_n^2/\Lambda^2)$ and ψ_n^- contributes $-f(\lambda_n^2/\Lambda^2)$ to $\text{Tr} \gamma_5 f(-\hat{\mathcal{D}}^2/\Lambda^2)$, and the sum of their contributions vanishes. Since this holds for all non-zero modes ($\lambda_n \neq 0$), only zero-modes possibly give a non-vanishing contribution to this trace.

Consider now the zero-modes, with $\lambda_n = 0$. Then ψ_n and $\gamma_5\psi_n$ have the same eigenvalue of $i\mathcal{D}$ (namely 0) and one no longer can conclude that they are linearly independent. If they are linearly dependent, i.e. if $\gamma_5\psi_n \sim \psi_n$, one necessarily has $\gamma_5\psi_n = \pm\psi_n$ and then ψ_n has definite chirality. If they are linearly independent, one can still form the $\psi_n^\pm = \psi_n \pm \gamma_5\psi_n$ which have definite and opposite chirality. This shows that within the zero-eigenspace of $i\mathcal{D}$, we can choose a basis of zero-modes with definite chirality. However, there is no reason any more to have (only) pairs of opposite chirality, and the number of positive chirality zero-modes may differ from the number of negative chirality zero-modes. As already said, this difference is called the index of $i\mathcal{D}$. Since $\lambda_n = 0$, all these zero-modes just contribute a $\pm f(0/\Lambda^2) = \pm 1$ to the trace $\text{Tr} \gamma_5 f(-\hat{\mathcal{D}}^2/\Lambda^2)$. Thus (writing $\mathcal{D}_E$ to insist that this is in Euclidean signature)

$$\text{index}(i\mathcal{D}_E) = \text{Tr} \gamma_5 f(-\mathcal{D}_E^2/\Lambda^2) , \quad (6.40)$$

independently of the value of Λ . Above we have computed this trace in the large Λ limit. As discussed below (6.33), our computation is reliable only in this limit, and we conclude that the integral of the abelian anomaly $\int d^4x \mathcal{A}(x)|_{t=1}$ is just (-2) times the quantity (6.40) (in the large Λ limit). In view of (6.34) we find that

$$\text{index}(i\mathcal{D}_E) = \frac{1}{32\pi^2} \int d^4x_E \epsilon_E^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} F_{\mu\nu}^E(x) F_{\rho\sigma}^E(x) . \quad (6.41)$$

This is a special case of the famous Atiyah-Singer index theorem [17, 26, 27].

6.3 Matter effective action and more Ward identities

In the next section we will do a Feynman diagram computation of one-loop triangle diagrams. In order to relate these computations to the non-conservation of the currents, we need to derive some further identities which usually go under the name of (generalized) Ward identities.

As mentioned above, it is useful to consider the partial functional integral done over the matter fields only :

$$e^{i\tilde{\Gamma}[A]} = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{i \int \mathcal{L}_{\text{matter}}[\psi, \bar{\psi}, A]} . \quad (6.42)$$

Integrating over the matter fields only, and not over the gauge fields, avoids all the complications related to the gauge fixing, Faddeev-Popov determinant and introduction of ghost fields. More importantly, for the discussion of anomalies under gauge transformations, the $\tilde{\Gamma}[A]$ is the appropriate object to consider. This $\tilde{\Gamma}[A]$ is some sort of quantum effective action. It is different from - but related to - the quantum effective action Γ discussed in subsection 1.5. Indeed, the matter Lagrangian $\mathcal{L}_{\text{matter}}[\psi, \bar{\psi}, A] = -\bar{\psi}(\mathcal{D} + m)\psi$ is quadratic in ψ and $\bar{\psi}$ and, at least formally, the integral is given by the determinant

$$\begin{aligned} \text{Det}(\mathcal{D} + m) &= \text{Det}(\mathcal{D} + m - i\mathcal{A}) = \text{Det}((\mathcal{D} + m)(\mathbf{1} - i(\mathcal{D} + m)^{-1}\mathcal{A})) \\ &= \text{Det}(\mathcal{D} + m) \text{Det}(\mathbf{1} - i(\mathcal{D} + m)^{-1}\mathcal{A}) = \mathcal{N} \times \text{Det}(\mathbf{1} - i(\mathcal{D} + m)^{-1}\mathcal{A}) , \end{aligned} \quad (6.43)$$

so that

$$i\tilde{\Gamma}[A] = \log \text{Det}(\mathcal{D} + m) = i\tilde{\Gamma}[0] + \text{Tr} \log(\mathbf{1} - i(\mathcal{D} + m)^{-1}\mathcal{A}) = i\tilde{\Gamma}[0] + \sum_{r \geq 1} \frac{1}{r} \text{Tr} (i(\mathcal{D} + m)^{-1}\mathcal{A})^r . \quad (6.44)$$

This sum is to be interpreted as a sum of one-loop diagrams with r Dirac propagators $-i(\mathcal{D} + m)^{-1}$ and r vertices $-\gamma^\mu t_\alpha$ where the A_μ^α are attached. These are all 1PI diagrams that contribute to the full quantum effective action Γ . The full Γ also contains higher-loop diagrams with internal gauge field propagators that arise when also doing the functional integral over the gauge fields. For our present purpose, it is enough to work with $\tilde{\Gamma}[A]$ and we simply refer to it as an effective action.

Taking functional derivatives of $e^{i\tilde{\Gamma}[A]}$ with respect to A_μ^α will insert into the functional integral a factor $i \frac{\partial \mathcal{L}}{\partial A_\mu^\alpha} = i j_\alpha^\mu$ where $j_\alpha^\mu = i\bar{\psi} \gamma^\mu t_\alpha \psi$ is the “gauge current”. Taking multiple functional derivatives produces multiple insertions of the corresponding currents and then the functional integral computes the vacuum expectation value of the corresponding product of quantum currents (up to a normalization constant that we drop) :

$$\frac{\delta}{\delta A_{\mu_1}^{\alpha_1}(x_1)} \cdots \frac{\delta}{\delta A_{\mu_r}^{\alpha_r}(x_r)} e^{i\tilde{\Gamma}[A]} = i^r \langle \text{vac, out} | T \{ J_{\alpha_1}^{\mu_1}(x_1) \cdots J_{\alpha_r}^{\mu_r}(x_r) \} | \text{vac, in} \rangle_A \quad (6.45)$$

This same result can also be obtained directly from Dyson's formula derived in subsection 1.3.2, without the need to rely on the functional integral. As discussed in subsection 1.4, we get the *connected* vacuum expectation value if we take functional derivatives of $i\tilde{\Gamma}[A]$ instead. We may furthermore set $A = 0$ in the end :

$$\frac{\delta}{\delta A_{\mu_1}^{\alpha_1}(x_1)} \cdots \frac{\delta}{\delta A_{\mu_r}^{\alpha_r}(x_r)} i\tilde{\Gamma}[A] \Big|_{A=0} = i^r \langle \text{vac, out} | T \{ J_{\alpha_1}^{\mu_1}(x_1) \dots J_{\alpha_r}^{\mu_r}(x_r) \} | \text{vac, in} \rangle_{\text{connected}} . \quad (6.46)$$

This corresponds exactly to a one-loop diagram with r currents (times i) inserted at the vertices and Dirac propagators in between. We will compute such diagrams in the next section.

6.3.1 Anomalous Ward identity for the abelian anomaly

We can apply similar manipulations to

$$\frac{\partial}{\partial x^\mu} \langle J_5^\mu(x) \rangle_A = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \partial_\mu j_5^\mu(x) e^{i \int \mathcal{L}_{\text{matter}}[\psi, \bar{\psi}, A]} , \quad (6.47)$$

which yields ³⁶

$$\frac{\delta}{\delta A_{\mu_1}^{\alpha_1}(x_1)} \cdots \frac{\delta}{\delta A_{\mu_r}^{\alpha_r}(x_r)} \frac{\partial}{\partial x^\mu} \langle J_5^\mu(x) \rangle_A = i^r \frac{\partial}{\partial x^\mu} \langle \text{vac, out} | T \{ J_5^\mu(x) J_{\alpha_1}^{\mu_1}(x_1) \dots J_{\alpha_r}^{\mu_r}(x_r) \} | \text{vac, in} \rangle_A^{\text{connected}} . \quad (6.48)$$

Now use (6.35) to replace $\partial_\mu \langle J_5^\mu(x) \rangle_A$ by $-\mathcal{A}(x)$ and, after taking the functional derivatives, set $A = 0$ to get

$$\frac{\delta}{\delta A_{\mu_1}^{\alpha_1}(x_1)} \cdots \frac{\delta}{\delta A_{\mu_r}^{\alpha_r}(x_r)} \mathcal{A}(x) \Big|_{A=0} = -i^r \frac{\partial}{\partial x^\mu} \langle \text{vac, out} | T \{ J_5^\mu(x) J_{\alpha_1}^{\mu_1}(x_1) \dots J_{\alpha_r}^{\mu_r}(x_r) \} | \text{vac, in} \rangle_{\text{connected}} . \quad (6.49)$$

The abelian anomaly (6.34) contains terms that are quadratic and higher in the gauge fields. Thus the first non-trivial relation (6.49) is for $r = 2$ and reads

$$\frac{\delta}{\delta A_\nu^\beta(y)} \frac{\delta}{\delta A_\rho^\gamma(z)} \mathcal{A}(x) \Big|_{A=0} = \frac{\partial}{\partial x^\mu} \langle \text{vac, out} | T \{ J_5^\mu(x) J_\beta^\nu(y) J_\gamma^\rho(z) \} | \text{vac, in} \rangle_{\text{connected}} \equiv -\frac{\partial}{\partial x^\mu} \Gamma_{5\beta\gamma}^{\mu\nu\rho} , \quad (6.50)$$

where $i\Gamma_{5\beta\gamma}^{\mu\nu\rho}$ is exactly the 1PI one-loop triangle diagram with one (abelian) chiral current and two (non-abelian) non-chiral currents at the vertices.³⁷ This makes the precise relation with the Feynman diagram computation given in the next section. Using the explicit form (6.34) of $\mathcal{A}$ and Fourier transforming from (x, y, z) to (k, p, q) and factoring the usual $(2\pi)^4 \delta^{(4)}(p+k+q)$, one can show that the previous relation is equivalent to

$$-i(p_\mu + q_\mu) \Gamma_{5\beta\gamma}^{\mu\nu\rho}(-p - q, p, q) = -\frac{1}{2\pi^2} \epsilon^{\nu\rho\lambda\sigma} p_\lambda q_\sigma \text{tr}_{\mathcal{R}} t t_{(\beta} t_{\gamma)} . \quad (6.51)$$

³⁶One might worry whether this is also justified if x coincides with one of the x_j . This corresponds exactly to the possible occurrence of the contact terms that originate from $[J_5^0, J_{\alpha_r}^{\mu_r}]$. But these commutators vanish again, and one does not need to worry.

³⁷More precisely, there is an insertion of $i \times J$ at each vertex, in agreement with the factors i^r in (6.48) or (6.49).

6.3.2 Anomalous Ward identity for the chiral anomaly

Somewhat misleadingly, the name chiral anomaly refers to an anomaly under gauge transformations. Let us first consider whether $\tilde{\Gamma}[A]$ is gauge invariant. Let $A'_\mu = A_\mu + D_\mu \epsilon$ be the gauge transformed gauge field. Then

$$e^{i\tilde{\Gamma}[A']} = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{i \int \mathcal{L}_{\text{matter}}[\psi, \bar{\psi}, A']} = \int \mathcal{D}\psi' \mathcal{D}\bar{\psi}' e^{i \int \mathcal{L}_{\text{matter}}[\psi', \bar{\psi}', A']} = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{i\hat{\mathcal{A}}} e^{i \int \mathcal{L}_{\text{matter}}[\psi, \bar{\psi}, A]} , \quad (6.52)$$

where we first renamed $(\psi, \bar{\psi})$ into $(\psi', \bar{\psi}')$, and then identified them with the gauge transformed fields (under the same transformation as A'_μ). Lastly, we used the gauge invariance of the matter Lagrangian and the possibly anomalous transformation of the integration measure. It turns out that $\hat{\mathcal{A}} = \int d^4x \epsilon^\alpha(x) \mathcal{A}_\alpha(x)$ only depends on the gauge fields A_μ , and not on ψ or $\bar{\psi}$, and can be taken out of the functional integral. Thus $\tilde{\Gamma}[A'] = \hat{\mathcal{A}} + \tilde{\Gamma}[A]$, or

$$\begin{aligned} \int d^4x \epsilon^\alpha(x) \mathcal{A}_\alpha(x) &= \delta\tilde{\Gamma}[A] = \tilde{\Gamma}[A'] - \tilde{\Gamma}[A] = \tilde{\Gamma}[A + \delta A] - \tilde{\Gamma}[A] = \int d^4x \delta A_\mu^\alpha(x) \frac{\delta\tilde{\Gamma}[A]}{\delta A_\mu^\alpha(x)} \\ &= \int d^4x (D_\mu \epsilon(x))^\alpha \langle J_\alpha^\mu(x) \rangle_A . \end{aligned} \quad (6.53)$$

Thus the anomaly here exactly equals the non-invariance of the effective action $\tilde{\Gamma}$. Upon integrating the gauge covariant derivative by parts, we relate again the anomaly to the, now covariant, current non-conservation :

$$(D_\mu \langle J^\mu(x) \rangle_A)_\alpha = -\mathcal{A}_\alpha(x) , \quad (6.54)$$

much analogous to the relation (6.35) for the abelian anomaly. As before, one can take successive functional derivatives of this relation with respect to the gauge fields. Since now the covariant derivative also involves the gauge field, there are some extra contributions involving the structure constants. In particular, taking two functional derivatives with respect to the gauge fields gives [14]

$$-\frac{\delta}{\delta A_\nu^\beta(y)} \frac{\delta}{\delta A_\rho^\gamma(z)} \mathcal{A}_\alpha(x) \Big|_{A=0} = \frac{\partial}{\partial x^\mu} \Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(x, y, z) + C_{\alpha\gamma\epsilon} \delta^{(4)}(x - z) \Gamma_{\epsilon\beta}^{\rho\nu}(x, y) + C_{\alpha\beta\epsilon} \delta^{(4)}(x - y) \Gamma_{\epsilon\gamma}^{\nu\rho}(x, z) . \quad (6.55)$$

Note that $-\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(x, y, z) = \langle T\{J_\alpha^\mu(x) J_\beta^\nu(y) J_\gamma^\rho(z)\} \rangle_{\text{connected}}$ and that, upon taking $\frac{\partial}{\partial x^\mu}$ one gets $\langle T\{\partial_\mu J_\alpha^\mu(x) J_\beta^\nu(y) J_\gamma^\rho(z)\} \rangle_{\text{connected}}$, plus two equal-time commutator terms of the form $\delta(x^0 - y^0) \times \langle T\{[J_\alpha^0(x), J_\beta^\nu(y)] J_\gamma^\rho(z)\} \rangle_{\text{connected}}$. The latter are non-vanishing at present, and will exactly cancel the two terms in (6.55) involving the structure constants, thus showing that the anomaly “is contained” in $\partial_\mu J_\alpha^\mu$.

In any case, (6.55) shows that the computation of the triangle diagram $\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}$ with a gauge current at each vertex will allow us to obtain the part of the anomaly that is bilinear in the gauge fields. (The $\Gamma_{\epsilon\beta}^{\rho\nu}$ are vacuum-polarization type 2-point functions.) Again, we will present this computation in the next section.

7 Computing the triangle diagrams

This section focuses on the Feynman diagram computation of the two anomalies we have discussed : the abelian anomaly under chiral transformations and the chiral anomaly which is an anomaly under gauge transformations in the presence of chiral fermions. As we have shown in the previous section via the anomalous Ward identities, these anomalies manifest themselves in one-loop triangle Feynman diagrams.

Firstly, we will compute the triangle diagram with one chiral current and two non-chiral gauge currents at the vertices. This will allow us to confirm exactly the relation (6.51), which in turn confirms the form and precise coefficient of the abelian anomaly (6.33) (or at least of the part that is bilinear in the gauge fields). Secondly, we will tackle the triangle diagram for chiral fermions coupled to non-abelian gauge fields. This corresponds to computing the terms on the right-hand side of the anomalous Ward identity (6.55), and hence will establish the form and coefficient of the part of the chiral (gauge) anomaly that is bilinear in the gauge fields. We will explain later-on how the higher order terms are determined by the Wess-Zumino consistency condition.

We will see throughout this section how anomalies originate in the impossibility to regulate the one-loop triangle diagrams in a way that manifestly preserves all the symmetries we would like to preserve. In particular, due to the prominent presence of the γ_5 matrix, dimensional regularization is not available. We will instead rely on Pauli-Villars regularization and will pay special attention to where the anomalies arise. The computations follow closely those in [14], see also [7].

7.1 Computations for the abelian anomaly

Let us then compute the so-called AVV triangle diagram, with one axial current and two vector currents and rederive the non-conservation of the axial current. We will carry out our computations in a slightly more general, non-abelian setting, thus replacing $j_5^\mu = i\bar{\psi}\gamma^\mu\gamma_5 t\psi$ with

$$j_{5\alpha}^\mu = i\bar{\psi}\gamma^\mu\gamma_5 t_\alpha\psi, \quad (7.1)$$

and compute the vertex function $\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}$. In the end, we substitute back $t_\alpha \rightarrow t$ to get the abelian anomaly.

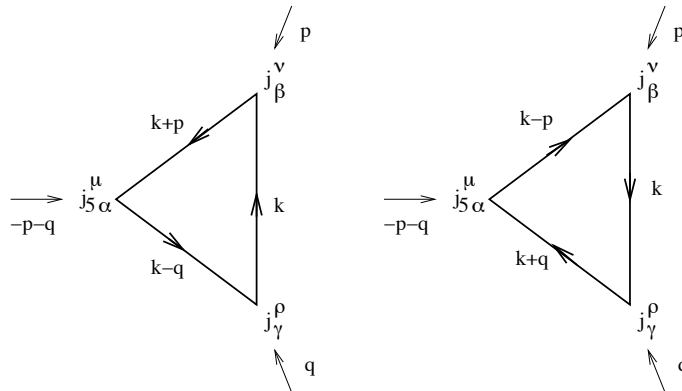


Figure 5: The two triangle diagrams contributing to the abelian anomaly

As shown in Fig. 5, there are two diagrams contributing to $\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}$, corresponding to the two ways to contract the fermion fields contained in the currents. The vertices contribute $i(2\pi)^4 \times$

$i\gamma^\mu\gamma_5 t_\alpha = -(2\pi)^4\gamma^\mu\gamma_5 t_\alpha$ and $-(2\pi)^4\gamma^\nu t_\beta$, resp. $-(2\pi)^4\gamma^\rho t_\gamma$, while the fermion loop contributes an extra minus sign. Recall also that a propagator for a fermion of momentum k and mass m is $\frac{-i}{(2\pi)^4} \frac{1}{i\cancel{k}+m-i\epsilon} = \frac{-i}{(2\pi)^4} \frac{-i\cancel{k}+m}{k^2+m^2-i\epsilon}$. Of course, here the fermions are massless. The different factors of $(2\pi)^4$ cancel between the vertices and propagators, but there is one factor $(2\pi)^4$ left since the Feynman diagrams compute ³⁸ $i(2\pi)^4\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}$. Then

$$\begin{aligned} & i\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q) \\ &= - \int \frac{d^4k}{(2\pi)^4} \text{tr} \left[(-\gamma^\mu\gamma_5 t_\alpha)(-i) \frac{-i(\cancel{k}+\not{p})}{(k+p)^2-i\epsilon} (-\gamma^\nu t_\beta)(-i) \frac{-i\cancel{k}}{k^2-i\epsilon} (-\gamma^\rho t_\gamma)(-i) \frac{-i(\cancel{k}-\not{q})}{(k-q)^2-i\epsilon} \right] \\ & \quad + \{(p, \nu, \beta) \leftrightarrow (q, \rho, \gamma)\} . \end{aligned} \quad (7.2)$$

Let us focus on the first term, which we denote $i\Gamma_{5\alpha\beta\gamma}^{(I)\mu\nu\rho}(-p-q, p, q)$ and which corresponds to the left diagram. The second term then is simply obtained through the permutations $\{(p, \nu, \beta) \leftrightarrow (q, \rho, \gamma)\}$. Simplifying the numerical factors and signs,³⁹ and specifying the trace, this yields:

$$\Gamma_{5\alpha\beta\gamma}^{(I)\mu\nu\rho}(-p-q, p, q) = -i \int \frac{d^4k}{(2\pi)^4} \text{tr}_D \left[\gamma_5 \gamma^\mu \frac{(\cancel{k}+\not{p})}{(k+p)^2-i\epsilon} \gamma^\nu \frac{\cancel{k}}{k^2-i\epsilon} \gamma^\rho \frac{(\cancel{k}-\not{q})}{(k-q)^2-i\epsilon} \right] \text{tr}_R t_\alpha t_\beta t_\gamma . \quad (7.3)$$

For large loop-momentum k the integrand behaves as $\sim d^4k k^{-3}$ and the integral is linearly divergent. Again, the presence of γ_5 prevents us from using dimensional regularization. As such, we revert to Pauli-Villars regularization, which consists of adding an artificial term with “fermions” of mass M and opposite statistics. Since it follows opposite statistics, the minus sign due to the fermion loop is not there and this term appears with an opposite sign. It should be noted that if the integral (7.3) had been convergent, the contribution of the regulator fields would have vanished as $M \rightarrow \infty$, as expected. However, since the integral is divergent, this translates into an M -dependence of the regularized integral. The regularized $\Gamma_{5\alpha\beta\gamma}^{(I)\mu\nu\rho}$ can be expressed as:

$$\Gamma_{5\alpha\beta\gamma}^{(I)\mu\nu\rho}(-p-q, p, q) \Big|_{\text{reg}} = i \int \frac{d^4k}{(2\pi)^4} \left(J_M^{\mu\nu\rho}(k, p, q) - J_0^{\mu\nu\rho}(k, p, q) \right) \text{tr}_R t_\alpha t_\beta t_\gamma, \quad (7.4)$$

where ($m = 0$ or $m = M$)

$$J_m^{\mu\nu\rho}(k, p, q) = \text{tr}_D \left[\gamma_5 \gamma^\mu \frac{\cancel{k}+\not{p}+im}{(k+p)^2+m^2-i\epsilon} \gamma^\nu \frac{\cancel{k}+im}{k^2+m^2-i\epsilon} \gamma^\rho \frac{\cancel{k}-\not{q}+im}{(k-q)^2+m^2-i\epsilon} \right] . \quad (7.5)$$

For a finite value of M , (7.4) is a convergent well-defined integral. As previously discussed, the Lagrangian for massive regulator Dirac fields lacks the chiral symmetry, and this is why an anomaly is found in the end. We are only interested in $(p_\mu + q_\mu)\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q) \Big|_{\text{reg}}$ and, hence, we only need to compute $(p_\mu + q_\mu)J_m^{\mu\nu\rho}(k, p, q)$. Rewriting:

$$\gamma_5(\not{p} + \not{q}) = \gamma_5(\not{p} + \cancel{k} - im) + (\cancel{k} - \not{q} - im)\gamma_5 + 2im\gamma_5 , \quad (7.6)$$

³⁸More precisely, we should take external momenta p, q and k and then the Feynman diagrams compute $i(2\pi)^4\delta^{(4)}(k+p+q)\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(k, p, q)$. The $(2\pi)^4\delta^{(4)}(k+p+q)$ arises upon Fourier transforming $\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(x, y, z)$. But as usual, we “factorize” and drop the overall $\delta^{(4)}(k+p+q)$ on both sides of the equation (7.2). We could have made similar remarks when we computed the vacuum polarization in (1.48).

³⁹Note that we have anticommutated the γ_5 with the γ^μ , resulting in another minus sign.

and using the cyclicity of the trace, certain terms cancel out between the numerator and denominator, and we end up with:

$$(p_\mu + q_\mu)J_m^{\mu\nu\rho}(k, p, q) = \frac{\text{tr}_D \gamma_5 \gamma^\nu (\not{k} + im) \gamma^\rho (\not{k} - \not{q} + im)}{[k^2 + m^2 - i\epsilon][(k - q)^2 + m^2 - i\epsilon]} + \frac{\text{tr}_D \gamma_5 (\not{k} + \not{p} + im) \gamma^\nu (\not{k} + im) \gamma^\rho}{[(k + p)^2 + m^2 - i\epsilon][k^2 + m^2 - i\epsilon]} \\ + 2im \frac{\text{tr}_D \gamma_5 (\not{k} + \not{p} + im) \gamma^\nu (\not{k} + im) \gamma^\rho (\not{k} - \not{q} + im)}{[(k + p)^2 + m^2 - i\epsilon][k^2 + m^2 - i\epsilon][(k - q)^2 + m^2 - i\epsilon]} . \quad (7.7)$$

For the Dirac trace involving γ_5 to be non vanishing, it should involve at least four γ^μ . Thus, the only non-vanishing contribution to the first term comes from $\text{tr}_D \gamma_5 \gamma^\nu \not{k} \gamma^\rho (\not{k} - \not{q}) = 4i\epsilon^{\nu\lambda\rho\sigma} k_\lambda (-q_\sigma)$. Then, when computing the regularized integral $\int d^4k$, the k_λ necessarily gets replaced by a four-vector proportional to q_λ , which is the only available one in this expression. When contracted with $\epsilon^{\nu\lambda\rho\sigma} q_\sigma$ this vanishes. Similarly, the second term in (7.7) only involves k and p and, after integration over d^4k , yields a contribution $\sim \epsilon^{\lambda\nu\sigma\rho} p_\lambda p_\sigma = 0$. Hence, only the third term in (7.7) will contribute and, since $\text{tr}_D \gamma_5 (\not{k} + \not{p} + im) \gamma^\nu (\not{k} + im) \gamma^\rho (\not{k} - \not{q} + im) = 4m \epsilon^{\nu\rho\lambda\sigma} p_\lambda q_\sigma$, we have

$$\int d^4k (p_\mu + q_\mu)J_m^{\mu\nu\rho}(k, p, q) = \int d^4k \frac{8im^2 \epsilon^{\nu\rho\lambda\sigma} p_\lambda q_\sigma}{[(k + p)^2 + m^2 - i\epsilon][k^2 + m^2 - i\epsilon][(k - q)^2 + m^2 - i\epsilon]} . \quad (7.8)$$

This term vanishes for $m = 0$, i.e. in $(p_\mu + q_\mu)J_0^{\mu\nu\rho}$, making it seem that the anomalous contribution solely comes from the regulator term. However, it should be noted that the first two terms in (7.7) only vanish after integration if we consider convergent integrals, and thus the combination $J_0 - J_M$. This is reassuring as it shows that the anomaly is not an artifact. Using (7.8) and (7.4) we arrive at

$$-i(p_\mu + q_\mu) \left[\Gamma_{5\alpha\beta\gamma}^{(I)\mu\nu\rho}(-p - q, p, q) \right]_{\text{reg}} = 8iM^2 \epsilon^{\nu\rho\lambda\sigma} p_\lambda q_\sigma J(p, q, M) \text{tr}_{\mathcal{R}} t_\alpha t_\beta t_\gamma , \quad (7.9)$$

with

$$J(p, q, M) = \int \frac{d^4k}{(2\pi)^4} \frac{1}{[(k + p)^2 + M^2 - i\epsilon][k^2 + M^2 - i\epsilon][(k - q)^2 + M^2 - i\epsilon]} . \quad (7.10)$$

The regulator M should be taken to ∞ in the end, so we only need the asymptotics of this integral for large M which is easily obtained by letting $k = Ml$:

$$J(p, q, M) \sim \frac{1}{M^2} \int \frac{d^4l}{(2\pi)^4} \frac{1}{[l^2 + 1 - i\epsilon]^3} = \frac{i}{32\pi^2 M^2} , \quad (7.11)$$

where the i comes from the Wick rotation.⁴⁰ Hence, we get a finite limit for $M^2 J(p, q, M)$ as we remove the regulator ($M \rightarrow \infty$). Adding the contribution from the second diagram with the appropriate permutations results in

$$-i(p_\mu + q_\mu) \Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(-p - q, p, q) = -\frac{1}{2\pi^2} \epsilon^{\nu\rho\lambda\sigma} p_\lambda q_\sigma \text{tr}_{\mathcal{R}} t_\alpha t_{(\beta} t_{\gamma)} . \quad (7.12)$$

Upon setting $t_\alpha = t$, this exactly reproduces the result (6.51), confirming once more that in the quantum theory the axial current J_5^μ is not conserved.

⁴⁰One has $\int \frac{d^4l}{(2\pi)^4} \frac{1}{[l^2 + 1 - i\epsilon]^3} = i \int \frac{d^4l_E}{(2\pi)^4} \frac{1}{[l_E^2 + 1]^3} = \frac{i}{(2\pi)^4} 2\pi^2 \int_0^\infty l^3 dl \frac{1}{[l^2 + 1]^3} = \frac{i}{8\pi^2} \frac{1}{2} \int_0^\infty d\xi \frac{\xi}{(\xi + 1)^3} = \frac{i}{32\pi^2}$.

At this stage, it is important to assess if this truly is a problem. In the present case, the anomaly is a mere indicator that the (anomalous) global chiral symmetry $\psi \rightarrow e^{i\epsilon\gamma_5}\psi$ is broken by a quantum effect. This means that certain selection rules no longer hold and otherwise forbidden (or suppressed) processes may occur. As previously mentioned, this precisely explains the observed disintegration rate $\pi^0 \rightarrow \gamma\gamma$. On the other hand, anomalies become problematic if they break gauge-invariance. In non-abelian gauge theories, vector currents couple to gauge fields and their non-conservation thus becomes a serious issue. Let us check that the vector currents are conserved. Because of the symmetry under exchange of (p, ν, β) and (q, ρ, γ) , it is enough to probe one of them and compute $p_\nu \left[\Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q) \right]_{\text{reg}}$. But now $p_\nu J_m^{\mu\nu\rho}(k, p, q)$ contains

$$\begin{aligned} (\not{k} + \not{p} + im)\not{p}(\not{k} + im) &= (\not{k} + \not{p} + im) \left[(\not{p} + \not{k} - im) - (\not{k} - im) \right] (\not{k} + im) \\ &= [(k+p)^2 + m^2] (\not{k} + im) - (\not{p} + \not{k} + im) [k^2 + m^2] , \end{aligned} \quad (7.13)$$

which now leads to

$$p_\nu J_M^{\mu\nu\rho}(k, p, q) = \frac{\text{tr}_D \gamma_5 \gamma^\mu (\not{k} + im) \gamma^\rho (\not{k} - \not{q} + im)}{[k^2 + m^2 - i\epsilon][(k-q)^2 + m^2 - i\epsilon]} - \frac{\text{tr}_D \gamma_5 \gamma^\mu (\not{k} + \not{p} + im) \gamma^\rho (\not{k} - \not{q} + im)}{[(k+p)^2 + m^2 - i\epsilon][(k-q)^2 + m^2 - i\epsilon]} . \quad (7.14)$$

As compared to (7.7) the third term is now absent. This difference is due to the absence of γ_5 in the vector current and hence in (7.14). As before, the two terms in (7.14) vanish after integration over d^4k , provided we always consider the convergent combination ⁴¹ $p_\nu (J_M^{\mu\nu\rho}(k, p, q) - J_0^{\mu\nu\rho}(k, p, q))$. We conclude that

$$p_\nu \Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q) = q_\rho \Gamma_{5\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q) = 0 . \quad (7.15)$$

7.2 Computations for the chiral anomaly

7.2.1 Chiral fermions

Let us begin by going over some facts about chiral fermions. Chirality projectors ⁴² P_L and P_R are defined as:

$$P_L = \frac{1 + \gamma_5}{2} , \quad P_R = \frac{1 - \gamma_5}{2} . \quad (7.16)$$

Denoting $\mathcal{H}$ the Hilbert space of all fermions, these projectors define the supplementary subspaces of left- and right-handed fermions, respectively $\mathcal{H}_L$ and $\mathcal{H}_R$. Left-handed and right-handed fermions are thus defined as follows:

$$\psi_L = P_L \psi , \quad \psi_R = P_R \psi , \quad \psi = \psi_L + \psi_R . \quad (7.17)$$

The projected fields ψ_L and ψ_R are eigenstates of the chirality matrix γ_5 :

$$\begin{aligned} \gamma_5 \psi_L &= \psi_L , & P_L \psi_L &= \psi_L , & P_R \psi_L &= 0 \\ \gamma_5 \psi_R &= -\psi_R , & P_R \psi_R &= \psi_R , & P_L \psi_R &= 0 . \end{aligned} \quad (7.18)$$

⁴¹For the first term the argument is exactly as before. For the second term, since the integral is convergent, we may shift $k \rightarrow k + q$ and then the integral can only depend on $p + q$ and there is no way to get a non-vanishing result when contracting with the ϵ -tensor.

⁴²There does not seem to be a unique choice of conventions used in the literature about what is called left-handed and right-handed. It would be better to talk about positive and negative chirality. But even then, the definitions used for γ_5 often differ by a sign. However, as long as we stick to one set of conventions, observable quantities will not depend on which one is chosen. Here, we follow the same conventions as [7, 14].

ψ_L has positive chirality and ψ_R has negative chirality.

Since $\gamma_5^\dagger = \gamma_5$ and $\{\gamma_5, \gamma^\mu\} = 0$, the projectors are hermitian and satisfy the following relations:

$$P_L \gamma^\mu = \gamma^\mu P_R \quad , \quad P_R \gamma^\mu = \gamma^\mu P_L \quad , \quad P_L P_R = P_R P_L = 0 \quad , \quad P_L \gamma^\mu P_L = 0 \quad (7.19)$$

$$\overline{\psi}_L P_L = 0 \quad , \quad \overline{\psi}_L P_R = \overline{\psi}_L \quad , \quad \overline{\psi}_R P_R = 0 \quad , \quad \overline{\psi}_R P_L = \overline{\psi}_R \quad . \quad (7.20)$$

As such,

$$\overline{\psi}_L \psi_L = \overline{\psi}_L (P_L \psi_L) = (\overline{\psi}_L P_L) \psi_L = 0 \quad , \quad \overline{\psi}_R \psi_R = (\overline{\psi}_R P_R) \psi_R = 0 \quad . \quad (7.21)$$

It follows that a mass term only couples opposite-chirality parts of a fermion,

$$\overline{\psi} \psi = \overline{\psi}_L \psi_R + \overline{\psi}_R \psi_L \quad , \quad (7.22)$$

and a chiral fermion (with either $\psi_L = 0$ or $\psi_R = 0$) cannot have a mass term.⁴³ On the other hand,

$$P_R \not{D} \psi_L = \not{D} P_L \psi_L = \not{D} \psi_L \quad , \quad P_L \not{D} \psi_R = \not{D} P_R \psi_R = \not{D} \psi_R \quad , \quad (7.23)$$

so that $\not{D} \psi_L$ is right-handed and $\not{D} \psi_R$ is left-handed, and $\overline{\psi}_L \not{D} \psi_L$ and $\overline{\psi}_R \not{D} \psi_R$ are non-vanishing.

7.2.2 Matter Lagrangian for chiral fermions

For a massless right-handed fermion, one can write a gauge-invariant kinetic term in the Lagrangian as

$$\mathcal{L}_{\text{matter}}^R = -\overline{\psi}_R \not{D} \psi_R = -\overline{\psi} \not{D} P_R \psi \quad . \quad (7.24)$$

The latter formulation, relying on non-chiral fermion fields and chirality projectors, allows us to continue using the usual Feynman rules. It follows that the propagators are now simply $\frac{-i}{(2\pi)^4} \times P_R \frac{-i \not{k}}{k^2 - i\epsilon}$ and that the vertices are $i(2\pi)^4 \times (i\gamma^\mu t_\alpha) P_R$.

Taking a closer look at the computation of the triangle diagram, one notices that there is always a projector P_R from a vertex next to a projector P_R from a propagator,

$$\dots (i\gamma^\mu t_\alpha) P_R P_R \frac{-i(\not{k} + \not{p})}{(k+p)^2 - i\epsilon} \dots = \dots (i\gamma^\mu t_\alpha) P_R \frac{-i(\not{k} + \not{p})}{(k+p)^2 - i\epsilon} \dots \quad (7.25)$$

so that this expression coincides exactly with the expression that would have been obtained if the vertices had been the same but the propagators were $\frac{-i}{(2\pi)^4} \times \frac{-i \not{k}}{k^2 - i\epsilon}$ i.e. without the chirality projector. These Feynman rules would be obtained from a matter Lagrangian

$$\mathcal{L}_{\text{matter}}^{R'} = -\overline{\psi} \not{D} \psi - \overline{\psi} (-i \not{A}) P_R \psi = -\overline{\psi} \not{D} P_L \psi - \overline{\psi} \not{D} P_R \psi \quad , \quad (7.26)$$

containing a left- and a right-handed (i.e. a non-chiral) propagating fermion, but only the right-handed part couples to the gauge field. This asymmetry causes the Lagrangian $\mathcal{L}_{\text{matter}}^{R'}$ to not be gauge-invariant.

⁴³One can still write a so-called Majorana mass term for chiral fermions, of the form $\overline{(\psi_L^c)} \psi_L$ using the charge conjugate spinor ψ_L^c (which is right-handed). But such terms violate fermion number conservation. Moreover, they are only compatible with gauge invariance if ψ_L is in a real representation of the gauge group, and, as we will see, for real representations the anomaly vanishes.

In the case of the gauge-invariant Lagrangian (7.24), the origin of the anomaly manifests itself through the definition of the following functional integral:

$$e^{i\tilde{\Gamma}[A]} = \int \mathcal{D}\psi_R \mathcal{D}\bar{\psi}_R e^{-i \int \bar{\psi}_R \not{D} \psi_R} \sim \text{Det}_R(\not{D}P_R) . \quad (7.27)$$

Since the functional integral is only over right-handed fermions, the corresponding determinant should be that of the restriction of $\not{D}P_R$ to $\mathcal{H}_R$. Since $\not{D}P_R$ is an operator that maps right-handed fermions to left-handed fermions, which live in supplementary subspaces of the Hilbert space, this restriction is not well defined and thus neither is the determinant. As we had seen in subsection 6.2.2, in the case of the abelian anomaly, the presence of an anomaly can again be related to the ill-definition of a determinant.

One may instead consider, in Euclidean signature, $(\not{D}_E P_R)^\dagger \not{D}_E P_R$ which maps $\mathcal{H}_R$ to itself and does have a well-defined and positive determinant. Taking its square root allows to define the modulus of what we would like to be the determinant of $\not{D}P_R$, but there remains an ambiguity in its phase. The phase corresponds to the imaginary part of the Euclidean effective action. If this ambiguity cannot be fixed in a gauge-invariant way one precisely has an anomaly. Thus the anomaly resides in the imaginary part of the Euclidean effective action. As we will discuss later-on, such an imaginary part necessarily involves the ϵ -tensor !

7.2.3 Feynman diagram computation of the triangle for chiral fermions

Let us now compute the three-point vertex function $\Gamma_{R,\alpha\beta\gamma}^{\mu\nu\rho}$ for a chiral right-handed fermion using the second formalism, i.e. using ordinary propagators and inserting a chirality projector P_R at each vertex. This is equivalent to computing the expectation value of three right-handed currents, $\Gamma_{R,\alpha\beta\gamma}^{\mu\nu\rho} = -\langle T(J_{R\alpha}^\mu J_{R\beta}^\nu J_{R\gamma}^\rho) \rangle$ where $J_{R\alpha}^\mu = i\bar{\psi}\gamma^\mu t_\alpha P_R \psi$. There are again two Feynman diagrams as shown in Fig. 6. The same computations as in the case of the abelian anomaly lead us to the following

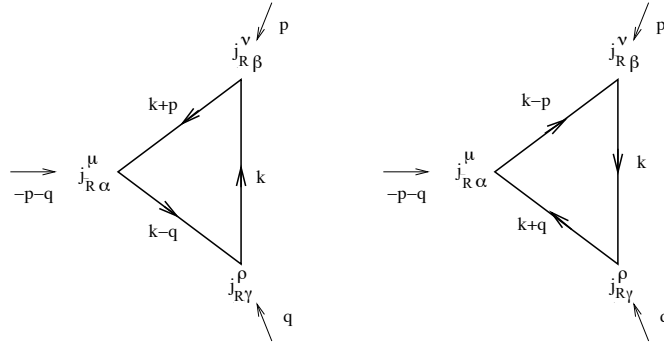


Figure 6: The two triangle diagrams contributing to the non-abelian chiral anomaly

expression ⁴⁴:

$$\Gamma_{R,\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q) = i \int \frac{d^4 k}{(2\pi)^4} \text{tr}_D \left\{ \gamma^\mu P_R \frac{(\not{k} + \not{p})}{(k+p)^2 - i\epsilon} \gamma^\nu P_R \frac{\not{k}}{k^2 - i\epsilon} \gamma^\rho P_R \frac{(\not{k} - \not{q})}{(k-q)^2 - i\epsilon} \right\} \text{tr}_R t_\alpha t_\beta t_\gamma + \{ (p, \nu, \beta) \leftrightarrow (q, \rho, \gamma) \} . \quad (7.28)$$

⁴⁴The sign difference with respect to (7.3) is only due to the fact that in (7.3) we have anti-commuted the γ_5 to the left of the γ^μ .

We will, once more, regularize this integral using the Pauli-Villars method. This seems quite counter intuitive as we had established that it does not make sense to consider a mass term in the case of chiral fermions. This obstacle is overcome through the alternative formulation (7.26) where we consider non-chiral fermions and chiral interactions. This allows us to add a mass term to the propagator, and thus to regularize using the Pauli-Villars method. The origin of the anomaly can then be traced back to the fact that the Lagrangian (7.26) used is not gauge invariant.

Since $\Gamma_{R,\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q)$ has a degree of divergence equal to one it is enough to add one regulator field of opposite statistics. However, at intermediate steps of the computation, we will encounter integrals of degree of divergence equal to two and, for this reason, we add one more pair of regulator fields (one fermionic and one bosonic). We use the notation $\eta_1 = \eta_3 = 1$ for the fermionic fields, and $\eta_2 = \eta_4 = -1$ for the bosonic ones, and denote their masses by M_n where, of course, the original field has $M_1 = 0$. Then we have

$$[\Gamma_{R,\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q)]_{\text{reg}} = i \int \frac{d^4k}{(2\pi)^4} \left(\sum_{n=1}^4 \eta_n K_{M_n}^{\mu\nu\rho}(k, p, q) \right) \text{tr}_{\mathcal{R}} t_\alpha t_\beta t_\gamma + \{(p, \nu, \beta) \leftrightarrow (q, \rho, \gamma)\} , \quad (7.29)$$

with

$$K_M^{\mu\nu\rho}(k, p, q) = \text{tr}_D \left[\gamma^\mu P_R \frac{\not{k} + \not{p} + iM}{(k+p)^2 + M^2 - i\epsilon} \gamma^\nu P_R \frac{\not{k} + iM}{k^2 + M^2 - i\epsilon} \gamma^\rho P_R \frac{\not{k} - \not{q} + iM}{(k-q)^2 + M^2 - i\epsilon} \right] . \quad (7.30)$$

Let us now simplify this expression using the properties of the chiral projectors. Firstly, due to the cyclicity of the trace, all mass terms of the numerator appear as

$$\dots \gamma^\lambda P_R iM \gamma^\sigma P_R \dots = \dots \gamma^\lambda iM P_R P_L \gamma^\sigma \dots = 0 . \quad (7.31)$$

Thus, they do not contribute. We can then further simplify this expression using the fact that $P_R^2 = P_R$:

$$K_M^{\mu\nu\rho}(k, p, q) = \frac{\text{tr}_D (\not{k} - \not{q}) \gamma^\mu (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R}{[(k+p)^2 + M^2 - i\epsilon][k^2 + M^2 - i\epsilon][(k-q)^2 + M^2 - i\epsilon]} . \quad (7.32)$$

Again, we are only interested in $(p_\mu + q_\mu) K_M^{\mu\nu\rho}(k, p, q)$. Thus we need

$$\begin{aligned} \text{tr}_D (\not{k} - \not{q}) (\not{p} + \not{q}) (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R &= \text{tr}_D (\not{k} - \not{q}) (\not{p} + \not{k} - \not{k} + \not{q}) (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R \\ &= (k+p)^2 \text{tr}_D (\not{k} - \not{q}) \gamma^\nu \not{k} \gamma^\rho P_R - (k-q)^2 \text{tr}_D (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R \\ &= [(k+p)^2 + M^2] \text{tr}_D (\not{k} - \not{q}) \gamma^\nu \not{k} \gamma^\rho P_R - [(k-q)^2 + M^2] \text{tr}_D (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R \\ &\quad + M^2 \text{tr}_D (\not{p} + \not{q}) \gamma^\nu \not{k} \gamma^\rho P_R , \end{aligned} \quad (7.33)$$

which, through cancellations between the numerators and denominators, leads us to:

$$\begin{aligned} (p_\mu + q_\mu) K_M^{\mu\nu\rho}(k, p, q) &= \frac{\text{tr}_D (\not{k} - \not{q}) \gamma^\nu \not{k} \gamma^\rho P_R}{[k^2 + M^2 - i\epsilon][(k-q)^2 + M^2 - i\epsilon]} - \frac{\text{tr}_D (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R}{[(k+p)^2 + M^2 - i\epsilon][k^2 + M^2 - i\epsilon]} \\ &\quad + M^2 \frac{\text{tr}_D (\not{p} + \not{q}) \gamma^\nu \not{k} \gamma^\rho P_R}{[(k+p)^2 + M^2 - i\epsilon][k^2 + M^2 - i\epsilon][(k-q)^2 + M^2 - i\epsilon]} . \end{aligned} \quad (7.34)$$

Again, in the first two terms, the γ_5 in the projectors P_R give vanishing contributions after doing the integration (since they only involve a single independent momentum), but now the 1 in P_R does

lead to a non-vanishing integral for each of these to terms. Hence,

$$\begin{aligned}
-i(p_\mu + q_\mu) [\Gamma_{R,\alpha\beta\gamma}^{\mu\nu\rho}(-p - q, p, q)]_{\text{reg}} &= \left(L^{\nu\rho}(-q) - L^{\nu\rho}(p) + \sum_{n=2}^4 \eta_n M_n^2 J_{M_n}^{\nu\rho}(p, q) \right) \text{tr}_{\mathcal{R}} t_\alpha t_\beta t_\gamma \\
&+ \{ (p, \nu, \beta) \leftrightarrow (q, \rho, \gamma) \} ,
\end{aligned} \tag{7.35}$$

where $L^{\nu\rho}(p)$ and $J_{M_n}^{\nu\rho}(p, q)$ are *convergent* integrals (at least as long as we keep the regulator masses M_N^2 finite), allowing us in particular to shift the integration variables:

$$L^{\nu\rho}(p) = \frac{1}{2} \int \frac{d^4 k}{(2\pi)^4} \sum_{n=1}^4 \eta_n \frac{\text{tr}_D (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho}{[(k+p)^2 + M_n^2 - i\epsilon][k^2 + M_n^2 - i\epsilon]} , \tag{7.36}$$

and, introducing Feynman parameters,

$$\begin{aligned}
J_{M_n}^{\nu\rho}(p, q) &= \int \frac{d^4 k}{(2\pi)^4} \frac{\text{tr}_D (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R}{[(k+p)^2 + M_n^2 - i\epsilon][k^2 + M_n^2 - i\epsilon][(k-q)^2 + M_n^2 - i\epsilon]} \\
&= 2 \int_0^1 dx \int_0^{1-x} dy \int \frac{d^4 k}{(2\pi)^4} \frac{\text{tr}_D (\not{k} + \not{p}) \gamma^\nu \not{k} \gamma^\rho P_R}{[(k+xp-yq)^2 + M_n^2 + r(p, q, x, y) - i\epsilon]^3} \\
&= 2 \int_0^1 dx \int_0^{1-x} dy \text{tr}_D (\not{p} + \not{q}) \gamma^\nu (y\not{q} - x\not{p}) \gamma^\rho P_R \int \frac{d^4 k'}{(2\pi)^4} \frac{1}{[k'^2 + M_n^2 + r(p, q, x, y) - i\epsilon]^3} ,
\end{aligned} \tag{7.37}$$

where we used that the odd power of k' in the numerator does not contribute, and where $r(p, q, x, y) = x(1-x)p^2 + y(1-y)q^2 + 2xy pq$. We only need the large M_n limits of $M_n^2 J_{M_n}^{\nu\rho}(p, q)$. They are all the same, (cf (7.11))

$$\begin{aligned}
\lim_{M_n \rightarrow \infty} M_n^2 J_{M_n}^{\nu\rho}(p, q) &= 2 \int_0^1 dx \int_0^{1-x} dy \text{tr}_D (\not{p} + \not{q}) \gamma^\nu (y\not{q} - x\not{p}) \gamma^\rho P_R \times \frac{i}{32\pi^2} \\
&= \frac{i}{16\pi^2} \frac{1}{6} \text{tr}_D (\not{p} + \not{q}) \gamma^\nu (\not{q} - \not{p}) \gamma^\rho P_R \\
&= \frac{i}{24\pi^2} \left[q^\nu q^\rho - p^\nu p^\rho + \frac{p^2 - q^2}{2} \eta^{\nu\rho} - i\epsilon^{\lambda\nu\sigma\rho} p_\lambda q_\sigma \right] .
\end{aligned} \tag{7.38}$$

We have mentioned above in subsection 7.2.2 (and explicitly observed for the abelian anomaly) that the anomaly is given by the terms involving $\epsilon^{\nu\rho\lambda\sigma}$. The $L^{\nu\rho}(p)$ and $L^{\nu\rho}(-q)$ do not give rise to any such terms. But we know from the (anomalous) Ward identity (6.55) that these additional pieces without ϵ -tensor are present. They correspond to the two terms in (6.55) that involve the structure constants⁴⁵ $C_{\alpha\gamma\epsilon}$ and $C_{\alpha\beta\epsilon}$. They are essentially vacuum-polarization diagrams, similar to the one computed above in QED. Just as the latter, they require the addition of a counterterm to yield a finite limit as the regulator masses are taken to infinity. We will not discuss this any further here.

⁴⁵One easily sees that $L^{\nu\rho}(p) = L^{\nu\rho}(-p) = L^{\rho\nu}(p)$, so that the combination $L^{\nu\rho}(-q) - L^{\nu\rho}(p)$ is odd under exchange of (p, ν) and (q, ρ) . The same is true for the terms in (7.38) that do not involve the ϵ -tensor. This means that upon adding the second diagram, one has $\text{tr}_{\mathcal{R}} t_\alpha t_\beta t_\gamma \rightarrow \text{tr}_{\mathcal{R}} t_\alpha (t_\beta t_\gamma - t_\gamma t_\beta) = iC_{\beta\gamma\delta} \text{tr}_{\mathcal{R}} t_\alpha t_\delta \sim C_{\beta\gamma\alpha}$, and this is how the structure constants arise from these terms.

Hence, inserting the $\epsilon^{\lambda\nu\sigma\rho}$ -terms of (7.38) into (7.35), we can already state the main result of this section

$$\begin{aligned} -i(p_\mu + q_\mu)\Gamma_{R,\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q)\Big|_{\epsilon\text{-terms}} &= \sum_{n=2}^4 \eta_n \frac{i}{24\pi^2} (-i)\epsilon^{\lambda\nu\sigma\rho} p_\lambda q_\sigma \text{tr}_{\mathcal{R}} t_\alpha t_\beta t_\gamma + \{(p, \nu, \beta) \leftrightarrow (q, \rho, \gamma)\} \\ &= \frac{1}{12\pi^2} \epsilon^{\nu\rho\lambda\sigma} p_\lambda q_\sigma D_{\alpha\beta\gamma}^{\mathcal{R}}, \end{aligned} \quad (7.39)$$

where we used $\sum_{n=2}^4 \eta_n = -\eta_0 = -1$ and $\text{tr}_{\mathcal{R}} t_\alpha t_{(\beta} t_{\gamma)} \equiv D_{\alpha\beta\gamma}^{\mathcal{R}}$, (and another sign comes from rearranging the indices on the ϵ -tensor). As compared to the analogous formula (6.51) for the abelian anomaly where the numerical factor was $-\frac{1}{2\pi^2}$, we now have a factor $\frac{1}{12\pi^2}$. The sign difference and one factor $\frac{1}{2}$ are easily understood as coming from the replacement of γ_5 by $P_R = \frac{1}{2}(1 - \gamma_5)$, but the presence of the 3 projectors also eliminated all mass terms in the numerator, implying further differences in the computation.

Using a reasoning similar to what we did for the abelian anomaly, equation (7.39) allows us to infer the part of the anomaly that is bilinear in the gauge fields :

$$\mathcal{A}_\alpha^R(x) = \frac{1}{24\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t_\alpha \partial_\mu A_\nu \partial_\rho A_\sigma + \mathcal{O}(A^3), \quad (7.40)$$

where we added a superscript R on $\mathcal{A}_\alpha(x)$ to remind us that this anomaly is computed for right-handed fermions. Note that our computation did not yield the $\mathcal{O}(A^3)$ contributions to the anomaly. To determine them one would need to evaluate similarly a four-point function which is given by the sum of 6 box diagrams. Of course, one only needs to compute one of them, the others correspond to permutations of the cyclic order of the vertices. Actually, this computation is not needed as the so-called Wess-Zumino consistency condition, to be discussed in the next section, completely fixes these higher-order contributions in terms of (7.40). The result is

$$\mathcal{A}_\alpha^R(x) = \frac{1}{24\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t_\alpha \partial_\mu \left(A_\nu \partial_\rho A_\sigma - \frac{i}{4} A_\nu [A_\rho, A_\sigma] \right). \quad (7.41)$$

Note the absence of terms quartic in A , which means that the 5-point function (pentagon diagrams) is not anomalous. Finally the A^3 -term can also be rewritten in terms of the symmetrized trace of 3 generators,⁴⁶ i.e. the $D_{abc}^{\mathcal{R}}$ -symbol :

$$\mathcal{A}_\alpha^R(x) = \frac{1}{24\pi^2} \epsilon^{\mu\nu\rho\sigma} \partial_\mu \left(A_\nu^\beta \partial_\rho A_\sigma^\gamma - \frac{i}{4} A_\nu^\beta [A_\rho, A_\sigma]^\gamma \right) D_{\alpha\beta\gamma}^{\mathcal{R}}. \quad (7.42)$$

If one does the same computations for a left-handed fermion, the only difference being $\gamma_5 \rightarrow -\gamma_5$, and then obviously the anomaly obtained is exactly the opposite :

$$\mathcal{A}_\alpha^L(x) = -\mathcal{A}_\alpha^R(x). \quad (7.43)$$

Obviously also, since a non-chiral fermion is composed of a left-handed and a right-handed fermion, both in the same representation of the gauge group, there is no anomaly for a non-chiral fermion.

⁴⁶Note that $\epsilon^{\mu\nu\rho\sigma} [A_\nu, [A_\rho, A_\sigma]]$ vanishes by the Jacobi identity. This allows us to rewrite $\epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t_\alpha A_\nu [A_\rho, A_\sigma] = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} \text{tr}_{\mathcal{R}} t_\alpha (A_\nu [A_\rho, A_\sigma] + [A_\rho, A_\sigma] A_\nu) = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} A_\nu^\beta i C_{\delta\epsilon}^\gamma A_\rho^\delta A_\sigma^\epsilon \text{tr}_{\mathcal{R}} t_\alpha (t_\beta t_\gamma + t_\gamma t_\beta) = \epsilon^{\mu\nu\rho\sigma} A_\nu^\beta i C_{\delta\epsilon}^\gamma A_\rho^\delta A_\sigma^\epsilon D_{\alpha\beta\gamma}^{\mathcal{R}}.$

8 General characterization and properties of the anomalies

8.1 Group structure

The possibility of having an anomaly is intrinsically linked to the structure of the gauge group. In fact, if $D_{\alpha\beta\gamma}^{\mathcal{R}} = \text{tr}_{\mathcal{R}} t_{\alpha} t_{\beta} t_{\gamma} = \text{str}_{\mathcal{R}} t_{\alpha} t_{\beta} t_{\gamma}$ vanishes, then so does the chiral anomaly (7.42). This is the case when we consider real or pseudo-real representations. Indeed, if the representation $\mathcal{R}$ is real, it is equal to its complex conjugate representation $\mathcal{R}^*$ which is given by $\exp(i\theta^{\alpha} t_{\alpha}^{\mathcal{R}*}) = (\exp(i\theta^{\alpha} t_{\alpha}^{\mathcal{R}}))^* = \exp(-i\theta^{\alpha} (t_{\alpha}^{\mathcal{R}})^*) = \exp(-i\theta^{\alpha} (t_{\alpha}^{\mathcal{R}})^T)$.⁴⁷ This translates into the following condition on the generators $t_{\alpha}^{\mathcal{R}*} = -(t_{\alpha}^{\mathcal{R}})^T$ and for a real representation $t_{\alpha}^{\mathcal{R}} = t_{\alpha}^{\mathcal{R}*} = -(t_{\alpha}^{\mathcal{R}})^T$, while for a pseudo-real one, this relation reads: $t_{\alpha}^{\mathcal{R}} = P t_{\alpha}^{\mathcal{R}*} P^{-1} = -P (t_{\alpha}^{\mathcal{R}})^T P^{-1}$, with some fixed matrix P . As such, for a real or pseudo-real representation

$$\begin{aligned} D_{\alpha\beta\gamma}^{\mathcal{R}} &= \text{str}_{\mathcal{R}} t_{\alpha} t_{\beta} t_{\gamma} = (-)^3 \text{str}_{\mathcal{R}} P (t_{\alpha})^T P^{-1} P (t_{\beta})^T P^{-1} P (t_{\gamma})^T P^{-1} \\ &= -\text{str}_{\mathcal{R}} t_{\alpha}^T t_{\beta}^T t_{\gamma}^T = -\text{str}_{\mathcal{R}} (t_{\gamma} t_{\beta} t_{\alpha})^T = -\text{str}_{\mathcal{R}} t_{\gamma} t_{\beta} t_{\alpha} = -\text{str}_{\mathcal{R}} t_{\alpha} t_{\beta} t_{\gamma} = -D_{\alpha\beta\gamma}^{\mathcal{R}}, \end{aligned} \quad (8.1)$$

(where $P = \mathbf{1}$ if the representation is real). This implies that $D_{\alpha\beta\gamma}^{\mathcal{R}} = 0$.

Thus, we cannot have anomalies in (four-dimensional) gauge theories with a gauge group which only has real and pseudo-real representations. This is the case for many well-known groups such as $SO(4n)$, $n \geq 2$ or $SO(2n+1)$, $n \geq 1$ as well as several others. Yet other groups do have non (pseudo) real representations, but still with all $D_{\alpha\beta\gamma}^{\mathcal{R}}$ vanishing. In fact, the only groups which have non-vanishing $D_{\alpha\beta\gamma}^{\mathcal{R}}$ and can lead to anomalies are $SU(n)$, $n \geq 3$, $U(1)$ and product groups involving these as factors. Note that this argument is specific to 4 mod 4 dimensions where the relevant D -symbol is the trace of an odd number of generators.

In particular, in the low-energy QFT limit of ten-dimensional superstring theories anomalies will involve the 6-index $D_{\alpha_1 \dots \alpha_6}$ -symbol, which has no reason to vanish for the $SO(n)$ groups.

8.2 Particles or antiparticles

It should be clear that if different chiral particles (fields) are present, one has to add the contributions of the different one-loop diagrams. As a result, all chiral particles contribute additively to the anomaly. Since the chiral fermions are massless, their contributions to the anomaly cannot depend on a mass parameter, but only on the group representation $\mathcal{R}$ they are in. As a result, their contribution to the anomaly is fully captured by the corresponding $D^{\mathcal{R}}$ -symbol. Since the fermionic propagator in the loop is the propagator of the particle *and* its anti-particle, it is clear that each particle-anti-particle pair only contributes once. But in four dimensions the antiparticle of a left-handed particle is right-handed and vice versa, so one might ask whether it is important which one of the pair has been designated as the particle. We will now show that it does not matter.

Suppose we have a left-handed fermionic particle in some representation $\mathcal{R}_L$ of the gauge group, with associated quantum field ψ . Then its antiparticle is described by the charge-conjugated $\psi^c = i\gamma^0 \mathcal{C} \psi^*$. The charge conjugation matrix $\mathcal{C}$ satisfies $\mathcal{C} \gamma_5^* = \gamma_5 \mathcal{C}$ so that

$$\gamma_5 \psi^c = \gamma_5 i\gamma^0 \mathcal{C} \psi^* = -i\gamma^0 \mathcal{C} (\gamma_5 \psi)^* = \mp i\gamma^0 \mathcal{C} \psi^* = \mp \psi^c \quad \text{if } \gamma_5 \psi = \pm \psi. \quad (8.2)$$

⁴⁷We take the generators to be hermitian.

We see that indeed, the charge conjugated field has opposite chirality. On the other hand, due to the complex conjugation, ψ^c transforms in the complex conjugate representation $(\mathcal{R}_L)^*$. As we have just seen, the D -symbol for the complex conjugate representation has the opposite sign. So if one starts out with the anti-particle, one gets one minus sign from having the opposite chirality and one gets another minus sign from being in the complex conjugate representation, so that overall the contribution to the anomaly is the same. Again, this argument is correct in $4 \bmod 4$ space-time dimensions. In $2 \bmod 4$ dimensions the complex conjugate representation has the same D -symbol. But one also has $\mathcal{C}\gamma^* = -\gamma\mathcal{C}$ (where γ is the chirality matrix $\sim \gamma^0 \dots \gamma^{d-1}$ in d dimensions), so that particles and anti-particles have the same chirality. Then switching from particles to anti-particles results in two plus signs rather than two minus signs, and again it does not matter.

Note that in 2, 6, 10, etc dimensions one can have a chiral particle that is its own anti-particle, usually called a Majorana-Weyl fermion.

8.3 Locality

Let us show that the anomaly is local i.e. that it can be written as $\int d^4x f(x)$ where $f(x)$ only depends on the values of the fields and their derivatives at x . In Fourier transform terms, this means that the anomaly is polynomial in the momenta.

Firstly, it should be underlined that anomalies arise from the regularization of divergent one-loop diagrams, and that convergent diagrams do not need to be regularized. Secondly, taking a derivative with respect to external momenta decreases the degree of divergence by one.⁴⁸

The three-point vertex function has a degree of divergence $\mathcal{D} = 1$, and two derivatives result in a convergent integral. Then three derivatives of $(p_1)_{\mu_1} \Gamma^{\mu_1 \dots \mu_3}(p_1, p_2, p_3)$ result in a convergent integral which doesn't need regularization and which then cannot have any anomalous part. As such, three derivatives kill the anomaly meaning that it is a polynomial of degree at most two in the momenta. This is indeed what we have observed. This argument is general for all anomalous one-loop diagrams in an arbitrary (even) space-time dimension : If the diagram has degree of divergence $\mathcal{D}$, then $\mathcal{D} + 1$ derivatives with respect to the external momenta result in a convergent integral that needs no regularization. As such, $\mathcal{D} + 2$ derivative kill the anomalous contribution to $(p_1)_{\mu_1} \Gamma^{\mu_1 \dots \mu_n}$ which must be a polynomial of degree $\mathcal{D} + 1$ in the p_j .

8.4 How many points necessary for an anomalous diagram?

This observation leads to an interesting constraint on the number of points (vertices) needed to observe an anomaly for a given space-time dimension d . Firstly, in odd dimensions, one cannot define a γ_5 matrix satisfying the ordinary properties. Thus, there are no anomalies.⁴⁹ Secondly, let us consider an even dimensional space-time with $d = 2n$. The quantity we are interested in is the contraction of the $(m+1)$ -point vertex function with one of the external momenta, $p_{\mu_1}^{(1)} \Gamma^{\mu_1 \dots \mu_{m+1}}$, and

⁴⁸In fact, the derivation yields a sum of terms where one of the internal propagators is replaced by $\frac{\partial}{\partial p_\mu} \left(-\frac{1}{\not{k} + \not{p} + \dots} \right) = \frac{(\not{k} + \not{p} + \dots) \gamma^\mu (\not{k} + \not{p} + \dots)}{[(k+p+\dots)^2]^2}$. For a one-loop diagram it is easy to see that all terms have their degree of divergence reduced by one.

⁴⁹More precisely, there are no anomalies of the kind we discuss here : abelian anomaly, anomaly under gauge transformations, or anomalies under diffeomorphisms / local Lorentz transformations. There are however anomalies under discrete symmetries like parity [28].

we expect to find a quantity proportional to the epsilon tensor $\epsilon^{\nu_1 \dots \nu_d}$. As just seen, we also expect the anomaly to be polynomial in the momenta. Thus, we need $d - m$ independent external momenta that we can contract this quantity with. This means that we need to have $d - m \leq m$, i.e. $m \geq n$. Thus, the minimal number of points that can lead to an anomalous one loop diagram is $n + 1$. In the case of 4 dimensional spacetime, this is in fact the triangle, as we have extensively discussed. In 6 dimensions the anomaly first shows up in a box diagram, in 8 dimension in the pentagon diagram, and in 10 dimensions one must compute the hexagon diagram. This requires way more work and quite a few tricks [6] !

8.5 Finiteness

Both in the case of the abelian and chiral anomaly, the anomalies we computed had finite, i.e. regulator independent, coefficients. (This was not the case for the non-anomalous part of the triangle amplitude which required addition of counterterms to cancel their divergent pieces.)

Let us now show why this in fact *has* to be the case because of locality and dimensionality. Firstly, the three-point vertex function has a degree of divergence one, which means, per Section 8.3, that the anomalous part is a polynomial of degree at most two in the external momenta. The most general form of such a polynomial is the following:

$$(p_\mu + q_\mu) \Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q) \Big|_{\text{anom part}} = D_1^{\nu\rho\lambda\sigma} p_\lambda p_\sigma + D_2^{\nu\rho\lambda\sigma} q_\lambda q_\sigma + D_3^{\nu\rho\lambda\sigma} p_\lambda q_\sigma + C_1^{\nu\rho\lambda} p_\lambda + C_2^{\nu\rho\lambda} q_\lambda + B_{\alpha\beta\gamma}^{\nu\rho}, \quad (8.3)$$

where $D_1, D_2, D_3, C_1, C_2, B$ are constant Lorentz tensors independent of momenta. First of all, there are no constant Lorentz tensors independent of momenta with three indices, so that C_1 and C_2 must vanish. Next, we expect the anomaly to be proportional to the epsilon tensor. Thus, D_1, D_2 and B have to vanish also, as the associated terms do not respect the required anti-symmetry. All that we are left with is the term $D_3^{\nu\rho\lambda\sigma} p_\lambda q_\sigma$.

Secondly, let us look at how the scaling dimension constrains the tensor $D_3^{\nu\rho\lambda\sigma}$. The three-point vertex function $\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q)$ has a scaling dimension 1 since each fermion propagator contributes a -1 and the integration contributes a $+4$. Thus, $(p_\mu + q_\mu) \Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(-p-q, p, q)$ has a scaling dimension 2, which implies that $D_3^{\nu\rho\lambda\sigma}$ has a scaling dimension 0. As it does not involve the momenta (and the fields are massless), by dimensional considerations, it cannot involve the Pauli-Villars regulator mass M . This means that the anomaly is necessarily finite, i.e. does not involve regularization-dependent coefficients.

The same type of argument can also be applied to the box-diagram responsible for the A^3 -part of the anomaly $\mathcal{A}_{\text{cubic}} \sim (p_\mu + q_\mu + k_\mu) \Gamma_{\alpha\beta\gamma\delta}^{\mu\nu\rho\sigma}(-p-q-k, p, q, k) \Big|_{\text{anom part}}$. This box diagram has one more fermion propagator and a degree of divergence 0, so that $\mathcal{A}_{\text{cubic}}$ is a polynomial of degree one. On the other hand the box diagram has scaling dimension 0, so that $\mathcal{A}_{\text{cubic}}$ has scaling dimension 1. Again, the only possible structure is $E_3^{\nu\rho\sigma\lambda}(p_\lambda + q_\lambda + k_\lambda)$ with a dimensionless E_3 proportional to the ϵ -tensor. So this term is also finite.

It is easy to see that these arguments generalize to arbitrary even dimensions : in $d = 2n$ dimensions, the $n + 1$ point one-loop diagram has degree of divergence $d - (n + 1) = n - 1$ so that $k_\mu \Gamma^{\mu \dots}$ is annihilated by $n + 1$ derivatives and must be a polynomial of degree n in the external momenta. But this expression has scaling dimension n and contains n momenta, hence the tensor-

coefficient is dimensionless and independent of any momentum and, again, cannot depend on the Pauli-Villars regulator mass M . The anomalous higher-point functions are treated similarly. This proves finiteness of the anomaly in any (even) dimension.

8.6 One-loop exactness

The Adler-Bardeen theorem [4] states that higher loop diagrams do not contribute to the anomaly, meaning that the one-loop computation we have done here is actually exact to all orders in perturbation theory. The original proof is complicated and well beyond the scope of these notes. Part of the reasoning is that higher loops necessarily contain bosonic (gauge field) propagators and these can be regulated in a gauge invariant way, thus effectively regulating the full diagram [6, 7]. The subtlety here is to show that regulating the gauge boson propagators sufficiently regulates all loop integrals and that this does not generate any non-gauge invariant (anomalous) terms.

Alternatively, we have argued that anomalies originate from a non-trivial transformation of the functional integral measure, giving rise to a non-trivial (functional) Jacobian determinant. But determinants always correspond to one-loop effects. This can be seen e.g. by introducing a loop counting parameter λ , as we did in subsection 2.5. The determinant cannot contain any λ , so it is of order $\lambda^{L-1} = \lambda^0$, confirming indeed $L = 1$, i.e. it contributes at the one-loop order only.

Below, we will sketch how the chiral anomaly can be related to the index of an appropriate extension of the Dirac operator. This shows again that the anomaly is independent of any loop-counting parameter and must correspond to a one-loop contribution only.

8.7 Relevant and irrelevant anomalies

We have shown that the anomaly is related to the non-invariance of the quantum effective action, see e.g. (6.53), $\int d^4x \epsilon^\alpha(x) \mathcal{A}_\alpha(x) = \delta_\epsilon \tilde{\Gamma}[A]$. Obviously also, since the anomaly arises as a one-loop effect, it is of “higher order” than the classical action. If we re-introduced $\hbar$ as a loop counting parameter, one would see that the anomaly has an extra $\hbar$ as compared to the classical action. Also, as obvious from the triangle, it involves 3 gauge coupling constants g (which we have hidden in the generators and hence in the D -symbol). As such the anomaly is of order g^3 where the tree-level couplings are of order g . We see that the anomaly is of the same order as the counterterms one adds to the action to cancel possible divergences (in other diagrams). One may then wonder whether it is possible to add some non gauge-invariant higher-order “counterterm” $\hat{S}_{ct}[A]$ to the classical action such that $\delta_\epsilon \hat{S}_{ct} = -\delta_\epsilon \tilde{\Gamma}[A]$. More precisely, $\hat{S}_{ct}[A]$ would be an order g^3 term to be added to the classical action and then would also appear as an additional term in $\tilde{\Gamma}[A]$:

$$\tilde{\Gamma}[A] \rightarrow \tilde{\Gamma}[A] + \hat{S}_{ct}[A] \quad , \quad \delta_\epsilon(\tilde{\Gamma}[A] + \hat{S}_{ct}[A]) = 0 \quad . \quad (8.4)$$

It looks as if one can simply take $\hat{S}_{ct}[A] = -\tilde{\Gamma}[A]$ and get rid of the anomaly. This is of course not true. In any local quantum field theory, we are only allowed to add local counterterms to the action, i.e. a space-time integral of a local expression involving only the fields at the point x as well as a finite number of derivatives acting on these fields. While we have shown that the anomaly $\delta_\epsilon \tilde{\Gamma}[A]$ is such a local functional, the effective action $\tilde{\Gamma}[A]$ itself is not. To understand this, note that the anomaly corresponds to $(p_1)_\mu \Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(p_1, p_2, p_3)$ which upon Fourier transforming is $\frac{\partial}{\partial x^\mu} \Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(x, y, z)$.

On the other hand, the effective action contains the $\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}$ itself, and “removing” the derivative with respect to x^μ is not a local operation.

Thus the question is whether or not we can find a *local functional* of the gauge fields $\hat{S}_{ct}[A]$ such that its gauge variation cancels the anomaly $\delta_\epsilon \tilde{\Gamma}[A]$. If we can, the anomaly is called irrelevant because it can be removed by adding such a counterterm to the classical action. If we cannot, the anomaly is called relevant :

$$\text{The anomaly is relevant} \iff - \int d^4x \epsilon^\alpha(x) \mathcal{A}_\alpha(x) \neq \delta_\epsilon \hat{S}_{ct}[A] \quad \text{with } \hat{S}_{ct}[A] \text{ a local functional.} \quad (8.5)$$

It is relatively easy to see that the chiral anomaly is a relevant anomaly. The $\hat{S}_{ct}[A]$ would need to give an additional contribution $\Delta\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(k, p, q)$ to the 3-point function that must be of scaling dimension 1, i.e. linear in the external momenta and totally symmetric under exchange of the external legs (i.e. $(k, \mu, \alpha) \leftrightarrow (p, \nu, \beta) \leftrightarrow (q, \rho, \gamma)$). This is usually referred to as the Bose symmetry, since this $\Delta\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}$ arises from three functional derivatives with respect to the gauge fields which are bosonic and commute. On the other hand, $\Delta\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}$ must involve the ϵ -tensor in order to possibly cancel the anomaly. To cancel the anomaly due to a right-handed fermion, the requirement is that

$$-i(p_\mu + q_\mu) \Delta\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(-p - q, p, q) = -\frac{1}{12\pi^2} \epsilon^{\nu\rho\lambda\sigma} p_\lambda q_\sigma D_{\alpha\beta\gamma} , \quad (8.6)$$

and similarly (with the opposite sign) for a left-handed fermion. It is easy to see that no such $\Delta\Gamma_{\alpha\beta\gamma}^{\mu\nu\rho}(k, p, q)$ with the required Bose symmetry can exist. We conclude that the chiral anomaly is indeed relevant.

8.8 Anomaly cancellation

As already mentioned, the contribution of fermions to the gauge anomaly is additive: the total anomaly is obtained by summing over all chiral species. Furthermore, the structure of the anomaly is universal and the only model dependence factors in through the group-theoretical factor $D_{\alpha\beta\gamma}^{\mathcal{R}}$, which depends only on the gauge group and representations $\mathcal{R}$ of the fermions.

For a general chiral gauge theory with left-handed fermions in representations $\mathcal{R}_i^L$ and right-handed ones in representations $\mathcal{R}_j^R$, the anomaly reads:

$$\mathcal{A}_\alpha = -\frac{1}{24\pi^2} \epsilon^{\mu\nu\rho\sigma} \partial_\mu \left(A_\nu^\beta \partial_\rho A_\sigma^\gamma - \frac{i}{4} A_\nu^\beta [A_\rho, A_\sigma]^\gamma \right) \left(\sum_i D_{\alpha\beta\gamma}^{\mathcal{R}_i^L} - \sum_j D_{\alpha\beta\gamma}^{\mathcal{R}_j^R} \right). \quad (8.7)$$

Breaking of the gauge invariance by anomalies ruins the consistency of the quantum theory. Hence, gauge invariance at the quantum level requires that the expression $(\sum_i D_{\alpha\beta\gamma}^{\mathcal{R}_i^L} - \sum_j D_{\alpha\beta\gamma}^{\mathcal{R}_j^R})$ vanishes, which is the condition for anomaly cancellation.

In the standard model the gauge group is the product $SU(3) \times SU(2) \times U(1)$. While we have seen that $SU(2)$ alone has a vanishing D -symbol, for product groups one can have different generators appear in the trace of the D -symbol. Since the generators for the different groups act on different representation spaces, the trace factorizes accordingly. If we call T_a the 8 generators of the $SU(3)$, t_α the 3 generators of $SU(2)$ and q the generator of the $U(1)$ then the different non-vanishing traces are (within any irreducible representation)

$$\text{tr } T_a T_b T_c \quad , \quad \text{tr } t_\alpha t_\beta q = q \text{tr } t_\alpha t_\beta \sim q \delta_{\alpha\beta} \quad \text{tr } T_a T_b q = q \text{tr } T_a T_b \sim q \delta_{ab} \quad , \quad \text{tr } q q q = q^3 . \quad (8.8)$$

For example, $\text{tr } t_\alpha t_\beta q$ will result form a triangle diagram with two vertices where one could attach $SU(2)$ gauge bosons and one vertex where a $U(1)$ gauge boson could be attached. q is the $U(1)$ (hyper) charge of the chiral fermion “circulating” in the loop. The different anomalies corresponding to the terms in (8.8) are referred to as $SU(3)^3$, $SU(2)^2 \times U(1)$, $SU(3)^2 \times U(1)$ and $U(1)^3$ -anomalies. For each of these anomalies one must check that they vanish when summed over all chiral fermion representations (with the signs according to the chirality). Remarkably, in the Standard Model this condition is satisfied separately within each fermion generation, leading to a non-anomalous gauge theory. Historically the requirement of anomaly cancellation had let to the prediction of the existence of the (then still unobserved) charm quark [21].

8.9 Wess-Zumino consistency condition

We have defined the anomaly $\mathcal{A}_\alpha$ as the gauge variation of the effective action after integration over the matter fields:

$$\delta_\epsilon \tilde{\Gamma}[A] = \int \epsilon^\alpha(x) \mathcal{A}_\alpha(x) \quad \Leftrightarrow \quad \mathcal{A}_\alpha(x) = - \left(D_\mu \frac{\delta}{\delta A_\mu(x)} \right)_\alpha \tilde{\Gamma}[A] . \quad (8.9)$$

This definition constrains the expression of $\mathcal{A}_\alpha$ and leads to a linear equation called the Wess-Zumino consistency condition. In fact, one can rewrite (8.9) using a functional derivative operator $\mathcal{F}$ defined as:

$$\mathcal{F}_\alpha(x) = - \left(D_\mu \frac{\delta}{\delta A_\mu(x)} \right)_\alpha = - \frac{\partial}{\partial x^\mu} \frac{\delta}{\delta A_\mu^\alpha(x)} - C_{\alpha\delta\gamma} A_\mu^\delta(x) \frac{\delta}{\delta A_\mu^\gamma(x)} , \quad (8.10)$$

which yields $\mathcal{A}_\alpha(x) = \mathcal{F}_\alpha(x) \tilde{\Gamma}[A]$. Using the Jacobi identity, one can show that this functional derivative satisfies the following algebra:

$$[\mathcal{F}_\alpha(x), \mathcal{F}_\beta(y)] = \delta^{(4)}(x - y) C_{\alpha\beta\gamma} \mathcal{F}_\gamma(x) , \quad (8.11)$$

Applying this to the effective action then yields the following relation for the anomaly :

$$\mathcal{F}_\alpha(x) \mathcal{A}_\beta(y) - \mathcal{F}_\beta(y) \mathcal{A}_\alpha(x) = C_{\alpha\beta\gamma} \delta^{(4)}(x - y) \mathcal{A}_\gamma(x). \quad (8.12)$$

This is Wess-Zumino’s consistency condition [15]. It constrains the form of the anomaly and, since $\mathcal{F}$ mixes terms of different degrees in the gauge fields, it allows us to determine the higher order ($\mathcal{O}(A^3)$) terms that contribute to the anomaly using only the term found in (7.40) from the triangle computation. This will be done in the next section after a reformulation of the Wess-Zumino consistency condition in terms of the nilpotency of the BRST transformation.

8.10 Reformulation in terms of differential forms

We have seen that the anomalies are intrinsically related to the presence of the γ_5 -matrix which in turn implies that the expression of the anomaly involves the $\epsilon^{\mu\nu\rho\sigma}$ -tensor. This will allow us to rewrite the anomaly $\int d^4x \epsilon^a \mathcal{A}_a$ as the integral of a differential 4-form. Differential p -forms can be naturally integrated over p -manifolds, and here the 4-manifold is the (flat) space-time. But the formulation in terms of differential forms immediately generalizes to curved manifolds. On a curved manifold we need to distinguish the antisymmetric symbol $\hat{\epsilon}^{\mu\nu\rho\sigma}$ which equals ± 1 or 0 and the “true” $\epsilon^{\mu\nu\rho\sigma}$ -tensor.

They are related as $\hat{\epsilon}^{\mu\nu\rho\sigma} = \sqrt{-g} \epsilon^{\mu\nu\rho\sigma}$ where $\sqrt{-g}$ is the square-root of (minus) the determinant of the space-time metric $g_{\mu\nu}$. Then

$$dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma = \hat{\epsilon}^{\mu\nu\rho\sigma} d^4x = \epsilon^{\mu\nu\rho\sigma} \sqrt{-g} d^4x . \quad (8.13)$$

The gauge field potentials $A_\mu = A_\mu^\alpha t_\alpha$ are combined into the Lie algebra-valued one forms

$$A = A_\mu dx^\mu = A_\mu^\alpha dx^\mu t_\alpha . \quad (8.14)$$

Recall that the exterior derivative is $d = dx^\mu \partial_\mu$. Using these relations, we can rewrite the anomaly for a left-handed (i.e. positive chirality) fermion, which is the opposite of (7.41), now on a curved space-time manifold, as

$$\begin{aligned} \int d^4x \sqrt{-g} \epsilon^a \mathcal{A}_a^L &= - \int d^4x \sqrt{-g} \frac{1}{24\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{tr} \epsilon \partial_\mu \left(A_\nu \partial_\rho A_\sigma - \frac{i}{4} A_\nu A_\rho A_\sigma + \frac{i}{4} A_\nu A_\sigma A_\rho \right) \\ &= - \frac{1}{24\pi^2} \int dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma \text{tr} \epsilon \partial_\mu \left(A_\nu \partial_\rho A_\sigma - \frac{i}{4} A_\nu A_\rho A_\sigma + \frac{i}{4} A_\nu A_\sigma A_\rho \right) \\ &= - \frac{1}{24\pi^2} \int \text{tr} \epsilon d \left(A \wedge dA - \frac{i}{2} A \wedge A \wedge A \right) . \end{aligned} \quad (8.15)$$

The final expression as the integral of a 4-form does not involve the metric, as usual with differential forms. Note that in the language of differential forms the field strength is given by

$$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu = (\partial_\mu A_\nu - i A_\mu A_\nu) dx^\mu \wedge dx^\nu = dA - iA^2 . \quad (8.16)$$

We can similarly rewrite the abelian anomaly (6.34) in terms of the differential forms resulting in ⁵⁰

$$\int d^4x \sqrt{-g} \epsilon \mathcal{A} = - \frac{1}{4\pi^2} \int \epsilon \text{tr} t F \wedge F = - \frac{1}{4\pi^2} \int \epsilon \text{tr} t (dA \wedge dA - 2i dA \wedge A \wedge A) . \quad (8.17)$$

It is worth noting that the chiral anomaly (8.15) is *not* $\sim \int \text{tr} \epsilon F \wedge F$.

⁵⁰Since t commutes with the A , it follows from the cyclicity of the trace that $\text{tr} tA \wedge A \wedge dA = \text{tr} t dA \wedge A \wedge A$. For the same reason $\text{tr} tA \wedge A \wedge A \wedge A = 0$.

9 BRST cohomology and descent equations

9.1 BRST transformations

We have discussed the anomalies under gauge transformations in terms of the effective action $\tilde{\Gamma}[A]$ obtained by doing the functional integral over the matter fields of $e^{iS_{\text{matter}}[\psi, \bar{\psi}, A]}$, cf (6.42). Since the gauge anomalies originated in the anomalous transformation of the functional integral measure for the chiral matter fermions, this was indeed the appropriate object to consider for studying the anomalies. However, in the full quantum theory one needs to go on and add the classical Yang-Mills action ⁵¹ $\int (-\frac{1}{4}F_{\mu\nu}^\alpha F^{\alpha\mu\nu})$ and also do the functional integral over the gauge fields, possibly with the relevant operator insertions, as described in Sect. 2. As is well known, one then needs to fix the gauge by some appropriate gauge-fixing action and compensate for the arbitrariness of this gauge fixing by inserting the Faddeev-Popov determinant. The latter can be rewritten in terms of a Gaussian functional integral over a so-called ghost and anti-ghost fields. This results in the full action being

$$S_{\text{full}} = \int d^4x \left(\underbrace{\mathcal{L}_{\text{matter}}[\psi, \bar{\psi}, A] - \frac{1}{4}F_{\mu\nu}^\alpha F^{\alpha\mu\nu}}_{\mathcal{L}_{\text{class}}} + \underbrace{\frac{\xi}{2}h_\alpha h_\alpha + h_\alpha f^\alpha[A]}_{\mathcal{L}_{\text{gauge fixing}}} + \underbrace{\omega_\alpha^* \rho^\alpha[A, \omega]}_{\mathcal{L}_{\text{ghost}}} \right) \quad (9.1)$$

The first two terms are referred to as the classical Lagrangian. $f^\alpha[A]$ is the gauge-fixing condition, e.g. $f^\alpha = \partial_\mu A^{\alpha\mu}$, and h_α is an (adjoint) auxiliary field. Since it appears quadratically, it can be trivially integrated out, resulting in an equivalent gauge-fixing Lagrangian $\frac{1}{2\xi}f^\alpha f^\alpha$. The parameter ξ is positive but otherwise arbitrary. The quantity ρ_α is given by

$$\rho^\alpha(x) = \int d^4y \frac{\delta f^\alpha[A(x)]}{\delta \epsilon^\beta(y)} \omega^\beta(y) , \quad (9.2)$$

and ω^α and ω_α^* are the anti-commuting ghost and anti-ghost fields. Integrating them out would give back the Faddeev-Popov determinant $\text{Det} \frac{\delta f^\alpha[A(x)]}{\delta \epsilon^\beta(y)}$. Note that the quantity ρ_α can be interpreted as the gauge transformed gauge-fixing function f^α with the parameter ϵ^α replaced by the ghost field ω^α . Despite its non-local appearance, for the standard choices of gauge-fixing functions f^α , the quantity $\rho^\alpha(x)$ only depends on $A_\mu^\beta(x)$ and $\omega^\beta(x)$ and their derivatives.

Of course, S_{full} is the gauge-fixed action and no longer is gauge invariant. But it is still invariant under some “remnant” of the gauge invariance which is invariance under BRST transformations. These BRST transformations are denoted by s and defined to act on the gauge and matter fields as ordinary gauge transformations but with ω^α replacing ϵ^α . Then the gauge-invariance of the classical Lagrangian implies its BRST invariance. Defining $\omega = \omega^\alpha t_\alpha$, $\omega^* = \omega_\alpha^* t_\alpha$ as well as $h = h_\alpha t_\alpha$, the full set of BRST transformations is defined as

$$\begin{aligned} sA_\mu &= \partial_\mu \omega - i[A_\mu, \omega] & , & \quad s\psi = i\omega\psi & , & \quad s\bar{\psi} = i\bar{\psi}\omega \\ s\omega &= i\omega\omega & , & \quad s\omega^* = -h & , & \quad sh = 0 & , \end{aligned} \quad (9.3)$$

⁵¹Here we assume that the Lie algebra is compact, i.e. that $\text{tr} t_\alpha t_\beta$ is a positive definite bilinear form. One can then adopt a basis where $\text{tr} t_\alpha t_\beta = C_{\mathcal{R}} \delta_{\alpha\beta}$ with the constant $C_{\mathcal{R}}$ depending on the representation. One can then use the “metric” $\delta_{\alpha\beta}$ and its inverse $\delta^{\alpha\beta}$ to raise and lower the indices $\alpha, \beta, \gamma, \dots$ which means that we can write them without distinction as upper or lower indices. Also, the structure constants $C_{\alpha\beta\gamma}$ then are totally antisymmetric - as we already implicitly assumed earlier in these notes.

together with the fact that $s(FG) = (sF)G \pm FsG$ with a minus sign if F is anti-commuting. Note for later use that this implies $s\omega^\alpha = -\frac{1}{2}C_{\alpha\beta\gamma}\omega^\beta\omega^\gamma$. The action of s on ω , ω^* and h is chosen such that this BRST transformation s is nilpotent, i.e. when s acts twice on any of the fields one gets zero :

$$s^2 = 0 . \quad (9.4)$$

It is now easy to see that also the remaining part of the Lagrangian is BRST invariant. First note that, by its definition, one has $\rho^\alpha = sf^\alpha$, so that by the nilpotency of s one has $s\rho^\alpha = 0$. Then

$$s(\mathcal{L}_{\text{gauge fixing}} + \mathcal{L}_{\text{ghost}}) = h_\alpha(sf^\alpha) + (s\omega_\alpha^*)\rho^\alpha - \omega_\alpha^*(s\rho^\alpha) = h_\alpha\rho^\alpha - h_\alpha\rho^\alpha = 0 . \quad (9.5)$$

One can actually write

$$\mathcal{L}_{\text{gauge fixing}} + \mathcal{L}_{\text{ghost}} = s \left(-\omega_\alpha^* \left(\frac{\xi}{2} h^\alpha + f^\alpha \right) \right) \equiv s G_{-1} , \quad (9.6)$$

and then (9.5) trivially follows from (9.4).

Since the ghost fields are anti-commuting, as are the differential forms of odd degree, it makes some sense to adopt the convention that the ghost (and anti-ghost) fields also anticommute with dx^μ . Accordingly then, one has $s d = -d s$. However, here we find it more straightforward *not* to adopt this convention, and rather have the ghost fields *commute* with dx^μ so that

$$s d = d s . \quad (9.7)$$

9.2 Reformulating relevant anomalies as BRST cohomologies

Since $s^2 = 0$, one introduces the notion of BRST-closed functionals F obeying $sF = 0$, as well as of BRST-exact functionals H obeying $H = sG$ for some G . This is much as for the exterior derivative d acting on differential forms. The role of the degree of the forms is played here by the ghost number, with ω and ω^* having ghost numbers $+1$ and -1 . Indeed, acting on any field, s increases the ghost number by one unit. We can now say that $\mathcal{L}_{\text{class}}$ is BRST-closed and that $\mathcal{L}_{\text{gauge fixing}} + \mathcal{L}_{\text{ghost}}$ is BRST-exact. Since gauge-invariant physical observables should not depend on the gauge-fixing function $f_a[A]$, changing it amounts to changing this BRST-exact term. The physics is then determined by the BRST-closed $\mathcal{L}_{\text{class}}$ up to addition of BRST-exact terms., i.e. by the BRST cohomology class at ghost number zero.

We now show that relevant anomalies can be understood as BRST-cohomology classes at ghost number one. First, since $\tilde{\Gamma}[A]$ only depends on A , its BRST variation simply is its gauge variation with ϵ replaced by ω , i.e.

$$\mathcal{A}[\omega, A] \equiv \int d^4x \omega^\alpha(x) \mathcal{A}_\alpha(x) = s\tilde{\Gamma}[A] . \quad (9.8)$$

From now on, we refer to the left-hand side as the anomaly. It clearly has ghost number one and is BRST-closed since $s\mathcal{A}[\omega, A] = s^2\tilde{\Gamma}[A] = 0$. Actually it looks as if the anomaly $\mathcal{A}[\omega, A]$ is also BRST-exact. But as with the discussion of relevant anomalies, we will only allow local functionals and, although the anomaly is a local functional, $\tilde{\Gamma}$ is not. Hence, the anomaly is BRST-closed, but not BRST-exact. As such it constitutes a BRST-cohomology class at ghost number one.

We have seen before that the Wess-Zumino consistency condition simply was a consequence of the anomaly being defined as the gauge variation of some (non-local) functional. At present this condition is a simple consequence of $s^2 = 0$. Indeed, applying s to (9.8), we get

$$\begin{aligned} 0 &= s^2 \tilde{\Gamma}[A] = s \int d^4x \omega^\alpha(x) \mathcal{A}_\alpha(x) = \int d^4x \left((s\omega^\alpha(x)) \mathcal{A}_\alpha(x) - \omega^\alpha(x) s\mathcal{A}_\alpha(x) \right) \\ &= -\frac{1}{2} \int d^4x C_{\alpha\beta\gamma} \omega^\beta(x) \omega^\gamma(x) \mathcal{A}_\alpha(x) - \int d^4x d^4y \omega^\alpha(x) \omega^\beta(y) \mathcal{F}_\beta(y) \mathcal{A}_\alpha(x), \end{aligned} \quad (9.9)$$

where we used the functional differential operator $\mathcal{F}$ introduced in (8.10) to express the BRST variation of $\mathcal{A}_\alpha$. Indeed, $\delta_\epsilon \mathcal{A}_\alpha(x) = \int d^4y \epsilon^\beta(y) \mathcal{F}_\beta(y) \mathcal{A}_\alpha(x)$ translates into $s\mathcal{A}_\alpha(x) = \int d^4y \omega^\beta(y) \mathcal{F}_\beta(y) \mathcal{A}_\alpha(x)$. Since the ghost fields anti-commute, we can anti-symmetrize the last term in (α, x) and (β, y) and then (9.9) can be written as

$$0 = \int d^4x d^4y \omega^\alpha(x) \omega^\beta(y) \left(\delta^{(4)}(x-y) C_{\alpha\beta\gamma} \mathcal{A}_\gamma(x) + \mathcal{F}_\beta(y) \mathcal{A}_\alpha(x) - \mathcal{F}_\alpha(x) \mathcal{A}_\beta(y) \right), \quad (9.10)$$

which vanishes by the Wess-Zumino condition (8.12). Hence, the fact that the anomaly is BRST closed is equivalent to satisfying the Wess-Zumino condition.⁵²

9.3 Solving the WZ-condition

The next task is to solve $s\mathcal{A}[\omega, A] = 0$ which will determine the terms cubic in A in terms of the quadratic terms. One way to do this is by “brute force”, i.e. one writes all possible terms compatible with the general constraints on the anomaly, i.e. $\mathcal{A}[\omega, A]$ must be the integral of a 4-form since, as discussed in subsection 8.10, it must necessarily involve the ϵ -tensor, and it must be made from d ’s and A ’s. Thus we may start with

$$\mathcal{A}^L[\omega, A] = -\frac{1}{24\pi^2} \int \text{tr} \omega \left(dAdA + \beta_1 dAA^2 + \beta_2 AdAA + \beta_3 A^2 dA + \gamma A^4 \right), \quad (9.11)$$

where the wedge product is understood, i.e. $dAA^2 = dA \wedge A \wedge A$, etc. We have already adjusted the coefficient of the $dAdA$ -term as appropriate for the anomaly due to a left-handed (positive chirality) fermion. In principle one could also add double-trace terms such as $\text{tr} \omega A \text{tr} AdA$. But there is no way to achieve a cancellation between such double-trace terms and the single trace terms written in (9.11).

Note that since s changes the scaling dimensions of all field uniformly, we can only achieve cancellation between the different terms if we start out with all terms in (9.11) having the same scaling dimension. One can then compute $s\mathcal{A}[\omega, A]$ and impose that it vanishes. Note that one assumes that $\omega = \omega^a(x) t_a$ vanishes at spatial infinity as well as at $t \rightarrow \pm\infty$ so that one is allowed to freely integrate by parts.⁵³ Said differently $\int d(\dots) = 0$. This means that s of the integrand must

⁵²At first sight it looks as if the Wess-Zumino condition only implies BRST-closedness. But the vanishing of the double integral in (9.10) for all ghost fields ω implies the vanishing of the term in parenthesis. Indeed the latter has exactly the same antisymmetry properties as $\omega^\alpha(x) \omega^\beta(y)$. Upon setting $\omega^a = \omega_{(1)}^a \pm \omega_{(2)}^a$ with arbitrary (anticommuting) $\omega_{(1,2)}^a$, a standard argument then implies the vanishing of the parenthesis. But this is exactly the Wess-Zumino consistency condition.

⁵³For topologically non-trivial space-times requiring several charts, one would need to further assume that the transition functions for the gauge-field are trivial or else that $\omega^\alpha(x)$ is non-vanishing only inside a given chart.

be an exact 4-form. A straightforward computation then yields $\gamma = 0$ and $\beta_1 = -\beta_2 = \beta_3 = -\frac{i}{2}$, so that $\beta_1 dAA^2 + \beta_2 AdAA + \beta_3 A^2 dA = -\frac{i}{2} dA^3$, and one gets exactly the expression (8.15) we already anticipated.

The previous method to solve the Wess-Zumino consistency condition is certainly sufficient in 4 dimensions, but when one wants to study anomalies in higher dimensions $d = 2n$, things quickly become painful. There is a much more elegant way to solve for the BRST-cohomology at ghost-number one on the space of local functionals that are integrals of local d -forms made from the gauge-field one-forms, exterior derivatives and one ghost field. It is based on the so-called descent equations which relate this d -form to a characteristic class which is a (formal) $(d+2)$ -form on a $(d+2)$ -dimensional space-time. This higher-dimensional space can be constructed as a product manifold $M = \mathcal{M} \times \mathcal{D}$ with $\mathcal{M}$ the space-time manifold and $\mathcal{D}$ a (two-dimensional) disc. The geometry of this $(d+2)$ -dimensional space-time, as well as the descent equations play quite an important role in the study of [18] which relates the anomaly on the d -dimensional $\mathcal{M}$ to the index of the Dirac operator $\mathcal{D}$ on the $(d+2)$ -dimensional space-time M , as we will briefly explain in the next section.

9.4 Descent equations : characteristic classes and Chern-Simons forms

9.4.1 Characteristic classes

Within the present context, we define a characteristic class P_r as a local $2r$ -form on a (compact) $2r$ -dimensional manifold M that is constructed from the gauge field strength F (or curvature 2-form R) such that its integral over the manifold is sensitive only to non-trivial transition functions of the gauge bundle. Well-known examples are instantons and magnetic monopoles. For the (four-dimensional) instantons $P = \text{tr } F \wedge F$ and the manifold is compactified $\mathbb{R}^4 \simeq S^4$. For the monopole $P = F$ and the manifold is S^2 . Since this is the simplest example let us exhibit the construction of the monopole bundle in some more detail.

For the (abelian) magnetic monopole one looks at static configurations with space being ordinary space $\mathbb{R}^3$ with a little ball of radius r_0 at the origin removed, so that one deals with $[r_0, \infty) \times S^2$. Since the gauge field strength 2-form here does not depend on the radial coordinate we may view it as defined on the sphere S^2 . The integral $\int_{S^2} F$ computes the flux of the magnetic field through the sphere and, by Gauss' theorem, this measures the total magnetic charge "inside" the sphere. Obviously, if $F = dA$ with a globally well-defined gauge-field one-form A , then by Stoke's theorem one has $\int_{S^2} F = \int_{S^2} dA = \int_{\partial S^2} A = 0$ since the boundary ∂S^2 of the sphere is empty, and there is no magnetic charge inside. If instead $F = g_m \sin \theta d\theta d\varphi$, then $\int_{S^2} F = 4\pi g_m$. Since this is independent of the radius of the sphere, we interpret g_m as the magnetic charge of a monopole sitting at the origin (or rather inside the little ball of radius r_0). Since F is closed, it must be locally exact, i.e. given locally by $F = dA$, but A can no longer be globally defined. One needs to introduce two (topologically trivial) charts which are the upper half and lower half spheres $S^2_{\pm}$ (or the product of these half-spheres with $[r_0, \infty)$), and then

$$A_+ = g_m(1 - \cos \theta)d\varphi \text{ on } S^2_+ \quad , \quad A_- = g_m(-1 - \cos \theta)d\varphi \text{ on } S^2_- \quad . \quad (9.12)$$

Each $A_{\pm}$ is well-defined and non-singular on its domain of definition, but would become ill-defined

when one tried to extend it to the full sphere.⁵⁴ Both $A_{\pm}$ yield the same F as $F = dA_{\pm}$ and are related by a $U(1)$ gauge transformation on the overlap $S^2_+ \cap S^2_- = S^1$:

$$A_+ - A_- = 2g_m d\varphi = ig_{+-}^{-1} dg_{+-} \quad , \quad g_{+-} = \exp(-2ig_m \varphi) . \quad (9.13)$$

For g_{+-} to be well-defined on S^1 one must have $g_m = \frac{N}{2}$ with integer N , so that $g_{+-} = \exp(-iN\varphi)$. We see that the magnetic charge is quantized, and equivalently

$$\frac{1}{2\pi} \int_{S^2} F = N . \quad (9.14)$$

One can change F by adding an exact term $F \rightarrow F + d\tilde{A}$ (with a globally well-defined $\tilde{A}$) and this does not change the integral of F over the manifold S^2 . The value of this integral only depends on the transition function g_{+-} , i.e. on N . Thus F is the characteristic class P_1 for the $U(1)$ gauge field strength on S^2 . It is instructive to compute $\int F$ again, in a way which clearly shows why the value of the integral is only sensitive to the transition function g_{+-}

$$\begin{aligned} \int_{S^2} F &= \int_{S^2_+} F + \int_{S^2_-} F = \int_{S^2_+} dA_+ + \int_{S^2_-} dA_- = \int_{\partial S^2_+} A_+ + \int_{\partial S^2_-} A_- \\ &= \int_{S^1} A_+ - \int_{S^1} A_- = \int_{S^1} ig_{+-}^{-1} dg_{+-} = 2\pi N , \end{aligned} \quad (9.15)$$

where the minus sign in the second line is due to the fact that the boundaries S^1 of the upper and lower half spheres have opposite orientation.

The general case proceeds exactly along the same line of argument. For an arbitrary gauge group, define

$$P_r(F) = \text{tr } F^r \equiv \text{tr } \underbrace{F \wedge \dots \wedge F}_{r \text{ times}} = F^{\alpha_1} \wedge F^{\alpha_r} \text{tr } t_{\alpha_1} \dots t_{\alpha_r} = F^{\alpha_1} \wedge F^{\alpha_r} \text{str } t_{\alpha_1} \dots t_{\alpha_r} , \quad (9.16)$$

since the 2-forms F^{α_i} and F^{α_j} all commute and effectively symmetrise the trace of the product of the t_{α} . The $P_r(F)$ are gauge invariant, since under a gauge transformation $F \rightarrow gFg^{-1}$, and by the cyclicity of the trace all g and g^{-1} cancel. Hence, $P_r(F)$ has trivial transition functions and is a globally well-defined $2r$ -form. Recall the Bianchi identity, $DF \equiv dF - i[A, F] = 0$, from which follows⁵⁵

$$dP_r(F) = d \text{tr } F^r = r \text{tr } (dF) F^{r-1} = ir \text{tr } AF^r - ir \text{tr } FAF^{r-1} = 0 , \quad (9.17)$$

so that $P_r(F)$ is closed, and hence locally (i.e. on each topologically trivial chart) exact. Moreover, as we now show, if A_1 and A_2 have the same transition functions, then $P_r(F_1) - P_r(F_2) = dQ$ with some *globally well-defined* $2r - 1$ form Q . Since we assume the manifold M to be compact it then follows from Stoke's theorem that

$$\int_M P_r(F_2) = \int_M P_r(F_1) , \quad (9.18)$$

⁵⁴The point is that $d\varphi$ is well-defined, except at the north pole ($\theta = 0$) and at the south pole ($\theta = \pi$). Now, in A_+ the factor $1 - \cos\theta$ vanishes at the north pole, and in A_- the factor $-1 - \cos\theta$ vanishes at the south pole. So, A_+ would be singular at the south pole and A_- at the north pole. If one tries to just use a single gauge potential, its singularity corresponds to the ‘‘Dirac string’’ in the earlier descriptions of the Dirac monopole.

⁵⁵Note that for Lie-algebra-valued p and q -forms ζ and ξ one has $\text{tr } \zeta \xi = (-)^{pq} \text{tr } \xi \zeta$.

as required for a characteristic class. We now construct the $2r - 1$ from Q . Let $A_t = A_1 + t(A_2 - A_1)$, $t \in [0, 1]$ be a gauge field interpolating between A_1 and A_2 . Clearly, A_t has the same transition functions as both A_1 and A_2 . Let F_t be the associated field strength $F_t = dA_t - iA_t^2$. One computes

$$\frac{\partial}{\partial t} F_t = d(A_2 - A_1) - i(A_2 - A_1)A_t - iA_t(A_2 - A_1) = D_t(A_2 - A_1) , \quad (9.19)$$

where D_t is the covariant derivative with gauge field A_t . In particular, the Binchi identity for F_t is $D_t F_t = 0$. Then

$$\frac{\partial}{\partial t} P_r(F_t) = r \operatorname{tr} \frac{\partial F_t}{\partial t} F_t^{r-1} = r \operatorname{tr} (D_t(A_2 - A_1)) F_t^{r-1} = r D_t (\operatorname{tr} (A_2 - A_1) F_t^{r-1}) . \quad (9.20)$$

Now while A_2 and A_1 each transform inhomogeneously under gauge transformations (on the overlap of two charts), their difference transforms homogeneously, just as does F_t . Hence $\operatorname{tr} (A_2 - A_1) F_t^{r-1}$ is invariant and the covariant derivative D_t in the last expression in (9.20) reduces to an ordinary exterior derivative. Integrating t from 0 to 1, one gets

$$P_r(F_2) - P_r(F_1) = dQ \quad , \quad Q = r \int_0^1 dt \operatorname{tr} (A_2 - A_1) F_t^{r-1} , \quad (9.21)$$

with a globally well defined $2r - 1$ form Q . This proves (9.18). Note that we always wrote tr , but we could have just as well written everything in terms of the symmetrized trace str , cf. the remark after (9.16)

9.4.2 Chern-Simons forms and descent equations

For reasons that will become clear in the next subsection, we now set $r = n + 1$. Since the $P_{n+1}(F)$ are closed, they must be locally exact, i.e. within each (topologically trivial) chart one must have $P_{n+1}(F) = dQ_{2n+1}$. The previous construction tells us how to construct these Q_{2n+1} . Within each chart, there is no issue of transition function, and we can simply take $A_2 = A$ and $A_1 = 0$, and then conclude from (9.21) that

$$P_{n+1}(F) = dQ_{2n+1} \quad , \quad Q_{2n+1} = (n + 1) \int_0^1 dt \operatorname{tr} A (tdA - it^2 A^2)^n . \quad (9.22)$$

Of course, $P_{n+1}(F) = dQ_{2n+1}$ defines Q_{2n+1} only up to adding an exact term $d\alpha_{2n}$. The Q_{2n+1} defined as in (9.22) are called the Chern-Simons forms. Explicitly one has ⁵⁶

$$\begin{aligned} Q_3 &= \operatorname{tr} \left(\operatorname{Ad} A - \frac{2i}{3} A^3 \right) , \\ Q_5 &= \operatorname{tr} \left(\operatorname{Ad} \operatorname{Ad} A - \frac{3i}{2} A^3 dA - \frac{3}{5} A^5 \right) . \end{aligned} \quad (9.23)$$

We should stress that these Chern-Simons forms are defined on each chart in terms of the gauge field on this chart.

⁵⁶One may e.g. want to check that indeed $dQ_3 = \operatorname{tr} F^2$. First note that while F^2 contains a term A^4 its trace vanishes. This is fortunate since when acting with d on Q_3 one cannot generate a term $\sim A^4$. It is then easy to check that $dQ_3 = \operatorname{tr} (d\operatorname{Ad} A - 2idAA^2) = \operatorname{tr} F^2$.

Unlike the $P_{n+1}(F)$, the Chern-Simons forms Q_{2n+1} are *not* gauge invariant, but one can use the gauge invariance of the $P_{n+1}(F)$ to constrain the gauge variation of the Q_{2n+1} as follows :

$$0 = \delta_\epsilon P_{n+1}(F) = \delta_\epsilon dQ_{2n+1} = d\delta_\epsilon Q_{2n+1} . \quad (9.24)$$

We see that the (infinitesimal) gauge variation of Q_{2n+1} is closed and, hence, locally exact. We conclude that on a given chart

$$\delta_\epsilon Q_{2n+1} = dQ_{2n}^1 . \quad (9.25)$$

The ambiguity of adding $d\alpha_{2n}$ to Q_{2n+1} results in an ambiguity of adding $\delta_\epsilon \alpha_{2n}$ to Q_{2n}^1 . Of course, (9.25) only defined Q_{2n}^1 up to adding an exact term. Thus, overall, Q_{2n}^1 is only defined up to adding $\delta_\epsilon \alpha_{2n} + d\beta_{2n-1}^1$. Altogether one has the following *descent equations* relating the $2n+2$ form P_{n+1} to the $2n+1$ form Q_{2n+1} and to the $2n$ form Q_{2n}^1 (locally on each chart) as :

$$dP_{n+1} = \delta_\epsilon P_{n+1} = 0 \Rightarrow P_{n+1} = dQ_{2n+1} , \quad \delta_\epsilon Q_{2n+1} = dQ_{2n}^1 , \quad Q_{2n}^1 \simeq Q_{2n}^1 + \delta_\epsilon \alpha_{2n} + d\beta_{2n-1}^1 . \quad (9.26)$$

By replacing the gauge variation parameter ϵ by the ghost field ω one immediately gets the BRST version of the descent equations :

$$dP_{n+1} = sP_{n+1} = 0 \Rightarrow P_{n+1} = dQ_{2n+1} , \quad sQ_{2n+1} = dQ_{2n}^1 , \quad Q_{2n}^1 \simeq Q_{2n}^1 + s\alpha_{2n} + d\beta_{2n-1}^1 . \quad (9.27)$$

The upper index now indicates the ghost number, i.e. Q_{2n}^1 contains one ghost field. Explicitly one has for $n = 1$ and $n = 2$

$$Q_2^1 = \text{tr } \omega dA \quad , \quad Q_4^1 = \text{tr } \omega d\left(AdA - \frac{i}{2}A^3\right) . \quad (9.28)$$

Again, we could have used the symmetrized trace throughout. Indeed, one can explicitly show that Q_2^1 and Q_4^1 as given here are equal to the same expressions but using the symmetrized trace.⁵⁷

We have seen that the $2n$ form Q_{2n}^1 is determined, up to the equivalence just mentioned, by the characteristic class P_{n+1} . The question then arises how many different characteristic classes there are. Gauge (or BRST) invariance implies that they must be constructed from traces of the field strengths F . Thus the only possible characteristic classes are of the form

$$\begin{aligned} P_{n+1} &= \text{tr } F^{n+1} \quad , \quad P_{n+1}^{(k, n+1-k)} = \text{tr } F^k \wedge \text{tr } F^{n+1-k} , \\ P_{n+1}^{(k_1, k_2, n+1-k_1-k_2)} &= \text{tr } F^{k_1} \wedge \text{tr } F^{k_2} \wedge \text{tr } F^{n+1-k_1-k_2} , \dots \end{aligned} \quad (9.29)$$

It is instructive to look at an example. Let $n = 4$ and consider the 10-form characteristic class $P_5^{(2,3)} = \text{tr } F^2 \wedge \text{tr } F^3$. Then $P_5^{(2,3)} = dQ_9^{(2,3)}$ with $Q_9^{(2,3)} = Q_3 \wedge \text{tr } F^3$ or $Q_9^{(2,3)} = \text{tr } F^2 \wedge Q_5$, and correspondingly $Q_8^{1,(2,3)} = Q_2^1 \wedge \text{tr } F^3$ or $Q_8^{1,(2,3)} = \text{tr } F^2 \wedge Q_4^1$. One says that one is doing the descent in the first or in the second factor. Obviously, any linear combination of the two with total coefficient

⁵⁷Of course, for Q_2^1 symmetrizing the trace follows from cyclicity. For Q_4^1 the first term $\text{tr } \omega dAdA$ is symmetric under exchange of the two factors dA which again symmetrizes the trace. The second term $\sim \text{tr } \omega dA^3$ seems to involve the trace of 4 generators, but one can easily see from the antisymmetry under exchanging two factors of A that one really has the trace of two generators with the commutator of two other generators, so that one again has a trace of only 3 generators. Actually it is not difficult to show that this trace of 3 generators also is a symmetrized trace, as one expected since the whole Q_4^1 is obtained through descent from $P_3(F)$ which involved a symmetrized trace.

one is also possible. As argued above, the difference of two solutions of the descent equations must be the sum of a BRST-exact and a d -exact term. Indeed, it is easy to see that the difference is given by

$$\begin{aligned}
& s(Q_3 \wedge Q_5) - d(Q_2^1 \wedge Q_5 - Q_3 \wedge Q_4^1) \\
&= (sQ_3) \wedge Q_5 + Q_3 \wedge (sQ_5) - (dQ_2^1) \wedge Q_5 - Q_2^1 \wedge dQ_5 + (dQ_3) \wedge Q_4^1 - Q_3 \wedge dQ_4^1 \\
&= -Q_2^1 \wedge dQ_5 + (dQ_3) \wedge Q_4^1 = -Q_2^1 \wedge \text{tr } F^2 + \text{tr } F^3 \wedge Q_4^1 .
\end{aligned} \tag{9.30}$$

9.5 BRST cohomology at ghost number one

We will now consider a $d = 2n$ dimensional space-time. We will show that starting with the characteristic class P_{n+1} defined on the $d + 2$ dimensional manifold $M = \mathcal{M} \times \mathcal{D}$, the quantity $\int Q_d^1$ is a non-trivial representative of the BRST cohomology at ghost number one, where $Q_d^1 \equiv Q_{2n}^1$ is obtained by the descent equations (9.27) from P_{n+1} . As such, it provides a solution of the Wess-Zumino consistency condition. This allows us to identify the anomaly as

$$\mathcal{A}[\omega, A] = c \int Q_d^1 , \tag{9.31}$$

where the only freedom resides in the coefficient c . While we have seen that the characteristic classes in $2n + 2$ dimensions can be the single-trace P_{n+1} or the multiple-trace $P_{n+1}^{k_1, \dots, n+1-k_1, \dots}$, only single traces can originate from the one-loop diagrams and we must keep with the P_{n+1} only.⁵⁸ A single computation of an $n + 1$ point one-loop diagram is needed to determine this coefficient c . Moreover, as we have seen, one does not need to compute the full diagram, only the part involving the ϵ -tensor.

Let us first note that the equivalence $Q_d^1 \equiv Q_d^1 + s\alpha_d + d\beta_{d-1}^1$ is exactly what we expect for the anomaly. Indeed, the d -exact term disappears upon integrating and the s -exact term amounts to adding a local counterterm $s \int \alpha_d$ to the anomaly, which we can always do, as discussed above. Hence, up to adding local counterterms, a relevant anomaly is determined by a non-vanishing coefficient c . This in turn corresponds to exactly one characteristic class in $d + 2$ dimensions.

Let us now prove that $\int Q_d^1 \equiv \int Q_{2n}^1$ is a non-trivial representative of the BRST cohomology at ghost number one. Let us recall that at ghost number zero the BRST-cohomology consists of the gauge-invariant functionals of the gauge field only.⁵⁹ Furthermore, there are no gauge-invariant forms of odd degree, and then for these forms the BRST-cohomology at ghost number zero is empty. Lastly, still at ghost number zero, a gauge invariant even form of the gauge field strengths like the P_{n+1} is a non-trivial representative of the BRST-cohomology. Indeed, the P_{n+1} are BRST-invariant, and cannot be $s\gamma_{2n+2}^{-1}$ since this γ_{2n+2}^{-1} must contain at least one anti-ghost field ω^* and $s\omega^* \sim h$ which is not present in P_{n+1} .

Now the descent equation $sQ_{2n+1} = dQ_{2n}^1$ and $s^2 = 0$ immediately give

$$0 = s(sQ_{2n+1}) = s(dQ_{2n}^1) = d(sQ_{2n}^1) , \tag{9.32}$$

⁵⁸This is true for the chiral gauge anomalies. For the gravitational anomalies to be discussed below, things are more complicated and the relevant characteristic classes are the Pontryagin classes that combine single trace and multiple trace classes in a well-defined way, as fixed by the index theorems.

⁵⁹One could also add gauge invariant functionals of the matter fields, but this is irrelevant here. The point is that any term that involves the h field or an equal number of ghost and anti-ghost fields necessarily is BRST-exact.

so that sQ_{2n}^1 is d-closed on the auxiliary $d + 2$ -dimensional space-time M . Now, even though the d -dimensional space-time may be topologically non-trivial and require several charts, as discussed before, we may assume that the ghost field is non-vanishing only within a single chart so that Q_{2n}^1 , as well as sQ_{2n}^1 are non-vanishing only within a single chart, and being d -closed then is d -exact. We conclude that $sQ_{2n}^1 = d\beta_{2n-1}^2$ and then integrating this $d = 2n$ form over the d -dimensional space-time $\mathcal{M}$ one gets zero :

$$s \int_{\mathcal{M}} Q_{2n}^1 = \int_{\mathcal{M}} d\beta_{2n-1}^2 = 0 , \quad (9.33)$$

and we have shown that $\int_{\mathcal{M}} Q_{2n}^1$ is BRST-closed, meaning that it satisfies the Wess-Zumino consistency condition.

We must still show that $\int_{\mathcal{M}} Q_{2n}^1$ is not BRST-exact. We begin by showing that Q_{2n}^1 is not BRST-exact. Suppose on the contrary that $Q_{2n}^1 = s\alpha_{2n}$. Then one would have $sQ_{2n+1} = dQ_{2n}^1 = ds\alpha_{2n} = sd\alpha_{2n}$ and $s(Q_{2n+1} - d\alpha_{2n}) = 0$, i.e. $Q_{2n+1} - d\alpha_{2n}$ is BRST-closed. Being a form of odd degree and ghost number 0 (where the BRST cohomology is empty), we must have $Q_{2n+1} - d\alpha_{2n} = s\gamma_{2n+1}^{-1}$. But by the same argument as above, there is no such γ_{2n+1}^{-1} , and $Q_{2n+1} - d\alpha_{2n} = 0$. This would imply $dQ_{2n+1} = 0$ which contradicts $dQ_{2n+1} = P_{n+1}$. Hence, Q_{2n}^1 is not BRST-exact. We must still show that $\int_{\mathcal{M}} Q_{2n}^1$ is not BRST-exact. Suppose on the contrary that $\int_{\mathcal{M}} Q_{2n}^1 = s \int \beta_{2n} = \int s\beta_{2n}$. This implies that $Q_{2n}^1 - s\beta_{2n} = d\gamma_{2n-1}^1$ for some γ_{2n-1}^1 . But Q_{2n}^1 was only defined modulo adding such an exact term $d\gamma_{2n-1}^1$ and then this redefined Q_{2n}^1 would be BRST-exact which cannot be as we have just shown. Then $\int_{\mathcal{M}} Q_{2n}^1$ cannot be BRST-exact, and is a non-trivial representative of the BRST-cohomology at ghost number one, which we wanted to show.

To conclude, for each characteristic class P_{n+1} there is a unique BRST-cohomology class at ghost number one, represented by $\int_{\mathcal{M}} Q_{2n}^1$, and up to a multiplicative coefficient c , this is the anomaly in $d = 2n$ dimensions. Conversely, the anomaly is unambiguously characterized by cP_{n+1} .

9.6 Gauge anomaly in $d = 2n$ dimensions

After all these formal developments let us discuss some more physical details. We are looking at the chiral anomaly, i.e. an anomaly under gauge transformations in the presence of chiral fermions, much as before, but now in an arbitrary even dimension $d = 2n$. As discussed in the previous section, this anomaly first manifests itself in the non-conservation an $n + 1$ -point one-loop function $\Gamma_{\alpha_1 \dots \alpha_{n+1}}^{\mu\nu_1 \dots \nu_n}$. One checks for conservation at the vertex with index μ . This involves $n!$ individual diagrams corresponding to the $n!$ permutations of the indices of the n other vertices. Each diagram gives rise to a trace $\text{tr}_{\mathcal{R}} t_{\alpha_1} \dots t_{\alpha_{n+1}}$, and adding all other permutations (together with the cyclicity of the trace) results in a totally symmetrized trace which is the corresponding D -symbol

$$D_{\alpha_1 \dots \alpha_{n+1}}^{\mathcal{R}} = \text{str}_{\mathcal{R}} t_{\alpha_1} \dots t_{\alpha_{n+1}} , \quad (9.34)$$

where the symmetrized trace is taken in the representation $\mathcal{R}$ of the chiral fermion. Exactly the same symmetrized trace emerges from the characteristic class, cf. (9.16)

$$P_{n+1}^{\mathcal{R}}(F) = F^{\alpha_1} \wedge \dots \wedge F^{\alpha_{n+1}} D_{\alpha_1 \dots \alpha_{n+1}}^{\mathcal{R}} . \quad (9.35)$$

More precisely, we had argued above that the anomalous contribution of the $n + 1$ point one-loop diagram had to be of the form $\epsilon^{\mu_1 \dots \mu_n \nu_1 \dots \nu_n} p_{\mu_1}^1 \dots p_{\mu_n}^n D_{\alpha_1 \dots \alpha_{n+1}}^{\mathcal{R}}$ which, upon multiplying with $\omega^{\alpha_{n+1}} A_{\mu_1}^{\alpha_1} \dots A_{\mu_n}^{\alpha_n}$ corresponds to a term in the anomaly proportional to

$$\text{str}_{\mathcal{R}} \omega \underbrace{\text{d}A \dots \text{d}A}_{n \text{ times}} . \quad (9.36)$$

But this is exactly the first term of Q_{2n}^1 as obtained by descent from $P_{n+1}^{\mathcal{R}} = \text{str}_{\mathcal{R}} F^{n+1}$. More precisely, we should write $Q_{2n}^{1, \mathcal{R}}$ to indicate in which representation the trace has to be taken. We then know from the previous discussion about solving the Wess-Zumino consistency condition and the BRST-cohomology at ghost number one, that the gauge anomaly in $d = 2n$ dimensions is

$$\mathcal{A}[\omega, A] = c_{2n} \int \left(\sum_j Q_{2n}^{1, \mathcal{R}_j^L} - \sum_i Q_{2n}^{1, \mathcal{R}_i^R} \right) , \quad (9.37)$$

where the first sum is over the different left-handed (positive chirality) fermions in representations $\mathcal{R}_i^L$ and the second sum over the different right-handed (negative chirality) fermions in representations $\mathcal{R}_i^R$. The coefficient c_{2n} can be determined by a careful generalization of our triangle diagram computation. One finds

$$c_{2n} = \frac{(-)^{n+1}}{(n+1)!(2\pi)^n} , \quad (9.38)$$

which for $n = 2$, i.e. $d = 4$ gives $c_4 = -\frac{1}{24\pi^2}$, in agreement with the coefficient in (8.15).

Above, we have discussed the issue of anomaly cancellation and the fact that this cancellation only involved the D -symbols for the different representations of the different chiral classes. We now see that the issue of anomaly cancellation can be equivalently discussed at the level of the characteristic classes computed in the representations of the different chiral fermions.

10 Gauge and gravitational anomalies and index theorems

In this last section, we want to briefly touch upon some further developments. Firstly, we will mention - but not prove - the relation between anomalies in $d = 2n$ dimensions and the index of an appropriate Dirac operator in $d + 2$ dimensions. Then we will briefly show how our discussion of gauge anomalies can be straightforwardly generalized to gravitational anomalies. Lastly, we discuss a prominent example of anomaly cancellation in a 10-dimensional supergravity theory.

10.1 Anomalies and index theorems

We have already seen in subsection 6.2.3 that the abelian anomaly in four dimensions was closely related to the index of the four-dimensional Dirac operator, and both were related in turn to the characteristic class $P_2(F) = \text{tr}_{\mathcal{R}} F \wedge F = F^\alpha \wedge F^\beta \text{tr}_{\mathcal{R}} t_\alpha t_\beta = c_{\mathcal{R}} F^\alpha \wedge F^\alpha$. This is very different from what we observed for the chiral anomaly under non-abelian gauge transformations where the anomaly is given by Q_4^1 which is related by descent to the characteristic class $P_3(F) = \text{tr} F^3 = F^\alpha \wedge F^\beta \wedge F^\gamma D_{\alpha\beta\gamma}^{\mathcal{R}}$ and which instead involves the D -symbol. There is, however, an index theorem in each even dimension, and the six-dimensional index theorem states that the index of an appropriate 6-dimensional Dirac operator equals some constant times $P_3(F)$. It is then tempting to speculate whether the chiral anomaly in 4 dimensions can be directly understood in terms of the index of some 6-dimensional extension of the 4-dimensional Dirac operator. More generally, the index theorem in $2n + 2$ (euclidean) dimensions states

$$\text{ind} (i\mathcal{D}_{2n+2}(A)) = \frac{c_{2n}}{2\pi} \int P_{n+1}(F) , \quad (10.1)$$

with the same coefficient c_{2n} as given in (9.38). The overall sign depends on the precise convention used to define the chirality γ -matrix. One sees that the anomaly as given in (9.37) is precisely obtained, up to a factor of 2π , by descent from the index of the $2n + 2$ dimensional Dirac operator.

While this is a very suggestive relation, it was the achievement of Alvarez-Gaumé and Ginsparg [18] to precisely construct a $2n + 2$ dimensional gauge-field and Dirac operator and show that the anomaly is indeed given by descent from the corresponding index. Here we will not spell out the details of their reasoning but only give a brief sketch.

The starting point is the observation that in euclidean signature the anomaly resides in the imaginary part of the euclidean effective action $\Gamma_E[A]$, i.e. in the phase of $e^{-\Gamma_E[A]}$. It is indeed the ambiguity in defining the phase of the ill-defined determinant $\text{Det} (i\mathcal{D}(A)P_L)$ which can be viewed as being at the origin of the anomaly. More precisely, one looks at the determinant of $\widehat{D}(A) = \mathcal{D}(A)P_L + \not{D}P_R$. Note that the euclidean (effective) action Γ_E and its continuation to Minkowski signature Γ_M differ by a factor of i as

$$\Gamma_E = -i\Gamma_M . \quad (10.2)$$

It is not difficult to see that the only terms in a Minkowski signature action that will become imaginary upon continuing to Euclidean signature are the terms involving the ϵ -tensor. This is very consistent with the fact that the anomaly resides in the imaginary part of the Euclidean effective

action ! Indeed, the terms involving the ϵ -tensor are integrals of d -forms over the d -dimensional manifold, $\int dx^0 \wedge \dots dx^{d-1} \xi_{0\dots(d-1)}$. Upon the continuation, both dx^0 and $\xi_{0\dots(d-1)}$ get a factor of $\pm i$, so overall this term remains real. But it also gets an i from the relation (10.2) and thus results in an imaginary contribution to the Euclidean action. On the other hand, in a term like $\int dx^0 dx^{d-1} x \sqrt{-g} \zeta_{\mu_1 \dots \mu_r} \zeta^{\mu_1 \dots \mu_r}$, the integrand remains real, but one gets an i from continuing dx^0 . So this term picks up an i , which together with the i in (10.2) results in a real contribution to the Euclidean action.

If the phase of the determinant can be unambiguously defined in a gauge-invariant way, then there is no anomaly. If not, an anomaly is present. Then [18] introduces a family of gauge transformations $g(x, \theta)$ depending on some parameter $\theta \in [0, 2\pi]$ such that $g(x, 0) = g(x, 2\pi) = \mathbf{1}$ and, denoting

$$A^\theta(x, \theta) = g(x, \theta)^{-1} A(x) g(x, \theta) - i g(x, \theta)^{-1} dg(x, \theta) \quad (10.3)$$

the corresponding gauge transformed field, they study how the phase of $e^{-\Gamma_E[A^\theta]}$ changes as θ goes from 0 to 2π . Since for both $\theta = 0$ and $\theta = 2\pi$ one has the original gauge field, this phase change must be $2\pi m$ with m being an integer. This integer m then is related to the anomaly as ⁶⁰ ($\hat{\epsilon} = -ig(\theta)^{-1} \frac{\partial g(\theta)}{\partial \theta}$)

$$\begin{aligned} m &= -\frac{1}{2\pi} \int_0^{2\pi} d\theta \frac{\partial \text{Im} \Gamma_E[A^\theta]}{\partial \theta} = -\frac{1}{2\pi i} \int_0^{2\pi} d\theta \frac{\partial \Gamma_E[A^\theta]}{\partial \theta} \\ &= -\frac{1}{2\pi i} \int_0^{2\pi} d\theta \int_{\mathcal{M}} d^{2n} x \sqrt{\det g} (D_\mu^\theta \hat{\epsilon}(x, \theta))^a \frac{\delta \Gamma_E[A^\theta]}{\delta A_\mu^{\theta, a}(x)} \\ &= -\frac{1}{2\pi i} \int_0^{2\pi} d\theta \delta_{\hat{\epsilon}} \Gamma_E[A^\theta] = \frac{1}{2\pi} \int_0^{2\pi} d\theta \delta_{\hat{\epsilon}} \Gamma_M[A^\theta] . \end{aligned} \quad (10.4)$$

The next step is to relate the integer m to the index of an appropriate Dirac operator. To this end, [18] extends A^θ to a two-parameter family of gauge fields $A^{\rho, \theta}(x)$ with (ρ, θ) parametrizing a disc ($\rho \leq 1$) and $A^{\rho=1, \theta}(x) = A^\theta(x)$. The integer m is then interpreted as some sort of winding number around the boundary of the disc. Then, in [18] it is shown that this winding number is the sum of local winding numbers around points in the interior of the disc where $\text{Det}(i\hat{D}(A^{\rho, \theta}))$ vanishes. Of course this happens whenever $\hat{D}(A^{\rho, \theta})$ has a zero-mode. The non-trivial point is that the local winding numbers are ± 1 according to the chirality of the zero-modes of a full $2n + 2$ dimensional Dirac operator $i\mathcal{D}_{2n+2}$, and one sees that this computes precisely the index of this operator. One concludes that

$$\text{ind}(i\mathcal{D}_{2n+2}(A^{\rho, \theta})) = m = \frac{1}{2\pi} \int_0^{2\pi} d\theta \delta_{\hat{\epsilon}} \Gamma_M[A^\theta] . \quad (10.5)$$

One still needs to specify how $A^{\rho, \theta}$ is determined, and how it is promoted to a true $2n + 2$ form, and what exactly is the $2n + 2$ dimensional manifold. This is done such that the manifold is $S^2 \times \mathcal{M}$, with $S^2 = S^2_+ \cup S^2_-$, and the previous disc is identified with S^2_+ . Then we have $A^{\rho, \theta} = A_+$ on $S^2_+ \times \mathcal{M}$, while on $S^2_- \times \mathcal{M}$ the gauge field $A^{\rho, \theta} = A_-$ is not extended (i.e. is missing the $d\rho$ and $d\theta$ -pieces).

⁶⁰We have $A_\mu^\theta = g(\theta)^{-1} A_\mu g(\theta) - ig(\theta)^{-1} \partial_\mu g(\theta)$ and under an infinitesimal variation of θ one has $\frac{\delta A^\theta}{\delta \theta} = \partial_\mu \frac{\partial \epsilon}{\partial \theta} - i[A_\theta, \frac{\partial \epsilon}{\partial \theta}] = D_\mu^\theta \frac{\partial \epsilon}{\partial \theta} = D_\mu^\theta \hat{\epsilon}$ where $\hat{\epsilon} = \frac{\partial \epsilon}{\partial \theta} = -ig(\theta)^{-1} \frac{\partial g(\theta)}{\partial \theta}$.

Specifically, one lets

$$A_+(x, \rho, \theta) = f(\rho) \left(A^\theta(x, \theta) - ig(x, \theta)^{-1} \partial_\theta g(x, \theta) d\theta \right) \quad , \quad A_-(x) = A(x) \quad , \quad (10.6)$$

with A^θ defined in (10.3) and $f(\rho)$ is a smooth function which is $f(\rho) \sim \rho$ for small ρ and $f(\rho) = 1$ for ρ close to 1 which corresponds to the fattened boundary of the disk. Then, on the overlap of $S_+^2 \times \mathcal{M}$ with $S_-^2 \times \mathcal{M}$, the A_+ and A_- are related by the $d + 1$ -dimensional gauge transformation $g(x, \theta)$. The “true” space manifold is taken to be compactified $\mathbb{R}^{2n} \simeq S^{2n}$. Since A_- is missing the $d\theta$ -piece, we have $Q_{2n+1}(A_-) = 0$. Then by the index theorem ⁶¹

$$\begin{aligned} \text{ind} (i\mathcal{D}_{2n+2}(A^{\rho, \theta})) &= \frac{c_{2n}}{2\pi} \int_{S^2 \times S^{2n}} \text{tr} F^{n+1} = \frac{c_{2n}}{2\pi} \left(\int_{S_+^2 \times S^{2n}} \text{tr} F^{n+1} + \int_{S_-^2 \times S^{2n}} \text{tr} F^{n+1} \right) \\ &= \frac{c_{2n}}{2\pi} \int_{S_+^2 \times S^{2n}} dQ_{2n+1}(A_+) = \frac{c_{2n}}{2\pi} \int_{S^1 \times S^{2n}} Q_{2n+1}(A_+) \quad . \end{aligned} \quad (10.7)$$

It is then shown [18] that the last integral equals

$$\text{ind} (i\mathcal{D}_{2n+2}(A^{\rho, \theta})) = \frac{c_{2n}}{2\pi} \int_0^{2\pi} d\theta \int_{S^{2n}} Q_{2n}^1(\hat{\epsilon}, A^\theta) \quad . \quad (10.8)$$

Comparing with (10.5) we see that

$$\int_0^{2\pi} d\theta \delta_{\hat{\epsilon}} \Gamma_M[A^\theta] = c_{2n} \int_0^{2\pi} d\theta \int_{S^{2n}} Q_{2n}^1(\hat{\epsilon}, A^\theta) \quad . \quad (10.9)$$

The θ dependence of $\hat{\epsilon}(x, \theta)$ is fairly arbitrary, and one concludes on the equality of the “integrand” at any θ and any $\hat{\epsilon}$. We may then fix some θ and simply rename $A^\theta(x) \rightarrow A(x)$ and $\hat{\epsilon}(x, \theta) \rightarrow \epsilon(x)$. The final result gives the anomalous variation of the Minkowskian effective action for a positive chirality (left-handed) chiral spin- $\frac{1}{2}$ fermion as

$$\mathcal{A}[\epsilon, A] = \delta_\epsilon \Gamma_M[A^\theta] = c_{2n} \int_{S^{2n}} Q_{2n}^1(\epsilon, A) \quad , \quad (10.10)$$

in perfect agreement with the above result (9.37).

In view of the generalization to include also gravitational anomalies to be discussed next, it is useful to formalize the relation just established as follows :

$$\text{ind} (i\mathcal{D}_{2n+2}(A^{\rho, \theta})) = \frac{1}{2\pi} \int I_{2n+2} \quad , \quad I_{2n+2} = d J_{2n+1} \quad , \quad \delta J_{2n+1} = d K_{2n}^1 \quad , \quad \mathcal{A}[\epsilon, A] = \int_{\mathcal{M}} K_{2n}^1 \quad , \quad (10.11)$$

where for the pure chiral gauge anomalies $I_{2n+2} = c_{2n} \text{tr} F^{n+1}$, $J_{2n+1} = c_{2n} Q_{2n+1}$ and $K_{2n}^1 = c_{2n} Q_{2n}^1$.

⁶¹On a general curved manifold M the index theorem involves an extra factor $\hat{A}(M)$ to be discussed below, but $\hat{A}(S^2 \times S^{2n}) = 1$, so this is irrelevant here.

10.2 Gravitational anomalies

Theories involving gravity bear some similarities with non-abelian gauge theories, but are rather more complicated. With the basic dynamical field being the space-time metric $g_{\mu\nu}$, the role of gauge-invariance is played by diffeomorphism invariance. Under an infinitesimal diffeomorphism $x^\mu \rightarrow x^\mu + \epsilon^\mu(x^\rho)$ the metric transforms as $\delta g_{\mu\nu} = -\nabla_\mu \epsilon_\nu - \nabla_\nu \epsilon_\mu$, where ∇ is the covariant derivative. This diffeomorphism invariance must be “gauge-fixed” when trying to do the gravitational functional integral. But again, one can do the functional integral in two steps, integrating over the matter fields first, resulting in an effective action $\Gamma[g_{\mu\nu}] \equiv \Gamma[g]$, with all the complications of gauge-fixing and ghosts deferred to later. The question of anomalies then is again whether this effective action $\Gamma[g]$ is still diffeomorphism invariant :

$$\delta_\epsilon \Gamma[g] = - \int (\nabla_\mu \epsilon_\nu + \nabla_\nu \epsilon_\mu) \frac{\delta \Gamma[g]}{\delta g_{\mu\nu}} . \quad (10.12)$$

The (matter) energy-momentum tensor is defined as

$$\delta_g S_{\text{matter}} = \frac{1}{2} \int \sqrt{g} T^{\mu\nu} \delta g_{\mu\nu} \quad \Rightarrow \quad T^{\mu\nu} = \frac{2}{\sqrt{g}} \frac{\delta S_{\text{matter}}}{\delta g_{\mu\nu}} . \quad (10.13)$$

Since $e^{i\Gamma[g]} = \int \mathcal{D}\psi e^{iS[g,\psi]}$ where ψ stands collectively for all matter fields, we have, much as for the gauge currents J_α^μ in the non-abelian gauge theories,

$$\frac{\delta \Gamma[g]}{\delta g_{\mu\nu}(x)} = \frac{1}{2} \sqrt{g} \langle T^{\mu\nu}(x) \rangle \quad (10.14)$$

Inserting this into (10.12) and integrating by parts ⁶² results in

$$\mathcal{A}_{\text{diff}}[\epsilon, g] \equiv \delta_\epsilon \Gamma[g] = \int \sqrt{g} \epsilon_\nu \nabla_\mu \langle T^{\mu\nu}(x) \rangle , \quad (10.15)$$

where we defined, much as before, the anomaly under diffeomorphisms as the non-invariance of the effective action. We see that diffeomorphism invariance of the matter effective action is equivalent to the covariant conservation of the expectation value of the energy-momentum tensor. Conversely, a gravitational anomaly, i.e. an anomaly under diffeomorphism invariance, shows up as the covariant non-conservation of this expectation value. Just as with the non-abelian gauge theories, one can take further functional derivatives with respect to the metric at different space-time points and generate, possibly anomalous, Ward identities for m -point functions of the energy-momentum tensor.

To couple spin- $\frac{1}{2}$ fields (or spin- $\frac{3}{2}$ fields) to gravity, one needs to introduce local orthonormal frames, or local Lorentz frames. This means that for each tangent space (at each point of the manifold) one makes a certain choice for an orthonormal basis of tangent vectors E_a with components E_a^μ such that $E_a^\mu E_b^\nu g_{\mu\nu} = \delta_{ab}$. (We will use Euclidean signature here. For Minkowski signature one would of course have $E_a^\mu E_b^\nu g_{\mu\nu} = \eta_{ab}$ instead.) Equivalently, the components e_μ^a of the dual forms $e^a = e_\mu^a dx^\mu$ (also called the viel-bein) obey

$$e_\mu^a E_b^\mu = \delta_b^a \quad , \quad e_\mu^a E_a^\nu = \delta_\mu^\nu \quad , \quad \delta_{ab} e_\mu^a e_\nu^b = g_{\mu\nu} . \quad (10.16)$$

⁶²Provided there are no boundary terms, covariant derivatives can be freely integrated by parts, e.g. $\int \sqrt{g} (\nabla_\mu a_\nu) b^{\mu\nu} = - \int \sqrt{g} a_\nu \nabla_\mu b^{\mu\nu}$, where, of course, $\nabla_\mu a_\nu = \partial_\mu a_\nu - \Gamma_{\mu\nu}^\rho a_\rho$ and $\nabla_\mu b^{\mu\nu} = \partial_\mu b^{\mu\nu} + \Gamma_{\mu\rho}^\mu b^{\rho\nu} + \Gamma_{\mu\rho}^\nu b^{\mu\rho}$.

Obviously, at each point, there is not a unique orthonormal frame, but a whole family related by orthogonal or pseudo-orthogonal (Lorentz) transformations $e'^a = (L^{-1})^a_b e^b$. One talks about local Lorentz transformations. One imposes however, that the frames change in a smooth way as one moves along the manifold, and then the local Lorentz transformations one is allowed to do must also depend smoothly on the point.

One can use the E^μ_a and e^a_μ to convert coordinate tensors like T^ν_μ into frame tensors as $T^b_a = E^\mu_a e^b_\nu T^\nu_\mu$, etc. One can then show that invariance under diffeomorphisms and invariance under local Lorentz transformations are essentially equivalent. This means that, as far as anomalies are concerned, we may probe for anomalies under local Lorentz transformations [29]. From the one-forms e^a one defines the so-called spin-connection one-form $\omega^{ab} = -\omega^{ba} = \omega^{ab}_\mu dx^\mu$ by requiring the absence of torsion :

$$de^a + \omega^{ab} \wedge e^b = 0 \quad , \quad \omega^{ab} = -\omega^{ba} . \quad (10.17)$$

Also, the curvature two-form $R^{ab} = \frac{1}{2} R^{ab}_{\mu\nu} dx^\mu \wedge dx^\nu$, where $R^{ab}_{\mu\nu} = e^a_\rho e^b_\sigma R^{\rho\sigma}_{\mu\nu}$ and $R^{\rho\sigma}_{\mu\nu}$ is the Riemann curvature tensor, is obtained from the spin-connection one-form as

$$R^{ab} = d\omega^{ab} + \omega^{ac} \wedge \omega^{cb} . \quad (10.18)$$

It will be useful to think of ω and R as matrix-valued one and two-forms and suppress the indices a, b , so that the previous relation simply reads $R = d\omega + \omega \wedge \omega \equiv d\omega + \omega^2$.

This spin connection one-form and curvature 2-form behave under local Lorentz transformations $e'^a = (L^{-1})^a_b e^b$ very much as the gauge field one-form A and gauge field strength 2-form F under gauge transformations :

$$\omega' = L^{-1} \omega L + L^{-1} dL \quad , \quad R' = L^{-1} R L , \quad (10.19)$$

or, for infinitesimal local Lorentz transformations $L = e^v$ with $v^{ab} = -v^{ba}$,

$$\delta_v \omega = dv + [\omega, v] \quad , \quad \delta_v R = [R, v] . \quad (10.20)$$

In particular, the whole discussion of characteristic classes and descent equations for the gauge theories can be transposed word by word to the gravitational case when formulated with the spin connection and curvature 2-form. While this is true, there is however an important restriction. Just as certain D -symbols vanish, depending on the representations, here, in $d = 2n$ dimensions we are dealing with the vector representation $\mathcal{R}_v$ of $SO(2n)$. This is a real representation. Indeed, One can choose the generators $J^{ab} = -J^{ba}$ of the vector representation to be hermitian, imaginary and antisymmetric :

$$(J^{ab}_v)_{cd} = i(\delta^a_c \delta^b_d - \delta^b_c \delta^a_d) . \quad (10.21)$$

Then the generators of the complex conjugate representation are $-(J^{ab})^T = J^{ab}$ and this is a real representation. As discussed before, for real representations the corresponding D -symbols with $n+1$ indices vanish whenever n is even. Said differently, the characteristic class $\text{tr}_{\mathcal{R}_v} R^{n+1} = R^{a_1 a_2} \wedge R^{a_2 a_3} \dots \wedge R^{a_{n+1} a_1}$ vanishes for even n by the antisymmetry of $R^{ab} = -R^{ba}$.

Now, the anomaly, if any, is due to chiral fermions and these fermions are not in the vector representation but in a spin representation of $SO(2n)$ with generators $J^{ab}_{\text{spin}} = \frac{i}{2} \gamma^{ab} = \frac{i}{4} (\gamma^a \gamma^b - \gamma^b \gamma^a)$. Then $R = \frac{1}{2} R^{ab} J^{ab}_{\text{spin}} = \frac{i}{4} R^{ab} \gamma^{ab} = \frac{i}{4} R^{ab} \gamma^a \gamma^b$, and the characteristic class $\text{tr}_{\text{spin}} R^{n+1}$ involves the trace

of $2n + 2$ matrices γ^{a_i} . This trace is the sum of products of $n + 1$ Kronecker deltas, with all possible pairings of the indices, times the dimension of the gamma-matrices. Those pairings which contract two indices on the same R drop out. For example, for $n = 1$, one has

$$\begin{aligned} \text{tr}_{\text{spin}} R \wedge R &= -\frac{1}{16} R^{ab} \wedge R^{cd} \text{tr}_{\text{spin}} \gamma^a \gamma^b \gamma^c \gamma^d = -\frac{1}{4} R^{ab} \wedge R^{cd} (\delta^{ab} \delta^{cd} - \delta^{ac} \delta^{bd} + \delta^{ad} \delta^{bc}) \\ &= -\frac{1}{4} (-R^{ab} \wedge R^{ab} + R^{ab} \wedge R^{ba}) = -\frac{1}{2} R^{ab} \wedge R^{ba} , \end{aligned} \quad (10.22)$$

which, up to the factor $-\frac{1}{2}$, is the same result as in the vector representation. This is easily seen to generalize as

$$\text{tr}_{\text{spin}} R^{2k} = \tilde{c}_k R^{a_1 a_2} \wedge R^{a_2 a_3} \wedge \dots \wedge R^{a_{2k} a_1} , \quad \text{tr}_{\text{spin}} R^{2k+1} = 0 , \quad (10.23)$$

with some combinatorial factor $\tilde{c}_k$. This being clarified we can now immediately state the gravitational descent equations :

$$P_{2k}(R) = \text{tr} R^{2k} , \quad P_{2k}(R) = dQ_{4k-1}^L(\omega) , \quad \delta_v Q_{4k-1}^L(\omega) = dQ_{4k-2}^{L,1}(v, \omega) , \quad k = 1, 2, \dots \quad (10.24)$$

where we added a superscript “L” for local Lorentz. Since the only non-vanishing characteristic classes are the $P_{2k}(R)$, the $Q_d^{L,1}$ are non-vanishing only for $d = 4k - 2$, i.e. $d = 2, 6, 10, \dots$. So only in these dimensions can we expect to have purely gravitational anomalies. We will also encounter characteristic classes which are (wedge) products of several traces, like e.g. $P_{2k,2l} = \text{tr} R^{2k} \wedge \text{tr} R^{2l}$. Then one can do the descent in one or the other factor, i.e. $\text{tr} R^{2k} \wedge Q_{4l-1}^L$ or $Q_{4k-1}^L \wedge \text{tr} R^{2l}$. Of course, both are equally good, since their difference is an exact form, namely $d\alpha_{4k+4l-2}$ with $\alpha_{4k+4l-2} = Q_{4k-1}^L \wedge Q_{4l-1}^L$. Continuing the descent, one gets either $\text{tr} R^{2k} \wedge Q_{4l-2}^{L,1}$ or $Q_{4k-2}^{L,1} \wedge \text{tr} R^{2l}$. Again they only differ by a term $d\beta_{4k+4l-3}^1 + \delta_v \gamma_{4k+4l-2}$, cf. the analogous discussion at the end of section 9.4.2.

One can now continue the discussion of gravitational anomalies in two directions. On the one hand, one can compute the anomalous part of one-loop diagrams with chiral fermions coupled to gravity, i.e. to metric perturbations $h_{\mu\nu} = g_{\mu\nu} - \delta_{\mu\nu}$ around flat space. Since metric perturbations couple to the energy-momentum tensor, this amounts to computing $(m + 1)$ -point one-loop functions with $m + 1$ energy-momentum tensors at the vertices, and check for conservation, cf. the Ward identities one could derive from (10.15). Again, one needs to saturate the $d = 4k - 2$ -dimensional ϵ -tensor with at least $2k - 1$ independent momenta and as many symmetric polarisation tensors. This means that $m \geq 2k - 1$. Thus in 2 dimensions ($k = 1$) we need to compute a 2-point function, in 6 dimensions ($k = 2$) a box diagram and in 10 dimensions ($k = 3$) a hexagon. Of course, this is the same counting as for the gauge anomalies. These computations have been done in [6] where, among others, the full two-point function in 2 dimensions is evaluated very explicitly, and the anomalous part of the hexagon diagram is determined - which is quite a *tour de force*.

On the other hand, one can rely again on the relation between the anomaly and the index theorem. The Dirac operator now includes a coupling to the spin connection, as well as possibly to (non-abelian) gauge fields ⁶³:

$$\mathcal{D} = \gamma^c E_c^\mu \left(\partial_\mu - i A_\mu^\alpha t_\alpha + \frac{1}{4} \omega_\mu^{ab} \gamma^{ab} \right) . \quad (10.25)$$

⁶³Note that $\frac{1}{4} \omega_\mu^{ab} \gamma^{ab} = -\frac{i}{2} \omega_\mu^{ab} J_{\text{spin}}^{ab}$, so that this term has exactly the same structure as the gauge term $-i A_\mu^\alpha t_\alpha$.

The relevant index theorem is a generalization of (10.1) to curved manifolds :

$$\text{ind} (i\mathcal{D}_{2n+2}(A)) = \int [\widehat{A}(R)\text{ch}(iF)]_{2n+2} , \quad (10.26)$$

where the Chern character $\text{ch}(iF)$ and the Dirac genus $\widehat{A}(R)$ are the formal sums of forms of different degrees, defined as

$$\text{ch}(iF) = \text{tr} \exp \left(-\frac{1}{2\pi} F \right) = \text{tr} \mathbf{1} - \frac{1}{2\pi} \text{tr} F + \frac{1}{8\pi^2} \text{tr} F^2 - \frac{1}{48\pi^3} \text{tr} F^3 + \dots = \text{tr} \mathbf{1} + \sum_{n \geq 0} \frac{c_{2n}}{2\pi} \text{tr} F^{n+1} , \quad (10.27)$$

where $\text{tr} \mathbf{1} = \dim(\mathcal{R})$ is just the dimension of the fermion representation, and

$$\begin{aligned} \widehat{A}(R) = & 1 + \frac{1}{(4\pi)^2} \frac{1}{12} \text{tr} R^2 + \frac{1}{(4\pi)^4} \left(\frac{1}{360} \text{tr} R^4 + \frac{1}{288} (\text{tr} R^2)^2 \right) \\ & + \frac{1}{(4\pi)^6} \left(\frac{1}{5670} \text{tr} R^6 + \frac{1}{4320} \text{tr} R^4 \text{tr} R^2 + \frac{1}{10368} (\text{tr} R^2)^3 \right) + \dots , \end{aligned} \quad (10.28)$$

and the subscript $[\dots]_{2n+2}$ in (10.26) indicates that one must extract the form of degree $2n+2$. We can then proceed exactly as formalized in (10.11), where here

$$I_{2n+2} = 2\pi [\widehat{A}(R)\text{ch}(iF)]_{2n+2} . \quad (10.29)$$

For example, in $d = 2n = 4$ dimensions one has

$$I_6 = 2\pi \left(-\frac{1}{48\pi^3} \text{tr} F^3 + \frac{1}{(4\pi)^2} \frac{1}{12} \text{tr} R^2 \left(-\frac{1}{2\pi} \text{tr} F \right) \right) . \quad (10.30)$$

The first term corresponds to the pure gauge anomaly extensively discussed before, while the second term corresponds to a so-called mixed gauge-gravitational anomaly. For any simple group the trace of a single generator vanishes and, hence, $\text{tr} F$ is non-zero only for an abelian $U(1)$ gauge group where it is proportional to the sum of all $U(1)$ charges. Hence, in a non-trivial gravitational background, the vanishing of this anomaly requires this sum to vanish. This is indeed the case for the hypercharges of any family in the standard model. We also see (again) that there is no pure gravitational anomaly in 4 dimensions.

10.3 Anomaly cancellation in ten-dimensional IIB supergravity

Supergravity theories contain the metric / viel-bein and its superpartner, the gravitino which is a spin- $\frac{3}{2}$ field which can be either chiral or real (Majorana), or both. In so-called extended supergravity theories there are two (or more) such gravitini, and they can have the same or different chiralities. In addition, anti-symmetric tensor fields are present, and if they satisfy a certain self-duality constraint they are also chiral fields. Indeed such fields can be constructed from a pair of chiral spinors. Finally, there can also be chiral spin- $\frac{1}{2}$ fields, as well as various (non-chiral) scalar and vector fields.

One of the most prominent supergravity theories is the ten-dimensional IIB supergravity which is the low-energy limit of IIB superstring theory.⁶⁴ It contains two positive chirality Majorana-Weyl

⁶⁴The “II” indicates that there are two (ten-dimensional) supersymmetries, and “B” indicates that the two supercharges are spinors of the same chirality, while in IIA they would have opposite chirality. Hence IIA supergravity is non-chiral and has no anomalies. The latter can be obtained by dimensional reduction from eleven-dimensional supergravity and, indeed, in eleven dimensions one cannot have anomalies.

gravitini, two negative chirality Majorana-Weyl spin- $\frac{1}{2}$ fields as well as one self-dual antisymmetric tensor field, in addition to the non-chiral fields. All these chiral fields contribute additively to the anomaly. It is enough to identify the corresponding characteristic classes in 12 dimensions that are given for each chiral field (up to the factor 2π) by the corresponding index. There is no gauge group here and, hence, $\text{ch}(iF) = 1$. We already know the index for the complex, chiral spin- $\frac{1}{2}$ field. Since here this is a real (Majorana) chiral field, one just needs to divide the previous result (10.26) by 2.

A spin- $\frac{3}{2}$ field can be obtained as a spin- $\frac{1}{2}$ field with an extra vector index, and subtracting the spin- $\frac{1}{2}$ part. The extra vector index can be treated as an $SO(d)$ gauge symmetry in the vector representation, resulting in a contribution to the index of ⁶⁵ $\text{tr}_{\text{vect}} \exp \left(\frac{1}{2\pi} \frac{1}{2} R^{ab} J_{\text{vect}}^{ab} \right) = \text{tr} \exp \left(\frac{i}{2\pi} R \right)$, so that the index is

$$\text{ind}(iD_{2n+2}^{3/2}) = \int \left[\hat{A}(R) \left(\text{tr} \exp \left(\frac{i}{2\pi} R \right) - 1 \right) \right]_{2n+2}, \quad (10.31)$$

where the -1 simply subtracts the spin- $\frac{1}{2}$ contribution. If such a spin- $\frac{3}{2}$ field would also couple to a gauge field, one would need to include a factor $\text{ch}(iF)$, but this is not the case here. Moreover, if it is Majorana-Weyl, one needs again to divide by 2. Note that $\text{tr} 1 = \delta_a^a = d = 2n$ just gives the dimension of the vector representation which is, of course, the space (-time) dimension. It follows that

$$\text{tr} \exp \left(\frac{i}{2\pi} R \right) - 1 = 2n - 1 - \frac{1}{2(2\pi)^2} \text{tr} R^2 + \frac{1}{24(2\pi)^4} \text{tr} R^4 - \frac{1}{720(2\pi)^6} \text{tr} R^6 + \dots \quad (10.32)$$

For the self-dual antisymmetric tensor the relevant index turns out to be

$$\text{ind}(iD_{2n+2}^A) = -\frac{1}{2} \int \left[\frac{1}{4} L(R) \right]_{2n+1}, \quad (10.33)$$

where $L(R)$ is the so-called Hirzebruch polynomial, given by

$$\begin{aligned} L(R) = & 1 - \frac{1}{(2\pi)^2} \frac{1}{6} \text{tr} R^2 + \frac{1}{(2\pi)^4} \left(-\frac{7}{180} \text{tr} R^4 + \frac{1}{72} (\text{tr} R^2)^2 \right) \\ & + \frac{1}{(2\pi)^6} \left(-\frac{31}{2835} \text{tr} R^6 + \frac{7}{1080} \text{tr} R^4 \text{tr} R^2 - \frac{1}{1296} (\text{tr} R^2)^3 \right) + \dots \end{aligned} \quad (10.34)$$

The factor $\frac{1}{4}$ in (10.33) originates as a $\frac{1}{2}$ from the reality constraint and another $\frac{1}{2}$ due to the second chirality projector. The sign reflects the fact that the antisymmetric tensor field is bosonic rather than fermionic. The extra factor $\frac{1}{2}$ takes into account the different sizes of the spinor representation between $2n+2$ and $2n$ dimensions. Let us then explicitly give the relevant 12-dimensional characteristic classes for the different anomalies in 10 dimensions.

Recall that the relevant I_{12} are given by the integrand on the rhs of the index theorems, times a factor 2π . Specifically, in ten dimensions, and in the absence of a gauge group, we have for a positive chirality Majorana spinor, positive chirality Majorana gravitino, and for a self-dual antisymmetric

⁶⁵Recall the generators (10.21) of the vector representation, so that $(R^{ab} J_{\text{vect}}^{ab})_{cd} = 2iR^{cd}$. Then e.g. $\text{tr}_{\text{vect}} (R^{ab} J_{\text{vect}}^{ab})^2 = (2i)^2 R^{cd} R^{dc}$, where $R^{ef} = \frac{1}{2} R_{\mu\nu}^{ef} dx^\mu \wedge dx^\nu$.

$$\begin{aligned}
I_{12}^{1/2, \text{Maj}} &= 2\pi \times \frac{1}{2} \times \frac{1}{(4\pi)^6} \left(\frac{1}{5670} \text{tr } R^6 + \frac{1}{4320} \text{tr } R^4 \text{tr } R^2 + \frac{1}{10368} (\text{tr } R^2)^3 \right) \\
I_{12}^{\text{gravitino}} &= 2\pi \times \frac{1}{2} \times \frac{1}{(4\pi)^6} \left(-\frac{11}{126} \text{tr } R^6 + \frac{5}{96} \text{tr } R^4 \text{tr } R^2 - \frac{7}{1152} (\text{tr } R^2)^3 \right) \\
I_{12}^{\text{antisym}} &= 2\pi \times \frac{1}{8(2\pi)^6} \left(\frac{31}{2835} \text{tr } R^6 - \frac{7}{1080} \text{tr } R^4 \text{tr } R^2 + \frac{1}{1296} (\text{tr } R^2)^3 \right) . \quad (10.35)
\end{aligned}$$

Then, summing the contributions from $n_{3/2}$ positive chirality Majorana gravitini, $n_{1/2}$ positive chirality Majorana spin- $\frac{1}{2}$ and n_A self-dual antisymmetric tensor fields, one gets the total anomaly polynomial

$$\begin{aligned}
I_{12}(n_{1/2}, n_{3/2}, n_A) &= \frac{1}{128(2\pi)^5} \left(\frac{n_{1/2} - 495n_{3/2} + 992n_A}{5670} \text{tr } R^6 \right. \\
&\quad \left. + \frac{n_{1/2} + 225n_{3/2} - 448n_A}{4320} \text{tr } R^4 \text{tr } R^2 + \frac{n_{1/2} - 63n_{3/2} + 128n_A}{10368} (\text{tr } R^2)^3 \right) \quad (10.36)
\end{aligned}$$

The theory has gravitational anomalies unless all 3 terms vanish. This condition constitutes a set of 3 homogeneous linear equations for the 3 unknowns $n_{1/2}$, $n_{3/2}$ and n_A which generically has no non-zero solution. However it is easy to see that here the system not only admits a solution, but the solution also is very simple, namely

$$n_A = 1 , \quad n_{3/2} = 2 , \quad n_{1/2} = -2 , \quad (10.37)$$

or any integer multiple of it. These numbers (10.37) correspond exactly to the field content of the ten-dimensional IIB supergravity : one self-dual antisymmetric tensor, two positive chirality Majorana-Weyl gravitini and two negative chirality Majorana-Weyl spin- $\frac{1}{2}$ fields!

⁶⁶Note that for the Majorana spin- $\frac{1}{2}$ without gauge group one simply has $I_{12}^{1/2, \text{Maj}} = 2\pi \frac{1}{2} \hat{A}(R)_{12}$, and that the coefficient $-\frac{11}{126}$ of the $\text{tr } R^6$ term in the bracket of $I_{12}^{\text{gravitino}}$ arises as $2\pi \frac{1}{2} \left(1 + \dots + \frac{1}{(4\pi)^6} \frac{1}{5670} \text{tr } R^6 \right) \left(10 - 1 + \dots - \frac{1}{720(2\pi)^6} \text{tr } R^6 \right) = 2\pi \frac{1}{2} \frac{1}{(4\pi)^6} \left(\frac{9}{5670} - \frac{64}{720} \right) \text{tr } R^6 + \dots = 2\pi \frac{1}{2} \frac{1}{(4\pi)^6} \left(-\frac{11}{126} \right) \text{tr } R^6 + \dots$

11 Conclusions of part II

We have studied anomalies in quantum field theory from several complementary perspectives. We have seen how classical symmetries can fail to survive quantization when regularization is required. This phenomenon manifests itself equivalently as the non-conservation of a quantum current, as anomalous Ward identities, or as the non-invariance of the functional integral measure. We have provided a clear dictionary how these different manifestations correspond to each other.

The abelian anomaly provides the simplest example. Through both Fujikawa's method to compute the Jacobian determinant in the functional integral, and through the explicit computation of the corresponding triangle diagrams, it can be seen that the anomaly is finite, local, and unavoidable. Its relation to the Atiyah–Singer index theorem in four dimensions already highlights its topological character. This abelian anomaly concerns global chiral symmetries, and only implies that certain selection rules no longer hold.

When chiral fermions are coupled to gauge fields, similar anomalies are present, but now they threaten the consistency of the theory itself. In this case, violation of gauge invariance would spoil unitarity and renormalizability, making anomaly cancellation a fundamental requirement. The remarkable fact that the anomaly is one-loop exact and constrained by the Wess-Zumino consistency conditions allows this requirement to be formulated as a purely group theoretical condition on the fermion representations.

The Standard Model offers a remarkable realization of an anomaly free theory : the assignments of quark and lepton quantum numbers are such that all gauge anomalies cancel within each fermion generation. This non-trivial cancellation, first identified by Bouchiat, Iliopoulos and Meyer, constitutes one of the most profound structural features of the theory.

On the more formal side, the Wess-Zumino consistency condition was reinterpreted as the BRST closedness of the anomaly, and then relevant anomalies correspond to BRST cohomology classes at ghost number one on the space of local functionals in d dimensions. This cohomology was shown to be uniquely characterized by formal characteristic classes in $d+2$ dimensions. This in turn hinted at a relation with index theorems in $d+2$ dimensions, and following [18] we have sketched how this allows to obtain the anomalies without doing a single loop computation, including also the generalization to gravitational anomalies. Even more amazing than the anomaly cancellation in the Standard Model is the highly non-trivial anomaly cancellation in ten-dimensional IIB supergravity.

Anomalies occupy a distinctive place in quantum field theory: they emerge from quantum effects but are governed by symmetry and topology, and they both violate classical invariance and enforce fundamental consistency conditions.

References

- [1] S.L. Adler, *Axial-vector vertex in spinor electrodynamics*, Phys. Rev. 177 (1969) 2426.
- [2] R. Jackiw and K. Johnson, *Anomalies of the axial-vector current*, Phys. Rev. 182 (1969) 1459.
- [3] J.S. Bell and R. Jackiw, *A PCAC puzzle : $\pi^0 \rightarrow \gamma\gamma$ in the σ -model*, Nuovo Cim. 60A (1969) 47.
- [4] S.L. Adler and W.A. Bardeen, *Absence of higher-order corrections in the anomalous axial-vector divergence equation*, Phys. Rev. 182 (1969) 1517.
- [5] S. L. Adler, *Anomalies to all orders*, doi:10.1142/9789812567147_0009 [arXiv:hep-th/0405040 [hep-th]].
- [6] L. Alvarez-Gaumé and E. Witten, *Gravitational anomalies*, Nucl. Phys. B234 (1983) 269.
- [7] S. Weinberg, *The quantum theory of fields*, vol 2, Cambridge University Press 1995.
- [8] S. Treiman, R. Jackiw, B Zumino and E. Witten, *Current algebra and anomalies*, Princeton University Press 1985.
- [9] L. Alvarez-Gaumé, *An introduction to anomalies*, Lectures given at Int. School on Mathematical Physics, Erice, Italy, 1985; preprint HUTP-85/A092, published in Erice School Math.Phys.1985:0093.
- [10] P. van Nieuwenhuizen, *Anomalies in quantum field theory: Cancellation of anomalies in $d = 10$ supergravity*, Leuven Notes in Math. and Theor. Phys., vol 3, Leuven University Press, 1989.
- [11] J. A. Harvey, *TASI 2003 lectures on anomalies*, [arXiv:hep-th/0509097 [hep-th]].
- [12] J. Zinn-Justin, *Chiral anomalies and topology*, Lect. Notes Phys. **659** (2005) 167 [arXiv:hep-th/0201220].
- [13] F. Bastianelli and P. van Nieuwenhuizen, *Path integrals and anomalies in curved space*, Cambridge University Press, 2006.
- [14] A. Bilal, *Lectures on anomalies*, arXiv:0802.0634 [hep-th].
- [15] J. Wess and B. Zumino, *Consequences of anomalous Ward identities*, Phys. Lett. 37B (1971) 95.
- [16] B. Zumino, *Chiral anomalies and differential geometry*, in : Relativity, Groups and Topology II (Les Houches Session XL, 1983), eds. B. S. DeWitt and R. Stora, North-Holland, 1984.
- [17] M. F. Atiyah and I. M. Singer, *The index of elliptic operators on compact manifolds*, Bull. Am. Math. Soc. **69** (1969), 422-433 ; *The Index of elliptic operators. 1*, Annals Math. **87** (1968), 484-530 ; *The Index of elliptic operators. 3*, Annals Math. **87** (1968), 546-604.
- [18] L. Alvarez-Gaumé and P. Ginsparg, *The structure of gauge and gravitational anomalies*, Ann. Phys. 161 (1985) 423, erratum-ibid 171 (1986) 233.

- [19] D. J. Gross and R. Jackiw, *Effect of anomalies on quasirenormalizable theories*, Phys. Rev. D **6** (1972), 477-493.
- [20] H. Georgi and S. L. Glashow, *Gauge theories without anomalies*, Phys. Rev. D **6** (1972), 429.
- [21] C. Bouchiat, J. Iliopoulos and Ph. Meyer, *An anomaly-free version of Weinberg's model*, Phys. Lett. **38B** (1972) 519.
- [22] M. B. Green and J. H. Schwarz, *Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory*, Phys. Lett. B **149** (1984), 117-122; *The Hexagon Gauge Anomaly in Type I Superstring Theory*, Nucl. Phys. B **255** (1985), 93-114.
- [23] A. Bilal, *Advanced quantum field theory : renormalization, non-abelian gauge theories and anomalies*, Lecture notes, Brussels 2024,
<http://www.solvayinstitutes.be/pdf/doctoral/2024/qft-M3-2024.pdf> .
- [24] K. Fujikawa, *Path Integral Measure for Gauge Invariant Fermion Theories*, Phys. Rev. Lett. **42** (1979), 1195-1198;
K. Fujikawa, *Path Integral for Gauge Theories with Fermions*, Phys. Rev. D **21** (1980), 2848
[erratum: Phys. Rev. D **22** (1980), 1499].
- [25] K. Fujikawa and H. Suzuki, *Path integrals and quantum anomalies*, Int. series of monographs on physics 122, Oxford university press 2004.
- [26] T. Eguchi, P.B. Gilkey and A.J. Hanson, *Gravitation, gauge theories and differential geometry*, Phys. Rept. **66** (1980) 213.
- [27] P.B. Gilkey, *Invariance theory, the heat equation and the Atiyah-Singer index theorem*, CRC press, 1995.
- [28] L. Alvarez-Gaume, S. Della Pietra and G. W. Moore, *Anomalies and Odd Dimensions*, Annals Phys. **163** (1985), 288.
- [29] W. A. Bardeen and B. Zumino, *Consistent and Covariant Anomalies in Gauge and Gravitational Theories*, Nucl. Phys. B **244** (1984), 421-453.