GAUGE THEORIES

Gauge theories — abelian or non-abelian — are quantum theories of vector field $A^a_\mu(x)$ whose interactions with each other and with other fields follows from a *local symmetry.* So let me start these notes by explaining the difference between local and global symmetries:

- ⋆ A *global symmetry* — also called a *rigid symmetry* — has similar transformation of the fields at all spacetime points x. For example, a global phase symmetry of a complex scalar field $\Phi(x)$ acts as

$$\Phi(x) \;\to\; \Phi'(x) \;=\; e^{i\theta}\Phi(x), \quad \text{same } \theta \text{ for all } x. \tag{1}$$

- ⋆ In a *local symmetry* — also called a *gauge symmetry* — the field transformations at different points x have independent parameters. For example, a local phase symmetry of a complex scalar field $\Phi(x)$ acts as

$$\Phi(x) \;\to\; \Phi'(x) \;=\; e^{i\theta(x)}\Phi(x), \quad \text{independent } \theta(x) \text{ at each } x. \tag{2}$$

- A point of terminology: What a physicist calls a global symmetry, a mathematician would call a local symmetry and vice verse — a local symmetry to a physicist is a global symmetry to a mathematician. The terms *rigid symmetry* and *gauge symmetry* help avoid the confusion — both physicists and mathematicians agree to their meaning.

Abelian Example: Local Phase Symmetry.

Before we delve into non-abelian gauge theory, let me start with an abelian example. Consider a complex scalar field $\Phi(x)$ with a classical Lagrangian

$$\mathcal{L} \;=\; \partial^\mu\Phi^*\partial_\mu\Phi \;-\; m^2\Phi^*\Phi \;-\; \frac{\lambda}{2}\,(\Phi^*\Phi)^2, \tag{3}$$

which has a global phase symmetry $\Phi'(x) = e^{i\theta}\Phi(x)$. In fact, the potential terms here $\Phi^*\Phi$ and $(\Phi^*\Phi)^2$ have a local phase symmetry $\Phi'(x) \;=\; e^{i\theta(x)}\Phi(x)$, but the kinetic term does not have this local symmetry. Indeed, under this would-be local symmetry

$$\begin{aligned}\partial_\mu \Phi'(x) \;&=\; e^{i\theta(x)} \times \partial_\mu \Phi(x) \;+\; \Phi(x) \times \left[\partial_\mu\left(e^{i\theta(x)}\right) \;=\; ie^{i\theta(x)}\partial_\mu \theta(x)\right] \\ &=\; e^{i\theta(x)} \times \left(\partial_\mu \Phi(x) \;+\; i\Phi(x)\partial_\mu \theta(x)\right),\end{aligned} \tag{4}$$

hence

$$\partial_\mu \Phi^{*\prime}(x)\partial^\mu \Phi'(x) \;=\; \left(\partial_\mu \Phi^*(x) - i\Phi^*(x)\partial_\mu \theta(x)\right)\left(\partial_\mu \Phi(x) + i\Phi(x)\partial_\mu \theta(x)\right) \;\neq\; \partial^\mu \Phi^*(x)\partial_\mu \Phi(x). \tag{5}$$

However, we may repair this problem by replacing the ordinary field derivatives $\partial_\mu \Phi$ and $\partial_\mu \Phi^*$ with the *covariant derivatives* $D_\mu \Phi$ and $D_\mu \Phi^*$ which transform under the local symmetry just like the field Φ and Φ^* themselves:

$$\begin{aligned}\Phi(x) \;&\to\; e^{+i\theta(x)}\Phi(x), \quad & D_\mu \Phi(x) \;&\to\; e^{+i\theta(x)} D_\mu \Phi(x), \\ \Phi^*(x) \;&\to\; e^{-i\theta(x)}\Phi^*(x), \quad & D_\mu \Phi^*(x) \;&\to\; e^{-i\theta(x)} D_\mu \Phi^*(x).\end{aligned} \tag{6}$$

Given such covariant derivatives, the Lagrangian

$$\mathcal{L} \;=\; D^\mu \Phi^* D_\mu \Phi \;-\; V(\Phi^* \Phi) \tag{7}$$

would be invariant under the local rather than global phase symmetry.

To make the covariant derivatives, we need a *connection* — a 4-vector field $\mathcal{A}^\mu(x)$ undergoing a gauge transform parametrized by the same $\theta(x)$ as the local phase symmetry, thus

$$\left.\begin{aligned}\Phi'(x) \;&=\; \exp(+i\theta(x)) \times \Phi(x), \\ \Phi^{*\prime}(x) \;&=\; \exp(-i\theta(x)) \times \Phi^*(x), \\ \mathcal{A}'_\mu(x) \;&=\; \mathcal{A}_\mu(x) \;-\; \partial_\mu \theta(x)\end{aligned}\right\} \quad \text{for the same } \theta(x). \tag{8}$$

Given such combined phase/gauge transformations of the fields, the covariant derivatives

$$\begin{aligned}D_\mu \Phi(x) \;&=\; \partial_\mu \Phi(x) \;+\; i\mathcal{A}_\mu(x)\Phi(x), \\ D_\mu \Phi^*(x) \;&=\; \partial_\mu \Phi^*(x) \;-\; i\mathcal{A}_\mu(x)\Phi^*(x),\end{aligned} \tag{9}$$

transform covariantly according to eq. (6). Indeed,

$$\begin{aligned} (D_\mu\Phi)' &= \partial_\mu\Phi' + i\mathcal{A}'_\mu \times \Phi' \\ &= \Big(\partial_\mu(e^{i\theta}\Phi) = e^{i\theta}\big(\partial_\mu\Phi + i\Phi\partial_\mu\theta\big)\Big) + i\big(\mathcal{A}_\mu - \partial_\mu\theta\big) \times e^{i\theta}\Phi \\ &= e^{i\theta}\left(\partial_\mu\Phi + \cancel{i\Phi\partial_\mu\theta} + i\mathcal{A}_\mu \times \Phi - \cancel{i\partial_\mu\theta \times \Phi}\right) \\ &= e^{i\theta}\big(\partial_\mu\Phi + i\mathcal{A}_\mu \times \Phi\big) = e^{i\theta} \times D_\mu\Phi, \end{aligned} \tag{10}$$

and likewise

$$\begin{aligned} (D_\mu\Phi^*)' &= \partial_\mu\Phi^{*\prime} - i\mathcal{A}'_\mu \times \Phi^{*\prime} \\ &= \Big(\partial_\mu(e^{-i\theta}\Phi^*) = e^{-i\theta}\big(\partial_\mu\Phi^* - i\Phi^*\partial_\mu\theta\big)\Big) - i\big(\mathcal{A}_\mu - \partial_\mu\theta\big) \times e^{i\theta}\Phi \\ &= e^{-i\theta}\left(\partial_\mu\Phi^* - \cancel{i\Phi^*\partial_\mu\theta} - i\mathcal{A}_\mu \times \Phi^* + \cancel{i\partial_\mu\theta \times \Phi*}\right) \\ &= e^{-i\theta}\big(\partial_\mu\Phi^* - i\mathcal{A}_\mu \times \Phi^*\big) = e^{-i\theta} \times D_\mu\Phi^*. \end{aligned} \tag{11}$$

More generally, consider a theory with multiple complex fields $\varphi_a(x)$; these fields may be scalar, fermionic, vector, whatever, as long as they have definite charges q_a WRT to the phase symmetry. Under the local phase symmetry, all these fields and the connection $\mathcal{A}_\mu(x)$ transform according to

$$\left.\begin{aligned} \varphi'_a(x) &= \exp(+iq_a\theta(x)) \times \varphi_a(x), \\ \varphi^{*\prime}_a(x) &= \exp(-iq_a\theta(x)) \times \varphi^*_a(x) \\ &\qquad \langle\!\langle\, \varphi^*_a \text{ has charge } -q_a \,\rangle\!\rangle, \\ \mathcal{A}'_\mu(x) &= \mathcal{A}_\mu(x) - \partial_\mu\theta(x), \end{aligned}\right\} \quad \text{all for the same } \theta(x). \tag{12}$$

Under these transformation laws, the derivatives

$$D_\mu\varphi_a = \partial_\mu\varphi_a + iq_a\mathcal{A}_\mu \times \varphi_a\,, \quad D_\mu\varphi^*_a = \partial_\mu\varphi^*_a - iq_a\mathcal{A}_\mu \times \varphi^*_a\,, \tag{13}$$

are covariant:

$$\big(D_\mu\varphi_a(x)\big)' = \exp(+iq_a\theta(x)) \times D_\mu\varphi_a(x), \quad \big(D_\mu\varphi^*_a(x)\big)' = \exp(-iq_a\theta(x)) \times D_\mu\varphi^*_a(x). \tag{14}$$

So let us identify the connection $\mathcal{A}_\mu(x)$ with the electromagnetic field $A_\mu(x)$ and let's couple it to a bunch of scalar fields $\Phi_a(x)$ having electric charges q_a. Thanks to the covariance of the derivatives (13), the Lagrangian

$$\mathcal{L} \;=\; -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} \;+\; \sum_a^{\text{scalars}} D_\mu\Phi_a^* D^\mu\Phi_a \;-\; V(\text{scalars}) \tag{15}$$

is invariant under local phase symmetry. Or rather, as long as the potential $V(\text{scalars})$ is invariant under the global $U(1)$ phase symmetry, the Lagrangian (15) is invariant under the local $U(1)$ symmetry.

Algebra of Covariant Derivatives

- Multiple covariant derivatives of charged fields are all covariant:

$$\begin{aligned}
\big(D_\mu D_\nu\varphi_a(x)\big)' \;&=\; \exp(iq_a\theta(x)) \times D_\mu D_\nu\varphi_a(x), \\
\big(D_\lambda D_\mu D_\nu\varphi_a(x)\big)' \;&=\; \exp(iq_a\theta(x)) \times D_\lambda D_\mu D_\nu\varphi_a(x), \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{16}$$

- Leibniz rule

$$D_\mu(\varphi_a \times \varphi_b) \;=\; (D_\mu\varphi_a)\times\varphi_b \;+\; \varphi_a\times(D_\mu\varphi_b) \quad \text{for } q(\varphi_a\times\varphi_b) \;=\; q_a \;+\; q_b\,. \tag{17}$$

Indeed,

$$\begin{aligned}
D_\mu(\varphi_a \times \varphi_b) \;&=\; \partial_\mu(\varphi_a\times\varphi_b) \;+\; i(q_a+q_b)\mathcal{A}_\mu\times\varphi_a\times\varphi_b \\
&=\; (\partial_\mu\varphi_a)\times\varphi_b \;+\; \varphi_a\times(\partial_\mu\varphi_b) \;+\; iq_a\mathcal{A}_\mu\varphi_a\times\varphi_b \;+\; \varphi_a\times iq_b\mathcal{A}_\mu\varphi_b \\
&=\; (D_\mu\varphi_a)\times\varphi_b \;+\; \varphi_a\times(D_\mu\varphi_b).
\end{aligned} \tag{18}$$

- ◦ In particular, for $q_a + q_b = 0$ the product $\varphi_a \times \varphi_b$ is neutral, thus

$$(D_\mu\varphi_a)\times\varphi_b \;+\; \varphi_a\times(D_\mu\varphi_b) \;=\; \text{ordinary } \partial_\mu(\varphi_a\times\varphi_b), \tag{19}$$

which allows us to integrate by parts:

$$\begin{aligned}
\int\! d^4x\,(D_\mu\varphi_a)\times\varphi_b \;+\int\! d^4x\,\varphi_a\times(D_\mu\varphi_b) \;&=\int\! d^4x\,\partial_\mu(\varphi_a\times\varphi_b) \\
&=\; \int\limits_{\text{boundary}} d^3x\, n_\mu(\varphi_a\times\varphi_b) \\
\langle\!\langle\,\text{usually}\,\rangle\!\rangle \;&=\; 0.
\end{aligned} \tag{20}$$

For example, the kinetic term for a charged scalar field Φ can be integrated by parts as

$$\int d^4x\,(D_\mu\Phi^*)(D^\mu\Phi) \;=\; -\int d^4x\,\Phi^*(D^2\Phi) \;=\; -\int d^4x\,(D^2\Phi^*)\Phi. \tag{21}$$

- Similarly, given a Lagrangian for the charged fields as an explicit function of fields and their covariant derivatives (rather than ordinary derivatives)

$$\mathcal{L}_{\text{charged}}(\varphi, D_\mu\varphi) \quad \text{where } \varphi_a \text{ run over all charged fields and their conjugates,} \tag{22}$$

we may derive manifestly-covariant Euler–Lagrange equations by integrating by parts the infinitesimal action variation:

$$\begin{aligned}\delta S &= \int d^4x \sum_a \left(\frac{\partial\mathcal{L}}{\partial\varphi_a}\times\delta\varphi_a + \frac{\partial\mathcal{L}}{\partial(D_\mu\varphi_a)}\times D_\mu(\delta\varphi_a)\right)\\ &\quad\langle\!\langle\,\text{integrating by parts}\,\rangle\!\rangle\\ &= \int d^4x \sum_a \delta\varphi_a(x)\times\left(\frac{\partial\mathcal{L}}{\partial\varphi_a} - D_\mu\left(\frac{\partial\mathcal{L}}{\partial(D_\mu\varphi_a)}\right)\right),\end{aligned} \tag{23}$$

hence the field configuration minimizing the classical action obeys

$$\forall a:\quad D_\mu\left(\frac{\partial\mathcal{L}}{\partial(D_\mu\varphi_a)}\right) - \frac{\partial\mathcal{L}}{\partial\varphi_a} = 0. \tag{24}$$

In particular, the charged scalar fields with the Lagrangian

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + \sum_a^{\text{scalars}} D_\mu\Phi^*_a D^\mu\Phi_a - V(\text{scalars}) \tag{15}$$

obey

$$\forall a:\quad D_\mu D^\mu\Phi^*_a + \frac{\partial V}{\partial\Phi_a} = 0 \quad\text{and}\quad D_\mu D^\mu\Phi_a + \frac{\partial V}{\partial\Phi^*_a} = 0. \tag{25}$$

Note however that writing the Lagrangian $\mathcal{L}(\varphi, D_\mu\varphi)$ as a function of fields and their *covariant* derivatives hides its dependence of the EM potential A^μ, which we need for the Maxwell equation

$$\partial_\mu F^{\mu\nu} \;=\; J^\nu \quad\text{where}\quad J^\nu \;=\; -\,\frac{\partial\mathcal{L}_{\rm net}}{\partial A_\nu}\bigg|^{\text{@fixed}}_{F_{\mu\nu},\varphi,\partial_\mu\varphi}. \tag{26}$$

In terms of the covariant derivatives of the charged fields

$$\frac{\partial(D_\mu\varphi_a)}{\partial A_\nu} \;=\; iq_a\delta^\nu_\mu\varphi_a\,, \tag{27}$$

hence

$$J^\nu \;=\; -\frac{\partial\mathcal{L}(\varphi,D\varphi)}{\partial A_\nu} \;=\; -i\sum_q \frac{\partial\mathcal{L}}{\partial(D_\nu\varphi_a)}\times q_a\varphi_a\,. \tag{28}$$

In particular, for the charged scalar fields with the Lagrangian (15), the electric current is

$$J^\nu \;=\; \sum_a\Bigl(-iq_a\Phi_a\times D^\nu\Phi^*_a \,+\, iq_a\Phi^*_a\times D^\nu\Phi_a\Bigr). \tag{29}$$

Note manifest invariance of this current under the local phase symmetry!

⋆ But the covariance of derivatives D_μ has its price: unlike the ordinary derivatives ∂_μ, *the covariant derivatives D_μ do not commute with each other*, $D_\mu D_\nu \neq D_\nu D_\mu$. Indeed,

$$\begin{aligned} D_\mu D_\nu\varphi \;&=\; (\partial_\mu+iq\mathcal{A}_\mu)(\partial_\nu+iq\mathcal{A}_\nu)\varphi \\ &=\; \partial_\mu\partial_\nu\varphi \;+\; iq\mathcal{A}_\mu\times\partial_\nu\varphi \;+\; iq\mathcal{A}_\nu\times\partial_\mu\varphi \;+\; iq(\partial_\mu\mathcal{A}_\nu)\times\varphi \;-\; q^2\mathcal{A}_\mu\mathcal{A}_\nu\times\varphi \end{aligned} \tag{30}$$

where the blue terms on the RHS are symmetric WRT $\mu\leftrightarrow\nu$ but the red term is not symmetric. Consequently,

$$D_\mu D_\nu\varphi \;-\; D_\nu D_\mu\varphi \;=\; iq(\partial_\mu\mathcal{A}_\nu-\partial_\nu\mathcal{A}_\mu)\times\varphi \;=\; iq\mathcal{F}_{\mu\nu}\times\varphi, \tag{31}$$

or in the operator language

$$[D_\mu,D_\nu] \;=\; i\mathcal{F}_{\mu\nu}\times\hat{Q} \tag{32}$$

where $\hat{Q}$ is the electric charge operator, $\hat{Q}\varphi = q\varphi$.

Non Abelian Example: Local SU(N) Symmetry

Take N free complex scalar fields $\phi^1, \ldots, \phi^N$ of the same mass. The Lagrangian

$$\mathcal{L} \;=\; \partial_\mu \phi_j^* \partial^\mu \phi^j \;-\; m^2 \phi_j^* \phi^j \qquad \langle\!\langle \,\text{implicit } \textstyle\sum_j \rangle\!\rangle \tag{33}$$

is invariant under global symmetries which mix the fields with each other,

$$\phi^{j\prime}(x) \;=\; U^j_{\ k} \phi^k(x), \qquad \phi_j^{*\prime}(x) \;=\; \left(U^\dagger\right)^k_{\ j} \phi_k^*(x) \qquad \langle\!\langle \,\text{implicit } \textstyle\sum_k \rangle\!\rangle \tag{34}$$

for a *unitary* $N \times N$ matrix $\|U^j_{\ k}\|$. Such matrices form a non-abelian group called $U(N)$, hence the $U(N)$ group of symmetries of the N complex fields. Actually, the *free* Lagrangian (33) has a bigger symmetry group $SO(2N)$ — real rotations of $2N$ real fields $\mathrm{Re}\,\Phi^j$ and $\mathrm{Im}\,\Phi^j$ into each other, but only the $U(N)$ symmetries (34) preserve the distinction between the particles (created by the $\hat{\phi}_j^\dagger$ fields) and the antiparticles (created by the $\hat{\phi}^j$ fields) as well as the Lagrangian (33).

To make our notations for the $U(N)$ symmetries (34) more compact, let's assemble the $\phi^j(x)$ fields into a column vector $\Phi(x)$ of length N while the complex conjugate fields $\phi_j^*(x)$ form a row vector of the same length,

$$\Phi(x) \;\stackrel{\text{def}}{=}\; \begin{pmatrix} \phi^1(x) \\ \vdots \\ \phi^N(x) \end{pmatrix}, \qquad \Phi^\dagger(x) \;\stackrel{\text{def}}{=}\; \left(\phi_1^*(x) \quad \cdots \quad \phi_N^*(x) \right). \tag{35}$$

In terms of these complex vectors, the global symmetries (34) act by matrix multiplication,

$$\Phi'(x) \;=\; U\Phi(x), \qquad \Phi^{\dagger\prime}(x) \;=\; \Phi^\dagger(x) U^\dagger, \tag{36}$$

and leave the Lagrangian

$$\mathcal{L} \;=\; \partial_\mu \Phi^\dagger \partial^\mu \Phi \;-\; m^2 \Phi^\dagger \Phi \tag{37}$$

invariant because

$$U^\dagger U \;=\; 1 \quad\Longrightarrow\quad \Phi'^\dagger \Phi' \;=\; \Phi^\dagger U^\dagger U \Phi \;=\; \Phi^\dagger \Phi \tag{38}$$

and likewise for the kinetic term.

If we want to promote the global symmetries (36) to local symmetries

$$\Phi'(x) \;=\; U(x)\Phi(x), \qquad \Phi^{\dagger\prime}(x) \;=\; \Phi^\dagger(x)U^\dagger(x), \qquad \text{independent } U(x)\in U(N) \text{ at each } x, \tag{39}$$

we would need to replace the ordinary derivatives ∂_μ in the Lagrangian with the covariant derivatives D_μ such that

$$D'_\mu\Phi'(x) \;=\; U(x)D_\mu\Phi(x), \qquad D'_\mu\Phi^{\dagger\prime}(x) \;=\; \bigl(D_\mu\Phi^\dagger(x)\bigr)U^\dagger(x). \tag{40}$$

Given such covariant derivatives, the Lagrangian

$$\mathcal{L} \;=\; (D_\mu\Phi^\dagger)(D^\mu\Phi) \;-\; m^2\Phi^\dagger\Phi \tag{41}$$

would be invariant under the local symmetries (39).

The derivatives covariant WRT local $U(N)$ symmetry have form

$$D_\mu\Phi(x) \;=\; \partial_\mu\Phi(x) \;+\; i\mathcal{A}_\mu(x)\Phi(x), \qquad D_\mu\Phi^\dagger(x) \;=\; \partial_\mu\Phi^\dagger(x) \;-\; i\Phi^\dagger(x)\mathcal{A}^\dagger_\mu(x) \tag{42}$$

for a *matrix-valued connection* $\mathcal{A}_\mu(x)$. In other words, the connection is an $N\times N$ matrix $\|\mathcal{A}^j_{\mu,k}(x)\|$ of vector fields, and the covariant derivatives (42) act on the component fields ϕ^j and ϕ^*_j as

$$D_\mu\phi^j(x) \;=\; \partial_\mu\phi^j(x) \;+\; i\mathcal{A}^j_{\mu,k}(x)\phi^k(x), \qquad D_\mu\phi^*_j(x) \;=\; \partial_\mu\phi^*_j(x) \;-\; i\mathcal{A}^{*k}_{\mu,j}(x)\phi^*_k(x). \tag{43}$$

Similar to the abelian case, the local unitary symmetry of the $\phi^j(x)$ and ϕ^*_j fields should be accompanied by the gauge transform of the vector fields $\mathcal{A}^j_{\mu,k}(x)$, but the specific form of this gauge transform is more complicated than its abelian counterpart. Indeed, to achieve the covariance of the derivatives (42), we need

$$\begin{aligned}
\bigl(D_\mu\Phi\bigr)' \;&=\; \partial_\mu(\Phi'=U\Phi) \;+\; i\mathcal{A}'_\mu(\Phi'=U\Phi) \;=\; U\partial_\mu\Phi \;+\; (\partial_\mu U)\Phi \;+\; i\mathcal{A}'_\mu U\Phi \\
&\parallel \\
U\,D_\mu\Psi \;&=\; U\partial_\mu\Phi \;+\; iU\mathcal{A}_\mu\Phi,
\end{aligned}$$

and hence

$$i\mathcal{A}'_\mu U\Phi \;=\; iU\mathcal{A}_\mu\Phi \;-\; (\partial_\mu U)\Phi. \tag{44}$$

To make sure this relation works for any complex N–vector $\Phi(x)$, we need

$$i\mathcal{A}'_\mu(x)U(x) \;=\; iU(x)\mathcal{A}_\mu(x) \;-\; \partial_\mu U(x), \tag{45}$$

so *the non-abelian gauge transform of the matrix-valued connection $\mathcal{A}_\mu(x)$ works according to*

$$\mathcal{A}'_\mu(x) \;=\; U(x)\mathcal{A}_\mu(x)U^{-1}(x) \;+\; i(\partial_\mu U(x))U^{-1}(x). \tag{46}$$

Note: the first term on the RHS is peculiar to the non-abelian gauge transforms — in the abelian case, it would be simply $\mathcal{A}_\mu(x)$ — while the second term generalizes the $-\partial_\mu\theta(x)$. Indeed, for $N=1$ a unitary 1×1 matrix is simply a unimodular complex number $u=e^{i\theta}$. Consequently, the $U(1)$ symmetry group is the abelian group of phase symmetries, while

$$i(\partial_\mu u)\times u^{-1} \;=\; i(\partial_\mu e^{i\theta})\times e^{-i\theta} \;=\; -\partial_\mu\theta, \tag{47}$$

hence

$$\mathcal{A}'_\mu(x) \;=\; \mathcal{A}_\mu(x) \;-\; \partial_\mu\theta(x). \tag{48}$$

Next, let's take a closer look at the non-abelian vector fields. A priori, the connection $\mathcal{A}_\mu(x)$ is a *complex* $N\times N$ matrix of vector fields, which is equivalent to $2N^2$ real vector fields. However, we only need the Hermitian part of that matrix, $\mathcal{A}_\mu^\dagger=\mathcal{A}_\mu$, which is equivalent to N^2 real vector fields. Indeed, the second term in eq. (46) is always Hermitian,

$$\begin{aligned}\big[i(\partial_\mu U)U^{-1}\big]^\dagger \;&=\; -i\big(U^{-1}\big)^\dagger\big(\partial_\mu U^\dagger\big)\\ &\qquad\langle\!\langle \text{ by unitarity of } U,\ U^\dagger=U^{-1}\,\rangle\!\rangle\\ &=\; -iU\big(\partial_\mu U^{-1}\big) \;=\; -iU\big(-U^{-1}(\partial_\mu U)U^{-1}\big)\\ &=\; +i(\partial_\mu U)U^{-1},\end{aligned} \tag{49}$$

hence IF $\mathcal{A}_\mu$ is Hermitian THEN so is $\mathcal{A}'_\mu$:

$$\begin{gathered}
\left[U\mathcal{A}_\mu U^{-1}\right]^\dagger \;=\; \left(U^{-1}\right)^\dagger \mathcal{A}_\mu^\dagger U^\dagger \;=\; U\mathcal{A}_\mu U^{-1} \\
\Downarrow \\
\left[\mathcal{A}'_\mu \;=\; U\mathcal{A}_\mu U^{-1} \;+\; i(\partial_\mu U)U^{-1}\right]^\dagger \;=\; U\mathcal{A}_\mu U^{-1} \;+\; i(\partial_\mu U)U^{-1} \;=\; \mathcal{A}'_\mu\,.
\end{gathered} \tag{50}$$

Moreover, the unitary symmetry group $U(N)$ is a direct product of $SU(N)$ — the group of unitary matrices with unit determinants — and the $U(1)$ group of overall phases,

$$\text{any } U\in U(N) \text{ is } U \;=\; e^{i\theta}\times\tilde{U} \quad\text{where}\quad \det(\tilde{U})=1 \quad\text{and}\quad \theta=\frac{\arg(\det(U))}{N}\,. \tag{51}$$

In terms of the scalar fields $\phi^j(x)$, the $U(1)$ is the common phase symmetry — with the same phase $e^{i\theta}$ for all the ϕ^j, — while the $SU(N)$ symmetries mix the fields with each other. Consequently, the $SU(N)$ and the $U(1)$ connections are completely independent from each other. Specifically, the $U(1)$ connection $\mathcal{A}_\mu^{U(1)}$ is proportional to the unit matrix, while the $SU(N)$ connection is a traceless matrix. Indeed,

$$\text{as long as } \det U(x)\equiv 1 \text{ and } \operatorname{tr}(\mathcal{A}_\mu(x))\equiv 0, \tag{52}$$

$$\operatorname{tr}\bigl(-i(\partial_\mu U)U^{-1}\bigr) \;=\; -i\partial_\mu \operatorname{tr}\bigl(\log(U)\bigr) \;=\; -i\partial_\mu \log\bigl(\det(U)=1\bigr) \;=\; 0, \tag{53}$$

$$\operatorname{tr}\bigl(U\mathcal{A}_\mu U^{-1}\bigr) \;=\; \operatorname{tr}\bigl(\mathcal{A}_\mu\bigr) \;=\; 0, \tag{54}$$

$$\text{hence } \operatorname{tr}\bigl(\mathcal{A}'_\mu(x)\bigr) \;=\; 0. \tag{55}$$

The complete independence of the $SU(N)$ and $U(1)$ factors of the unitary group $U(N)$ means that either factor may be a local or a global symmetry independently of the other factor. In particular, a theory may have a local $SU(N)$ symmetry while the $U(1)$ remains a global phase symmetry, and that's what I am going to assume through the rest of this section. Consequently, there is no $U(1)$ connection, while *the $SU(N)$ connection $\mathcal{A}_\mu$ is a traceless Hermitian matrix* equivalent to N^2-1 real vector fields $\mathcal{A}_\mu^a(x)$, $a=1,\ldots,(N^2-1)$.

For example, for $N = 2$ there are 3 independent traceless Hermitian matrices, namely the Pauli matrices

$$\tau^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{56}$$

Consequently, the $SU(2)$ connection $\mathcal{A}_\mu(x)$ can be written as

$$\left[\mathcal{A}_\mu(x)\right]^j{}_k = \sum_{a=1,2,3} \mathcal{A}_\mu^a(x) \times \left(\frac{\tau^a}{2}\right)^j{}_k \tag{57}$$

in terms of 3 ordinary real vector fields $\mathcal{A}_\mu^a(x)$.

For $N \geq 3$, there are $N^2 - 1$ independent traceless Hermitian matrices, for example the Gell-Mann matrices λ^a. Here is their explicit forms for $N = 3$:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ +i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \begin{pmatrix} +1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ +i & 0 & 0 \end{pmatrix}, \quad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & +i & 0 \end{pmatrix},$$
$$\lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{58}$$

Consequently, the $SU(N)$ connection expands into $N^2 - 1$ ordinary real vector fields as

$$\left[\mathcal{A}_\mu(x)\right]^j{}_k = \sum_{a=1}^{N^2-1} \mathcal{A}_\mu^a(x) \times \left(\frac{\lambda^a}{2}\right)^j{}_k \tag{59}$$

For future reference, here are some properties of the Gell-Mann matrices:

- Similar to the Pauli matrices τ^a, the Gell-Mann matrices λ^a are Hermitian, traceless, and normalized to $\mathrm{tr}(\lambda^a \lambda^b) = 2\delta^{ab}$.
- $[\lambda^a, \lambda^b] = 2i \sum_c f^{abc} \lambda^c$ for some totally antisymmetric *structure constants* $f^{[abc]}$ of the $SU(N)$ Lie algebra. This commutation relation generalizes the isospin commutation relation $[\tau^a, \tau^b] = 2i \sum_c \epsilon^{abc} \tau^c$ for the Pauli matrices.

◦ Unlike the Pauli matrices, the Gell-Mann matrices do not anticommute with each other and do not square to unit matrices, $\{\lambda^a, \lambda^b\} \neq 2\delta^{ab} 1_{N\times N}$. Instead, for $N \geq 3$ we have

$$\{\lambda^a, \lambda^b\} \;=\; \frac{4\delta^{ab}}{N} \mathbf{1}_{N\times N} \;+\; \sum_c 2d^{abc} \lambda^c \tag{60}$$

for some totally symmetric coefficients $d^{(abc)}$.

Now let's go back to the component vector fields $\mathcal{A}^a_\mu(x)$. Earlier in this section I wrote down the non-abelian gauge transform of the vector fields in the matrix language, but translating it in terms of the component fields is rather painful. Or rather, it is quite painful for finite local symmetries $U(x)$, but it becomes much easier for the *infinitesimal symmetries*: In matrix language,

$$U(x) \;=\; \exp(i\Lambda(x)) \;=\; 1 \;+\; i\Lambda(x) \;+\; O(\Lambda^2) \tag{61}$$

for some infinitesimal matrix-valued $\Lambda(x)$. To keep the $U(x)$ unitary and $\det(U) = 1$, the $\Lambda(x)$ matrix should be Hermitian and traceless, hence

$$\Lambda(x) \;=\; \Lambda^a(x) \times \frac{\lambda^a}{2} \quad \langle\!\langle \text{implicit } \textstyle\sum_a \rangle\!\rangle \tag{62}$$

for some infinitesimal real numbers $\Lambda^a(x)$. Under such infinitesimal local symmetries, the scalar fields $\phi^j(x)$ and $\phi^*_j(x)$ transform into

$$\begin{aligned} \phi^{j\prime}(x) \;&=\; \phi^j(x) \;+\; i\Lambda^a(x) \left(\frac{\lambda^a}{2}\right)^j_{\;k} \phi^k(x) \;+\; O(\Lambda^2 \phi), \\ \phi^{*\prime}_j(x) \;&=\; \phi^*_j(x) \;-\; i\Lambda^a(x) \phi^*_k(x) \left(\frac{\lambda^a}{2}\right)^k_{\;j} \;+\; O(\phi^* \Lambda^2). \end{aligned} \tag{63}$$

At the same time, for the vector fields we have

$$i(\partial_\mu U)\,U^{-1} \;=\; -\partial_\mu\Lambda(x) \;+\; O(\Lambda^2), \tag{64}$$

$$\begin{aligned} U\,\mathcal{A}_\mu\,U^{-1} \;&=\; \bigl(1+i\Lambda+O(\Lambda)^2\bigr)\mathcal{A}_\mu\bigl(1-i\Lambda+O(\Lambda^2)\bigr)\bigr) \\ &=\; \mathcal{A}_\mu \;+\; i[\Lambda,\mathcal{A}_\mu] \;+\; O(\mathcal{A}\Lambda^2), \end{aligned} \tag{65}$$

and hence to first order in Λ,

$$\mathcal{A}'_\mu(x) \;=\; \mathcal{A}_\mu(x) \;+\; i[\Lambda(x),\mathcal{A}_\mu(x)] \;-\; \partial_\mu\Lambda(x). \tag{66}$$

In components,

$$\begin{aligned} i[\Lambda(x),\mathcal{A}_\mu(x)] \;&=\; \Lambda^b(x)\times\mathcal{A}^c_\mu(x)\times i\left[\frac{\lambda^b}{2},\frac{\lambda^c}{2}\right] \\ &=\; \Lambda^b(x)\times\mathcal{A}^c_\mu(x)\times\left(-f^{bca}\,\frac{\lambda^a}{2} \;=\; -f^{abc}\,\frac{\lambda^a}{2}\right) \\ &=\; -\left(f^{abc}\Lambda^b(x)\mathcal{A}^c_\mu(x)\right)\times\frac{\lambda^a}{2}\,, \end{aligned} \tag{67}$$

hence

$$\mathcal{A}'_\mu(x) \;=\; \frac{\lambda^a}{2}\times\left(\mathcal{A}^a_\mu(x) \;-\; f^{abc}\Lambda^b(x)\mathcal{A}^c_\mu(x) \;-\; \partial_\mu\Lambda^a(x)\right) \tag{68}$$

and therefore

$$\mathcal{A}^{a\prime}_\mu(x) \;=\; \mathcal{A}^a_\mu(x) \;-\; f^{abc}\Lambda^b(x)\mathcal{A}^c_\mu(x) \;-\; \partial_\mu\Lambda^a(x). \tag{69}$$

Non Abelian Tension Fields

In an abelian $U(1)$ gauge theory such as QED, the covariant derivatives D_μ do not commute with each other, and their commutators are related to the EM tensions fields as $[D_\mu,D_\nu]\phi(x) = iqF_{\mu\nu}(x)\phi(x)$. In non-abelian gauge theories, there is a similar relation in the matrix language,

$$[D_\mu,D_\nu]\Phi(x) \;=\; i\mathcal{F}_{\mu\nu}(x)\Phi(x) \tag{70}$$

where $\mathcal{F}_{\mu\nu}(x)$ is the matrix-valued tensor of tension fields. But the relation of this tensor to the connection $\mathcal{A}_\mu(x)$ is more complicated than in the abelian case. To see how it works, let's spell out the double covariant derivative

$$
\begin{aligned}
D_\mu D_\nu \Phi &= (\partial_\mu + i\mathcal{A}_\mu)(\partial_\nu + i\mathcal{A}_\nu)\Phi \\
&= \partial_\mu\partial_\nu\Phi \;+\; i\mathcal{A}_\mu \times \partial_\nu\Phi \;+\; i\mathcal{A}_\nu \times \partial_\mu\Phi \;+\; i(\partial_\mu\mathcal{A}_\nu) \times \Phi \;-\; \mathcal{A}_\mu\mathcal{A}_\nu \times \Phi.
\end{aligned} \tag{71}
$$

On the second line here I have color-coded in blue the terms which are symmetric WRT to the $\mu \leftrightarrow \nu$ interchange, and in red the terms which are not symmetric. Note that the last term is not symmetric because the matrices $\mathcal{A}_\mu$ and $\mathcal{A}_\nu$ generally do not commute with each other. Consequently,

$$
D_\mu D_\nu \Phi \;-\; D_\nu D_\mu \Phi \;=\; i(\partial_\mu\mathcal{A}_\nu) \times \Phi \;-\; i(\partial_\nu\mathcal{A}_\mu) \times \Phi \;-\; \mathcal{A}_\mu\mathcal{A}_\nu \times \Phi \;+\; \mathcal{A}_\nu\mathcal{A}_\mu \times \Phi, \tag{72}
$$

or in other words,

$$
[D_\mu, D_\nu]\Phi(x) \;=\; i\mathcal{F}_{\mu\nu}(x) \times \Phi(x) \tag{73}
$$

where

$$
\mathcal{F}_{\mu\nu}(x) \;=\; \partial_\mu\mathcal{A}_\nu(x) \;-\; \partial_\nu\mathcal{A}_\mu(x) \;+\; i[\mathcal{A}_\mu(x), \mathcal{A}_\nu(x)]. \tag{74}
$$

Or in components,

$$
\mathcal{F}_{\mu\nu}(x) \;=\; \mathcal{F}^a_{\mu\nu}(x) \times \frac{\lambda^a}{2} \quad \text{for} \quad \mathcal{F}^a_{\mu\nu}(x) \;=\; \partial_\mu\mathcal{A}^a_\nu(x) \;-\; \partial_\nu\mathcal{A}^a_\mu(x) \;-\; f^{abc}\mathcal{A}^b_\mu(x)\mathcal{A}^c_\nu(x). \tag{75}
$$

Unlike their abelian counterparts, the non-abelian tensions (74) are not gauge invariant. Instead, they transform covariantly under the local $SU(N)$ symmetries: In matrix language

$$
\mathcal{F}'_{\mu\nu}(x) \;=\; U(x)\mathcal{F}_{\mu\nu}(x)U^{-1}(x), \tag{76}
$$

while in components, the $\mathcal{F}^a_{\mu\nu}(x)$ form an *adjoint multiplet* of the $SU(N)$ symmetry,

$$
\mathcal{F}^{a\prime}_{\mu\nu}(x) \;=\; R^{ab}_{\text{adj}}(U(x)) \times \mathcal{F}^b_{\mu\nu}(x) \tag{77}
$$

where

$$
R^{ab}_{\text{adj}}(U) \;=\; \tfrac{1}{2}\,\mathrm{tr}\big(\lambda^a U \lambda^b U^{-1}\big) \tag{78}
$$

is the *adjoint representation* of $U \in SU(N)$.

Eq. (76) for the non-abelian tension fields may be derived directly from eq. (74) and the non-abelian gauge transform (46) of the vector field $\mathcal{A}_\mu(x)$ — this is a part of your next homework set#6. But it is much easier to derive eq. (76) from the commutator (73) and the covariance of the derivative D_μ. Indeed, multiple derivatives like $D_\mu D_\nu \Phi(x)$ are just as covariant as single derivatives,

$$\begin{aligned} \text{for } \Phi'(x) &= U(x)\Phi(x), \quad D'_\mu D'_\nu \Phi'(x) = U(x) D_\mu D_\nu \Phi(x) \implies \\ &\implies \big[D'_\mu, D'_\nu\big]\Phi'(x) = U(x)\big[D_\mu, D_\nu\big]\Phi(x), \end{aligned} \tag{79}$$

hence in light of eq. (73),

$$i\mathcal{F}'_{\mu\nu}(x) \times U(x)\Phi(x) = U(x) \times i\mathcal{F}_{\mu\nu}(x) \times \Phi(x), \tag{80}$$

and to make sure this relation works for any $\Phi(x)$ we need

$$\mathcal{F}'_{\mu\nu}(x) = U(x) \times \mathcal{F}_{\mu\nu}(x) \times U^{-1}(x). \tag{76}$$

As to the component form (77) of this transformation, using $\frac{1}{2}\operatorname{tr}(\lambda^a\lambda^b) = \delta^{ab}$ we get

$$\begin{aligned} \mathcal{F}^{a\prime}_{\mu\nu}(x) &= \operatorname{tr}\big(\lambda^a \mathcal{F}'_{\mu\nu}(x)\big) = \operatorname{tr}\big(\lambda^a U(x)\mathcal{F}_{\mu\nu}(x)U^{-1}(x)\big) \\ &= \tfrac{1}{2}\operatorname{tr}\big(\lambda^a U(x)\lambda^b U^{-1}(x)\big) \times \mathcal{F}^b_{\mu\nu}(x) \\ &= R^{ab}_{\text{adj}}(U(x)) \times \mathcal{F}^b_{\mu\nu} \end{aligned} \tag{77}$$

where

$$R^{ab}_{\text{adj}}(U) = \tfrac{1}{2}\operatorname{tr}\big(\lambda^a U \lambda^b U^{-1}\big). \tag{78}$$

As a matrix, the $\|R^{ab}_{\text{adj}}(U)\|$ is a real orthogonal $(N^2-1)\times(N^2-1)$ matrix, and as a function of U it's the *adjoint representation* of the $SU(N)$ symmetry group,

$$\forall U_1, U_2 \in SU(N): \quad R_{\text{adj}}(U_2U_1) = R_{\text{adj}}(U_2) \times R_{\text{adj}}(U_1). \tag{81}$$

Proof of reality: For any matrices $A, B, \ldots, Z$, $\big[\operatorname{tr}(AB\cdots Z)\big]^* = \operatorname{tr}(Z^\dagger \cdots B^\dagger A^\dagger)$, hence for hermitian matrices λ^a and λ^b and a unitary matrix U

$$\big[\operatorname{tr}(\lambda^a U \lambda^b U^{-1})\big]^* = \operatorname{tr}\big((U^{-1})^\dagger (\lambda^b)^\dagger U^\dagger (\lambda^a)^\dagger\big) = \operatorname{tr}(U\lambda^b U^{-1}\lambda^a) = \operatorname{tr}(\lambda^a U \lambda^b U^{-1}),$$

which means $\big[R^{ab}_{\text{adj}}(U)\big]^* = R^{ab}_{\text{adj}}(U)$.

Lemma: for any $N \times N$ matrices A and B,

$$\sum_a \operatorname{tr}(\lambda^a A) \times \operatorname{tr}(\lambda^a B) \;=\; 2\operatorname{tr}(AB) \;-\; \frac{2}{N}\operatorname{tr}(A) \times \operatorname{tr}(B). \tag{82}$$

Proof or orthogonality:

$$\begin{aligned}
\left(R^\top_{\text{adj}}(U) \times R_{\text{adj}}(U)\right)^{bc} &= \sum_a R^{ab}_{\text{adj}}(U) \times R^{ac}_{\text{adj}}(U)\\
&= \sum_a \tfrac{1}{2}\operatorname{tr}\bigl(\lambda^a(U\lambda^b U^{-1})\bigr) \times \tfrac{1}{2}\operatorname{tr}\bigl(\lambda^a(U\lambda^c U^{-1})\bigr)\\
\langle\!\langle \text{ by Lemma (82) } \rangle\!\rangle \quad &= \tfrac{1}{2}\operatorname{tr}(U\lambda^b U^{-1} \times U\lambda^c U^{-1}) \;-\; \frac{1}{2N}\operatorname{tr}(U\lambda^b U^{-1}) \times \operatorname{tr}(U\lambda^c U^{-1})\\
&= \tfrac{1}{2}\operatorname{tr}(\lambda^b\lambda^c) \;-\; \frac{1}{2N}\operatorname{tr}(\lambda^b) \times \operatorname{tr}(\lambda^c) \;=\; \delta^{bc} \;-\; 0,
\end{aligned} \tag{83}$$

which means $R^\top_{\text{adj}}(U) \times R_{\text{adj}}(U) = 1$.

Proof of the group law (81)**:**

$$\begin{aligned}
\left(R_{\text{adj}}(U_2) \times R_{\text{adj}}(U_1)\right)^{ab} &= \sum_c R^{ac}_{\text{adj}}(U_2) \times R^{cb}_{\text{adj}}(U_1)\\
&= \sum_c \tfrac{1}{2}\operatorname{tr}\bigl(\lambda^c(U_2^{-1}\lambda^a U_2)\bigr) \times \tfrac{1}{2}\operatorname{tr}\bigl(\lambda^c(U_1\lambda^b U_1^{-1})\bigr)\\
\langle\!\langle \text{ by Lemma (82) } \rangle\!\rangle \quad &= \tfrac{1}{2}\operatorname{tr}\bigl((U_2^{-1}\lambda^a U_2) \times (U_1\lambda^b U_1^{-1})\bigr)\\
&\qquad - \frac{1}{2N}\operatorname{tr}(U_2^{-1}\lambda^a U_2) \times \operatorname{tr}(U_1\lambda^b U_1^{-1})\\
&= \tfrac{1}{2}\operatorname{tr}\bigl(\lambda^a U_2 U_1 \lambda^b U_1^{-1} U_2^{-1}\bigr) \;-\; \frac{1}{2N}\operatorname{tr}(\lambda^a) \times \operatorname{tr}(\lambda^b)\\
&= \tfrac{1}{2}\operatorname{tr}\bigl(\lambda^a (U_2U_1)\lambda^b (U_2U_1)^{-1}\bigr) \;-\; 0 \times 0\\
&= R^{ab}_{\text{adj}}(U_2U_1),
\end{aligned} \tag{84}$$

thus $R_{\text{adj}}(U_2) \times R_{\text{adj}}(U_1) = R_{\text{adj}}(U_2U_1)$.

Example: for the $SU(2)$ isospin symmetry, U is the iso-doublet representation of some iso-space rotation while $R^{ab}_{\text{adj}}(U)$ is the iso-vector representation of the same rotation.

Later in class I shall tell you more about the adjoint multiplets as well as other kinds of multiplets of various symmetries, and in your next homework#6 you will learn more about the fields in adjoint multiplets of $SU(N)$ — and in particular about the tension fields $\mathcal{F}^a_{\mu\nu}(x)$. But meanwhile, we may use orthogonality of the $\|R^{ab}_{\text{adj}}\|$ matrices to form a gauge-invariant quadratic combination of the tension fields, namely

$$\operatorname{tr}\big(\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}\big) \;=\; \mathcal{F}^a_{\mu\nu}\mathcal{F}^{b\mu\nu}\times\left(\operatorname{tr}\left(\frac{\lambda^a}{2}\,\frac{\lambda^b}{2}\right)\;=\;\frac{\delta^{ab}}{2}\right)\;=\;\frac{1}{2}\mathcal{F}^a_{\mu\nu}\mathcal{F}^{a\mu\nu}. \tag{85}$$

The invariance of this combination follows from

$$\big(\mathcal{F}^a_{\mu\nu}\mathcal{F}^{a\mu\nu}\big)' \;=\; R^{ab}_{\text{adj}}(U)\mathcal{F}^b_{\mu\nu}\times R^{ac}_{\text{adj}}(U)\mathcal{F}^{c\mu\nu} \;=\; \delta^{bc}\times\mathcal{F}^b_{\mu\nu}\mathcal{F}^{c\mu\nu} \;=\; \mathcal{F}^a_{\mu\nu}\mathcal{F}^{a\mu\nu}, \tag{86}$$

or in matrix form

$$\operatorname{tr}\big(\mathcal{F}'_{\mu\nu}\mathcal{F}^{\mu\nu\prime}\big) \;=\; \operatorname{tr}\big(U\mathcal{F}_{\mu\nu}U^{-1}\times U\mathcal{F}^{\mu\nu}U^{-1}\big) \;=\; \operatorname{tr}\big(\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}\big). \tag{87}$$

Yang–Mills Theory

Yang–Mills theory is the theory of non-abelian gauge fields $\mathcal{A}^a_\mu(x)$ interacting with each other; there are no other fields. The physical Lagrangian of the theory is simply

$$\mathcal{L} \;=\; -\frac{1}{2g^2}\operatorname{tr}\big(\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}\big) \;=\; -\frac{1}{4g^2}\,\mathcal{F}^a_{\mu\nu}\mathcal{F}^{a\mu\nu} \tag{88}$$

for

$$\mathcal{F}^a_{\mu\nu} \;\stackrel{\text{def}}{=}\; \partial_\mu\mathcal{A}^a_\nu \;-\; \partial_\nu\mathcal{A}^a_\mu \;-\; f^{abc}\mathcal{A}^b_\mu\mathcal{A}^c_\nu. \tag{89}$$

The $1/g^2$ factor in the Yang–Mills Lagrangian (88) makes for a non-canonical normalization of the gauge fields $\mathcal{A}^a_\mu$. To get the canonically normalized vector fields, we rescale

$$A^a_\mu(x) \;=\; \frac{1}{g}\,\mathcal{A}^a_\mu(x) \quad\text{and}\quad F^a_{\mu\nu}(x) \;=\; \frac{1}{g}\,\mathcal{F}^a_{\mu\nu}(x), \tag{90}$$

hence

$$\mathcal{L} \;=\; -\tfrac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} \tag{91}$$

for

$$F^a_{\mu\nu} \;=\; \partial_\mu A^a_\nu \;-\; \partial_\nu A^a_\mu \;-\; g f^{abc} A^b_\mu A^c_\nu. \tag{92}$$

For small $g \ll 1$, we may treat the non-abelian parts of $F^a_{\mu\nu}$ as small perturbation, hence

$$\mathcal{L} \;=\; -\tfrac{1}{4}\big(\partial_\mu A^a_\nu \;-\; \partial_\nu A^a_\mu\big)^2 \;+\; \frac{g}{2}\big(\partial_\mu A^a_\nu \;-\; \partial_\nu A^a_\mu\big) \times f^{abc} A^{b\mu} A^{c\nu} \;-\; \frac{g^2}{4} f^{abc} f^{ade} A^b_\mu A^c_\nu A^{d\mu} A^{e\nu} \tag{93}$$

where the quadratic term (marked blue) describes $N^2 - 1$ species of free photon-like gluons, while the cubic and the quartic terms (marked red) describe the interactions between the gluon fields.

Adding Matter

As an example of a more general gauge theory, let's couple the Yang–Mills vector fields $A^a_\mu(x)$ to N complex scalar fields $\phi^j(x)$ subject to the same local $SU(N)$ symmetry. The overall Lagrangian is

$$\mathcal{L}_{\rm net} \;=\; \mathcal{L}_{\rm YM} \;+\; \mathcal{L}_\Phi \tag{94}$$

where $\mathcal{L}_{\rm YM}$ is the Yang–Mills Lagrangian exactly as in eqs. (88) or (93), while

$$\mathcal{L}_\Phi \;=\; D_\mu\Phi^\dagger\, D^\mu\Phi \;-\; m^2\Phi^\dagger\Phi. \tag{95}$$

In terms of the canonically normalized vector fields $A^a_\mu(x)$ we have

$$D_\mu\Phi \;=\; \partial_\mu\Phi \;+\; igA^a_\mu(\tfrac{1}{2}\lambda^a)\Phi, \quad D_\mu\Phi^\dagger \;=\; \partial_\mu\Phi^\dagger \;-\; igA^a_\mu\Phi^\dagger(\tfrac{1}{2}\lambda^a), \tag{96}$$

hence expanding the scalar fields' Lagrangian $\mathcal{L}_\Phi$ in powers of the gauge coupling g, we get

$$\begin{aligned}\mathcal{L}_\Phi \;=\; & \partial_\mu\Phi^\dagger\partial^\mu\Phi \;-\; m^2\Phi^\dagger\Phi \\ & +\; gA^a_\mu \times \Big(i\partial^\mu\Phi^\dagger(\tfrac{1}{2}\lambda^a)\Phi \;-\; \Phi^\dagger(\tfrac{1}{2}\lambda^a)\partial^\mu\Phi\Big) \;+\; g^2 A^a_\mu A^{b\mu} \times \Phi^\dagger\big(\tfrac{1}{2}\{\tfrac{1}{2}\lambda^a, \tfrac{1}{2}\lambda^b\}\big)\Phi.\end{aligned} \tag{97}$$

Again, the blue color marks the quadratic terms describing N free complex fields while red marks the cubic and quartic terms describing the interactions of the scalars with the gauge fields. Note that the same coupling g which governs how strongly the gauge fields interact with each other also governs the strength of their interactions with the scalar fields. Actually, *for any kind of a field — scalar, fermion, vector, whatever, — which happens to interact with the gauge fields of a particular local symmetry, the strength of all such interactions is governed by the same parameter* g.

General Gauge Symmetries

A Bit of Group Theory

Thus far we have focused on the non-abelian gauge theories with $SU(N)$ groups of local symmetries. But there are other kinds of gauge symmetry groups G; in general, any G that is continuous, *semisimple*, and *compact* can be a gauge symmetry of some QFT. But before we learn about such general gauge theories, we should learn a bit of group theory relevant to such theories.

Let's start with a few definitions: Suppose two symmetry groups G_1 and G_2 do not overlap (or rather, have no common elements except 1) and commute with each other:

$$\forall g_1 \in G_1,\ \forall g_2 \in G_2 : \quad g_1 g_2 \;=\; g_2 g_1\,. \tag{98}$$

Then we may form the *direct product* $G = G_1 \otimes G_2$ comprised of combined symmetries $g = (g_1 \in G_1)(g_2 \in G_2)$. And since G_1 and G_2 do not overlap, for any g is this direct product G, its decomposition into the product $(g_1 \in G_1)(g_2 \in G_2)$ is unique. This definition has a straightforward generalization to direct products of more than two groups, $G = G_1 \otimes G_2 \otimes G_3 \cdots$. For example, the translation group in 3D space is the direct product of 3 1D translation groups in x, y, and z directions, $T(3d) = T(1d)_x \otimes T(1d)_y \otimes t(1d)_z$.

Next, a subgroup H of group G is called a *normal subgroup* if it's invariant under all inner automorphisms of G; that is, if

$$\forall h \in H \text{ and } \forall g \in G : \quad g^{-1} h g \;\in H. \tag{99}$$

For example, let G be the group ISO(2) of continuous isometries of a 2D plane; this group is comprised of 2D rotations, 2D translations, and their products. Its subgroup H comprised of pure translations (without a rotation) is normal. Indeed, let $h = T(\mathbf{a}$ be a pure translation, while a general isometry is a product $g = T(\mathbf{b}R(\phi)$ or a translation and a rotation. Then

$$g^{-1} = R^{-1}(\phi)T^{-1}(\mathbf{b}) = R(-\phi)T(-\mathbf{b}) \tag{100}$$

and hence

$$g^{-1}hg = R(-\phi)T(-\mathbf{b})T(\mathbf{a})T(\mathbf{b})R(\phi) = R(-\phi)T(-\mathbf{b}+\mathbf{a}+\mathbf{b}=\mathbf{a})R(\phi) = T\big(R(-\phi)\mathbf{a}\big) \tag{101}$$

is indeed a pure translation by vector $R(-\phi)\mathbf{a}$.

Now, the simple and the semisimple continuous groups. A continuous group G is called *simple* if it has no continuous normal subgroups (except G itself). And a *semisimple* group G is a direct product of simple groups. For example:

- All the $SU(N)$ groups are simple. Likewise, all the nonabelian $SO(N)$ groups are simple.
- The $SU(2) \times SU(3)$ group is not simple but semi-simple. Likewise, all the $SU(N) \otimes SU(M)$ or $SU(N) \otimes SU(M) \otimes SU(K)$ groups are not simple but semisimple.
- OOH, the 2D isometry group ISO(2) is neither simple nor semisimple: It does nave a normal subgroup of pure translations, but it is not a direct product of the 2D rotations and 2D rotation groups, $\mathrm{ISO}(2) \neq SO(2) \otimes T(2d)$ because the rotations and the translations do not commute with each other. Likewise, the isometry groups $\mathrm{ISO}(d)$ in higher dimensions $d > 2$ are neither simple nor semisimple. In particular, the Poincare group in $3+1$ dimensions is neither simple nor semisimple.

Finally, a Lie group is called *compact* if its group manifold is compact.

- For example, the $SU(2)$ group is compact because it's group manifold is the 3D unit sphere S^3 in 4D. Indeed, a 2×2 matrix U is an $SU(2)$ matrix if an only if

$$U = a + ib_1\sigma_1 + ib_2\sigma_2 + ib_3\sigma_3 \tag{102}$$

for a real 4-vector (b_1, b_2, b_3, a) of unit norm, $b_1^2 + b_2^2 + b_3^2 + a^2 = 1$.

- Likewise, all the $SU(N)$ or $SO(N)$ groups are compact.
- On the other hand, the continuous Lorentz group $SO^+(3,1)$ is not compact. Indeed, if you Lorentz-boost many times in the same direction, you would get a net boost by a velocity asymptotically approaching c, but you would never get back to the original un-boosted frame. So the subgroup of boosts in $1+1$ dimensions is non-compact, hence the whole Lorentz group in $3+1$ dimensions is also non-compact.

Now consider the simplicity and/or compactness of a Lie group G from the point of view of Lie algebra $\mathbf{G}$ of the generators of G,

$$[\hat{T}^a, \hat{T}^b] = if^{ab}{}_c \hat{T}^c \qquad \langle\!\langle \text{implicit } \textstyle\sum_c \rangle\!\rangle. \tag{103}$$

A *quadratic Casimir operator* of such algebra is a quadratic polynomial of the generators

$$\hat{C}_2 = g_{ab} \hat{T}^a \hat{T}_b \qquad \langle\!\langle \text{implicit } \textstyle\sum_{a,b} \rangle\!\rangle \tag{104}$$

for some metric g_{ab} that happens to commute with all the generators,

$$\hat{C}_2 \hat{T}^c = \hat{T}^c \hat{C}_2 \quad \text{for all } \hat{T}^c. \tag{105}$$

For example, the quadratic Casimir of the $SO(3)$ algebra of angular momenta $\hat{J}^{x,y,z}$ is $\hat{\mathbf{J}}^2 = \hat{J}^x \hat{J}^x + \hat{J}^y \hat{J}^y + \hat{J}^z \hat{J}^z$, which indeed commutes with all 3 generators $\hat{J}^{x,y,z}$.

Theorem 1: *A Lie group G is simple if and only if its Lie algebra has a unique quadratic Casimir* (up to an overall factor). For example, the $SO(3)$ groups is simple, and its algebra has a unique quadratic Casimir $\hat{\mathbf{J}}^2$. On the other hand, the 3D isometry group ISO(3) is non-simple, and its algebra has 3 linearly independent quadratic Casimirs $\hat{\mathbf{J}}^2$, $\hat{\mathbf{P}}^2$, and $\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}$.

Theorem 2: *A simple group G is compact is and only if the metric matrix g_{ab} of its quadratic Casimir is positive-definite.* For example, the compact $SO(3)$ group has positive-definite metric $g_{ab} = \delta_{ab}$. On the other hand, simple but non-compact Lorentz group $SO^+(3,1)$ has quadratic Casimir operator

$$\hat{C}_2 = \tfrac{1}{2} \hat{J}^{\mu\nu} \hat{J}_{\mu\nu} = \hat{\mathbf{J}}^2 - \hat{\mathbf{K}}^2 \tag{106}$$

where the $\hat{J}^{x,y,z}$ generate the 3-space rotations while the $\hat{K}^{x,y,z}$ generate the Lorenz boosts. The metric g_{ab} in this quadratic Casimir (106) has mixed signature $(+++---)$, which agrees with the Lorenz group being non-compact.

Any non-degenerate Casimir metric g_{ab} and its inverse g^{ab} can be used to raise and lower the generator indices, for example

$$f^{ab}_{\ \ c} \rightarrow f^{abc} = f^{ab}_{\ \ d} g^{dc}. \tag{107}$$

Theorem 3: *For the algebra of any simple group, the structure functions f^{abc} (with all indices raised) are totally antisymmetric in all 3 indices a, b, c.* For example, for the $SO(3)$ algebra $f^{abc} = \epsilon^{abc}$. *Conversely, if f^{abc} are totally antisymmetric, then the group is either simple or a direct product of simple or abelian factors.*

In a moment, we shall see that the total antisymmetry of the f^{abc} as well as positive definiteness of the g_{ab} metric are necessary to make a well-behaved gauge theory. Consequently,

the local symmetry group G of any QFT must be one of the following:
- The abelian $U(1)$ group of phase symmetries, or
- A non-abelian simple compact group, or
- A direct product $G = G_1 \otimes G_2 \otimes \cdots$, where each factor G_i is either the abelian $U(1)$ or a nonabelian simple compact group.

(108)

⋆ For example, the gauge group of the Standard Model is the direct product

$$G = SU(3)_C \otimes SU(2)_W \otimes U(1)_Y \tag{109}$$

where the $SU(3)_C$ factor (where C stands for 'color' of the quarks) is responsible by the strong interactions while the $SU(2)_W \times U(1)_Y$ are responsible for the weak and EM interactions.

General Gauge Fields

In a general gauge theory, *the gauge connection $\mathcal{A}^\mu$ is Lie-algebra valued.* That is, for each generator $\hat{T}^a$ of the gauge group's Lie algebra $\mathbf{G}$ there is a vector field $\mathcal{A}^\mu_a(x)$ which acts as a component of the Lie-algebra-valued connection

$$\mathcal{A}^\mu \;=\; \mathcal{A}^\mu_a \times \hat{T}^a. \tag{110}$$

The curvature for this connection is the Lie-algebra-valued antisymmetric tensor field

$$\mathcal{F}^{\mu\nu}(x) \;=\; \partial^\mu \mathcal{A}^\nu(x) \;-\; \partial^\nu \mathcal{A}^\mu(x) \;+\; i\big[\mathcal{A}^\mu(x), \mathcal{A}^\nu(x)\big], \tag{111}$$

or in components

$$\mathcal{F}^{\mu\nu}_a(x) \;=\; \partial^\mu \mathcal{A}^\nu_a(x) \;-\; \partial^\nu \mathcal{A}^\mu_a(x) \;-\; f^{bc}{}_a \times \mathcal{A}^\mu_b(x)\mathcal{A}^\nu_c(x). \tag{112}$$

The local symmetries are parametrized by $u(x) \in G$ — for each x there is an element of the gauge group G. For infinitesimal symmetries

$$u(x) \;=\; \exp\big(i\Lambda_a(x)\hat{T}^a\big) \;=\; 1 \;+\; i\Lambda_a(x) \times \hat{T}^a \;+\; O(\Lambda^2) \tag{113}$$

for some infinitesimal real parameters $\Lambda^a(x)$. Under such infinitesimal symmetries, the gauge fields $\mathcal{A}^a_\mu(x)$ transform inhomogeneously as

$$\delta\mathcal{A}^\mu_a(x) \;=\; -\partial^\mu \Lambda^a(x) \;-\; f^{bc}{}_a \Lambda_b(x) \mathcal{A}^\mu_c(x) \tag{114}$$

while the tension fields $F^{\mu\nu}_a(x)$ transform homogeneously as

$$\delta\mathcal{F}^{\mu\nu}_a(x) \;=\; -f^{bc}{}_a \Lambda_b(x) \mathcal{F}^{\mu\nu}_c(x) \tag{115}$$

Now let's construct the Yang–Mills theory of the non-abelian gauge fields. The Lagrangian of this theory should be quadratic in the derivatives $\partial^\mu \mathcal{A}^\nu_a$ and gauge invariant, so it should be some invariant quadratic polynomial of the tension fields,

$$\mathcal{L} \;=\; (\text{coeff})g^{ab} \times \mathcal{F}_{a\mu\nu}\mathcal{F}_b^{\mu\nu} \tag{116}$$

for some invariant metric g_{ab}. But let's make sure this Lagrangian is indeed gauge invariant: Under an infinitesimal gauge transform (115), we have

$$\delta\mathcal{L} \;=\; (\text{coeff})g^{ab} \times \mathcal{F}_{a\mu\nu} \times f^{cd}{}_b\Lambda_c(x)\mathcal{F}_d^{\mu\nu} \;=\; (\text{coeff})\big(\mathcal{F}_{a\mu\nu}\mathcal{F}_d^{\mu\nu}\big) \times f^{cd}{}_b g^{ba} \times \Lambda_c \tag{117}$$

where the first factor $(\cdots)$ is symmetric in $a \leftrightarrow d$. Consequently, to make sure this $\delta\mathcal{L}$ vanishes for any Λ_c, the

$$f^{cda} \;\stackrel{\text{def}}{=}\; f^{cd}{}_b g^{ba} \tag{118}$$

should be antisymmetric WRT last 2 indices, $f^{cda} = -f^{cad}$. Since the same f is always antisymmetric in its first two indices, $f^{cda} = -f^{dca}$, this means total antisymmetry WRT all 3 of its indices. And as we saw in theorem 3, this happens only for simple gauge groups G or direct products of simple or abelian factors. But for a group G that includes a non-semisimple factor like ISO(3), the Lagrangian (116) would not be gauge invariant.

Furthermore, the kinetic-energy terms $(\partial^0 A^i_a)$ of all the vector field must have positive signs, so the invariant metric g^{ab} in eq. (116) must be positive definite. By theorem 2, this means that if G is simple then it must be compact, and if G is a direct product of simple and abelian factors, then each simple factor must be compact.

Altogether, this explains the rules (108) for the allowed gauge symmetries.

For a simple gauge group, the invariant metric g_{ab} is unique (up to an overall factor) and positive definite. Consequently, we may always choose a basis $\hat{T}^a$ for the group's generator such that for that basis $g_{ab} = \delta_{ab}$. Normally, one always works in such a basis, which allows to keep all the generator indices $a, b, \ldots$ upstairs, thus

$$\begin{aligned} \mathcal{A}_\mu \;&=\; \mathcal{A}^a_\mu \hat{T}^a, \qquad \mathcal{F}_{\mu\nu} \;=\; \mathcal{F}^a_{\mu\nu}\hat{T}^a, \\ \mathcal{F}^a_{\mu\nu} \;&=\; \partial_\mu \mathcal{A}^a_\nu \;-\; \partial_\nu \mathcal{A}^a_\mu \;-\; f^{abc}\mathcal{A}_\mu\mathcal{A}_\nu\,, \end{aligned} \tag{119}$$

and

$$\mathcal{L} \;=\; \frac{-1}{4g^2}\,\mathcal{F}^a_{\mu\nu}\mathcal{F}^{a\mu\nu} \tag{120}$$

where g is the gauge coupling. In terms of the canonically normalized vector potentials

$A^a_\mu(x)$ and tensions $F^a_{\mu\nu}$,

$$\mathcal{A}^a_\mu(x) \;=\; g \times A^a_\mu, \qquad \mathcal{F}^a_{\mu\nu}(x) \;=\; g \times F^a_{\mu\nu}(x), \tag{121}$$

hence

$$\text{connection } \mathcal{A}_\mu(x) \;=\; gA^a_\mu(x) \times \hat{T}^a, \tag{122}$$

$$\text{curvature } \mathcal{F}_{\mu\nu}(x) \;=\; gF^a_{\mu\nu}(x) \times \hat{T}^a, \tag{123}$$

$$F^a_{\mu\nu}(x) \;=\; \partial_\mu A^a_\nu(x) \;-\; \partial_\nu A_\mu(x) \;-\; gf^{abc}A^b_\mu(x)A^c_\nu(x), \tag{124}$$

$$\mathcal{L} \;=\; -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu}. \tag{125}$$

For a product gauge group $G = G_1 \otimes G_2 \otimes \cdots$, each simple or abelian factor G_i may have its own gauge coupling g_i. I shall explain how this works in a later section, but for now let's consider the matter fields in gauge theories with simple gauge groups.

Matter Fields in Gauge Theories

Besides the gauge fields, most gauge theories also have some kinds of matter fields: scalars, fermions, whatever. Most generally, all such fields must form complete multiplets of the gauge symmetry group G. In each such multiplet (m), the generators $\hat{T}^a$ of G are represented by $|m| \times |m|$ matrices $T^a_{(m)}$ obeying the same commutation relations as the generators themselves,

$$\left[T^a_{(m)}, T^b_{(m)}\right] \;=\; if^{abc} \times T^c_{(m)}\,. \tag{126}$$

Note: all such representations must be finite and unitary — in order to allow gauge-invariant kinetic terms that's positive for all the fields. A **theorem** of Lie group theory says that *for a simple compact group, all finite representations are unitary; but a simple non-compact group does not have any finite unitary representations except for the trivial singlet.* So the matter representations provide another reason why the gauge groups G should be compact.

Under infinitesimal gauge symmetries, a field Ψ^α belonging to some multiplet (m) is mixed with other fields Ψ^β belonging to the same multiplet — but not with fields in any other multiplets, even if they are of the same type — according to

$$\delta\Psi^\alpha(x) \;=\; i\Lambda^a(x)\left[T^a_{(m)}\right]^\alpha_{\ \beta}\Psi^\beta(x). \tag{127}$$

The covariant derivatives $D_\mu\Psi^\alpha$ also mix up fields belonging to the same multiplet (m), specifically

$$D_\mu\Psi^\alpha(x) \;=\; \partial_\mu\Psi^\alpha(x) \;+\; igA^a_\mu(x)\left[T^a_{(m)}\right]^\alpha_{\ \beta}\Psi^\beta(x). \tag{128}$$

Note different matrices $T^a_{(m)}$ for covariant derivatives of fields belonging to different multiplet types; this is similar to different fields having different electric charges in QED.

Let's verify the covariance of the derivatives (128) WRT infinitesimal gauge symmetries. In matrix language — where we treat the whole multiplet of fields Ψ^α as a column vector Ψ, we have

$$\begin{aligned}
\delta D_\mu\Psi \;&=\; D_\mu(\delta\Psi) \;+\; (\delta D_\mu)\Psi \\
&=\; \partial_\mu\delta\Psi \;+\; igA^a_\mu T^a_{(m)}\times\delta\Psi \;+\; ig\delta A^a_\mu\times T^a_{(m)}\Psi \\
&=\; i\Lambda^a T^a_{(m)}\times\partial_\mu\Psi \;+\; \cancel{i(\partial_\mu\Lambda^a)\times T^a_{(m)}}\Psi \;-\; gA^a_\mu T^a_{(m)}\times\Lambda^b T^b_{(m)}\Psi \\
&\qquad -\; \cancel{i(\partial_\mu\Lambda^a)\times T^a_{(m)}}\Psi \;-\; igf^{abc}\Lambda^b A^c_\mu\times T^a_{(m)}\Psi \\
&\qquad \langle\!\langle \text{relabeling indices} \rangle\!\rangle \\
&=\; i\Lambda^a T^a_{(m)}\times\partial_\mu\Psi \;-\; gA^c_\mu T^c_{(m)}\times\Lambda^a T^a_{(m)}\Psi \;-\; igf^{bac}\Lambda^a A^c_\mu\times T^b_{(m)}\Psi \\
&=\; i\Lambda^a\times\Big(T^a_{(m)}\partial_\mu\Psi \;+\; igA^c_\mu\times\Big(T^c_{(m)}T^a_{(m)}\Psi \;+\; if^{bac}T^b_{(m)}\Psi\Big)\Big)
\end{aligned} \tag{129}$$

where

$$if^{bac}T^b_{(m)} \;=\; if^{acb}T^b_{(m)} \;=\; \left[T^a_{(m)},T^c_{(m)}\right] \quad\Longrightarrow\quad T^c_{(m)}T^a_{(m)}\Psi \;+\; if^{bac}T^b_{(m)}\Psi \;=\; T^a_{(m)}T^c_{(m)}\Psi, \tag{130}$$

hence

$$\begin{aligned}
\delta D_\mu\Psi \;&=\; i\Lambda^a\times\Big(T^a_{(m)}\partial_\mu\Psi \;+\; igA^c_\mu T^a_{(m)}T^c_{(m)}\Psi\Big) \;=\; i\Lambda^a T^a_{(m)}\times\Big(\partial_\mu\Psi \;+\; igA^c_\mu T^c_{(m)}\Psi\Big) \\
&=\; i\Lambda^a T^a_{(m)}\times D_\mu\Psi,
\end{aligned} \tag{131}$$

quod erat demonstrandum.

To save time, I am not going to prove the covariance of D_μ under finite gauge transforms $u(x)$. Instead, let me simply summarize how such finite gauge transforms act on various fields. In general,

$$\text{any finite } u \in G \text{ is } u \;=\; \exp(i\Lambda^a \hat{T}^a) \text{ for some finite } \Lambda^a, \tag{132}$$

and the representation of this finite group element in a multiplet type (m) is a finite matrix

$$R_{(m)}(u) \;=\; \exp\bigl(i\Lambda^a T^a_{(m)}\bigr). \tag{133}$$

Consequently, under a finite gauge transform $u(x) = \exp(i\Lambda^a(x)\hat{T}^a)$, the matter fields $\Psi^\alpha(x)$ belonging to a multiplet (m) mix with each other — but only with the members of the same multiplet — as

$$\Psi^{\alpha\prime}(x) \;=\; \bigl[\exp\bigl(i\Lambda^a(x)T^a_{(m)}\bigr)\bigr]^\alpha_{\ \beta}\Psi^\beta(x). \tag{134}$$

As to the gauge fields, it is best to write their transformation laws in terms of the Lie-algebra-valued connection $\mathcal{A}_\mu(x)$ and curvature $\mathcal{F}_{\mu\nu}(x)$:

$$\mathcal{A}'_\mu(x) \;=\; i(\partial_\mu u(x))u^{-1}(x) \;+\; u(x)\mathcal{A}_\mu(x)u^{-1}(x), \tag{135}$$

$$\mathcal{F}'_{\mu\nu}(x) \;=\; u(x)\mathcal{F}_{\mu\nu}(x)u^{-1}(x). \tag{136}$$

Consequently, for any representation (r) of the gauge symmetry group G

$$\begin{aligned} A'^a_\mu(x)T^a_{(r)} \;&=\; \frac{i}{g}\,\partial_\mu\bigl(R_{(r)}(u(x))\bigr) \times R^{-1}_{(r)}(u(x)) \\ &\qquad +\; R_{(r)}(u(x)) \times A^b_\mu(x)T^b_{(r)} \times R^{-1}_{(r)}(u(x)), \end{aligned} \tag{137}$$

$$F'^a_{\mu\nu}(x)T^a_{(r)} \;=\; R_{(r)}(u(x)) \times F^b_{\mu\nu}(x)T^b_{(r)} \times R^{-1}_{(r)}(u(x)), \tag{138}$$

for the same $A'^a_\mu(x)$ and $F'^a_{\mu\nu}(x)$ for any representation (r). In components, eq. (135) becomes rather unwieldy, but eq. (136) amounts to the tension fields $F^a_{\mu\nu}(x)$ forming an *adjoint* *multiplet* of G, thus

$$F_{\mu\nu}^{a\prime}(x) \;=\; R_{\rm adj}^{ab}(u(x)) \times F_{\mu\nu}^{b}(x). \tag{139}$$

Note: any simple Lie group has an adjoint representation where the generators $\hat{T}^a$ are represented by

$$\left[T^a_{\rm adj}\right]^{bc} \;=\; -if^{abc}; \tag{140}$$

the commutation relations $\left[T^a_{\rm adj}, T^b_{\rm adj}\right] = if^{abc}T^c_{\rm adj}$ between these $\dim(G)\times\dim(G)$ matrices follow from the Jacobi identity

$$\forall a,b,c: \quad [\hat{T}^a,[\hat{T}^b,\hat{T}^c]] \;+\; [\hat{T}^b,[\hat{T}^c,\hat{T}^a]] \;+\; [\hat{T}^c,[\hat{T}^a,\hat{T}^b]] \;=\; 0 \tag{141}$$

for the Lie algebra **G**. Proof: in terms of the structure constants f^{abc},

$$[\hat{T}^a,[\hat{T}^b,\hat{T}^c]] \;=\; [\hat{T}^a, if^{bce}\hat{T}^e] \;=\; if^{bce}[\hat{T}^a,\hat{T}^e] \;=\; if^{bce}\times if^{aed}\times\hat{T}^d, \tag{142}$$

so the Jacobi identity (141) amounts to

$$-f^{bce}f^{aed}\times\hat{T}^d \;-\; f^{cae}f^{bed}\times\hat{T}^d \;-\; f^{abe}f^{ced}\times\hat{T}^d \;=\; 0 \tag{143}$$

and hence

$$f^{bce}f^{aed} \;+\; f^{cae}f^{bed} \;+\; f^{abe}f^{ced} \;=\; 0. \tag{144}$$

Now let's apply this identity to the adjoint representation's generators (140):

$$\begin{aligned}
\left[T^a_{\rm adj}, T^b_{\rm adj}\right]^{cd} \;&=\; \left(T^a_{\rm adj}\right)^{ce}\left(T^b_{\rm adj}\right)^{ed} \;-\; (a\leftrightarrow b)\\
&=\; \left(-if^{ace}\right)\left(-if^{bed}\right) \;-\; \left(-if^{bce}\right)\left(-if^{aed}\right)\\
&=\; +f^{cae}f^{bed} \;+\; f^{bce}f^{aed}\\
\langle\!\langle \text{ by eq. (144) } \rangle\!\rangle \qquad &=\; -f^{abe}f^{ced} \;=\; if^{abe}\times\left(if^{ced} = -if^{ecd}\right)\\
&=\; if^{abe}\times\left(T^e_{\rm adj}\right)^{cd},
\end{aligned} \tag{145}$$

thus indeed

$$\left[T^a_{\rm adj}, T^b_{\rm adj}\right] \;=\; if^{abe}T^e_{\rm adj}. \tag{146}$$

Fields $\Phi^a(x)$ in an adjoint multiplet transform under infinitesimal gauge symmetries as

$$\delta\Phi^a(x) \;=\; -f^{abc}\Lambda^b(x)\Phi^c(x) \tag{147}$$

and the covariant derivatives D_μ act on them as

$$D_\mu\Phi^a(x) \;=\; \partial_\mu\Phi^a(x) \;-\; gf^{abc}A^b_\mu(x)\Phi^c(x). \tag{148}$$

Or in matrix form — or rather Lie algebra form — $\hat{\Phi}(x) = \Phi^a(x)\hat{T}^a$,

$$\delta\hat{\Phi}(x) \;=\; i[\hat{\Lambda}(x),\hat{\Phi}(x)], \quad D_\mu\hat{\Phi} \;=\; \partial_\mu\hat{\Phi}(x) \;+\; i[\mathcal{A}_\mu(x),\hat{\Phi}(x)]. \tag{149}$$

The Lie algebra form also makes it easy to write down the finite gauge transform of an adjoint multiplet,

$$\hat{\Phi}'(x) \;=\; u(x)\hat{\Phi}(x)u^{-1}(x). \tag{150}$$

In particular, the tension fields $F^a_{\mu\nu}(x)$ — which transform according to eq. (150) — form an adjoint multiplet of the gauge symmetry.

Killing–Cartan classification

All the simple compact Lie algebra have been classified by Wilhelm Killing and Élie Cartan back in 1888–94. In modern terminology (Eugene Dynkin, 1947), there 4 infinite series A_n, B_n, C_n, and D_n, and 5 exceptional algebras G_2, F_4, E_6, E_7, and E_8. The index n here is the *rank* of the Lie algebra — the maximal number of independent generators that commute with each other. The 4 infinite series — sometimes called the *classical Lie algebras* correspond to the familiar unitary, orthogonal, or symplectic matrix groups. Specifically:

- The A_n algebras correspond to the special unitary groups, $A_n = SU(n+1)$, $n = 1,2,3,\ldots$.
- The B_n algebras correspond to the real orthogonal groups in odd dimensions, $B_n = SO(2n+1)$, $n = 1,2,3,\ldots$.

- The C_n algebra correspond to the unitary symplectic groups $USp(2n)$, *cf.* Wikipedia article on the subject. Briefly, the $USp(2n)$ group comprises unitary $2n \times 2n$ matrices U which also preserve a given antisymmetric tensor Ω_{ij}; in matrix notations,

$$\Omega = \begin{pmatrix} 0_{n\times n} & -1_{n\times n} \\ -1_{\mathrm{n}\times\mathrm{n}} & 0_{n\times n} \end{pmatrix}, \qquad U^\top \Omega U = \Omega. \tag{151}$$

- The D_n algebras correspond to the real orthogonal groups in even dimensions, $D_n = SO(2n)$, $n = 2, 3, 4, \ldots$.

∘ Alas, the 5 exceptional algebras G_2, F_4, E_6, E_7, and E_8 do not correspond to any classical matrix groups.

Combined Gauge Symmetries

A gauge symmetry group G does not have to be simple. It may also be a direct product of several simple or abelian $U(1)$ factors,

$$G = G_1 \otimes G_2 \otimes G_3 \otimes \cdots, \tag{152}$$

where each factor G_i comes with its own gauge fields — one for each generator of G_i — and its own gauge coupling g_i, thus

$$\mathcal{L} = \sum_i \frac{-1}{2g_i^2} \operatorname{tr}\bigl(\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}\bigr)_{G_i} + \mathcal{L}[\text{matter}]. \tag{153}$$

For example, the Standard Model has $G = SU(3) \times SU(2) \times U(1)$; the $SU(3)$ — which acts on the quark's colors — comes with 8 gluon fields $\mathcal{G}^a_\mu$ which are responsible for the strong interactions; while the 3 gauge fields $\mathcal{W}^a_\mu$ of the $SU(2)$ and 1 gauge field $\mathcal{B}_\mu$ of the $U(1)$ are responsible for the weak and the electromagnetic interactions. The three factors of the gauge group have rather different couplings,

$$\mathcal{L}_{\text{SM}} = -\frac{1}{2g_3^2} \operatorname{tr}\bigl(\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu}\bigr) - \frac{1}{2g_2^2} \operatorname{tr}\bigl(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\bigr) - \frac{1}{4g_1^2} \mathcal{B}_{\mu\nu}\mathcal{B}^{\mu\nu} + \mathcal{L}[\text{matter}], \tag{154}$$

for

$$\frac{4\pi}{g_3^2} \approx 8.45, \quad \frac{4\pi}{g_2^2} \approx 29.59, \quad \frac{4\pi}{g_1^2} \approx 98.36. \tag{155}$$

(Renormalized $\overline{\text{MS}}$ couplings at energy scale $E = M_Z = 91.19$ GeV.)

The matter multiplets of product gauge groups (152) are products of multiplets of the individual factors,

$$(m)\ =\ (m_1)\otimes(m_2)\otimes(m_3)\otimes\cdots,\quad (m_1)\text{ of }G_1,\quad (m_2)\text{ of }G_2,\quad (m_3)\text{ of }G_3,\ \ldots. \tag{156}$$

For the abelian factors of G (if any), all multiplets are singlets but they may have different $U(1)$ charges (which we need to specify). For example, the fermionic fields of the Standard Model form 5 kinds of the $SU(3)\times SU(2)\times U(1)$ multiplets:

- The left-handed quarks form triplets of $SU(3)$, doublets of $SU(2)$ — (u,d), (c,s), and (t,b), — and have $U(1)$ *hypercharge* $y=+\frac{1}{6}$. Consequently, each such multiplet has 6 members labeled by a color index $j=1,2,3$ and an $SU(2)$ flavor index $\alpha=1,2$, and the covariant derivatives act on the member fields $\Psi_Q^{j,\alpha}$ as

$$\begin{aligned} D_\mu\Psi_Q^{j,\alpha}\ &=\ \partial_\mu\Psi_Q^{j,\alpha}\ +\ \frac{ig_3}{2}\,G_\mu^a(\lambda^a)^j{}_k\Psi_Q^{k,\alpha}\ +\ \frac{ig_2}{2}\,W_\mu^a(\tau^a)^\alpha{}_\beta\Psi_Q^{j,\beta}\ +\ \frac{ig_1}{6}\,B_\mu\Psi_Q^{j,\alpha} \\ &\text{for}\quad j,k=1,2,3,\quad \alpha,\beta=1,2. \end{aligned} \tag{157}$$

- The right-handed quarks of flavors u, c, and t also form $SU(3)$ triplets, but they are singlets of $SU(2)$ and have hypercharge $y=+\frac{2}{3}$. Each such multiplet has 3 members Ψ_U^j distinguished by their colors $j=1,2,3$, and their covariant derivatives are

$$D_\mu\Psi_U^j\ =\ \partial_\mu\Psi_U^j\ +\ \frac{ig_3}{2}\,G_\mu^a(\lambda^a)^j{}_k\Psi_U^k\ +\ \frac{2ig_1}{3}\,B_\mu\Psi_Q^j\,. \tag{158}$$

- The right-handed quarks of flavors d, s, and b are also $SU(3)$ triplets and $SU(2)$ singlets, but they have a different hypercharge $y=-\frac{1}{3}$. Each such multiplet has 3 members Ψ_D^j similar to the Ψ_U^j, but their covariant derivatives have a different coupling of the $U(1)$ gauge field B_μ, namely

$$D_\mu\Psi_D^j\ =\ \partial_\mu\Psi_D^j\ +\ \frac{ig_3}{2}\,G_\mu^a(\lambda^a)^j{}_k\Psi_D^k\ -\ \frac{ig_1}{3}\,B_\mu\Psi_D^j\,. \tag{159}$$

- The left-handed leptons are $SU(3)$ singlets but $SU(2)$ doublets (ν_e,e^-), (ν_μ,μ^-), and (ν_τ,τ^-) of hypercharge $y=-\frac{1}{2}$. Each of these multiplets has 2 members Ψ_L^α distinguished by their flavors $\alpha=1,2$, but there are no color indices so they do not couple to the $SU(3)$ gauge fields. The covariant derivatives of the LH lepton fields are

$$D_\mu \Psi_L^\alpha \;=\; \partial_\mu \Psi_L^\alpha \;+\; \frac{ig_2}{2}\, W_\mu^a (\tau^a)^\alpha{}_\beta \Psi_L^\beta \;-\; \frac{ig_1}{2}\, B_\mu \Psi_L^\beta\,. \tag{160}$$

- The right-handed charged leptons e^-, μ^-, and τ^- are singlets of both $SU(3)$ and $SU(2)$ — hence only one member Ψ_E per multiplet, without any color or flavor indices, — and have hypercharge $y = -1$. The covariant derivatives of the RH charged lepton fields are

$$D_\mu \Psi_E \;=\; \partial_\mu \Psi_E \;-\; ig_1 B_\mu \Psi_E\,. \tag{161}$$

◦ Finally, the right-handed neutrinos. Presently, we do not know where these fields exist at all; but if they do exist, they do not couple to any of the Standard Model's gauge fields. Thus, they are singlets of both $SU(3)$ and $SU(2)$ and also have zero hypercharges, so their covariant derivatives are simply

$$D_\mu \Psi_N \;=\; \partial_\mu \Psi_N \;+\; 0. \tag{162}$$