5. Local gauge symmetry with matter coupling

Now we shall return to the RHS of sourced gauge field equations like Maxwell's equations. The currents that we need are indeed Noether currents, but to incorporate them properly, we need to generalize the rigid symmetries that we have been considering to local symmetries, where the transformation parameters are allowed to depend upon the location in spacetime: $\Theta^i(x)$.

— One immediately then runs into the problem that $\partial_\mu \phi_A$ and $\partial_\mu \varphi^{*A}$ no longer transform simply according to their A-index representation type, because one picks up $\partial_\mu \Theta^i$ terms.

Consider this problem in the simple case of a complex scalar field of charge eq, where e is the electric charge unit and $q \in \mathbb{Z}$ determines the multiple of that unit carried by excitations of the field $\varphi(x)$. The local $U(1)$ symmetry transformation is now

$$\varphi(x) \longrightarrow \varphi'(x) = e^{-iq\alpha(x)}\varphi(x)$$
$$\varphi^*(x) \longrightarrow \varphi^{*\prime}(x) = e^{iq\alpha(x)}\varphi^*(x)$$

where $\alpha(x)$ is now a local, i.e. spacetime-dependent, parameter. Clearly, the derivatives now transform as

$$\partial_\mu \varphi(x) \longrightarrow \partial_\mu(e^{-iq\alpha}\varphi) = e^{-iq\alpha}\partial_\mu\varphi - iq\,\partial_\mu\alpha\, e^{-iq\alpha}\varphi$$
$$\partial_\mu \varphi^*(x) \longrightarrow \partial_\mu(e^{iq\alpha}\varphi) = e^{iq\alpha}\partial_\mu\varphi^* + iq\,\partial_\mu\alpha\, e^{iq\alpha}\varphi^*.$$

The first terms in these transformations have the desired covariant form taken from the transformations of φ and φ^*. If these were all we had, then the kinetic term $-\frac{1}{2}\partial_\mu\varphi^*\partial^\mu\varphi$ would still be invariant. However, the second terms involving $\partial_\mu\alpha(x)$ spoil this covariance, and hence spoil the invariance of the kinetic term.

The resolution of this problem is to include a coupling to the Maxwell vector potential A_μ, with a coordinated gauge transformation.

To do this, we define the covariant derivatives of φ and φ^*:

$$D_\mu\varphi = \partial_\mu\varphi + ieqA_\mu\varphi\,, \quad D_\mu\varphi^* = \partial_\mu\varphi^* - ieqA_\mu\varphi^*.$$

This is where the charge unit, or coupling constant, e comes in. As we have seen, the integer q determines the rate of phase transformations of φ and φ^*. We may consider q to determine the representations of the group $U(1)$ carried by (φ, φ^*).

In order to make the covariant derivative combinations transform covariantly, i.e. according to the patterns set by $\varphi(x)$ and $\varphi^*(x)$, we need to coordinate the Maxwell local gauge parameter $\Lambda(x)$ with the local parameter $\alpha(x)$ of the $U(1)$ transformations of $(\varphi(x), \varphi^*(x))$. We choose $\Lambda(x) = \frac{1}{e}\alpha(x)$, so $A_\mu(x) \longrightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x)$. Then one has

$$D_\mu\varphi \rightarrow e^{-iq\alpha}\partial_\mu\varphi - iq\cancel{\partial_\mu\alpha\, e^{-iq\alpha}\varphi} + ieqA_\mu e^{-iq\alpha}\varphi + iq\cancel{\partial_\mu\alpha\, e^{-iq\alpha}\varphi}$$
$$= e^{-iq\alpha}D_\mu\varphi$$

and

$$D_\mu\varphi^* \longrightarrow e^{iq\alpha}\partial_\mu\varphi^* + iq\cancel{\partial_\mu\alpha\, e^{iq\alpha}\varphi^*} - ieqA_\mu e^{iq\alpha}\varphi^* - iq\cancel{\partial_\mu\alpha\, e^{iq\alpha}\varphi^*}$$
$$= e^{iq\alpha}D_\mu\varphi^*.$$

Thus the covariant derivative combinations restore the covariant transformation patterns taken from undifferentiated φ and φ^*. Note that, in order to achieve this result, the covariant derivative operator D_μ needs to be built specifically for the field it is operating on, taking a different form when acting on φ^* from its form when acting on φ.

The gauge-field-coupled action for φ and φ^* is then

$$I_{\substack{\text{coupled}\\ \text{scalars}}} = \int d^4x\left(-D^\mu\varphi^* D_\mu\varphi - V(\varphi^*, \varphi)\right)$$

$$= \int d^4x\left(-\partial^\mu\varphi^*\partial_\mu\varphi + ieqA^\mu(\varphi^*\partial_\mu\varphi - \partial_\mu\varphi^*\varphi) - e^2q^2A^\mu A_\mu\varphi^*\varphi - V(\varphi^*, \varphi)\right)$$

Invariance of $V(\varphi, \varphi^*)$ insensitive to local transformation character.

I_{coupled} is now fully locally $U(1)$ invariant provided $V(\varphi^*, \varphi)$ is an invariant potential.

Note the structure of the term linear in A^μ in this gauge coupled action: $\int d^4x\,(-eA^\mu J_\mu)$ where $J_\mu = -iq\,\varphi^*\overleftrightarrow{\partial_\mu}\varphi$ is the Noether current for the $e^{-iq\alpha}$ $U(1)$ transformation.

Noether coupling

The structure of the gauge-coupled scalar field action can be arrived at another way. Start with the uncoupled action $I_{\text{uncoupled}} = \int d^4x \left(-\partial^\mu\varphi^* \partial_\mu\varphi - V(\varphi^*,\varphi)\right)$. Provided $V(\varphi^*,\varphi)$ is invariant, this action is invariant under rigid $\varphi \to e^{-iq\alpha}\varphi$, $\varphi^* \to e^{iq\alpha}\varphi^*$ transformations, but it ceases to be invariant if one lets the transformation parameter α become local, $\alpha \to \alpha(x)$. Restricting attention to infinitesimal transformations with parameter $\delta\alpha(x)$, the action varies by

$$\Delta I_{\text{uncoupled}} = \int d^4x \left(-iq\,\partial^\mu\delta\alpha\,\varphi^*\partial_\mu\varphi + iq\,\partial^\mu\varphi^*\varphi\,\partial_\mu\delta\alpha\right)$$

since $\Delta\varphi = -iq\delta\alpha\varphi$, $\Delta\varphi^* = iq\delta\alpha\varphi^*$ (dropping $\mathcal{O}(\delta\alpha^2)$ terms). We recognise this as $\Delta I_{\text{uncoupled}} = \int d^4x \left(\partial^\mu\delta\alpha\, J_\mu\right)$, yielding the Noether current as the cofactor of $\partial^\mu\delta\alpha$.

Proceeding order by order in the coupling constant e, it now becomes clear why including the linear gauge field coupling $\int d^4x\left(-eA^\mu J_\mu\right)$ cancels the leading variation $\int d^4x\left(\partial^\mu\delta\alpha\, J_\mu\right)$. — However, the coupling process doesn't stop there, since the Noether current J_μ contains a derivative and so is not itself invariant under the local symmetry transformations: $\Delta J_\mu = -2q^2\varphi^*\varphi\,\partial_\mu\delta\alpha$. Accordingly, in order to obtain full invariance under the $\alpha(x)$ local symmetry, it is necessary to add an $\mathcal{O}(e^2)$ term to the coupled action: $\int d^4x\left(-e^2q^2A^\mu A_\mu\varphi^*\varphi\right)$. After including this term, the coupling is complete and one has full invariance of I_{coupled} for infinitesimal $\delta\alpha(x)$ local transformations. One can then check that this Noether coupled action is invariant also under finite $\alpha(x)$ local transformations. For connected symmetry groups, invariance of the gauge coupled action for finite $\alpha(x)$ transformations follows inevitably from invariance under infinitesimal $\delta\alpha(x)$ transformations.

Field Strength from covariant derivative commutator

Coupling scalar fields by promoting the partial derivative ∂_μ in the uncoupled action's kinetic term to a covariant derivative is called minimal coupling.

The covariant derivative D_μ has the important property of producing the Maxwell field strength from its commutator: $[D_\mu, D_\nu]\varphi = eF_{\mu\nu}(iq\varphi)$, $[D_\mu, D_\nu]\varphi^* = eF_{\mu\nu}(-iq\varphi^*)$ or, in general, $[D_\mu, D_\nu]\Xi = -eF_{\mu\nu}\frac{\Delta\Xi}{\partial\alpha}$ where $\Xi = \varphi$ or φ^*.

As ~~before~~, we find $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Since the covariant derivative was constructed precisely so as to preserve the covariant transformation property of whatever field Ξ it acts upon, it is natural that the commutator $[D_\mu, D_\nu]$ should produce a covariant object. In the case of $U(1)$ transformations, the transformation property of $F_{\mu\nu}$ under local gauge transformations is simply invariance: $\Delta F_{\mu\nu} = \frac{1}{e}(\partial_\mu\partial_\nu\alpha - \partial_\nu\partial_\mu\alpha) = 0$.

Now let's use the covariant derivative commutator to learn something else about the field strength $F_{\mu\nu}$. Consider the triple derivative operation

$$\begin{aligned} D_\mu([D_\nu, D_\rho]\varphi) &= D_\mu(eF_{\nu\rho}(iq\varphi)) \\ &= \partial_\mu(eF_{\nu\rho}(iq\varphi)) + ieqA_\mu eF_{\nu\rho}(iq\varphi) \\ &= e(\partial_\mu F_{\nu\rho})(iq\varphi) + eF_{\nu\rho}(iqD_\mu\varphi) \\ &= e(\partial_\mu F_{\nu\rho})(iq\varphi) + [D_\nu, D_\rho]D_\mu\varphi \end{aligned}$$

So $[D_\mu, [D_\nu, D_\rho]]\varphi = e(\partial_\mu F_{\nu\rho})(iq\varphi)$

or, more generally, $[D_\mu, [D_\nu, D_\rho]]\Xi = e(\partial_\mu F_{\nu\rho})\frac{\Delta\Xi}{\delta\alpha}$, $\Xi = \varphi$ or φ^*

Now use the Jacobi identity:

$$[D_\mu, [D_\nu, D_\rho]] + [D_\nu, [D_\rho, D_\mu]] + [D_\rho, [D_\mu, D_\nu]] \equiv 0$$

So

$$(\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu})e\frac{\Delta\Xi}{\delta\alpha} = 0 \quad \text{and}$$

so, provided Ξ is not itself an invariant, we learn

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0 .$$

One directly verifies that $\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu}$ is totally $[\mu\nu\rho]$ antisymmetric. So this $[\mu\nu\rho]$ antisymmetric set of equations has only $\frac{4\cdot3\cdot2}{3\cdot2\cdot1} = 4$ independent equations. This $[\mu\nu\rho]$ equation is thus equivalent to

$$\partial_\mu F_{\nu\rho}\,\epsilon^{\mu\nu\rho\lambda} \equiv 0 \quad , \quad \text{or} \quad \partial_\mu {}^*F^{\mu\lambda} \equiv 0 \, ,$$

Bianchi identity in GR – actually derived by Ricci-Curbastro

i.e. the sourceless Maxwell equations, which appear here as the "Bianchi identity" for the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Scalar electrodynamics

Now combine the minimally-coupled complex scalar field action with the Maxwell action, but generalizing the single scalar φ to a set φ_i $i = 1 \ldots N$, each with its corresponding $q_{\hat\imath} \in \mathbb{Z}$. We will use the $\hat\imath$ notation to indicate an index value i, which is however excused from Einstein summation.

$\hat\imath$ same value as i

Thus one writes $\varphi_i \to e^{-iq_{\hat\imath}\alpha}\varphi_i$, $\varphi_i^* \to e^{iq_{\hat\imath}\alpha}\varphi_i^*$ in which the $\hat\imath$ does not trigger Einstein summation. The combined action is

$$I_{\substack{\text{scalar}\\ \text{E.D.}}} = \int d^4x \left(-\tfrac{1}{4} F_{\mu\nu}F^{\mu\nu} - D^\mu \varphi_i^* D_\mu \varphi_i - V(\varphi_i, \varphi_i^*) \right)$$

$$i = 1, \ldots, N$$

$$D_\mu \Xi = \partial_\mu \Xi - e A_\mu \frac{\Delta \Xi}{\delta\alpha} \quad \text{where } \Xi \text{ is any } \varphi_i \text{ or } \varphi_i^* \text{ field}$$

In detail, we have $\frac{\Delta\varphi_i}{\delta\alpha} = -iq_{\hat\imath}\varphi_i$, $\frac{\Delta\varphi_i^*}{\delta\alpha} = iq_{\hat\imath}\varphi_i^*$

(again, recall that $\hat\imath$ does not trigger Einstein summation).

Provided $V(\varphi_i, \varphi_i^*)$ is chosen to be $U(1)$ invariant, $I_{\text{scalar E.D.}}$ is fully locally $U(1)$ invariant with parameter $\alpha(x)$. Note that while we have chosen a reducible $U(1)$ representation with N pairs (φ_i, φ_i^*) each with its own $q_{\hat\imath}$, there is only one local $U(1)$ symmetry group with one local parameter $\alpha(x)$.

Consider now the field equations for this coupled system. There are three types of field equations, from varying $\delta\varphi_i$, $\delta\varphi_i^*$ and δA_μ.

Vary independently the $\delta\varphi_i$ and $\delta\varphi_i^*$:

$$\frac{\delta I}{\delta\varphi_i}: \left[\partial^\mu(\partial_\mu - ieq_i A_\mu) - ieq_i A^\mu(\partial_\mu - ieq_i A_\mu)\right]\varphi_i^* - \frac{\partial V}{\partial\varphi_i} = 0$$

this can be regrouped to give simply $D^\mu D_\mu \varphi_i^* - \frac{\partial V}{\partial\varphi_i} = 0$

$$\frac{\delta I}{\delta\varphi_i^*}: \left[\partial^\mu(\partial_\mu + ieq_i A_\mu) + ieq_i A^\mu(\partial_\mu + ieq_i A_\mu)\right]\varphi_i - \frac{\partial V}{\partial\varphi_i^*} = 0$$

which can be regrouped to give $D^\mu D_\mu \varphi_i - \frac{\partial V}{\partial\varphi_i^*} = 0$.

Effectively can integrate D_μ by parts

Note that, although the covariant derivatives $D_\mu\varphi_i$ and $D_\mu\varphi_i^*$ in the action are composed of derivative and A_μ parts which behave differently under $\delta\varphi_i$ or $\delta\varphi_i^*$ variation, the net result of the variation, integration by parts and regrouping is as if the covariant derivative could simply be integrated by parts. This happens because the sum of the $U(1)$ charges for φ_i and φ_i^*, $(eq_i - eq_i)$, adds up zero as is required for invariance of the action.

The third type of field equation comes from varying the Maxwell vector potential A_μ:

$$\frac{\delta I}{\delta A^\nu}: \partial^\mu F_{\mu\nu} + ieq_i \varphi_i^* D_\nu \varphi_i - ieq_i D_\nu\varphi_i^* \varphi_i = 0$$

(N.B. repeated i indices do trigger Einstein summation)

thus $\quad \partial^\mu F_{\mu\nu} = -ieq_i \varphi_i^* \overleftrightarrow{D_\nu} \varphi_i = e J_\nu^{tot}(\varphi_i, \varphi_i^*, A_\mu)$.

This is the Maxwell equation for the coupled system, now including a current source J_ν^{tot} on the RHS. This source generalizes the rigid symmetry's Noether current by the replacement $\partial_\nu \to D_\nu$, as is necessary for the current on the RHS to be a $U(1)$ invariant, matching the $U(1)$ invariance of $\partial^\mu F_{\mu\nu}$ on the LHS.

Now consider what happens if one takes a ∂^ν divergence of the LHS: $\partial^\nu \partial^\mu F_{\mu\nu} \equiv 0$, since $F_{\mu\nu} = F_{[\mu\nu]}$ is antisymmetric. Consistency of the field equation set requires the same ∂^ν divergence of the RHS to vanish as well: one requires $-ieq_i \partial^\nu (\varphi_i^* \overleftrightarrow{D}_\nu \varphi_i) \stackrel{!}{=} 0$, i.e. one requires <u>conservation</u> $\partial^\nu J_\nu^{tot} \stackrel{!}{=} 0$ of the total current on the RHS. Look how this is achieved — one finds

$$-ieq_i \partial^\nu (\varphi_i^* \overleftrightarrow{D}_\nu \varphi_i) = ieq_i (D^\nu D_\nu \varphi_i^*)\varphi_i - ieq_i \varphi_i^* (D^\nu D_\nu \varphi_i) \stackrel{!}{=} 0$$

But using the φ_i and φ_i^* field equations, one has $D^\nu D_\nu \varphi_i^* = \frac{\partial V}{\partial \varphi_i}$, $D^\nu D_\nu \varphi_i = \frac{\partial V}{\partial \varphi_i^*}$, so the RHS of this conservation condition becomes

$$ieq_i \frac{\partial V}{\partial \varphi_i}\varphi_i - ieq_i \varphi_i^* \frac{\partial V}{\partial \varphi_i^*} .$$

Now one needs to recall that $V(\varphi_i, \varphi_i^*)$ must be an <u>invariant</u> potential, and under a $U(1)$ symmetry transformation one has

$$\Delta V = \frac{\partial V}{\partial \varphi_i}\Delta\varphi_i + \frac{\partial V}{\partial \varphi_i^*}\Delta\varphi_i^*$$

$$= \left(-iq_i \frac{\partial V}{\partial \varphi_i}\varphi_i + iq_i \frac{\partial V}{\partial \varphi_i^*}\varphi_i^*\right)\delta\alpha(x)$$

So $\frac{\Delta V}{\delta\alpha} = 0$ implies the necessary $\partial^\nu J_\nu^{tot} = 0$ current conservation.

The construction of couplings between massless spin-one fields and scalar or spinor "matter" fields by ensuring local gauge invariance is motivated on two fronts. One is purely theoretical: this turns out to be the only way to stabilize such couplings against uncontrollable "nonrenormalizable" ultraviolet divergences in quantum field theory. The other is purely experimental: gauge theories have proved to be extraordinarily accurate in describing the phenomena of elementary particle physics.

6. Non-abelian gauge fields

We have so far constructed actions like $\int d^4x\,(-\partial^\mu \varphi^{*a} \partial_\mu \varphi_a)$ which are invariant under rigid $SU(n)$ transformations such as $\varphi_a \to \mathcal{U}_a{}^b \varphi_b$, $\varphi^{*a} \to \varphi^{*b} \mathcal{U}^{\dagger}{}_b{}^a$, $\partial_\mu \mathcal{U}_a{}^b = 0$, for φ_a transforming in the fundamental representation. Now we wish to follow the example set by the $U(1)$ scalar electrodynamics coupling and promote this rigid symmetry to a <u>local</u> symmetry.

One runs into the same sort of problem with $\partial_\mu \varphi_a$ derivatives as in the $U(1)$ Abelian case:

$$\partial_\mu \varphi_a \to \mathcal{U}_a{}^b \partial_\mu \varphi_b + (\partial_\mu \mathcal{U}_a{}^b)\varphi_b ,$$

and the $\partial_\mu \mathcal{U}_a{}^b$ terms spoil the invariance of the action. We cure this as in the $U(1)$ case by introducing an $SU(n)$ <u>gauge field</u> $A_\mu^k(x)$. For $\mathcal{U}_a{}^b(x) = [\exp(i\theta^k(x) T^k)]_a{}^b$, we require the covariant derivative

$$D_\mu \varphi_a = \partial_\mu \varphi_a - ig A_\mu{}^k (T^k)_a{}^b \varphi_b$$

to transform covariantly, i.e. $D_\mu \varphi_a \to \mathcal{U}_a{}^b D_\mu \varphi_b$. The constant g is the <u>gauge coupling constant</u>, analogous to the electric charge unit e in the scalar electrodynamics construction. In order to achieve $D_\mu \varphi \to \mathcal{U} D_\mu \varphi$ with $\varphi \to \mathcal{U}\varphi$, we require, for $A_\mu^k(x) \to A'^k_\mu(x)$, that

$$-ig A'^k_\mu (T^k)_a{}^b \mathcal{U}_b{}^c \varphi_c = -ig\, \mathcal{U}_a{}^b A_\mu{}^k (T^k)_b{}^c \varphi_c - (\partial_\mu \mathcal{U}_a{}^b)\varphi_b$$

• Now, we require this to hold for <u>any</u> $\varphi_a(x)$ transforming in the standard $\varphi_a \to \mathcal{U}_a{}^b \varphi_b$ covariant fashion. It then follows that

$$A'^k_\mu (T^k)_a{}^b = \mathcal{U}_a{}^d A_\mu{}^k (T^k)_d{}^e\, \mathcal{U}^{-1}{}_e{}^b - \frac{i}{g}(\partial_\mu \mathcal{U}_a{}^c)\, \mathcal{U}^{-1}{}_c{}^b$$

For $SU(n)$, recall $\mathcal{U}^{-1} = \mathcal{U}^{\dagger}$

$$= \Big[\underbrace{\mathcal{U}(A_\mu{}^k T^k}_{\text{homogeneous term in } A_\mu{}^k} \underbrace{- \frac{i}{g}\mathcal{U}^{-1}\partial_\mu \mathcal{U})\,\mathcal{U}^{-1}}_{\text{inhomogeneous term}}\Big]_a{}^b$$

As written, the transformation $A_\mu^k(x) \to A_\mu'^k(x)$ appears to depend on the specific choice of representation T^k for the matter field $\psi(x)$ acted upon in the covariant derivative.

$\Delta A_\mu^k = A_\mu'^k - A_\mu^k$

Let's investigate this for an infinitesimal transformation: $\theta^k(x) \to \delta\theta^k(x)$. Then $T^k \Delta A_\mu^k = \frac{1}{g}\partial_\mu \delta\theta^k T^k + i\left[\delta\theta^k T^k, A_\mu^r T^r\right]$, and we can rewrite this using the Lie algebra $[T^k, T^r] = i f^{krm} T^m$, so $T^k \Delta A_\mu^k = \frac{1}{g}\partial_\mu \delta\theta^k T^k - \delta\theta^k A_\mu^r f^{krm} T^m$. Then, since the T^k Lie algebra generators are linearly independent (for a non-trivial representation), one has, after changing summation indices,

$$\Delta A_\mu^i = \frac{1}{g}\partial_\mu \delta\theta^i - f^{ijk}\delta\theta^j A_\mu^k + O(\delta\theta^2).$$

The second term here, linear in A_μ^k, is precisely of the form that one would expect for a field transforming in the adjoint representation. Recalling that $(T_{adj}^i)^{jk} = -i f^{ijk}$, one can rewrite the infinitesimal A_μ^i gauge transformation as

$$\Delta A_\mu^i = \frac{1}{g}\partial_\mu \delta\theta^i + i\,\delta\theta^j (T_{adj}^j)^{ik} A_\mu^k .$$

The same is true for finite rigid transformations: if $\partial_\mu \theta^i = 0$, then $A_\mu^k T^k \to \mathcal{U}(A_\mu^k T^k)\mathcal{U}^+$, which is a finite linear adjoint representation of A_μ^k.

Thus, the homogeneous term in $A_\mu'^k$ corresponds to a finite adjoint representation transformation, but the inhomogeneous term involving $\partial_\mu \mathcal{U}$ changes the overall A_μ^k gauge field transformation to a different, nonlinear type.

Despite the inhomogeneous character of the A_μ^k transformation, this is still a good group realization. In order for a transformation to realize a group action, the key requirement is that it respect the group composition law. Let us thus consider a group element combination $\mathcal{U}(\theta''(x)) = \mathcal{U}(\theta'(x))\,\mathcal{U}(\theta(x))$, local at the spacetime point x^μ. We then also have $\mathcal{U}^{-1}(\theta'') = \mathcal{U}^{-1}(\theta)\,\mathcal{U}^{-1}(\theta')$.

Letting $A_\mu \cdot T = A_\mu{}^k T^k$, etc, we have first

$$A'_\mu \cdot T = \mathcal{U}(\theta)\left[A_\mu \cdot T - \frac{i}{g}\mathcal{U}^{-1}(\theta)\partial_\mu \mathcal{U}(\theta)\right]\mathcal{U}^{-1}(\theta)$$ and then

$$A''_\mu \cdot T = \mathcal{U}(\theta')\left[A'_\mu \cdot T - \frac{i}{g}\mathcal{U}^{-1}(\theta')\partial_\mu \mathcal{U}(\theta')\right]\mathcal{U}^{-1}(\theta')$$

inserting for $A'_\mu \cdot T$

$$= \mathcal{U}(\theta')\left[\mathcal{U}(\theta)\left(A_\mu \cdot T - \frac{i}{g}\mathcal{U}^{-1}(\theta)\partial_\mu \mathcal{U}(\theta)\right)\mathcal{U}^{-1}(\theta) - \frac{i}{g}\mathcal{U}^{-1}(\theta')\partial_\mu \mathcal{U}(\theta')\right]\mathcal{U}^{-1}(\theta')$$

$$= \mathcal{U}(\theta'')\left[A_\mu \cdot T - \frac{i}{g}\mathcal{U}^{-1}(\theta'')\partial_\mu \mathcal{U}(\theta'')\right]\mathcal{U}^{-1}(\theta'') .$$

Thus, although inhomogenous and hence nonlinear, the transformation $A_\mu^k \to A'^k_\mu$ still respects the group composition law and hence forms a good group realization.

Non-abelian field strength

As in the Abelian $\mathcal{U}(1)$ case, the commutator of covariant derivatives yields a covariant quantity antisymmetric in $[\mu\nu]$ spacetime indices, but now carrying a gauge group adjoint representation. Applying $[D_\mu, D_\nu]$ to a field $\varphi_a(x)$ produces by construction a covariant quantity,

$$\left([D_\mu, D_\nu]_a{}^b \varphi_b\right)' = \mathcal{U}_a{}^b\, [D_\mu, D_\nu]_b{}^c \varphi_c ,$$ so

$[D_\mu, D_\nu]_b{}^c$ must itself be some covariant quantity generalizing the Maxwell field strength. Working directly from the definition $(D_\mu \varphi)_a = ((\partial_\mu - ig A_\mu \cdot T)\varphi)_a$, one finds

$$([D_\mu, D_\nu]\varphi)_a = -ig\left(\partial_\mu A_\nu^i - \partial_\nu A_\mu^i + g f^{ijk} A_\mu^j A_\nu^k\right)(T^i)_a{}^b \varphi_b$$

$$= -g\left(\partial_\mu A_\nu^i - \partial_\nu A_\mu^i + g f^{ijk} A_\mu^j A_\nu^k\right)\frac{\Delta \varphi_a}{\delta \theta^i}$$

$\frac{\Delta \varphi_a}{\delta \theta^i} = i (T^i)_a{}^b \varphi_b$

So we define the non-abelian, or Yang-Mills, field strength

$$F_{\mu\nu}{}^i = \partial_\mu A_\nu^i - \partial_\nu A_\mu^i + g f^{ijk} A_\mu^j A_\nu^k$$

The field strength $F_{\mu\nu}{}^i$ carries an adjoint representation index and transforms covariantly in the adjoint rep. Unlike the gauge field $A_\mu{}^i$, which transforms nonlinearly and inhomogeneously, the field strength $F_{\mu\nu}{}^i$ is a standard tensorial quantity with a linear transformation law. For an infinitesimal gauge transformation with parameter $\delta\theta^k(x)$, the transformation of $F_{\mu\nu}{}^i$ is $\Delta F_{\mu\nu}{}^i = i\,\delta\theta^k T_{adj}^{k\,i}{}_j F_{\mu\nu}{}^j = \delta\theta^k f^{ki}{}_j F_{\mu\nu}{}^j$ which is linear in $F_{\mu\nu}{}^i$.

$T_{adj}^{k\,i}{}_j = -i f^{k\,i}{}_j$

• Note a key difference between the Abelian $U(1)$ and the non-abelian field strengths. The Abelian $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is invariant, while the non-abelian $F_{\mu\nu}{}^i$ transforms covariantly in the adjoint representation. Actually, the invariance of the Abelian $F_{\mu\nu}$ can be expressed as a special case of covariance, since for an Abelian group the f^{ijk} structure constants all vanish, so the adjoint rep. is just the singlet, invariant representation – for $U(1)$, the adjoint generator vanishes.

Yang-Mills action and field equations

Using $F_{\mu\nu}{}^i$, one straightforwardly constructs an invariant action for the gauge fields of an arbitrary Lie group G:

$$I_{YM} = \int d^4x \left(-\frac{1}{4} F_{\mu\nu}{}^i F^{i\mu\nu}\right)$$

note: sum on i gives an invariant – in adjoint rep., δ^{ij} is a numerically invariant tensor

For $SU(n)$ $(n \geq 2)$ gauge groups, a standard normalization of the fundamental representation generators T^i is such that $\mathrm{tr}(T^i T^j) = \frac{1}{2}\delta^{ij}$, so another common way of writing the Yang-Mills action is in terms of $F_{\mu\nu} = F_{\mu\nu}{}^i T^i$, for which

$$I_{YM} = -\frac{1}{2}\,\mathrm{tr}\int d^4x\, F_{\mu\nu} F^{\mu\nu}.$$

In varying the Yang-Mills action to produce the gauge field equations, it is convenient first to consider the variation of $F_{\mu\nu}{}^i$ resulting from a general variation $\delta A_\mu{}^i$ of the gauge field:

$$\delta F_{\mu\nu}{}^i = \partial_\mu \delta A_\nu{}^i - \partial_\nu \delta A_\mu{}^i + g f^{ijk} \delta A_\mu{}^j A_\nu{}^k + g f^{ijk} A_\mu{}^j \delta A_\nu{}^k.$$

Recognizing once again $T_{adj}{}^{kij} = -i f^{kij}$, one can write

$$\begin{aligned} D_\mu \delta A_\nu{}^i &= \partial_\mu \delta A_\nu{}^i - i g A_\mu{}^k T_{adj}{}^{kij} \delta A_\nu{}^j \\ &= \partial_\mu \delta A_\nu{}^i - g A_\mu{}^k f^{kij} \delta A_\nu{}^j \\ &= \partial_\mu \delta A_\nu{}^i + g f^{ijk} A_\mu{}^j \delta A_\nu{}^k \end{aligned}$$

So one can write

"Palatini identity" $\delta F_{\mu\nu}{}^i = D_\mu \delta A_\nu{}^i - D_\nu \delta A_\mu{}^i$, where $\delta A_\mu{}^i$ is treated as an adjoint-representation tensorial quantity.

Why has this covariant derivative structure appeared? The answer is that $\delta A_\mu{}^i$ actually <u>is</u> a linearly transforming tensorial quantity. Consider the infinitesimal gauge variations of $A_\mu{}^i$ and $(A_\mu{}^i + \delta A_\mu{}^i)$:

$$\Delta A_\mu{}^i = \frac{1}{g} \partial_\mu \delta\theta^i - f^{ijk} \delta\theta^j A_\mu{}^k \quad + \mathcal{O}(\delta\theta^2)$$

$$\Delta(A_\mu{}^i + \delta A_\mu{}^i) = \frac{1}{g} \partial_\mu \delta\theta^i - f^{ijk} \delta\theta^j (A_\mu{}^k + \delta A_\mu{}^k) + \mathcal{O}(\delta\theta^2)$$

So $\Delta \delta A_\mu{}^i = \Delta(A_\mu{}^i + \delta A_\mu{}^i) - \Delta A_\mu{}^i = - f^{ijk} \delta\theta^j \delta A_\mu{}^k + \mathcal{O}(\delta\theta^2)$

i.e. the infinitesimal variation $\delta A_\mu{}^i$ transforms <u>homogeneously</u> as an adjoint representation quantity, since the inhomogeneous piece in the $A_\mu{}^i$ transformation cancels out. Thus, it is completely appropriate to construct $D_\mu \delta A_\nu{}^i$ as an adjoint representation covariant derivative.

Now vary I_{YM} under a general gauge-field variation $\delta A_\mu{}^i$:

$$\begin{aligned} \delta I_{YM} &= -\frac{1}{2} \int d^4x\, F^{i\mu\nu} (D_\mu \delta A_\nu{}^i - D_\nu \delta A_\mu{}^i) \\ &= - \int d^4x\, F^{i\mu\nu} D_\mu \delta A_\nu{}^i \end{aligned}$$

Strictly speaking, one should now pull apart $D_\mu \delta A_\nu^i$ into its $\partial_\mu \delta A_\nu^i$ and $g f^{ijk} A_\mu^j \delta A_\nu^k$ and integrate the $\partial_\mu \delta A_\nu^i$ term by parts. However, just as we found in scalar electrodynamics, when one does this and then recombines terms, the net result is <u>as if</u> one had simply integrated D_μ by parts:

$$\delta I_{YM} = \int d^4x \, D_\mu F^{i\mu\nu} \delta A_\nu^i$$

Thus, requiring $\delta I_{YM} = 0$ for δA_ν^i subject to the usual initial & final time and asymptotic $r \to \infty$ restrictions, one finds the Yang-Mills field equation, so far without sources:

really should write $(D_\mu F^{\mu\nu})^i$, etc.

$$D_\mu F^{i\mu\nu} = 0 \qquad D_\mu: \text{adjoint cov. der.}$$

We shall in due course supply a matter sector that will yield a non-abelian current for the RHS of this equation. But we can already investigate what kind of conservation condition this current should have. Taking an ordinary partial derivative ∂_ν of the LHS doesn't yield anything special. But look what happens if we take a <u>covariant</u> divergence:

$\Delta F^{i\mu\nu} = -f^{ijk} \delta\theta^j F^{k\mu\nu}$

$$\begin{aligned} D_\nu D_\mu F^{i\mu\nu} &= -\tfrac{1}{2} [D_\mu, D_\nu] F^{i\mu\nu} \\ &= \frac{g}{2} F_{\mu\nu}{}^j (-f^{ijk}) F^{k\mu\nu} \\ &= \frac{-g}{2} f^{ijk} F_{\mu\nu}{}^j F^{k\mu\nu} \\ &\equiv 0 \quad \text{since } F_{\mu\nu}{}^j F^{k\mu\nu} \text{ is } (jk) \text{ symmetric} \\ &\qquad \text{but } f^{ijk} \text{ is } [jk] \text{ antisymmetric} \end{aligned}$$

Thus, source currents on the RHS of Yang-Mills field equations will need to be <u>covariantly conserved</u>.

Non-abelian Bianchi identity

Now repeat the analysis we made of the $U(1)$ field strength stemming from the covariant derivative Jacobi identity, but in the non-abelian case: Consider $[D_\rho,[D_\mu,D_\nu]]$ + cycle terms. Write $(D_\mu)_A{}^B = \partial_\mu \delta_A{}^B - ig A_\mu\cdot T_A{}^B$ for $(T^i)_A{}^B$ in some arbitrary representation of the gauge group G. So acting on a field Ξ_A, one has $(D_\mu)_A{}^B \Xi_B = \partial_\mu \Xi_A - ig A_\mu\cdot T_A{}^B \Xi_B$. Now consider

$$\partial_\rho\left([D_\mu,D_\nu]_A{}^B \Xi_B\right) = \partial_\rho\left([D_\mu,D_\nu]_A{}^B\right)\Xi_B + [D_\mu,D_\nu]_A{}^B \partial_\rho(\Xi_B)$$

So $[\partial_\rho,[D_\mu,D_\nu]]_A{}^B \Xi_B = \partial_\rho\left([D_\mu,D_\nu]_A{}^B\right)\Xi_B$.

Then, since the field Ξ_B is arbitrary, one must have

$$[\partial_\rho,[D_\mu,D_\nu]]_A{}^B = \partial_\rho\left([D_\mu,D_\nu]_A{}^B\right).$$

Thus, $$[D_\rho,[D_\mu,D_\nu]]_A{}^B = -ig\,\partial_\rho(F_{\mu\nu}\cdot T)_A{}^B - g^2 A_\rho\cdot T_A{}^D F_{\mu\nu}\cdot T_D{}^B + g^2 F_{\mu\nu}\cdot T_A{}^D A_\rho\cdot T_D{}^B$$

$$= -ig\,\partial_\rho(F_{\mu\nu}{}^j)(T^j)_A{}^B - g^2 A_\rho{}^k F_{\mu\nu}{}^j [T^k,T^j]_A{}^B$$

$$= -ig\,\partial_\rho(F_{\mu\nu}{}^j)(T^j)_A{}^B - ig^2 A_\rho{}^k F_{\mu\nu}{}^j f^{kjl}(T^l)_A{}^B$$

$$= -ig\, D_\rho(F_{\mu\nu}{}^j)(T^j)_A{}^B$$

D_ρ in adjoint

Hence, the Jacobi identity $[D_\rho,[D_\mu,D_\nu]] + [D_\mu,[D_\nu,D_\rho]] + [D_\nu,[D_\rho,D_\mu]] \equiv 0$ together with linear independence of the generators $(T^i)_A{}^B$ imply the non-abelian Bianchi identity

$$D_\rho(F_{\mu\nu}{}^j) + D_\mu(F_{\nu\rho}{}^j) + D_\nu(F_{\rho\mu}{}^j) \equiv 0 .$$

As in the Abelian $U(1)$ case, the LHS is totally $[\rho\mu\nu]$ antisymmetric, so the number of independent $[\rho\mu\nu]$ equations is $\frac{4.3.2}{3!} = 4$. The $[\rho\mu\nu]$ equation is equivalent to $\frac{1}{2}\epsilon^{\sigma\rho\mu\nu} D_\rho(F_{\mu\nu}{}^j) \equiv 0$, i.e. $D_\rho(*F^{j\sigma\rho}) \equiv 0$.

Unlike the Abelian field strength identity $\partial_\rho(*F^{\sigma\rho}) \equiv 0$ for $U(1)$, the non-abelian Bianchi identity can only be written using the gauge field $A_\rho{}^j$, which appears in D_ρ.

Non-abelian matter coupling

It is now straightforward to make gauge-coupled matter models for scalar fields φ_A transforming in some representation (possibly reducible) $(T^i)_A{}^B$ of the local symmetry group G: just generalize the $U(1)$ minimal coupling procedure, $\partial_\mu \varphi_A \to D_\mu \varphi_A$. For complex scalar fields $(\varphi_A, \varphi^{*A})$, we have

$$D_\mu \varphi_A = \partial_\mu \varphi_A - ig A_\mu \cdot T_A{}^B \varphi_B \,, \quad D_\mu \varphi^{*A} = \partial_\mu \varphi^{*A} + ig\varphi^{*B} A_\mu \cdot T_B{}^A$$

and one can construct a locally G-invariant matter action:

$$I_{\text{matter}} = \int d^4x \left(-D_\mu \varphi_A D^\mu \varphi^{*A} - V(\varphi_C, \varphi^{*D})\right)$$

where $V(\varphi_C, \varphi^{*D})$ is some G-invariant potential.

In writing this action and set of covariant derivatives, we are considering a model with a <u>simple</u> gauge group G, for which there is just one coupling constant g. For non-simple product groups $G_1 \times G_2 \times \cdots$, one can have a different coupling constant g_i for each simple G_i group factor, since symmetry transformations of such a direct product can never rotate between reps of the distinct simple factors.

Now vary. The matter field equations are

$$\delta I_{\text{matter}}/\delta\varphi_A : \; D^\mu D_\mu \varphi^{*A} - \frac{\partial V}{\partial \varphi_A} = 0 \qquad \delta I_{\text{matter}}/\delta\varphi^{*A} : \; D^\mu D_\mu \varphi_A - \frac{\partial V}{\partial \varphi^{*A}} = 0$$

where D_μ is of the appropriate form for φ^{*A} or φ_A.

Vary with respect to $A_\nu{}^i$ after including also I_{YM} gives:

$$\delta/\delta A_\nu{}^i (I_{YM} + I_{\text{matter}}) : \; D_\mu F^{i\mu\nu} = ig\left(\varphi^{*B} (T^i)_B{}^A D^\nu \varphi_A - D^\nu \varphi^{*A} (T^i)_A{}^B \varphi_B\right)$$

$$(D_\mu \text{ in adjoint here}) \qquad = g J^{i\nu}(\varphi, \varphi^*, A)$$

where $J^{i\nu}(\varphi, \varphi^*, A) = i(\varphi^{*B} \overleftrightarrow{D}^\nu \varphi_A)(T^i)_B{}^A$ is the full non-abelian current. As we have seen, the covariant identity $D_\nu D_\mu F^{i\mu\nu} \equiv 0$ requires <u>covariant conservation</u> $D_\nu J^{i\nu} = 0$ of the full current, by virtue of the matter field equations and the G-invariance of the scalar field potential $V(\varphi_C, \varphi^{*D})$.

7. Spontaneous Symmetry Breaking of rigid symmetries

Return once more to our Abelian $SO(2) \cong U(1)$ scalar doublet model, but choose now to give $\Phi_{a=1,2}$ different names σ and π

$$I_{SO(2)} = \int d^4x \left(-\tfrac{1}{2}\partial_\nu \sigma \partial^\nu \sigma - \tfrac{1}{2}\partial_\nu \pi \partial^\nu \pi - V(\sigma^2+\pi^2) \right).$$

For the $SO(2)$ symmetric potential V, pick

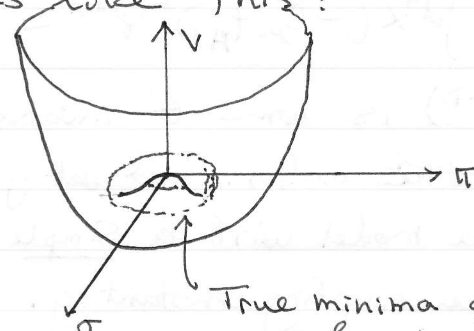
"wrong sign" μ^2 term

$$V(\sigma^2+\pi^2) = -\frac{\mu^2}{2}(\sigma^2+\pi^2) + \frac{\lambda}{4}(\sigma^2+\pi^2)^2 \quad \text{with } \mu^2 \text{ and } \lambda > 0.$$

This potential looks like this:

"wine bottle" or "Mexican hat"

The point $(\sigma,\pi)=(0,0)$ is an extremum but not a minimum.



The potential V decreases as one goes away from (0,0).

True minima of V lie on a circle away from (0,0).

The potential V is extremized when

$$\frac{\partial V}{\partial \sigma} = \sigma\left(-\mu^2 + \lambda(\sigma^2+\pi^2)\right) = 0 \,, \quad \frac{\partial V}{\partial \pi} = \pi\left(-\mu^2 + \lambda(\sigma^2+\pi^2)\right) = 0$$

So the extrema are at $(\sigma,\pi)=(0,0)$ (a local maximum) and on the circle in (σ,π) space where $\sigma^2+\pi^2 = v^2$, $v = (\mu^2/\lambda)^{1/2}$. Thus, the true minima of the potential V are degenerate, lying somewhere on this circle. However, the system must fall into a vacuum state corresponding to some particular point on this circle, however arbitrary. The vacuum value of (σ,π) must be constant in x^ν. The total energy is

$$E = \int d^3x \left(\tfrac{1}{2}(\dot\sigma)^2 + \tfrac{1}{2}(\dot\pi)^2 + \tfrac{1}{2}\partial_i\sigma\partial_i\sigma + \tfrac{1}{2}\partial_i\pi\partial_i\pi + V(\sigma,\pi) \right)$$

gradient terms: sum of squares, same sign

So the gradient terms will cause $E > E_{min}$ unless $\partial_\nu \sigma = 0 = \partial_\nu \pi$.

Thus the vacuum must correspond to (σ,π) at some point on the circle minimizing V, but precisely which point is not determined by the field equations.

This situation is known as <u>spontaneous symmetry breaking</u>: although the action and field equations respect the SO(2) symmetry, the choice of vacuum breaks it.

— For sake of definiteness, let the vacuum values be $\bar{\sigma} = v = (\mu^2/\lambda)^{1/2}$, $\bar{\pi} = 0$. One can then analyse the dynamics of fields about this vacuum by defining a shifted field $\tilde{\sigma} = \sigma - \bar{\sigma} = \sigma - v$. In terms of $\tilde{\sigma}$ and π, one has

$$I_{\text{shifted}} = \int d^4x \left(-\tfrac{1}{2}\partial_\nu\tilde{\sigma}\partial^\nu\tilde{\sigma} - \tfrac{1}{2}\partial_\nu\pi\partial^\nu\pi - V(\tilde{\sigma},\pi)\right)$$

where $V(\tilde{\sigma},\pi) = \mu^2(\tilde{\sigma})^2 + \lambda v\tilde{\sigma}(\tilde{\sigma}^2+\pi^2) + \frac{\lambda}{4}(\tilde{\sigma}^2+\pi^2)^2 - \frac{\mu^4}{4\lambda}$.

The constant $(-\frac{\mu^4}{4\lambda})$ gives the minimum value for V, but will not affect the field equations and may be disregarded.

— Note the change in sign of the quadratic $\tilde{\sigma}^2$ term: this now corresponds to a regular mass $\tilde{m} = \sqrt{2\mu^2}$ (a mass term would be $\frac{\tilde{m}^2}{2}\tilde{\sigma}^2$, so $\frac{\tilde{m}^2}{2} = \mu^2$), unlike the "tachyonic" $-\frac{\mu^2}{2}\sigma^2$ mass term about the $(\sigma,\pi) = (0,0)$ extremum.

— The other characteristic feature of the system of fields about the $(\sigma,\pi) = (v,0)$ vacuum is the <u>disappearance</u> of the mass for the π field: after the vacuum shift, π becomes <u>massless</u>. The massless π mode corresponds to fluctuations away from the $(\sigma,\pi) = (v,0)$ vacuum around the circle of equivalent vacua, in which direction $V(\sigma,\pi)$ is <u>flat</u>. Of course, the $(\tilde{\sigma},\pi)$ system does have higher-order terms in the potential V, corresponding to 3-point and 4-point interactions, but the mass is identified just from the quadratic term.

π mode is massless

This appearance of a massless mode after shifting the vacuum to a true minimum of the potential is an instance of <u>Goldstone's Theorem</u>: every symmetry generator, of a theory with a rigid symmetry, that is <u>broken</u> by the choice of a vacuum corresponds to a <u>massless</u> field with respect to that vacuum.

Non-abelian example

Consider now an $SO(n)$ invariant model with n real scalar fields ϕ_a, $a=1,\dots,n$ and the action

$$\int d^4x \left(-\tfrac{1}{2}\partial_\mu\phi_a\partial^\mu\phi_a + \tfrac{\mu^2}{2}\phi_a\phi_a - \tfrac{\lambda}{4}(\phi_a\phi_a)^2\right)$$

where again μ^2 and $\lambda > 0$. The minimum of the potential $V = -\frac{\mu^2}{2}\phi_a\phi_a + \frac{\lambda}{4}(\phi_a\phi_a)^2$ is at points $\bar\Phi_a$ where $\bar\Phi_a\bar\Phi_a = \mu^2/\lambda$.

Hence the surface in ϕ_a target space where V is minimized is an $(n-1)$ sphere $S^{(n-1)}$. Again, the system must fall spontaneously into some vacuum on this surface, but the particular point on the $S^{(n-1)}$ sphere is arbitrary. About the vacuum point $\phi_a = \bar\Phi_a$ there is still an unbroken subgroup of $SO(n)$ that leaves the point $\bar\Phi_a$ invariant. This subgroup is $SO(n-1)$ (generalizing the $SO(2)$ subgroup of $SO(3)$ leaving, eg, the north pole of an S^2 invariant). For convenience, think of aligning the $\bar\Phi_a$ vacuum with the $\hat n$ basis vector. The unbroken subgroup $SO(n-1)$ then rotates among the $(n-1)$ directions where $\hat n$ has vanishing components, corresponding to $a=1,2,\dots,n-1$.

– Count the number of <u>broken</u> generators of $SO(n)$, i.e. those that do <u>not</u> leave $\bar\Phi_a$ invariant:

$$\dim(SO(n)) - \dim(SO(n-1)) = \tfrac{1}{2}n(n-1) - \tfrac{1}{2}(n-1)(n-2) = n-1.$$

Now consider the $(\text{mass})^2$ matrix for fluctuations about the $\bar\Phi_a$ vacuum. This is given by the matrix of second partial derivatives of V, evaluated at $\phi_a = \bar\Phi_a$:

$$M^2_{ab} = \frac{\partial^2 V(\phi)}{\partial\phi_a\partial\phi_b}\Big|_{\phi=\bar\Phi} = -\mu^2\delta_{ab} + \lambda\delta_{ab}\bar\Phi_c\bar\Phi_c + 2\lambda\bar\Phi_a\bar\Phi_b$$

$$= 2\lambda\bar\Phi_a\bar\Phi_b \quad \text{since} \quad \bar\Phi_c\bar\Phi_c = \mu^2/\lambda.$$

This has just one eigenvector, $\bar\Phi_a$, that has a non-zero eigenvalue $m^2 = 2\lambda\bar\Phi_a\bar\Phi_a = 2\mu^2 > 0$. The other $(n-1)$ eigenvectors have eigenvalue zero: number of <u>zero modes</u> equals the number of <u>broken generators</u>.

Goldstone's Theorem

Consider a theory with n real scalar fields invariant under a rigid symmetry group G. Even for a complex G representation, one can always do this, by considering real and imaginary parts of the scalar fields separately. Let the Lagrangian be

$a=1,\dots,n$ $$\mathcal{L} = -\tfrac{1}{2}\partial_\mu\phi_a\partial^\mu\phi_a - V(\phi)$$ where $V(\phi)$ is a polynomial in the scalar fields ϕ_a which is G-invariant:

$$\Delta V = \frac{\partial V}{\partial\phi_a}\Delta\phi_a = 0$$ for infinitesimal G-transformations $\Delta\phi_a$.

Let the Lie algebra generators for the representation carried by the ϕ_a be $(T^i)_{ab}$. Note that we are not bothering to raise & lower indices here because we are considering this to be a real representation – e.g. the complex $\underline{3}$ of $SU(3)$ may be considered as a real 6-vector.

Then $\Delta\phi_a = i\,\delta\theta^i (T^i)_{ab}\phi_b$, $i=1,2,\dots,\dim(G)$.

Hence $\delta\theta^i\left(\frac{\partial V}{\partial\phi_a}(T^i)_{ab}\phi_b\right) = 0$ for arbitrary infinitesimal $\delta\theta^i$.

Consequently one must have $\frac{\partial V}{\partial\phi_a}(T^i)_{ab}\phi_b = 0$ for all $i=1,\dots,\dim G$.

Differentiate $\partial/\partial\phi_c$:

$$\frac{\partial^2 V}{\partial\phi_c\partial\phi_a}(T^i)_{ab}\phi_b + \frac{\partial V}{\partial\phi_a}(T^i)_{ac} = 0,$$

then evaluate at a field value $\bar\Phi_a$ that minimizes V:

$$\left(\frac{\partial^2 V}{\partial\phi_c\partial\phi_a}\right)_{|\phi=\bar\Phi}(T^i)_{ab}\bar\Phi_b = 0 \qquad \text{since} \left(\frac{\partial V}{\partial\phi_a}\right)_{|\phi=\bar\Phi} = 0.$$

Now, one identifies the $(\text{mass})^2$ matrix from an expansion of V about the vacuum $\phi_a = \bar\Phi_a$:

$$V = \text{const.} + \tfrac{1}{2}M^2_{ab}(\phi_a-\bar\Phi_a)(\phi_b-\bar\Phi_b) + \text{higher order terms}.$$

Note that there is no linear term in $(\phi_a-\bar\Phi_a)$ since V is minimized at $\bar\Phi_a$.

Thus, $\left(\frac{\partial^2 V}{\partial\phi_a\partial\phi_b}\right)_{|\phi=\bar\Phi} = M^2_{ab}$ and hence $M^2_{ab}(T^i)_{bc}\bar\Phi_c = 0$

$a,b=1,\dots,n$; $i=1\dots\dim G$

Now let the stability subgroup (or little group) of the vacuum $\bar\Phi_a$ be denoted H, with Lie algebra generators $(t^{\hat i})_{ab}$, $\hat i = 1,2,\dots,\dim H$, $H\subset G$.

The stability subgroup generators $(t^{\hat{i}})_{ab}$ are distinguished among the general $(T^i)_{ab}$ of G by the vacuum stability condition

$(t^{\hat{i}})_{ab}\bar{\Phi}_b = 0$: H-action leaves $\bar{\Phi}_a$ invariant.

Now, the condition $M^2_{ab}(T^i)_{bc}\bar{\Phi}_c = 0$ clearly does not give any information about M^2_{ab} for $(T^i)_{bc} = (t^{\hat{i}})_{bc}$, $i = \hat{i}$ dim H values. However, for each nonvanishing vector $(T^i)_{bc}\bar{\Phi}_c$, we learn from $M^2_{ab}(T^i)_{bc}\bar{\Phi}_c = 0$ that M^2_{ab} has a zero eigenvalue.

We now need to count the number of independent nonvanishing vectors $(T^i)_{bc}\bar{\Phi}_c$. Note that for the representation $(T^i)_{ab}$ written in real form, the matrices $i(T^i)_{ab}$ must be real. But $(T^i)_{ab}$ must also be Hermitean for the compact groups G that we consider. Hence, $(T^i)_{ab}$ must be purely antisymmetric $(T^i)_{ab} = -(T^i)_{ba}$. Now define $S^{ij} = (T^i\bar{\Phi}, T^j\bar{\Phi})$, where (v,w) is the usual inner product, $(v,w) = v^*_a w_a = v_a w_a$ for real v & w.

$T^i = T^{i\dagger}$ Since T^i is Hermitean, $S^{ij} = (\bar{\Phi}, T^i T^j \bar{\Phi})$, so

$S^{ij} - S^{ji} = (\bar{\Phi}, [T^i, T^j]\bar{\Phi}) = i f^{ijk}(\bar{\Phi}, T^k\bar{\Phi})$. But $(\bar{\Phi}, T^k\bar{\Phi}) = 0$ since $(T^k)_{ab} = -(T^k)_{ba}$. So S^{ij} is symmetric.

Now, any real symmetric matrix S^{ij} can be diagonalized by an orthogonal transformation, yielding a diagonal $\tilde{S}^{ij} = (O^T S O)^{ij}$. The nonzero diagonal components (eigenvalues) of $\tilde{S}$ correspond to independent nonvanishing $T^i\bar{\Phi}$. As for the zero eigenvalues, suppose $\tilde{S}^{ij}d^j = 0$. Then $d^i\tilde{S}^{ij}d^j = (\tilde{d}^i T^i\bar{\Phi}, \tilde{d}^j T^j\bar{\Phi}) = 0$, where $\tilde{d}^j = O^{jk}d^k$.

Rank of S^{ij} is dim G − dim H

Hence every zero eigenvalue of S^{ij} corresponds to a combination $\tilde{d}^i T^i$ of generators that are in the stability subgroup H, because the positive-definiteness of the (v,w) inner product then implies $\tilde{d}^i T^i\bar{\Phi} = 0$. So the independent nonvanishing $T^i\bar{\Phi}$ span a space of dimension $\dim(G) - \dim(H)$, the complement to the space annihilated by S^{ij}.

This is Goldstone's Theorem: the $(\text{mass})^2$ matrix M^2_{ab} has $\dim(G) - \dim(H)$ zero eigenvalues corresponding to the $\dim(G) - \dim(H)$ dimensional vector space spanned by the nonvanishing $(T^i)_{ab}\bar{\Phi}_b$. Thus, for each broken generator $\tilde{d}^i T^i$ of G, there is a massless mode.