Exercise Sheet 12

Non-Abelian gauge theories

In 1954 Yang and Mills [YM54] generalized the Abelian gauge invariance of QED to non-Abelian gauge symmetries, in the quoted paper especially to SU(2) isospin symmetry. Here we consider a general SU(N) group for N Dirac fields with the *same* mass, written in a column vector $(\psi_1, \cdots, \psi_N)^{\mathrm{T}}$. The Lagrangian for the free fields reads

$$\mathscr{L}_0 = \overline{\psi}(\mathrm{i}\not{\partial} - m)\psi. \tag{1}$$

(a) Show that the free Lagrangian is invariant under global SU(N)-transformations,

$$\psi'(x) = \hat{U}\psi(x), \quad \overline{\psi}' = \overline{\psi}\hat{U}^\dagger. \tag{2}$$

Solution:

$$\mathscr{L}' = \overline{\psi'}(\mathrm{i}\not{\partial} - m)\psi' = \overline{\psi}\hat{U}^\dagger(\mathrm{i}\not{\partial} - m)\hat{U}\psi = \overline{\psi}(\mathrm{i}\not{\partial} - m)\hat{U}^\dagger\hat{U}\psi = \overline{\psi}(\mathrm{i}\not{\partial} - m)\psi = \mathscr{L}. \tag{3}$$

(b) We define the infinitesimal generators of the group as $\hat{T}^a$, such that the global transformation reads

$$\hat{U}(\vec{\chi}) = \exp(-\mathrm{i}g\hat{T}^a\chi^a). \tag{4}$$

with real parameters $\vec{\chi} = (\chi^a)$. What follows for an infinitesimal transformation from unitarity and unimodularity,

$$\hat{U}\hat{U}^\dagger = \mathbb{1}, \quad \det\hat{U} = 1? \tag{5}$$

Hint: For the latter constraint, show first, that for an arbitrary infinitesimal matrix $\delta\hat{M}$

$$\det(\mathbb{1} + \delta\hat{M}) = 1 + \operatorname{tr}\hat{M}. \tag{6}$$

Solution: Setting $\hat{U} = \mathbb{1} - \mathrm{i}g\delta\chi^a\hat{T}^a$ and expanding up to $\mathscr{O}(\delta)$ we get

$$\begin{aligned} &\hat{U}\hat{U}^\dagger = (\mathbb{1} - \mathrm{i}g\delta\chi^a\hat{T}^a)(\mathbb{1} + \mathrm{i}g\delta\chi^{a\dagger}) = \mathbb{1} - \mathrm{i}g_\delta\chi^a(\hat{T}^a - \hat{T}^{a\dagger}) + \mathscr{O}(\delta^2) \stackrel{!}{=} \mathbb{1} + \mathscr{O}(\delta^2) \\ &\Rightarrow \hat{T}^a - \hat{T}^{a\dagger} = 0 \Rightarrow \hat{T}^a = \hat{T}^{a\dagger}, \end{aligned} \tag{7}$$

i.e., the $\hat{T}^a$ must be Hermitian. For the unimodularity condition we find

$$\det\hat{U} = \det(\mathbb{1} - \mathrm{i}\delta\chi^a\hat{T}^a) = 1 - \mathrm{i}\delta\chi^a\operatorname{tr}\hat{T}^a + \mathscr{O}(\delta^2) \stackrel{!}{=} 1 + \mathscr{O}(\delta^2) \Rightarrow \operatorname{tr}\hat{T}^a = 0, \tag{8}$$

i.e., the matrices $\hat{T}^a$ must be traceless. The traceless hermitian $\mathbb{C}^{N\times N}$ matrices obviously form a real vector space. A hermitian $\mathbb{C}^{N\times N}$ matrix $\hat{T}$ has N real diagonal elements. There are $N^2 - N$ off-diagonal elements, of which the $(N^2-N)/2$ elements in the upper triangle of the matrix can be choosen as arbitrary real numbers, while the corresponding $N(N-1)/2$ elements in the lower triangle are the conjugate complex elements. Together there are $N + N(N-1) = N^2$ indepdendent real parameters for an arbitrary Hermitian $\mathbb{C}^{n\times n}$ matrix. If the matrix is supposed to be also traceless, i.e., $T_{jj} = 0$, one has only $(N-1)$ independent diagonal elements, i.e., a Hermitean traceless matrix can be uniquely parametrized with (N^2-1) real numbers, i.e., this real vector space of $\mathbb{C}^{N\times N}$ matrices is (N^2-1)-dimensional.

(c) Now we want to generalize the symmetry to a local gauge invariance, i.e., we make $\chi = \chi(\underline{x})$ in (4). Then, as in electrodynamics, we attempt to define a gauge-covariant derivative D_μ by introducing vector fields $\hat{A}_\mu = A_\mu^a\hat{T}^a$, $A_\mu^a \in \mathbb{R}$, and defining

$$\mathrm{D}_\mu\psi(\underline{x}) = [\partial_\mu + \mathrm{i}g\hat{A}_\mu(\underline{x})]\psi(\underline{x}). \tag{9}$$

Derive the gauge-transformation rule for $\hat{A}_\mu$ such that with

$$\psi'(\underline{x}) = \hat{U}(\underline{x})\psi(x) \tag{10}$$

the covariant derivative transforms in the same way,

$$\mathrm{D}'_\mu \psi'(\underline{x}) = \hat{U}(\underline{x})\mathrm{D}_\mu \psi(\underline{x}). \tag{11}$$

Solution:

$$\mathrm{D}'_\mu \psi' = (\partial_\mu + \mathrm{i}g\hat{A}'_\mu)\hat{U}\psi = (\hat{U}\partial_\mu + \partial_\mu \hat{U} + \mathrm{i}g\hat{A}'_\mu \hat{U})\psi. \tag{12}$$

On the other hand

$$\hat{U}\hat{D}_\mu \psi = \hat{U}(\partial_\mu + \mathrm{i}g\hat{A}_\mu)\psi. \tag{13}$$

Comparing (12) and (13) the transformation law (11) is fulfilled if

$$\begin{aligned} &(\hat{U}\partial_\mu + \partial_\mu \hat{U} + \mathrm{i}g\hat{A}'_\mu \hat{U})\psi = \hat{U}(\partial_\mu + \mathrm{i}g\hat{A}_\mu)\psi \\ \Rightarrow &(\partial_\mu \hat{U} + \mathrm{i}g\hat{A}'_\mu \hat{U})\psi = \mathrm{i}g\hat{U}\hat{A}_\mu\psi \\ \Rightarrow &\hat{A}'_\mu \hat{U}\psi = \hat{U}\hat{A}_\mu\psi - \frac{1}{\mathrm{i}g}(\partial_\mu \hat{U})\psi. \end{aligned} \tag{14}$$

Since this must hold for all ψ we get

$$\hat{A}'_\mu \hat{U} = \hat{U}\hat{A}_\mu + \frac{\mathrm{i}}{g}\partial_\mu \hat{U}. \tag{15}$$

Multiplying this equation with $\hat{U}^\dagger$ from the right, finally yields the desired transformation law

$$\hat{A}'_\mu = \hat{U}\hat{A}_\mu \hat{U}^\dagger + \frac{\mathrm{i}}{g}(\partial_\mu \hat{U})\hat{U}^\dagger. \tag{16}$$

(d) Show that the transformation rule is compatible with $A^a_\mu \in \mathbb{R}$. **Solution:** In (16) the first term is obviously a Hermitian traceless matrix if $\hat{A}_\mu$ is such a matrix, which is true by assumption, cf. 9. For the 2nd term we use

$$\hat{U}\hat{U}^\dagger = \mathbb{1} \Rightarrow \partial_\mu(\hat{U}\hat{U}^\dagger) = (\partial_\mu \hat{U})\hat{U}^\dagger + \hat{U}\partial_\mu \hat{U}^\dagger = 0 \Rightarrow (\partial_\mu \hat{U})\hat{U}^\dagger = -\hat{U}\partial_\mu U^\dagger = -[(\partial_\mu \hat{U})\hat{U}^\dagger]^\dagger, \tag{17}$$

from which

$$\left[\frac{\mathrm{i}}{g}(\partial_\mu \hat{U})\hat{U}^\dagger\right]^\dagger = -\frac{\mathrm{i}}{g}\hat{U}\partial_\mu \hat{U}^\dagger \stackrel{(17)}{=} +\frac{\mathrm{i}}{g}(\partial_\mu \hat{U})\hat{U}^\dagger, \tag{18}$$

i.e., the 2nd term in (16) is Hermitian.

For the proof that this term is also traceless we use the Jacobi formula for the derivative of a determinant and $\hat{U}^{-1} = \hat{U}^\dagger$:

$$\det \hat{U} = 1 \Rightarrow \partial_\mu \det \hat{U} = \det \hat{U}\,\mathrm{tr}(\hat{U}^{-1}\partial_\mu \hat{U}) = \mathrm{tr}(\partial_\mu \hat{U})\hat{U}^\dagger = 0. \tag{19}$$

You find a derivation of the Jacobi formula below in the appendix.

(e) Calculate the commutator $\left[\mathrm{D}_\mu, \mathrm{D}_\nu\right] = \mathrm{i}g\hat{F}_{\mu\nu} = \mathrm{i}gF^a_{\mu\nu}T^a$ with $F^a_{\mu\nu} \in \mathbb{R}$ by applying it to an arbitrary field ψ, which transforms under local gauge transformations as given in (9).

Note: The commutation relations of the generators are defined via the structure constants $f^{abc} \in \mathbb{R}$,

$$\left[\hat{T}^a, \hat{T}^b\right] = \mathrm{i}f^{abc}\hat{T}^c. \tag{20}$$

Solution: First we evaluate

$$\mathrm{D}_\mu \mathrm{D}_\nu \psi = (\partial_\mu + \mathrm{i}g\hat{A}_\mu)(\partial_\nu + \mathrm{i}\hat{A}_\nu)\psi = \partial_\mu\partial_\nu\psi + \mathrm{i}g(\partial_\mu \hat{A}_\nu)\psi + \mathrm{i}gA_\nu\partial_\mu\psi + \mathrm{i}g\hat{A}_\mu\partial_\nu\psi - g^2\hat{A}_\mu\hat{A}\nu\psi. \tag{21}$$

Subtracting from this the expression with μ and ν interchanged leads to

$$\hat{F}_{\mu\nu}\psi = \frac{1}{\mathrm{i}g}\left[\mathrm{D}_\mu, \mathrm{D}_\nu\right]\psi = \left(\partial_\mu \hat{A}_\nu - g^2\left[\hat{A}_\mu, \hat{A}_\nu\right]\right)\psi \tag{22}$$

and thus with (20)

$$\hat{F}_{\mu\nu} = F^a_{\mu\nu}\hat{T}^a = (\partial_\mu A^a_\nu - \partial_\nu A^a_\mu - gf^{abc}A^b_\mu A^c_\nu)\hat{T}^a. \tag{23}$$

(f) Using the fact that the generators $\hat{T}^a$ of the Lie algebra su(N) can be chosen such that

$$\mathrm{tr}(\hat{T}^a\hat{T}^b) = \frac{1}{2}\delta^{ab}\mathbb{1}, \tag{24}$$

show that the Lagrangian

$$\mathscr{L} = -\frac{1}{2}\,\mathrm{tr}(\hat{F}_{\mu\nu}\hat{F}^{\mu\nu}) + \overline{\psi}(\mathrm{i}\not{D} - m)\psi = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \overline{\psi}(\mathrm{i}\not{D} - m)\psi \tag{25}$$

is gauge invariant.

Solution: To show that the first term is gauge invariant, we note that ψ and $\mathrm{D}_\nu\psi$ both transform as described by (2) and (11), so that

$$\left[\mathrm{D}'_\mu, \mathrm{D}'_\nu\right]\psi' = \left[\mathrm{D}'_\mu, \mathrm{D}'_\nu\right]\hat{U}\psi = \hat{U}\left[\mathrm{D}_\mu, \mathrm{D}_\nu\right]\psi = \hat{U}\left[\mathrm{D}_\mu, \mathrm{D}_\nu\right]\hat{U}^{\mathrm{T}}\psi' \tag{26}$$

Since this holds for any ψ' by using (22) we get

$$\hat{F}'_{\mu\nu} = \hat{U}F_{\mu\nu}\hat{U}^\dagger. \tag{27}$$

Now, because one can commute matrices in matrix products under the trace we have

$$\mathrm{tr}(\hat{F}'_{\mu\nu}\hat{F}'^{\mu\nu}) = \mathrm{tr}(\hat{U}\hat{F}_{\mu\nu}\hat{U}^\dagger\hat{U}\hat{F}^{\mu\nu}\hat{U}) = \mathrm{tr}(\hat{U}\hat{F}_{\mu\nu}\hat{F}^{\mu\nu}\hat{U}^\dagger) = \mathrm{tr}(\hat{F}_{\mu\nu}\hat{F}^{\mu\nu}\hat{U}^\dagger\hat{U}) = \mathrm{tr}(\hat{F}_{\mu\nu}\hat{F}^{\mu\nu}). \tag{28}$$

So the first term in (25) is gauge invariant.

For the 2nd term the invariance follows from (2) and (11):

$$\overline{\psi'}(\mathrm{i}\not{D}' - m)\psi' = \overline{\psi}\hat{U}^\dagger(\mathrm{i}\hat{U}\not{D} - m\hat{U})\psi = \overline{\psi}(\mathrm{i}\not{D} - m)\psi. \tag{29}$$

(g) How about a mass term

$$\mathscr{L}_M = M^2\,\mathrm{tr}(\hat{A}_\mu\hat{A}^\mu) = \frac{1}{2}M^2A^a_\mu A^{a\mu}? \tag{30}$$

Solution: First we note that

$$\hat{U}\hat{U}^\dagger = \mathbb{1} \Rightarrow \partial_\mu(\hat{U}\hat{U}^\dagger) = (\partial_\mu\hat{U})\hat{U}^\dagger + \hat{U}\partial_\mu\hat{U}^\dagger = 0. \tag{31}$$

With (16) we get

$$\begin{aligned}
\mathrm{tr}(\hat{A}'_\mu\hat{A}'^\mu) &= \mathrm{tr}\left[\left(\hat{U}\hat{A}_\mu\hat{U}^\dagger + \frac{\mathrm{i}}{g}(\partial_\mu\hat{U})\hat{U}^\dagger\right)\left(\hat{U}\hat{A}^\mu\hat{U}^\dagger + \frac{\mathrm{i}}{g}(\partial^\mu\hat{U})\hat{U}^\dagger\right)\right] \\
&\overset{(31)}{=} \mathrm{tr}\left[\left(\hat{U}\hat{A}_\mu\hat{U}^\dagger + \frac{\mathrm{i}}{g}(\partial_\mu\hat{U})\hat{U}^\dagger\right)\left(\hat{U}\hat{A}^\mu\hat{U}^\dagger - \frac{\mathrm{i}}{g}\hat{U}\partial^\mu\hat{U}^\dagger\right)\right] \\
&= \mathrm{tr}\left[\hat{A}_\mu\hat{A}^\mu + \frac{\mathrm{i}}{g}(\partial_\mu\hat{U})\hat{A}^\mu\hat{U}^\dagger - \frac{\mathrm{i}}{g}\hat{U}\hat{A}_\mu(\partial^\mu\hat{U}^\dagger) + \frac{1}{g^2}(\partial_\mu\hat{U})(\partial^\mu\hat{U})\right] \\
&\neq \mathrm{tr}(\hat{A}_\mu\hat{A}^\mu),
\end{aligned} \tag{32}$$

i.e., a naive massterm for the gauge field is *not* gauge invariant.

(h) Why must the coupling g for all fields ψ the same (universality of the interaction strength in non-Abelian gauge theories)?

Solution: Since $\hat{F}_{\mu\nu}$ contains the gauge-coupling constant, g, all "matter fields" must transform with the same constant g. In the Abelian case, like in QED, $F_{\mu\nu}$ does not contain the coupling constant, and thus one can have different coupling constants for any matter field, i.e., in QED from the point of view of gauge invariance each sort of particle could have its own electric charge.

Note: The gauge groups, occuring in the standard model, are

- the Abelian case (QED) with gauge group U(1). Then one has only one generator $\hat{T} = 1$ (i.e., simply the real number 1. For any charged fermion (and its antiparticle), described by a Dirac field, with charge number Q the gauge-invariant matter Lagrangian reads

$$\mathscr{L}_{\text{QED}} = \overline{\psi}(\mathrm{i}\not{D} - m)\psi, \quad \mathrm{D}_\mu = \partial_\mu + \mathrm{i}QeA_\mu. \tag{33}$$

- $\mathrm{SU}(3)_\mathrm{C}$ (QCD) with $\hat{T}_\mathrm{C}^a = \hat{\lambda}^a/2$, where $\hat{\lambda}^a$ are the 8 Gell-Mann matrices acting on color space of the quarks, i.e., each quark field has 3 color components $\psi = (\psi_r, \psi_g, \psi_b)^\mathrm{T}$. There are 6 quarks $\psi = (\psi_f)$, $f \in \{\mathrm{u, d, c, s, t,b}\}$, with each ψ_f being a color triplet, gauge-covariant Lagrangian for the quarks, neglecting weak and electromagnetic interactions, then reads

$$\mathscr{L} = \overline{\psi}(\mathrm{i}\not{D} - \hat{M})\psi, \tag{34}$$

with the gauge-covariant derivative, $\mathrm{D}_\mu = \partial_\mu + \mathrm{i}g_\mathrm{s}\hat{T}_\mathrm{C}^a G_\mu^a$, and the mass matrix $\hat{M} = \mathrm{diag}(m_\mathrm{u}, \cdots, m_\mathrm{b})$.

- $\mathrm{SU}(2)_\text{wiso} \times \mathrm{U}(1)_\text{whyper}$ (QFD). This gauge group is more special since it's a chiral gauge group, acting differently on the left- and righthanded parts of the fermion fields.

  For a simplified model with only 1 family the fermion-gauge-field and keeping the neutrino massless we have a wiso-doublet of left-handed fermions,

$$\begin{pmatrix} \nu_{\mathrm{e\,L}} \\ e_\mathrm{L} \end{pmatrix}, \quad \begin{pmatrix} u_\mathrm{L} \\ d_\mathrm{L} \end{pmatrix} \tag{35}$$

  and wiso-singlets

$$e_\mathrm{R}, \quad u_\mathrm{R}, \quad d_\mathrm{R}. \tag{36}$$

  In the following we express all fermions in terms of Dirac fields, defining a wiso-doublet for both the leptons and the quarks,

$$\psi_\ell = \begin{pmatrix} \nu_\mathrm{e} \\ e \end{pmatrix}, \quad \psi_\mathrm{q} = \begin{pmatrix} u \\ d \end{pmatrix}. \tag{37}$$

  We denote the three $\mathrm{SU}(2)_\text{wiso}$ gauge fields with $W_\mu^a \in \mathbb{R}$ and the $\mathrm{U}(1)_\text{whyper}$ gauge field with B_μ and the SU(2)-generators as $\hat{T}_\text{wiso}^a = \hat{\sigma}^a/2$ (with the Pauli matrices, $\hat{\sigma}^a$ acting on flavor space).

  Further we denote the hyper-charge diagonal matrices, acting in flavor space, as $\hat{Y}_{\mathrm{L}\ell} = -\mathbb{1}_2/2$, $\hat{Y}_{\mathrm{R}\ell} = \mathrm{diag}(0, -1)$, $\hat{Y}_\mathrm{L,q} = \mathbb{1}_2/6$, and $\hat{Y}_\mathrm{R,q} = \mathrm{diag}(2/3, -1/3)$. Then the gauge-covariant piece for the fermion fields (including couplings with the gauge fields) reads

$$\mathscr{L} = \overline{\psi}_\ell(\mathrm{i}\not{\partial} + \mathrm{i}g\not{W}^a\hat{T}^a P_L + \mathrm{i}g'\not{B}(\hat{Y}_{\mathrm{L}\ell}P_L + \hat{Y}_{\mathrm{R}\ell}P_R)\psi_\ell + \overline{\psi}_\mathrm{q}(\mathrm{i}\not{\partial} + \mathrm{i}g\not{W}^a\hat{T}^a P_L + \mathrm{i}g'\not{B}(\hat{Y}_\mathrm{L,q}P_L + \hat{Y}_\mathrm{R,q}P_R)\psi_\mathrm{q}. \tag{38}$$

  It is also understood that each quark comes in a color triplet.

  In QFD the fermions get their mass via couplings to the Higgs field, which is a wiso-doublet $\Phi = (\Phi_1, \Phi_2)^\mathrm{T}$ with hyper-charge matrix $\hat{Y}_\mathrm{H} = \mathbb{1}/2$.

A Derivation of the Jacobi-formula

We want to prove the Jacobi formula for any invertible matrix $\hat{A}(t) \in \mathbb{C}^{n\times n}$, depending on some real parameter, t:

$$\mathrm{d}_t \det\hat{A} = \det A \,\mathrm{tr}\left(\hat{A}^{-1}\mathrm{d}_t\hat{A}\right). \tag{39}$$

The determinant can be written in a form, treating the row and column indices in a symmetric fashion by

$$\det\hat{A} = \epsilon_{a_1\ldots a_n}A_{a_1 1}\cdots A_{a_n n} = \frac{1}{n!}\epsilon_{a_1\ldots a_n}\epsilon_{b_1\ldots b_n}A_{a_1 b_1}\cdots A_{a_n b_n}. \tag{40}$$

Taking the derivative gives, via the product rule,

$$\begin{aligned}
\mathrm{d}_t \det\hat{A} &= \frac{1}{n!}\epsilon_{a_1\ldots a_n}\epsilon_{b_1\ldots b_n}\left(\dot{A}_{a_1b_1}\cdots A_{a_nb_n} + A_{a_1b_1}\dot{A}_{a_2b_2} + \cdots + A_{a_1b_1}\cdots\dot{A}_{a_nb_n}\right) \\
&= \frac{1}{(n-1)!}\epsilon_{a_1\ldots a_n}\epsilon_{b_1\ldots b_n}\dot{A}_{a_1b_1}A_{a_2b_2}\cdots A_{a_nb_n} \\
&= \dot{A}_{a_1b_1}\tilde{A}_{a_1b_1} = \mathrm{tr}\left(\dot{\hat{A}}\hat{\tilde{A}}^{\mathrm{T}}\right).
\end{aligned} \tag{41}$$

with the matrix

$$\tilde{A}_{a_1b_1} = \frac{1}{(n-1)!}\epsilon_{a_1\ldots a_n}\epsilon_{b_1\ldots b_n}\dot{A}_{a_1b_1}A_{a_2b_2}\cdots A_{a_nb_n}. \tag{42}$$

From this and the definition of the determinant we get

$$\tilde{A}_{a_1b_1}A_{a_1k} = \frac{1}{(n-1)!}\det\hat{A}\epsilon_{kk_2\ldots k_n}\epsilon_{k_1\ldots k_n} = \det\hat{A}\delta_{kk_1}. \tag{43}$$

This can be written in matrix form as

$$\hat{\tilde{A}}^{\mathrm{T}}\hat{A} = \det\hat{A}\mathbb{1} \;\Rightarrow\; \hat{\tilde{A}}^{\mathrm{T}} = (\det\hat{A})\hat{A}^{-1}. \tag{44}$$

Using this in (41) we finally get

$$\mathrm{d}_t \det\hat{A} = \mathrm{tr}\left[\dot{A}(\det\hat{A})\hat{A}^{-1}\right] = \det\hat{A}\,\mathrm{tr}\left(\dot{A}\hat{A}^{-1}\right) = \det\hat{A}\,\mathrm{tr}\left(\hat{A}^{-1}\dot{\hat{A}}\right). \tag{45}$$

References

[YM54] C.-N. Yang and R. L. Mills, Conservation of Isotopic Spin and Isotopic Gauge Invariance, Phys. Rev. **96**, 191 (1954), `https://doi.org/10.1103/PhysRev.96.191`.