Renormalizability of Gauge Theories

For a general renormalizable QFT, the bare Lagrangian

$$\mathcal{L}_{\text{bare}} = \mathcal{L}_{\text{ren}} + \mathcal{L}^{\text{counter}}_{\text{terms}} \tag{1}$$

includes *all* operators of dimension $\Delta \leq D$ that respect the *manifest* symmetries of the theory in question. By manifest symmetries I mean the symmetries of the action which remain unbroken by the gauge-fixing, by the UV regularization, or by any other part of the quantization procedure. In perturbation theory, these symmetries are manifest in the Feynman rules (including the UV cutoff scheme), hence the name *manifest*.

Not all good symmetries of QFTs are manifest. For example, in QED the local $U(1)$ symmetry is *not manifest* because it's broken by the gauge fixing. (Which alas is necessary for constructing the photon's propagator.) At first blush, this allows for gauge-non-invariant counterterms in the bare QED Lagrangian — such as the photon mass^2 term $\delta_m^\gamma A^\mu A_\mu$ or the 4-photon coupling $\delta^{4\gamma}(A_\mu A^\mu)^2$ — which would then spoil QED's renormalizability.

Fortunately, the gauge-fixed QED still has a manifest *global* $U(1)$ symmetry and hence a manifestly conserved electric current J^μ_{el}. In the Feynman diagram language, this conserved current gives rise to the Ward–Takahashi identities which in turn eliminate all the unwanted counterterms. And that's how QED remains a renormalizable theory in 4D.

In QCD or other non-abelian gauge theories the situation is more difficult. The non-abelian gauge currents $J^{\mu a}$ are covariantly conserved,

$$D_\mu J^{\mu a} = \partial_\mu J^{\mu a} - g f^{abc} A^b_\mu J^{\mu c} = 0, \tag{2}$$

rather than straightforwardly conserved, so they do not give rise to the Ward–Takahashi–like identities. Consequently, there does not seem to be any protection against gauge-not-invariant counterterms such as gluon mass^2, ghost mass^2, or 4-gluon couplings with wrong gauge-index structures. And that's why until 1970 most physicists did not believe the non-abelian gauge theories to be renormalizable.

But in 1971/72, Martinus J. G. Veltman and his then-student Gerard 't Hooft showed that dimensionally regulated Feynman diagrams for non-abelian gauge theories obey some rather complicated identities to all loop orders, and thanks to these identities all the unwanted counterterms actually vanish. In other words, Veltman and 't Hooft proved that the non-abelian gauge theories are in fact renormalizable, and even the Higgsed-down non-abelian gauge theories are renormalizable. Historically, this proof is what made high-energy physicists of the day interested in the non-abelian gauge theories such as Glashow–Weinberg–Salam $SU(2) \times U(1)$ or QCD.

The diagrammatic proof of 't Hooft and Veltman was rather complicated, but soon afterward Andrei Slavnov and John C. Taylor came up with a more formal proof of Ward–Takahashi–like (but somewhat weaker) identities for the non-abelian gauge theory, and by 1976 Carlo Becchi, Alain Rouet, Raymond Stora, and (independently) Igor Tyutin discovered the symmetry giving rise to these identities. Thanks to this BRST symmetry, the modern proof of non-abelian gauge theories renormalizability is *much* easier than the original 't Hooft–Veltman proof.

Let me outline the modern proof without getting too deeply into technical details. For simplicity I shall focus on QCD, although very similar arguments would apply to any other kind a non-abelian gauge theory.

First of all, QCD is power-counting renormalizable, so all its UV divergences can be canceled by a finite number of counterterms, and all such counterterms are operators of dimension $\Delta \leq 4$. Moreover, the divergences — and hence the counterterms — must respect all the manifest symmetries of the gauge-fixed and UV-regulated theory. The only question is whether the bare QCD Lagrangian actually includes all the necessary counterterms: if yes then QCD is renormalizable, but if not than it isn't.

Thus, to prove the renormalizability, I'll start with a list of exact symmetries of the gauge-fixed QCD and its renormalized Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{ren}} &= \mathcal{L}_{\text{phys}} + \mathcal{L}_{\text{g.f.}} + \mathcal{L}_{\text{gh}} \\ &= -\frac{1}{4} F^a_{\mu\nu} F^{a\,\mu\nu} + \sum_f \overline{\Psi}_{if}(i\not{D} + m_f)\Psi^{if} + \frac{\xi}{2} b^a b^a - b^a \partial_\mu A^{\mu a} + \partial_\mu \bar{c}^a D^\mu c^a. \end{aligned} \tag{3}$$

Next, I'll get a UV cutoff which preserves all those symmetries, which will keep them manifest in the perturbation theory. And then I shall prove the **Theorem:** *All operators of* *dimension $\Delta \leq 4$ allowed by the by the manifest symmetries are operators already present in the renormalized QCD Lagrangian* (3).

So let's start with the list of manifest symmetries:

1. Spacetime symmetries: Lorentz, translations, and parity.

2. Global $SU(3)_{\text{color}}$ symmetry: in matrix notations,

$$\begin{aligned} \Psi(x) &\to U\Psi(x), \quad \overline{\Psi}(x) \to \overline{\Psi}(x)U^\dagger, \quad A_\mu(x) \to UA_\mu(x)U^\dagger, \\ c(x) &\to Uc(x)U^\dagger, \quad \bar{c}(x) \to U\bar{c}(x)U^\dagger, \quad b(x) \to Ub(x)U^\dagger, \end{aligned} \tag{4}$$

   all for the same $SU(3)$ matrix U at all x.

   * For a different gauge theory with gauge group G, we have the global G as a manifest symmetry.

3. Baryon number symmetry (or equivalently the quark number symmetry) and the charge conjugation.

4. Ghost number symmetry $c^a(x) \to e^{i\theta} c^a(x)$, $\bar{c}^a(x) \to e^{-i\theta}\bar{c}^a(x)$. Thanks to this symmetry, every Lagrangian terms — including the counterterms — must involve equal numbers of ghost and antighost fields.

5. Antighost shift symmetry $\bar{c}^a(x) \to \bar{c}^a(x) + \eta^a$ for constant odd Grassmann numbers η^a. Because of this symmetry, any Lagrangian term or counterterm involving an antighost field must involve its derivative $\partial_\mu \bar{c}^a$ rather than the $\bar{c}^a$ field itself.

6. The BRST symmetry, which acts on the canonically normalized component fields as

$$\delta\Psi^{fi}(x)) = \epsilon\{Q, \Psi^{fi}(x)\} = \tilde{g}\epsilon\, c^a(x)(t^a)^i{}_j\Psi^{fj}(x), \tag{5.a}$$

$$\delta\overline{\Psi}_{fi}(x)) = \epsilon\{Q, \overline{\Psi}_{fi}(x)\} = \tilde{g}\epsilon\,\overline{\Psi}_{fj}(x)(t^a)^j{}_i c^a(c), \tag{5.b}$$

$$\delta A^a_\mu(x) = \epsilon[Q, A^a_\mu(x)] = i\epsilon\,\partial_\mu c^a(x) - i\tilde{g}\epsilon f^{abc} A^b_\mu(x)c^c(x), \tag{5.c}$$

$$\delta c^a(x) = \epsilon\{Q, c^a(x)\} = i\tilde{g}\epsilon\, f^{abc}\, c^b(x)c^c(x), \tag{5.d}$$

$$\delta\bar{c}^a(x) = \epsilon\{Q, \bar{c}^a(x)\} = -i\epsilon\, b^a, \tag{5.e}$$

$$\delta b^a(x) = \epsilon[Q, b^a(x)] = 0. \tag{5.f}$$

   Caveat: the coupling $\tilde{g}$ multiplying the non-linear terms on the RHS of these formulae is subject to quantum corrections, so it may be different from the renormalized coupling g in the physical Lagrangian of the theory.

All of the symmetries **1–6** are manifest in the QCD Feynman rules. Moreover, they are preserved by the dimensional regularization, so all the UV-regulated Feynman amplitudes must be invariant WRT all these symmetries. In particular, all the divergent amplitudes must be invariant, so the counterterms which cancel the divergences — and hence the entire bare Lagrangian — must be invariant under all the symmetries 1–6.

To clarify the symmetry restrictions on the bare QCD Lagrangian

$$\mathcal{L}_{\text{bare}} \;=\; \mathcal{L}_{\text{ren}} \;+\; \mathcal{L}^{\text{counter}}_{\text{terms}}, \tag{6}$$

let's reorganize it into two types of operators according to whether they involve any unphysical fields, *i.e.* the ghosts, the antighosts, or the auxiliary fields, thus

$$\mathcal{L}_{\text{bare}} \;=\; \mathcal{L}_{b1}(A_\mu, \Psi, \overline{\Psi} \text{ only}) \;+\; \mathcal{L}_{b2}(c, \bar{c}, b; A_\mu). \tag{7}$$

Note that the second term here does not involve the quark fields Ψ and $\overline{\Psi}$: By the quark number conservation, any term involving quarks must involve both Ψ and $\overline{\Psi}$, and if such term also involves the unphysical fields c, $\bar{c}$, or b, then it has scaling dimension $\Delta > 4$, so it should not appear in the bare Lagrangian.

Let's start with the $\mathcal{L}_{b1}$ terms involving only the physical Ψ, $\overline{\Psi}$, and A fields. The BRST symmetry (5) acts on these fields as an infinitesimal gauge transform parametrized by $\Lambda^a(x) = -i\epsilon c^a(x)$. Since the ghost fields $c^a(x)$ are arbitrary functions of x, this makes for completely general infinitesimal gauge transforms of the physical fields. Hence, the BRST invariance of the $\mathcal{L}_{b1}$ requires it to be gauge invariant!

Combined with the other symmetries 1–3 of the physical fields, demanding the $\mathcal{L}_{b1}$ to be gauge invariant limits the the operators of dimension $\Delta \leq 4$ comprising the $\mathcal{L}_{b1}$ to the operators already present in the physical QCD Lagrangian, thus

$$\begin{aligned}
\mathcal{L}_{b1} \;=\; & -\frac{Z_3}{4}\left(\partial_\mu A^a_\nu \,-\, \partial_\nu A^a_\mu \,-\, \tilde{g} f^{abc} A^b_\mu A^c_\nu\right)^2 \\
& +\sum_f Z_2^{(q_f)} \overline{\Psi}_f \Big(i\gamma^\mu\big(\partial_\mu \,+\, i\tilde{g} A^a_\mu t^a\big) \,-\, m^{\text{bare}}_f\Big)\Psi^f \\
& + \text{ nothing else.}
\end{aligned} \tag{8}$$

Now consider the $\mathcal{L}_{b2}$ terms involving ghosts c^a, antighosts $\bar{c}^a$, and/or auxiliary fields b^a. Because of the ghost number and antighost shift symmetries 4–5, all operators involving the (anti)ghosts must involve products $(\partial_\mu\bar{c})\times c$ of dimension $\Delta=3$, so an operator of dimension $\Delta\leq 4$ has room for one more derivative or a bosonic field A_μ but not anything else. As to operators involving the auxiliary fields, the b^a has dimension $\Delta=2$, so the operators of dimension $\Delta\leq 4$ are limited to b^2, $b\partial_\mu A^\mu$, and $bA_\mu A^\mu$. Altogether, the allowed $\mathcal{L}_{b2}$ terms are limited to

$$\begin{aligned}\mathcal{L}_{b2} \;&=\; Z_2^{(\mathrm{gh})}(\partial_\mu\bar{c}^a)(\partial^\mu c^a) \;+\; T^{abc}(\partial_\mu\bar{c}^a)A^{\mu b}c^c \\ &\quad+\; \frac{\tilde{\xi}}{2}\,b^a b^a \;+\; \tilde{Z}\,b^a(\partial_\mu A^{\mu a}) \;+\; R^{abc}b^a A_\mu^b A^{\mu c}\end{aligned} \tag{9}$$

for some kinds of constants T^{abc} and R^{abc}. By the global $SU(3)$ invariance, these constants must be proportional to the Clebsch–Gordan coefficients for a singlet product of 3 adjoint multiplets, $\mathbf{1}\subset\mathbf{8}\times\mathbf{8}\times\mathbf{8}$.

Further restrictions on the bare Lagrangian terms (9) comes from the BRST symmetry. Demanding $[Q,\mathcal{L}_{b2}]=0$ and working through the algebra — which I leave as an optional exercise for the students — we find that we need

$$R^{abc} \;=\; 0, \quad T^{abc} \;=\; -\tilde{g}Z_2^{(\mathrm{gh})}\times f^{abc}, \quad \text{and} \quad \tilde{Z} \;=\; Z_2^{(\mathrm{gh})}, \tag{10}$$

hence

$$\mathcal{L}_{b2} \;=\; Z_2^{(\mathrm{gh})}\big(\partial_\mu\bar{c}^a\big)\big(\partial^\mu c^a \;-\; \tilde{g}f^{abc}A^{\mu b}c^c\big) \;+\; \frac{\tilde{\xi}}{2}\,b^a b^a \;+\; Z_2^{(\mathrm{gh})}\,b^a\big(\partial_\mu A^{\mu a}\big), \tag{11}$$

which includes only the terms already present in $\mathcal{L}_{\mathrm{g.f.}}+\mathcal{L}_{\mathrm{gh}}$. And this completes our proof of the Theorem and hence of QCD renormalizability.

Slavnov–Taylor Identities

Besides establishing QCD renormalizability, eqs. (8) and (11) for the net bare Lagrangian impose several relations on the renormalization factors $Z_1^{\cdots}$, $Z_2^{\cdots}$, and Z_3. Specifically,

$$\left.\begin{aligned} gZ_1^{(3g)} &= \tilde{g}Z_3\,, \\ g^2Z_1^{(4g)} &= \tilde{g}^2Z_3\,, \\ gZ_1^{(q)} &= \tilde{g}Z_2^{(q)}, \\ gZ_1^{(\mathrm{gh})} &= \tilde{g}Z_2^{(\mathrm{gh})}, \end{aligned}\right\} \quad \text{all for the same } \frac{\tilde{g}}{g} \text{ ratio.} \tag{12}$$

Unlike QED which has the Ward–Takahashi identity $Z_2 = Z_1$, QCD has weaker Slavnov–Taylor identities (12). Nevertheless, they suffice to keep the gauge coupling renormalization universal for all types of gauge interactions,

$$\left(\frac{Z_1}{Z_2}\right)_{\mathrm{quarks}} = \left(\frac{Z_1}{Z_2}\right)_{\mathrm{ghosts}} = \left(\frac{Z_1^{(3g)}}{Z_3}\right) = \left(\frac{Z_1^{(4g)}}{Z_3}\right)^{1/2} = \text{same } \frac{\tilde{g}}{g_{\mathrm{phys}}}\,. \tag{13}$$

By the way, the $\tilde{g}$ in these formulae is not the bare coupling g_{bare} but rather $g_{\mathrm{bare}} \times \sqrt{Z_3}$. Indeed, the bare and the physical couplings are related to each other as

$$\begin{aligned} gZ_1^{(3g)} &= g_{\mathrm{bare}}\,Z_3^{3/2}, \\ g^2Z_1^{(4g)} &= g_{\mathrm{bare}}^2\,Z_3^2, \\ gZ_1^{(q)} &= g_{\mathrm{bare}}\,Z_2^{(q)}\sqrt{Z_3}, \\ gZ_1^{(\mathrm{gh})} &= g_{\mathrm{bare}}\,Z_2^{(\mathrm{gh})}\sqrt{Z_3}; \end{aligned} \tag{14}$$

comparing these formulae to the Slavnov–Taylor identities (12) we immediately see that they agree for

$$\tilde{g} = g_{\mathrm{bare}} \times \sqrt{Z_3}\,. \tag{15}$$

Non-abelian gauge theories with several different types of matter multiplets have more Slavnov–Taylor identities. For example, Dirac fermions in different multiplets of the gauge group have exactly the same Z_1/Z_2 ratios:

$$\forall \text{ multiplet } (m): \quad \frac{Z_1^{(m)}}{Z_2^{(m)}} = \text{same } \frac{\tilde{g} = \sqrt{Z_3}g_{\mathrm{bare}}}{g_{\mathrm{phys}}}\,. \tag{16}$$

So if we add an exotic $SU(3)$ multiplet of fermions to QCD, then the exotic fermions would have exactly the same Z_1/Z_2 ratio as the regular quarks. Note: in general $Z_2^{\mathrm{exotic}} \neq Z_2^{\mathrm{regular}}$

and $Z_1^{\rm exotic} \neq Z_1^{\rm regular}$, but the Z_1/Z_2 ratio is universal for all multiplets. Similar universality applies to scalars in any non-singlet multiplet of the gauge group:

$$\forall\binom{\text{scalar}}{\text{multiplet}}(m):\quad \frac{Z_1^{(1g)}(m)}{Z_2(m)} = \left(\frac{Z_1^{(2g)}(m)}{Z_2(m)}\right)^{1/2} = \text{same}\ \frac{\tilde g = \sqrt{Z_3}g_{\rm bare}}{g_{\rm phys}}\,. \tag{17}$$

Going back to QCD and its counterterms, rewriting the Slavnov–Taylor identities in terms of the counterterm coefficients $\delta = Z - 1$ produces rather messy formulae. However, such formulae become much simpler in the minimal-subtraction renormalization scheme (MS) (*cf.* my notes on the subject) where each

$$Z = 1 + \delta = 1 + \frac{a_1}{\epsilon} + \frac{a_2}{\epsilon^2} + \cdots, \tag{18}$$

and all the higher-order pole coefficients follow from the *residue* $\operatorname{Res}\delta = a_1$. In terms of the counterterm residues, the identities (12) become

$$\operatorname{Res}\bigl[\delta_1^{(q)} - \delta_2^{(q)}\bigr] = \operatorname{Res}\bigl[\delta_1^{(gh)} - \delta_2^{(gh)}\bigr] = \operatorname{Res}\bigl[\delta_1^{(3g)} - \delta_3\bigr] = \tfrac{1}{2}\operatorname{Res}\bigl[\delta_1^{(4g)} - \delta_3\bigr]. \tag{19}$$

Moreover, each one of these differences in combinations with the (residue of the) δ_3 counterterm may be used to calculate the β function of the gauge theory. In light of eqs. (14),

$$\frac{dg(\mu)}{d\log\mu} = \beta(g) = g\hat L\ \operatorname{Res}\bigl[2\delta_1^{(q)} - 2\delta_2^{(q)} - \delta_3\bigr] \tag{20.1}$$

$$= g\hat L\ \operatorname{Res}\bigl[2\delta_1^{(gh)} - 2\delta_2^{(gh)} - \delta_3\bigr] \tag{20.2}$$

$$= g\hat L\ \operatorname{Res}\bigl[2\delta_1^{(3g)} - 3\delta_3\bigr] \tag{20.4}$$

$$= \frac{g}{2}\,\hat L\ \operatorname{Res}\bigl[2\delta_1^{(4g)} - 4\delta_3\bigr], \tag{20.4}$$

where $\hat L = g^2(\partial/\partial g^2)$ is the number-of-loops operator.

Back in 1973, David Gross and Frank Wilczek calculated all the relevant counterterms to one-loop order and verified that all 4 eqs. (20) indeed give the same negative answer for the QCD β function,

$$\beta(g) \;=\; (2N_f - 11N_c) \times \frac{g^3}{48\pi^2} \;+\; O(g^5). \tag{21}$$

In this class, we shall limit ourselves to calculating the $\delta_1^{(q)}$, $\delta_2^{(q)}$, and δ_3 counterterms and hence using eq. (20.1) for the β function. These calculations are presented in my next set of notes.