Problem 1(a):
In matrix notations, the non-abelian gauge symmetries act on vector potentials $A_\mu(x)$ according to

$$A'_\mu(x) = U(x)A_\mu(x)U^\dagger(x) + i\partial_\mu U(x)U^\dagger(x). \quad (S.1)$$

Taking

$$F_{\mu\nu}(x) \overset{\text{def}}{=} \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) + i[A_\mu(x), A_\nu(x)] \quad (S.2)$$

as the definition of the tension fields $F_{\mu\nu}(x)$, we then have

$$F'_{\mu\nu}(x) = \partial_\mu A'_\nu(x) - \partial_\nu A'_\mu(x) + i[A'_\mu(x), A'_\nu(x)], \quad (S.3)$$

whatever that evaluates to. For the first term on the right hand side, we have

$$\partial_\mu A'_\nu = \partial_\mu \left(U A_\nu U^\dagger + i(\partial_\nu U)U^\dagger\right)$$

$$= U(\partial_\mu A_\nu)U^\dagger + \left[(\partial_\mu U)U^\dagger, U A_\nu U^\dagger\right] + i(\partial_\mu \partial_\nu U)U^\dagger - i(\partial_\nu U)U^\dagger \times (\partial_\mu U)U^\dagger \quad (S.4)$$

since

$$\partial_\mu \left(U A_\nu U^\dagger\right) = U(\partial_\mu A_\nu)U^\dagger + (\partial_\mu U)A_\nu U^\dagger - U A_\nu (\partial_\mu U^\dagger)$$

$\langle$ using $U^\dagger U = 1$ and $\partial_\mu U^\dagger = -U^\dagger(\partial_\mu U)U^\dagger$$\rangle$

$$= U(\partial_\mu A_\nu)U^\dagger + (\partial_\mu U)U^\dagger A_\nu U^\dagger - U A_\nu U^\dagger(\partial_\mu U)U^\dagger$$

$$= U(\partial_\mu A_\nu)U^\dagger + \left[(\partial_\mu U)U^\dagger, U A_\nu U^\dagger\right] \quad (S.5)$$

and

$$\partial_\mu \left((\partial_\nu U)U^\dagger\right) = (\partial_\mu \partial_\nu U)U^\dagger + (\partial_\nu U)(\partial_\mu U^\dagger) = (\partial_\mu \partial_\nu U)U^\dagger - (\partial_\nu U)U^\dagger(\partial_\mu U)U^\dagger. \quad (S.6)$$

Likewise

$$\partial_\nu A'_\mu = U(\partial_\nu A_\mu)U^\dagger + \left[(\partial_\nu U)U^\dagger, U A_\mu U^\dagger\right] + i(\partial_\nu \partial_\mu U)U^\dagger - i(\partial_\nu U)U^\dagger \times (\partial_\mu U)U^\dagger \quad (S.7)$$
and hence
\[
\partial_\mu A'_\nu - \partial_\nu A'_\mu = U (\partial_\mu A_\nu - \partial_\nu A_\mu) U^\dagger + \left[ (\partial_\mu U) U^\dagger, U A_\nu U^\dagger \right] - \left[ (\partial_\nu U) U^\dagger, U A_\mu U^\dagger \right] \\
+ 0 + i \left[ (\partial_\mu U) U^\dagger, (\partial_\nu U) U^\dagger \right].
\]

(S.8)

At the same time, the commutator part of the tension fields transforms into
\[
i [A'_\mu, A'_\nu] = i \left[ (U A_\mu U^\dagger + i(\partial_\mu U) U^\dagger), (U A_\nu U^\dagger + i(\partial_\nu U) U^\dagger) \right] \\
= i \left[ U A_\mu U^\dagger, U A_\nu U^\dagger \right] - \left[ (\partial_\mu U) U^\dagger, U A_\nu U^\dagger \right] \\
- \left[ U A_\mu U^\dagger, (\partial_\nu U) U^\dagger \right] - i \left[ (\partial_\mu U) U^\dagger, (\partial_\nu U) U^\dagger \right],
\]

(S.9)

Combining eqs. (S.8) and (S.9) leads to massive cancellation of 6 out of terms on the combined right hand side. Only the first terms on right hand sides of (S.8) and (S.9) survive the cancellation, thus
\[
\partial_\mu A'_\nu - \partial_\nu A'_\mu + i [A'_\mu, A'_\nu] = U (\partial_\mu A_\nu - \partial_\nu A_\mu) U^\dagger + i \left[ U A_\mu U^\dagger, U A_\nu U^\dagger \right] \\
= U (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]) U^\dagger,
\]

(S.10)
or in other words,
\[
\mathcal{F}'_{\mu\nu}(x) = U(x) \mathcal{F}_{\mu\nu}(x) U^\dagger(x).
\]

(S.11)

Q.E.D.

Problem 1(b):
Under the gauge transform (3), the ordinary derivative \( \partial_\mu \Phi(x) \) transforms to
\[
\partial_\mu \Phi' = \partial_\mu \left( U \Phi U^\dagger \right) \\
= U(\partial_\mu \Phi) U^\dagger + (\partial_\mu U) \Phi U^\dagger + U \Phi (\partial_\mu U^\dagger) \\
= U(\partial_\mu \Phi) U^\dagger + (\partial_\mu U) U^\dagger \Phi U^\dagger - U \Phi U^\dagger (\partial_\mu U) U^\dagger \\
= U(\partial_\mu \Phi) U^\dagger + \left[ (\partial_\mu U) U^\dagger, U \Phi U^\dagger \right].
\]

(S.12)

At the same time, the commutator term in the covariant derivative (4) transforms to
\[
i [A'_\mu, \Phi] = i \left[ U A_\mu U^\dagger, U \Phi U^\dagger \right] - \left[ (\partial_\mu U) U^\dagger, U \Phi U^\dagger \right].
\]

(S.13)
When we combine these 2 formulae, the terms containing $\partial_\mu U$ cancel out, and we are left with

$$\partial_\mu \Phi' + i [A'_\mu, \Phi'] = U(\partial_\mu \Phi)U^\dagger + i \left[ U A_\mu U^\dagger, U \Phi U^\dagger \right] = U \left( \partial_\mu \Phi + [A_\mu, \Phi] \right) U^\dagger. \quad (S.14)$$

In other words, the covariant derivative (4) indeed transforms like the $\Phi$ field itself,

$$D'_\mu \Phi'(x) = U(x) \times D_\mu \Phi(x) \times U^\dagger(x). \quad (S.15)$$

**Problem 1(c):**

Applying the covariant derivative (4) twice, we obtain

$$D_\mu D_\nu \Phi = \partial_\mu (D_\nu \Phi) + i [A_\mu, D_\nu \Phi]$$

$$= \partial_\mu (\partial_\nu \Phi + i [A_\nu, \Phi]) + i [A_\mu, (\partial_\nu \Phi + i [A_\nu, \Phi])]$$

$$= \partial_\mu \partial_\nu \Phi + i [\partial_\mu A_\nu, \Phi] + i [A_\nu, \partial_\mu \Phi] + i [A_\mu, \partial_\nu \Phi] - [A_\mu, [A_\nu, \Phi]]. \quad (S.16)$$

On the bottom line here, the first and the third + fourth terms are symmetric under exchange $\mu \leftrightarrow \nu$ but the second and the fifth terms are not symmetric. Consequently, subtracting $D_\nu D_\mu \Phi$ from $D_\mu D_\nu \Phi$ we get

$$[D_\mu, D_\nu] \Phi = i [\partial_\mu A_\nu, \Phi] - i [\partial_\nu A_\mu, \Phi] - [A_\mu, [A_\nu, \Phi]] + [A_\nu, [A_\mu, \Phi]] \quad (S.17)$$

Thanks to the Jacobi identity

$$[A_\mu, [A_\nu, \Phi]] + [A_\nu, [\Phi, A_\mu]] + [\Phi, [A_\mu, A_\nu]] = 0, \quad (S.18)$$

we may combine the last two terms in eq. (S.17) to

$$- [A_\mu, [A_\nu, \Phi]] + [A_\nu, [A_\mu, \Phi]] = - [A_\mu, [A_\nu, \Phi]] - [A_\nu, [\Phi, A_\mu]]$$

$$= + [\Phi, [A_\mu, A_\nu]] = - [[A_\mu, A_\nu], \Phi]. \quad (S.19)$$

Therefore,

$$[D_\mu, D_\nu] \Phi = i [\partial_\mu A_\nu, \Phi] - i [\partial_\nu A_\mu, \Phi] - [[A_\mu, A_\nu], \Phi]$$

$$= i \left[ (\partial_\mu A_\nu - \partial_\nu A_\mu + i [A_\mu, A_\nu]), \Phi \right]$$

$$= i [F_{\mu\nu}, \Phi]. \quad (S.20)$$
Problem 1(d):
There are two ways to prove the non-abelian Bianchi identity: using part (c) and the Jacobi identity for commutators, or the hard calculation based directly on eq. (S.2). Let me start with the easier proof.

Consider two adjoint fields $\Phi(x)$ and $\Psi(x)$. Thanks to the Jacobi identity

$$[A_\lambda, [\Phi, \Psi]] = -[[A_\lambda, \Phi], \Psi] + [[A_\lambda, \Phi], \Psi] = +[[A_\lambda, \Phi], \Psi] + [\Phi, [A_\lambda, \Psi]] \quad (S.21)$$

we have a Leibniz rule for the covariant derivative (4):

$$D_\lambda([\Phi, \Psi]) = \partial_\lambda([\Phi, \Psi]) + i[A_\lambda, [\Phi, \Psi]]$$

$$= [\partial_\lambda \Phi, \Psi] + [\Phi, \partial_\lambda \Psi] + [[A_\lambda, \Phi], \Psi] + [\Phi, [A_\lambda, \Psi]] \quad (S.22)$$

In particular, for $\Phi = F_{\mu\nu}$ and arbitrary $\Psi$, we have

$$D_\lambda([F_{\mu\nu}, \Psi]) = [D_\lambda F_{\mu\nu}, \Psi] + [F_{\mu\nu}, D_\lambda \Psi] \quad (S.23)$$

But in part (c) we saw that $[F_{\mu\nu}, \Psi] = -i[D_\mu, D_\nu] \Psi$; likewise, $[F_{\mu\nu}, D_\lambda \Psi] = -i[D_\mu, D_\nu] D_\lambda \Psi$. Consequently, eq. (S.23) becomes

$$-iD_\lambda[D_\mu, D_\nu] \Psi = [D_\lambda F_{\mu\nu}, \Psi] - i[D_\mu, D_\nu] D_\lambda \Psi \quad (S.24)$$

and hence

$$[D_\lambda F_{\mu\nu}, \Psi] = -i[D_\lambda, [D_\mu, D_\nu]] \Psi \quad (S.25)$$

Now, let’s sum 3 such formulae, one for each cyclic permutations of the indices $\lambda, \mu, \nu$. On the left hand side, this gives us

$$[((D_\lambda F_{\mu\nu} + D_\mu F_{\nu\lambda} + D_\nu F_{\lambda\mu}), \Psi)] = \cdots$$

while on the right hand side we obtain

$$\cdots = -i([D_\lambda, [D_\mu, D_\nu]] + [D_\mu, [D_\nu, D_\lambda]] + [D_\nu, [D_\lambda, D_\mu]]) \Psi = 0 \quad (S.26)$$
due to Jacobi identity for the three covariant derivatives $D_\lambda$, $D_\mu$, and $D_\nu$. Thus

$$\left[ \left( D_\lambda F_{\mu\nu} + D_\mu F_{\nu\lambda} + D_\nu F_{\lambda\mu} \right), \Psi \right] = 0,$$

and this must be true for any adjoint field $\Psi(x)$. Moreover, for any $x, \lambda, \mu, \nu$, the $N \times N$ matrix

$$D_\lambda F_{\mu\nu} + D_\mu F_{\nu\lambda} + D_\nu F_{\lambda\mu}$$

is traceless, and the only way it may commute with all traceless hermitian matrices $\Psi(x)$ is by being zero, thus

$$D_\lambda F_{\mu\nu} + D_\mu F_{\nu\lambda} + D_\nu F_{\lambda\mu} = 0. \quad (S.27)$$

This is my first proof of the non-abelian Bianchi identity.

The second proof of the Bianchi identity follows directly from the definition (S.2) of the non-abelian tension fields and the covariant derivatives (4). Let’s spell out $D_\lambda F_{\mu\nu}$ in detail:

$$D_\lambda F_{\mu\nu} = \partial_\lambda F_{\mu\nu} + i[A_\lambda, F_{\mu\nu}]$$

$$= \partial_\lambda (\partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]) + i [A_\lambda, (\partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu])]$$

$$= \partial_\lambda \partial_\mu A_\nu - \partial_\lambda \partial_\nu A_\mu + i[\partial_\lambda A_\mu, A_\nu] + i[A_\mu, \partial_\lambda A_\nu]$$

$$+ i[A_\lambda, \partial_\mu A_\nu] - i[A_\lambda, \partial_\nu A_\mu] - [A_\lambda, [A_\mu, A_\nu]]$$

$$= \left( \partial_\lambda \partial_\mu A_\nu - \partial_\lambda \partial_\nu A_\mu \right) + i \left( [\partial_\lambda A_\mu, A_\nu] - [\partial_\mu A_\nu, A_\lambda] \right)$$

$$+ i \left( [A_\mu, \partial_\lambda A_\nu] - [A_\lambda, \partial_\nu A_\mu] \right) - \left( [A_\lambda, [A_\mu, A_\nu]] \right). \quad (S.28)$$

On the bottom two lines here I have grouped terms in () so that after summing over cyclic permutations of the indices $\lambda, \mu, \nu$, we get zero sum separately for each group. Indeed,

$$\left( \partial_\lambda \partial_\mu A_\nu - \partial_\lambda \partial_\nu A_\mu \right) + \text{cyclic} = \left( \partial_\lambda \partial_\mu A_\nu - \partial_\nu \partial_\lambda A_\mu \right) + \text{cyclic}$$

$$= 0 \quad \langle \text{by inspection} \rangle,$$

$$\left( [\partial_\lambda A_\mu, A_\nu] - [\partial_\mu A_\nu, A_\lambda] \right) + \text{cyclic} = 0 \quad \langle \text{by inspection} \rangle,$$

$$\left( [A_\mu, \partial_\lambda A_\nu] - [A_\lambda, \partial_\nu A_\mu] \right) + \text{cyclic} = 0 \quad \langle \text{by inspection} \rangle,$$

and

$$[A_\lambda, [A_\mu, A_\nu]] + \text{cyclic} = 0 \quad \langle \text{by Jacobi identity} \rangle.$$
Therefore,
\[ D_{\lambda} F_{\mu\nu} + \text{cyclic} \equiv D_{\lambda} F_{\mu\nu} + D_{\mu} F_{\nu\lambda} + D_{\nu} F_{\lambda\mu} = 0. \]  
(S.30)

**Problem 1(e):**

\[ \delta F_{\mu\nu} = \delta (\partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} + i[A_{\mu}, A_{\nu}]) \]
\[ = \partial_{\mu} \delta A_{\nu} - \partial_{\nu} \delta A_{\mu} + i[\delta A_{\mu}, A_{\nu}] + i[A_{\mu}, \delta A_{\nu}] \]
\[ = \left( \partial_{\mu} \delta A_{\nu} + i[A_{\mu}, \delta A_{\nu}] \right) \quad \left( \partial_{\nu} \delta A_{\mu} + i[A_{\nu}, \delta A_{\mu}] \right) \]
\[ \equiv D_{\mu} \delta A_{\nu} - D_{\nu} \delta A_{\mu}. \]  
(S.31)

**Problem 1(f):**

Under infinitesimal variations \( \delta A_{\nu} \) of the gauge fields,
\[ \delta \text{tr} \left( F^{\mu\nu} F_{\mu\nu} \right) = 2 \text{tr} \left( F^{\mu\nu} \delta F_{\mu\nu} \right) = 2 \text{tr} \left( F^{\mu\nu} (D_{\mu} \delta A_{\nu} - D_{\nu} \delta A_{\mu}) \right) = 4 \text{tr} \left( F^{\mu\nu} D_{\mu} \delta A_{\nu} \right), \]  
(S.32)

so the action \( \int d^{4} x \mathcal{L} \) of the gauge theory varies by
\[ \delta S = \int d^{4} x \delta \mathcal{L} \]
\[ = \int d^{4} x \left( \frac{-2}{g^{2}} F^{\mu\nu} \times D_{\mu} \delta A_{\nu} \times 2 J^{\nu} \times \delta A_{\nu} \right) \]
\[ = \int d^{4} x \left( 2 \delta A_{\nu} \times \left( \frac{1}{g^{2}} D_{\mu} F^{\mu\nu} - J^{\nu} \right) \right). \]  
(S.33)

On the last line here I have integrated the first term by parts using
\[ \text{tr} (F^{\mu\nu} \times D_{\mu} \delta A_{\nu}) + \text{tr} (D_{\mu} F^{\mu\nu} \times \delta A_{\nu}) = \partial_{\mu} \text{tr} (F^{\mu\nu} \times \delta A_{\nu}). \]  
(S.34)

Indeed, for any two adjoint fields \( \Phi \) and \( \Psi \) on which the covariant derivatives act according to eq. (4),
\[ \text{tr} (D_{\mu} \Phi \times \Psi) + \text{tr} (\Phi \times D_{\mu} \Psi) = \text{tr} \left( (\partial_{\mu} \Phi) \Psi + i[A_{\mu}, \Phi] \Psi + \Phi (\partial_{\mu} \Psi) + i \Phi [A_{\mu}, \Psi] \right) \]
\[ = \partial_{\mu} \text{tr} (\Phi \Psi) + i \text{tr} ([A_{\mu}, \Phi \Psi]) \]
\[ = \partial_{\mu} \text{tr} (\Phi \Psi) \]  
(S.35)

because a commutator like \([ A_{\mu}, \Phi \Psi] \) has zero trace.
The Euler–Lagrange equations for the non-abelian gauge fields follow from the action variation (S.33):

\[ D_\mu F^{\mu\nu} = g^2 J^\nu, \]  
(S.36)

or in components, \( D_\mu F^{a\mu\nu} = g^2 J^{a\nu} \). The \( g^2 \) in this formula is due to non-canonical normalization of the fields and currents; For the canonically-normalized \( F^{a\mu\nu} = g^{-1} F^{a\mu\nu} \) and \( J^{a\nu}_{\text{can}} = g J^{a\nu} \), the Euler–Lagrange equation (S.36) becomes simply \( D_\nu F^{a\mu\nu} = J^{a\nu}_{\text{can}} \).

Together, the Bianchi identities (S.27) and the Euler–Lagrange equations (S.36) generalize the Maxwell equations to non-abelian gauge fields \( F^{\mu\nu} \). However, unlike the Maxwell equations for the EM which involve only the tension fields \( F^{\mu\nu} \) but not the potentials \( A^\mu \), in the non-abelian case the potential fields \( A^\mu \) do enter eqs. (S.27) and (S.36) through the covariant derivatives \( D_\mu \).

Problem 1(g):

For the abelian EM fields, Maxwell equations \( \partial_\mu F^{\mu\nu} = J^\nu \) require the electric current to be conserved, \( \partial_\nu J^\nu = \partial_\nu \partial_\mu F^{\mu\nu} = 0 \) since \( F^{\mu\nu} = -F^{\nu\mu} \) and the derivatives commute with each other. The non-abelian tension fields \( F^{\mu\nu} \) are also antisymmetric in \( \mu \leftrightarrow \nu \), but the covariant derivatives do not commute, \( D_\mu D_\nu \neq D_\nu D_\mu \). Therefore,

\[ g^2 D_\nu J^\nu = D_\nu D_\mu F^{\mu\nu} = \frac{1}{2} [D_\mu, D_\nu] F^{\mu\nu} = \frac{i}{2} [F^{\mu\nu}, F^{\mu\nu}] \]  
(S.37)

where the last equality works exactly as in part (c) — the \( F^{a\mu\nu} \) fields form an adjoint multiplet of fields, and for any such multiplet packed into an hermitian \( N \times N \) matrix \( \Phi \), \( [D_\mu, D_\nu] \Phi = i [F^{\mu\nu}, \Phi] \). However, unlike a generic matrix \( \Phi \) which may commute or not commute with the \( F^{\mu\nu} \), for any \( \mu \) and \( \nu \) the \( F^{\mu\nu} \) matrix always commutes with itself. Thus,

\[ [F^{\mu\nu}, F^{\mu\nu}] = 0 \quad \text{even before summing over } \mu \text{ and } \nu. \]  
(S.38)

Of course, after the summing over \( \mu \) and \( \nu \) we still have a zero, thus \( D_\nu D_\mu F^{\mu\nu}(x) \equiv 0 \). In light of the Euler–Lagrange equations (S.36), this require the non-abelian currents \( J^{a\mu} \) to be
covariantly conserved:

\[ D_\nu J^\nu = 0, \]  

(S.39)

or in components

\[ \partial_\nu J^{a\nu} - f^{abc} A_{\nu}^b J^{c\nu} = 0. \]  

(S.40)

Note: because of the covariantizing term here, we do not have conserved net charges; alas,

\[ \frac{d}{dt} \int d^3 x J^{a0}(x, t) \neq 0. \]  

(S.41)

Problem 2(a):

Under a local symmetry (7) with infinitesimal parameters \( \Lambda^a(x) \) the gauge fields \( A_{\mu}^a(x) \) change by

\[ \Delta A_{\mu}^a(x) = A_{\mu}^a(x) - \partial_\mu \Lambda^a(x) - f^{abc} \Lambda^b(x) A_{\mu}^c(x). \]  

(S.42)

At the same time, the ‘matter fields’ \( \Phi_\alpha(x) \) in some multiplet \( (m) \) of the gauge group \( G \) change by

\[ \Delta \Phi_\alpha(x) = i \Lambda^a(x) \left( T^a_{(m)} \right)_{\alpha\beta} \Phi_\beta(x), \]  

(S.43)

cf. eq. (9). Consequently, the covariant derivatives (10) of the matter fields change by

\[
\begin{align*}
\Delta D_\mu \Phi_\alpha(x) & \equiv \Delta \left( \partial_\mu \Phi_\alpha(x) + i A_{\mu}^a(x) \left( T^a_{(m)} \right)_{\alpha\beta} \Phi_\beta(x) \right) \\
& = \partial_\mu (\Delta \Phi_\alpha(x)) + i A_{\mu}^a(x) \left( T^a_{(m)} \right)_{\alpha\beta} \times \Delta \Phi_\beta(x) + i \Delta A_{\mu}^a(x) \times \left( T^a_{(m)} \right)_{\alpha\beta} \Phi_\beta(x) \\
& = i \Lambda^a(x) \left( T^a_{(m)} \right)_{\alpha\beta} \times \partial_\mu \Phi_\beta(x) + i(\partial_\mu \Lambda^a(x)) \times \left( T^a_{(m)} \right)_{\alpha\beta} \Phi_\beta(x) \\
& \quad + i A_{\mu}^a(x) \left( T^a_{(m)} \right)_{\alpha\beta} \times \Lambda^b(x) \left( T^b_{(m)} \right)_{\beta\gamma} \Phi_\gamma(x) \\
& \quad - i(\partial_\mu \Lambda^a(x)) \times \left( T^a_{(m)} \right)_{\alpha\beta} \Phi_\beta(x) - i f^{abc} \Lambda^b(x) A_{\mu}^c(x) \times \left( T^a_{(m)} \right)_{\alpha\beta} \Phi_\beta(x) \\
& = i \Lambda^a(x) \left( T^a_{(m)} \right)_{\alpha\beta} \times \partial_\mu \Phi_\beta(x) + i(\partial_\mu \Lambda^a(x)) \times 0 \\
& \quad - \Lambda^b(x) A_{\mu}^c(x) \left( T^c_{(m)} T^b_{(m)} + i f^{abc} T^a_{(m)} \right)_{\alpha\gamma} \Phi_\gamma(x)
\end{align*}
\]  

(S.44)

In any multiplet of Lie group \( G \), the matrices \( T^a_{(m)} \) satisfy the same commutation relations as the Lie algebra’s generators, \( [T^b_{(m)}, T^c_{(m)}] = i f^{bca} T^a_{(m)} = i f^{abc} T^a_{(m)} \). Hence, on the last line of
eq. (S.44)

\[ T^c_{(m)} T^d_{(m)} + i f^{abc} T^a_{(m)} = T^c_{(m)} T^b_{(m)} + \left[ T^b_{(m)}, T^c_{(m)} \right] = T^b_{(m)} T^c_{(m)} \]  

(S.45)

and consequently

\[
\Delta D_\mu \Phi_\alpha(x) = i \Lambda^a(x) \left( T^a_{(m)} \right)_\alpha_\beta \partial_\mu \Phi_\beta(x) - \Lambda^b(x) A^{c}_{\mu}(x) \left( T^b_{(m)} \right)_\alpha_\beta \left( T^c_{(m)} \right)_\beta_\gamma \Phi_\gamma(x) \\
= i \Lambda^a(x) \left( T^a_{(m)} \right)_\alpha_\beta \times \left( \partial_\mu \Phi_\beta(x) + i A^{c}_{\mu}(x) \left( T^c_{(m)} \right)_\beta_\gamma \Phi_\gamma(x) \right) \\
= i \Lambda^a(x) \left( T^a_{(m)} \right)_\alpha_\beta \times D_\mu \Phi_\beta(x).
\]

(S.46)

Therefore, the derivatives \( D_\mu \Phi_\alpha(x) \) transform under infinitesimal gauge symmetries (7) in exactly similar manner to the \( \Phi_\alpha(x) \) fields themselves. In other words, the covariant derivatives (10) are indeed covariant.

Problem 2:

Under finite gauge symmetries \( G(x) \), the fields \( \Phi_\alpha(x) \) in multiplet \( (m) \) transform to

\[ \Phi_\alpha(x) \to R^{(m)}_{\alpha\beta}(G(x)) \Phi_\beta(x) \]  

(S.47)

where \( R^{(m)}_{\alpha\beta}(G) \) is the unitary matrix representing group element \( G \) in multiplet \( (m) \); for \( G = \exp(-i \Theta^a \hat{T}^a) \) for some finite real parameters \( \Theta^a \),

\[ R^{(m)}_{\alpha\beta}(G) = \exp(-i \Theta^a T^a_{(m)})_{\alpha\beta}. \]  

(S.48)

Let’s write the covariant derivatives (10) in \( \text{dim}(m) \times \text{dim}(m) \) matrix form as

\[ D_\mu \Phi(x) = \partial_\mu \Phi(x) + i A^{(m)}_{\mu}(x) \Phi(x) \quad \text{where} \quad A^{(m)}_{\mu}(x) \overset{\text{def}}{=} \sum_a A^a_{\mu}(x) \times \left( T^a_{(m)} \right). \]  

(S.49)

To make sure these derivatives transform covariantly under finite local symmetries (S.47), the matrix-valued gauge fields \( A^{(m)}_{\mu}(x) \) should transform as

\[ A^{(m)}_{\mu}(x) \to R^{(m)}(G(x)) \times A^{(m)}_{\mu}(x) \times \left( R^{(m)}(G(x))^{-1} \right) + i \partial_\mu R^{(m)}(G(x)) \times \left( R^{(m)}(G(x))^{-1} \right). \]  

(S.50)

This works similarly to the \( SU(N) \) symmetry I have discussed in class:

\[
D_\mu \Phi \to D'_\mu \Phi' = \partial_\mu \left( R^{(m)} \Phi \right) + i \left( R^{(m)} A_{\mu} \left( R^{(m)} \right)^{-1} + i \partial_\mu R^{(m)} \times \left( R^{(m)} \right)^{-1} \right) \times R^{(m)} \Phi \\
= \partial_\mu R^{(m)} \times \Phi + R^{(m)} \times \partial_\mu \Phi + i R^{(m)} A_{\mu} \times \Phi - \partial_\mu R^{(m)} \times \Phi \\
= R^{(m)} \times D_\mu \Phi. 
\]

(S.51)

But the real problem here is to make sure that transforms (S.50) are consistent with having the same gauge fields \( A^a_{\mu}(x) \) for all multiplets of the gauge group.
Before I write down the transformation law for the $A_\mu^a(x)$ fields in a multiplet-independent manner, let me note that the symmetries $G(x)$ should be continuous functions of $x$. Consequently, for an infinitesimal displacement $\epsilon^\mu$, $G(x + \epsilon) \times G^{-1}(x) = 1 + O(\epsilon)$. But for any Lie group member infinitesimally close to unity, its displacement from unity is a linear combination of the Lie algebra generators $\hat{T}^a$, thus

$$G(x + \epsilon) \times G^{-1}(x) = 1 + i\epsilon^\mu C_\mu^a(x)\hat{T}^a + O(\epsilon^2) \quad (S.52)$$

for some real coefficients $C_\mu^a(x)$. In terms of derivatives of $G(x)$,

$$\partial_\mu G(x) \times G^{-1}(x) = iC_\mu^a(x)\hat{T}^a. \quad (S.53)$$

In terms of the coefficients $C_\mu^a(x)$ in this formula, the gauge fields $A_\mu^a(x)$ transform to

$$A_\mu^a(x) = R_{\text{adj}}^{ab}(G(x)) \times A_\mu^b(x) - C_\mu^a(x) \quad (S.54)$$

where $R_{\text{adj}}^{ab}(G)$ represents $G$ in the adjoint multiplet of the Lie group $G$.

Now let me show that the gauge transform (S.54) leads to eqs. (S.50) for all multiplets $(m)$ of the Lie group $G$. Any representation of $G$ must respect the group product,

$$R^{(m)}(G_2 \times G_1) = R^{(m)}(G_2) \times R^{(m)}(G_1). \quad (S.55)$$

Also, in the infinitesimal neighborhood of the unity,

$$R^{(m)}(1 + i\epsilon^\alpha \hat{T}^\alpha) = 1 + i\epsilon^\alpha T^\alpha_{(m)}. \quad (S.56)$$

Consequently, for any multiplet $(m)$,

$$R^{(m)}(G(x + \epsilon) \times \left(R^{(m)}(G(x))\right)^{-1} = R^{(m)}(G(x + \epsilon) \times G^{-1}(x))$$

$$= R^{(m)}(1 + i\epsilon^\mu C_\mu^a\hat{T}^a)$$

$$= 1 + i\epsilon^\mu C_\mu^a T^a_{(m)} \quad (S.57)$$

and hence

$$\partial_\mu R^{(m)}(G(x)) \times \left(R^{(m)}(G(x))\right)^{-1} = iC_\mu^a T^a_{(m)} \quad (S.58)$$

with exactly the same coefficients $C_\mu^a(x)$ as in eq. (S.53). Therefore, the second term in eq. (S.50) for the transformation of the $A_{\mu}^{(m)}(x) = A_\mu^a(x) T^a_{(m)}$ agrees with the $-C_\mu^a(x)$ term in eq. (S.54) for the transformation of the component fields $A_\mu^a(x)$.
As to the first term in eq. (S.50), it agrees with the first term in eq. (S.54) because for any multiplet

\[ R^{(m)}(G) \times T^b_{(m)} \times \left( R^{(m)}(G) \right)^{-1} = \sum_a T^a_{(m)} R^{ab}_{\text{adj}}(G), \]  

hence

\[ R^{(m)}(G) \times A^b_{\mu} T^b_{(m)} \times \left( R^{(m)}(G) \right)^{-1} = A^b_{\mu} T^a_{(m)} R^{ab}_{\text{adj}}(G) = T^a_{(m)} \left( R^{ab}_{\text{adj}}(G) A^b_{\mu} \right). \]  

Finally, let me prove the Lemma (S.59). I assume \( \mathcal{G} = \exp(-i\Theta^a T^a) \) for some real parameters \( \Theta^a \) and hence \( R^{(m)}(G) = \exp(-i\Theta^a T^a_{(m)}) \). Consequently,

\[ R^{(m)}(G) T^a_{(m)} \left( R^{(m)}(G) \right)^{-1} = \exp \left( -i\Theta^b T^b_{(m)} \right) T^a_{(m)} \exp \left( +i\Theta^b T^b_{(m)} \right) \]

\[ = T^a_{(m)} - i \left[ \Theta^b T^b_{(m)}, T^a_{(m)} \right] + \frac{(-i)^2}{2} \left[ \Theta^c T^c_{(m)}, \left[ \Theta^b T^b_{(m)}, T^a_{(m)} \right] \right] \]

\[ + \frac{(-i)^3}{3!} \left[ \Theta^d T^d_{(m)}, \left[ \Theta^c T^c_{(m)}, \left[ \Theta^b T^b_{(m)}, T^a_{(m)} \right] \right] \right] + \cdots \]  

All the commutators here follow from the Lie algebra

\[ (-i)^2 \left[ \Theta^c T^c_{(m)}, \left[ \Theta^b T^b_{(m)}, T^a_{(m)} \right] \right] = \Theta^b f^{bac} T^c_{(m)}, \]

\[ (-i)^3 \left[ \Theta^d T^d_{(m)}, \left[ \Theta^c T^c_{(m)}, \left[ \Theta^b T^b_{(m)}, T^a_{(m)} \right] \right] \right] = \Theta^b f^{bac} \Theta^e f^{ced} T^a_{(m)}, \]

All the contractions of \( \Theta^a \) with structure constants on the right hand sides here may be interpreted in terms of the adjoint multiplet of the group \( G \) where \( (T^a_{\text{adj}})^{bc} = -i f^{abc} \):

\[ \Theta^b f^{bac} = (i\Theta^b T^a_{\text{adj}})^{ac}, \]

\[ (\Theta^b f^{bad}) (\Theta^e f^{ced}) = (i\Theta^b T^a_{\text{adj}})^{ad} (i\Theta^e T^a_{\text{adj}})^{de} = \left( i\Theta^b T^a_{\text{adj}} \right)^{2} = \Theta^b f^{bac} \Theta^e f^{ced} \]

\[ \Theta^b f^{bac} \Theta^e f^{ced} (\Theta^d f^{def}) (\Theta^g f^{gfh}) = \left( i\Theta^b T^a_{\text{adj}} \right)^{3} = \Theta^b f^{bac} \Theta^e f^{ced} \Theta^d f^{def} \Theta^g f^{gfh} \]  

Combining eqs. (S.61) through (S.63), we obtain

\[ R^{(m)}(G) T^a_{(m)} \left( R^{(m)}(G) \right)^{-1} = T^a_{(m)} + (i\Theta^b T^b_{\text{adj}})^{ac} T^c_{(m)} + \frac{1}{2} \left( (i\Theta^b T^b_{\text{adj}})^{2} \right)^{ae} T^e_{(m)} \]

\[ + \frac{1}{6} \left( (i\Theta^b T^b_{\text{adj}})^{3} \right)^{ag} T^g_{(m)} + \cdots \]

\[ = \left( \exp \left( i\Theta^b T^b_{\text{adj}} \right) \right)^{ac} T^c_{(m)} \]

\[ = T^c_{(m)} \left( \exp \left( -i\Theta^b T^b_{\text{adj}} \right) \right)^{ca} \]

\[ = T^c_{(m)} \times R^{ca}_{\text{adj}}(G), \]

which proves the Lemma (S.59).
Problem 2(b):
The covariant derivatives $D_\mu$ acting on the fermionic fields $\Psi_\alpha$ in the Lagrangian (11) according to eq. (10). Splitting those derivatives into $\partial_\mu$ and the $A_\mu^a$ part gives us

$$
\mathcal{L}_\Psi \equiv \overline{\Psi}(i\gamma^\mu D_\mu - m)\Psi = \overline{\Psi}^a(i\gamma^\mu \partial_\mu - m)\Psi_\alpha - A_\mu^a \times \overline{\Psi}^\alpha \gamma^\mu T^a_{(m)} \Psi_\beta \quad (S.65)
$$

In terms of eq. (5), the second term in (S.65) defines the currents

$$
J^{a\mu} = \overline{\Psi}^\mu T^a_{(m)} \Psi \equiv \left(T^a_{(m)}\right)_\alpha^\beta \overline{\Psi}^\alpha \gamma^\mu \Psi_\beta. \quad (S.66)
$$

Under infinitesimal gauge transforms (7), we have

$$
\Delta \Psi_\alpha(x) = i\Lambda^a(x) \left(T^a_{(m)}\right)_\alpha^\beta \Psi_\beta(x), \quad \Delta \overline{\Psi}^a(x) = -\Lambda^a(x) \overline{\Psi}^\beta(x) \left(T^a_{(m)}\right)_\beta^\alpha, \quad (S.67)
$$
or in $\text{dim}(m) \times \text{dim}(m)$ matrix notations,

$$
\Delta \Psi(x) = i\Lambda^a(x) T^a_{(m)} \Psi(x), \quad \Delta \overline{\Psi}(x) = -i\Lambda^a \overline{\Psi}(x) T^a_{(m)}. \quad (S.68)
$$

The second equation here follows from the first by hermitian (or rather Dirac) conjugation; note that the $T^a_{(m)}$ matrices are hermitian.

Consequently,

$$
\Delta J^{a\mu}(x) = \Delta \overline{\Psi} \times \gamma^\mu T^a_{(m)} \Psi + \overline{\Psi} \gamma^\mu T^a_{(m)} \times \Delta \Psi \\
= -i\Lambda^b(\overline{\Psi}(x)T^b_{(m)} \times \gamma^\mu T^a_{(m)} \Psi + \overline{\Psi} \gamma^\mu T^a_{(m)} \times i\Lambda^b T^b_{(m)} \Psi \\
= i\Lambda^b(x) \times \overline{\Psi}(x)\gamma^\mu \left[T^a_{(m)}, T^b_{(m)}\right] \Psi(x) \\
= i\Lambda^b(x) \times \overline{\Psi}(x)\gamma^\mu \left[i f^{abc} T^c_{(m)}\right] \Psi(x) \\
= -f^{abc} \Lambda^b(x) \times \overline{\Psi}(x)\gamma^\mu T^c_{(m)} \Psi(x) \\
= -f^{abc} \Lambda^b(x) \times J^{c\mu}(x) \\
= i\Lambda^b(x) \left(T^b_{\text{adj}}\right)^{ac} J^{c\mu}(x), \quad (S.69)
$$

which means that the currents $J^{a\mu}$ transform into each other as members of the adjoint multiplet, at least under infinitesimal gauge transforms.
Dealing with finite gauge transforms is an optional part of this exercise. Under a finite gauge transform \( G(x) \),
\[
\Psi(x) \rightarrow R^{(m)}(G(x))\Psi(x), \quad \overline{\Psi}(x) \rightarrow \overline{\Psi}(x) \left( R^{(m)}(G(x)) \right)^{-1},
\]
(S.70)
hence the currents (S.66) transform into
\[
J^{a\mu}(x) \rightarrow \overline{\Psi}(x) \left( R^{(m)}(G(x)) \right)^{-1} \gamma^\mu T^a_{(m)} R^{(m)}(G(x))\Psi(x).
\]
(S.71)

In light of Lemma (S.59),
\[
\left( R^{(m)}(G(x)) \right)^{-1} T^a_{(m)} R^{(m)}(G(x)) = \left( R^{-1}_{\text{adj}}(G) \right)^{ca} T^c_{(m)} = R^{ac}_{\text{adj}}(G) T^c_{(m)},
\]
(S.72)
hence
\[
J^{a\mu}(x) \rightarrow R^{ac}_{\text{adj}}(G) \times \overline{\Psi} \gamma^\mu T^c_{(m)} \Psi = R^{ac}_{\text{adj}}(G) \times J^{c\nu}.
\]
(S.73)

**Problem 2(c):**

First, let me prove the Leibniz rule for the covariant derivatives acting on the currents (S.66):
\[
D_\mu \left( J^{a\mu} = \overline{\Psi} \gamma^\mu T^a_{(m)} \Psi \right) = (D_\mu \overline{\Psi}) \gamma^\mu T^a_{(m)} \Psi + \overline{\Psi} \gamma^\mu T^a_{(m)} (D_\mu \Psi).
\]
(S.74)

More generally, for any matrix \( \Gamma \) acting only on the Dirac indices — and hence commuting with the \( T^a_{(m)} \) —
\[
D_\mu \left( \overline{\Psi} \Gamma T^a_{(m)} \Psi \right) = (D_\mu \overline{\Psi}) \Gamma T^a_{(m)} \Psi + \overline{\Psi} \Gamma T^a_{(m)} (D_\mu \Psi).
\]
(S.75)

Indeed,
\[
D_\mu \left( \overline{\Psi} \Gamma T^a_{(m)} \Psi \right) = \partial_\mu \left( \overline{\Psi} \Gamma T^a_{(m)} \Psi \right) - f^{abc} A^b_\mu \left( \overline{\Psi} \Gamma T^c_{(m)} \Psi \right)

= (\partial_\mu \overline{\Psi}) \Gamma T^a_{(m)} \Psi + \overline{\Psi} \Gamma T^a_{(m)} (\partial_\mu \Psi)

+ i A^b_\mu \times \overline{\Psi} \Gamma \left( f^{abc} T^c_{(m)} = \left[ T^a_{(m)}, T^b_{(m)} \right] \right) \Psi

= (\partial_\mu \overline{\Psi} - i A^b_\mu \overline{\Psi} T^b_{(m)}) T^a_{(m)} \Gamma \Psi + \overline{\Psi} \Gamma T^a_{(m)} \left( \partial_\mu \Psi + i A^b_\mu T^b_{(m)} \Psi \right)

= (D_\mu \overline{\Psi}) \Gamma T^a_{(m)} \Psi + \overline{\Psi} \Gamma T^a_{(m)} (D_\mu \Psi).
\]
(S.76)

Now let’s specialize to the currents \( J^{a\mu} \) and their divergences (S.74). The equations of
motion for the fermionic fields $\Psi(x)$ and $\overline{\Psi}(x)$ are the \textit{covariant} Dirac equations

$$i\gamma^\mu D_\mu \Psi(x) - m\Psi(x) = 0 \quad \text{and} \quad -iD_\mu \overline{\Psi}(x)\gamma^\mu - m\overline{\Psi}(x) = 0. \quad (S.77)$$

When $\Psi(x)$ and $\overline{\Psi}(x)$ obey these equations, we have

$$D_\mu J^{a\mu} = (D_\mu \overline{\Psi}(x))_{(m)} T^a_{(m)}(\gamma^\mu D_\mu \Psi) \quad \text{cf. eq. (S.74)}$$

$$= (+im\overline{\Psi}) T^a_{(m)} \Psi + \overline{\Psi} T^a_{(m)} (-im\Psi) \quad (S.78)$$

--- the currents $J^{a\mu}$ are \textit{covariantly conserved}.

Note that unlike the ordinary conserved currents which satisfy $\partial_\mu J^\mu = 0$, the covariantly conserved currents do not give rise to time-independent net charges. Indeed, spelling out $D_\mu J^{a\mu} = 0$ in components, we have

$$\partial_0 J^a_0 - f^{abc} A^b_0 J^c_0 - \nabla \cdot J^a + f^{abc} A^b \cdot J^c = 0, \quad (S.79)$$

hence

$$\frac{d}{dt} \int d^3x J^a_0 = \frac{d}{dt} \int d^3x f^{abc} (A^b_0 J^c_0 - A^b \cdot J^c) \neq 0. \quad (S.80)$$

\underline{Problem 3(a)}:
In the Weyl basis

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad (i\not\partial - m) = \begin{pmatrix} -m & i\sigma^\mu \partial_\mu \\ i\sigma^\mu \partial_\mu & -m \end{pmatrix}. \quad (S.81)$$

In class I showed that decomposing the 4-component Dirac spinor field $\Psi(x)$ into 2-component left-handed and right-handed Weyl spinor fields $\Psi_L(x)$ and $\Psi_R(x)$ turns the Dirac Lagrangian to

$$\mathcal{L} \equiv \Psi^\dagger \gamma^0 (i\not\partial - m) \Psi = (\Psi^\dagger_L, \Psi^\dagger_R) \begin{pmatrix} i\sigma^\mu \partial_\mu & -m \\ -m & i\sigma^\mu \partial_\mu \end{pmatrix} \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} \quad (S.82)$$

$$= i\Psi^\dagger_L \sigma^\mu \partial_\mu \Psi_L + i\Psi^\dagger_R \sigma^\mu \partial_\mu \Psi_R - m\Psi^\dagger_L \Psi_R - \Psi^\dagger_R \Psi_L.$$
I am also re-naming the left-handed spinor $\Psi_L(x) \to \chi(x)$. Thus, I have two left-handed Weyl spinor fields $\chi(x)$ and $\bar{\chi}(x)$, and

$$
\Psi_L(x) = \chi(x), \quad \Psi_R(x) = \sigma_2 \bar{\chi}^*(x) \implies \Psi(x) = \begin{pmatrix} \chi(x) \\ -\sigma_2 \bar{\chi}^*(x) \end{pmatrix}.
$$

Plugging these substitutions into the Lagrangian (S.82), we immediately obtain

$$
\mathcal{L} = i\chi^\dagger \bar{\sigma}^\mu \partial_\mu \chi + i\bar{\chi}^\dagger \sigma_2 \bar{\sigma}^\mu \partial_\mu \bar{\chi}^* + m\chi^\dagger \sigma_2 \bar{\chi}^* + m\bar{\chi}^\dagger \sigma_2 \chi.
$$

The second term here may be simplified using $\sigma_2 \sigma_2 = 1$ but $\sigma_2 \bar{\sigma} \sigma_2 = -\sigma^\dagger = -\bar{\sigma}^\dagger$; in terms of $\sigma^\mu$ and $\bar{\sigma}^\mu$, these Pauli-matrix relations become $\sigma_2 \sigma^\mu \sigma_2 = (\bar{\sigma}^\mu)^* = (\bar{\sigma}^\mu)^\dagger$. Consequently

$$
\bar{\chi}^\dagger \sigma_2 \sigma^\mu \sigma_2 \partial_\mu \bar{\chi}^* = \bar{\chi}^\dagger (\bar{\sigma}^\mu)^\dagger (\partial_\mu \bar{\chi})^\dagger = -(\partial_\mu \bar{\chi}^\dagger) \bar{\sigma}^\mu \bar{\chi} + \bar{\chi}^\dagger \bar{\sigma}^\mu \partial_\mu \bar{\chi} - \partial_\mu (\bar{\chi}^\dagger \sigma^\mu \bar{\chi}),
$$

where the second equality obtains by transposing (without conjugation) the whole product of spinors and matrices; the $-$ sign at this step is due to exchange of fermionic fields $\bar{\chi}$ and $\bar{\chi}^\dagger$.

The last term on the right-hand side of eq. (S.84) is a total derivative, hence

$$
\mathcal{L} = i\chi^\dagger \bar{\sigma}^\mu \partial_\mu \chi + i\bar{\chi}^\dagger \bar{\sigma}^\mu \partial_\mu \bar{\chi}^* + m\chi^\dagger \sigma_2 \bar{\chi}^* + m\bar{\chi}^\dagger \sigma_2 \chi + \text{a total derivative}.
$$

**Problem 3(b):**

Under a vector symmetry, the whole Dirac spinor $\Psi$ is multiplied by a common phase factor $\exp(-i\theta_v)$. Under an axial symmetry, the LH components $\Psi_L$ and the RH components $\Psi_R$ get opposite phases,

$$
\Psi_L \to \exp(+i\theta_a) \Psi_L, \quad \Psi_R \to \exp(-i\theta_a) \Psi_R, \quad \Psi \to \exp(-i\theta_a \gamma^5) \Psi.
$$

In terms of the LH Weyl spinors $\chi$ and $\bar{\chi}$ the situation is different because $\bar{\chi}$ transforms in the opposite way to the $\Psi_R$. Thus, $\chi$ and $\bar{\chi}$ have opposite charges under the vector symmetry but the same charges under the axial symmetry,

$$
\text{Vector:} \quad \chi \to \exp(-i\theta_v) \chi, \quad \bar{\chi} \to \exp(+i\theta_v) \bar{\chi},
$$

$$
\text{Axial:} \quad \chi \to \exp(+i\theta_a) \chi, \quad \bar{\chi} \to \exp(+i\theta_a) \bar{\chi}.
$$

The kinetic terms in the Lagrangian (13) are invariant under both vector and axial symmetries, but the mass terms are more finicky. Indeed, $\bar{\chi}^\dagger \sigma_2 \chi$ and $\chi^\dagger \sigma_2 \bar{\chi}^*$ are remain invariant.
only when $\chi$ and $\bar{\chi}$ get opposite phases, so the mass terms respect the vector symmetry but not the axial symmetry! Specifically, under the axial symmetry

$$m \times \chi^\dagger \sigma_2 \chi^* + m \times \bar{\chi}^\dagger \sigma_2 \bar{\chi} \rightarrow \exp(-2i\theta_a)m \times \chi^\dagger \sigma_2 \chi^* + \exp(+2i\theta_a)m \times \bar{\chi}^\dagger \sigma_2 \bar{\chi}. \quad (S.87)$$

Thus, the axial symmetry is a good symmetry of the Lagrangian only when $m = 0$.

BTW, we may generalize the Lagrangian (13) to allow for a complex mass parameter $m$,

$$\mathcal{L} = i\chi^\dagger \gamma^\mu \partial_\mu \chi + i\bar{\chi}^\dagger \gamma^\mu \partial_\mu \bar{\chi} + m^* \times \chi^\dagger \sigma_2 \chi^* + m \times \bar{\chi}^\dagger \sigma_2 \bar{\chi}. \quad (S.88)$$

In Dirac spinor terms, this means

$$\mathcal{L} = \bar{\Psi} \left(i\frac{\partial}{\partial x} - \Re(m) - \Im(m)\gamma^5\right) \Psi. \quad (S.89)$$

However, we may always change the phase of $m$ by an axial symmetry; indeed, eq. (S.87) can be interpreted as $m \rightarrow m \times \exp(+2i\theta_a)$. In particular, we may always render $m$ real and positive (unless it was zero to begin with), so physicists rarely bother with complex masses for free fermions.

But for interacting fermions, we may have less choice. Sometimes, an axial symmetry that would make the fermion mass $m$ real would also screw up the interaction part of the Lagrangian. In such cases, it’s often more convenient to keep the interactions simple at the expense of a complex mass $m$ — or for multiple fermions, a complex mass matrix as in eq. (14).

Problem 3(c):
Back in homework set #6 we saw that space reflection $P$ and charge conjugation $C$ act on the Dirac spinor field $\Psi(x)$ according to

$$P : \Psi'(x') = P\gamma^0 \Psi(x), \quad x' = (-bx, +t),$$

$$C : \Psi'(x) = C\gamma^2 \Psi^*(x), \quad (S.90)$$

where $P = \pm 1$ and $C = \pm 1$ are the intrinsic $P$-parity and $C$-parity of the fermionic species in question. Let’s split the Dirac spinor $\Psi$ into its left-handed and right-handed components
\( \Psi_L \) and \( \Psi_R \). In the Weyl basis
\[
\Psi(x) = \begin{pmatrix} \Psi_L(x) \\ \Psi_R(x) \end{pmatrix}, \quad \gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix},
\]
hence the symmetries (S.90) amount to
\[
P : \Psi'_L(x') = P \Psi_R(x), \quad \Psi'_R(x') = P \Psi_L(x),
\]
\[
C : \Psi'_L(x) = C \sigma_2 \Psi_R^*(x), \quad \Psi'_R(x) = -C \sigma_2 \Psi_L^*(x).
\]
Now we need to translate this action in terms of \( \chi(x) = \Psi_L(x) \) and \( \tilde{\chi}(x) = \sigma_2 \Psi_R^*(x) \). (Please note that \( \tilde{\chi} = +\sigma_2 \Psi_R^* \) but \( \Psi_R = -\sigma_2 \tilde{\chi}^* \) because \( \sigma_2^* = -\sigma_2 \).) Thus,
\[
P : \chi'(x') = \Psi'_L(x') = P \Psi_R(x) = -\sigma_2 \tilde{\chi}^*(x),
\]
\[
P : \tilde{\chi}'(x') = \sigma_2 \Psi_R^*(x') = P \sigma_2 \Psi_L^*(x) = P \sigma_2 \chi^*(x),
\]
\[
C : \chi'(x) = \Psi'_L(x) = C \sigma_2 \Psi_R^*(x) = C \tilde{\chi}(x),
\]
\[
C : \tilde{\chi}'(x) = \sigma_2 \Psi_R^*(x) = \sigma_2 (-C \sigma_2 \Psi_L^*(x))^* = +C \Psi_L(x) = C \chi(x).
\]
Note that both symmetries \( P \) and \( C \) exchange the two Weyl spinors \( \chi \) and \( \tilde{\chi} \). Consequently, these symmetries are ill-defined for stand-alone LH Weyl spinor fields. We may define \( P \) and \( C \) symmetries for a system of several LH Weyl spinors, but that requires organizing the spinor fields in \((\chi, \tilde{\chi})\) pairs. Moreover, for charged Weyl spinor fields, each \((\chi, \tilde{\chi})\) pair must have opposite charges to make sure that the symmetries (S.92) respect the charges.

However, the combined \( CP \) symmetry — reversal of all charges at the same time as exchanging left and right — does not exchange \( \chi \leftrightarrow \tilde{\chi} \), so it may be defined for standalone LH Weyl spinors. (Or for multiple Weyl spinor fields with asymmetric charges.) Indeed, combining eqs. (S.92) for the action of \( P \) and \( C \) symmetries on \( \chi \) and \( \tilde{\chi} \), we obtain
\[
\chi(x) \xrightarrow{P} -P \sigma_2 \tilde{\chi}^*(x') \xrightarrow{C} -CP \sigma_2 \chi^*(x'),
\]
\[
\tilde{\chi}(x) \xrightarrow{P} +P \sigma_2 \chi^*(x') \xrightarrow{C} +CP \sigma_2 \tilde{\chi}^*(x'),
\]
or in other words,
\[
CP : \chi'(x') = -CP \sigma_2 \chi^*(x),
\]
\[
CP : \tilde{\chi}'(x') = +CP \sigma_2 \tilde{\chi}^*(x).
\]
Note that the overall signs here are different for the \( \chi \) and the \( \tilde{\chi} \) — that’s because the fermions and the antifermions have opposite intrinsic P-parities, cf. homework set #4.
For the same reason — opposite intrinsic P-parities of fermions and antifermions — the C and P symmetries of fermions anticommute with each other instead of commuting. Consequently, while \( C^2 = 1 \) and \( P^2 = 1 \), the combined \( CP \) symmetries squares to \(-1\); or rather to \(-1\) for fermions but \(+1\) for bosons, so in effect \( (CP)^2 \) is a rotation by \( 2\pi \). Indeed, thanks to \( \sigma^2 \sigma^* = -1 \),

\[
\chi(x) \xrightarrow{CP} \mp \sigma_2 \chi^*(x') \xrightarrow{CP} + \sigma_2 \sigma_2^* \chi(x) = -\chi(x),
\]

\[
\tilde{\chi}(x) \xrightarrow{CP} \pm \sigma_2 \tilde{\chi}^*(x') \xrightarrow{CP} + \sigma_2 \sigma_2^* \tilde{\chi}(x) = -\tilde{\chi}(x),
\]

regardless of the specific signs \( \mp CP \) in eqs. (S.94).

**Problem 3(d):**

As written, the Lagrangian (14) involves derivatives of the \( \chi_j(x) \) fields but not of the conjugate fields \( \chi_j^\dagger(x) \); this makes the Lagrangian not-quite real, but its imaginary part is a total derivative, so it does not matter. On the other hand, not having the \( \partial_\mu \chi_j^\dagger \) in the Lagrangian makes it particularly simple to write down the equations of motion for the \( \chi_j \) fields:

\[
\frac{\partial L}{\partial \chi_j^\dagger} = i \bar{\sigma}^\mu \partial_\mu \chi_j + \sum_k M^*_{jk} \sigma_2 \chi^k = 0.
\]

Equations of motion for the \( \chi_j^\dagger \) fields obtain from the usual Euler–Lagrange formalism,

\[
\frac{\partial L}{\partial \partial_\mu \chi_j} = i \chi_j^\dagger \bar{\sigma}^\mu, \quad \frac{\partial L}{\partial \chi_j} = \sum_k M^{jk} \chi_k^\dagger \sigma_2,
\]

hence

\[
- i \partial_\mu \chi_j^\dagger \bar{\sigma}^\mu + \sum_k M^{jk} \chi_k^\dagger \sigma_2 = 0.
\]

Alternatively, we get the same equations by simply taking the hermitian conjugates of eqs. (S.96). In \( N \times N \) matrix notations, eqs. (S.96) and (S.98) become

\[
\bar{\sigma}^\mu \partial_\mu \chi = +i M^* \sigma_2 \chi^* \quad \text{and} \quad \partial_\mu \chi^\dagger \bar{\sigma}^\mu = -i \chi^\dagger \sigma_2 M.
\]

(Note that the mass matrix \( M \) is symmetric, \( M_{jk} = M_{kj} \).) Transposing the second equation
here, we get

\[(\bar{\sigma}^\mu)^\top \partial_\mu \chi^* = -iM(\sigma_2)^\top \chi = +iM\sigma_2\chi,\]  

(S.100)

then using \((\bar{\sigma}^\mu)^\top = \sigma_2\sigma^\mu\), we may re-write this equation as

\[\sigma^\mu \partial_\mu (\sigma_2 \chi^*) = +iM\chi.\]  

(S.101)

To obtain the Klein–Gordon equations for the spinor fields, let’s combine eq. (S.101) with the first eq. (S.99):

\[(\sigma^\mu \partial_\mu)(\bar{\sigma}^\nu \partial_\nu)\chi = (\sigma^\mu \partial_\mu)(iM^*\sigma_2 \chi^*) = iM^*(\sigma^\mu \partial_\mu)(\sigma_2 \chi^*) = iM^*(iM)\chi = -M^*M\chi.\]  

(S.102)

On the left hand side here, the derivatives \(\partial_\mu\) and \(\partial_\nu\) commute with each other, hence

\[(\sigma^\mu \partial_\mu)(\bar{\sigma}^\nu \partial_\nu) = \partial_\mu \partial_\nu \times \sigma^\mu \bar{\sigma}^\nu = \partial_\nu \partial_\mu \times \sigma^\mu \bar{\sigma}^\nu = \partial_\mu \partial_\nu \times \sigma^\nu \bar{\sigma}^\mu = \partial_\mu \partial_\nu \times \frac{1}{2}(\sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu).\]  

(S.103)

By inspection, \(\sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu = 2g^{\mu\nu}\):

- For \(\mu = \nu = 0\), \(\sigma^0 = \bar{\sigma}^0 = 1\), hence \(\sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu = 2 \times 1 \times 1 = 2g^{00}\).
- For \(\mu = 0\) and \(\nu = i = 1, 2, 3\), \(\sigma^\mu = \bar{\sigma}^\mu = 1\), \(\sigma^i = \sigma^i\), \(\bar{\sigma}^i = -\sigma^i\), hence \(\sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu = 1 \times (-\sigma^i) + \sigma^i \times 1 = 0 = g^{0i}\).
- For \(\mu = i = 1, 2, 3\) and \(\nu = j = 1, 2, 3\), \(\sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu = \sigma^i(-\sigma^j) + \sigma^j(-\sigma^i) = -\{\sigma^i, \sigma^2\} = -2\delta^{ij} = 2g^{ij}\).

Consequently,

\[(\sigma^\mu \partial_\mu)(\bar{\sigma}^\nu \partial_\nu) = \partial_\mu \partial_\nu \times \frac{1}{2}(\sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu) = \partial_\mu \partial_\nu \times g^{\mu\nu} = \partial^2.\]  

(S.104)

Plugging this formula into eq. (S.102) immediately gives us

\[\partial^2 \chi(x) = -M^*M\chi(x),\]  

(S.105)

which is the matrix form of the combined Klein–Gordon equations for multiple fields \(\chi_j(x)\).
Problem 3(e):
First, let’s check the CP invariance of the kinetic terms in the Lagrangian (14). Since the kinetic term for each \( \chi_j(x) \) is CP-invariant by itself, let me drop the species index \( j \) from this calculation. Under CP,

\[
x' = (-bx, +t), \quad \chi'(x') = \pm i\sigma_2 \chi^*(x), \quad \chi'^\dagger(x') = \mp i\chi^\dagger(x)\sigma_2,
\]

hence

\[
L_{\text{kin}}(x) \rightarrow L'_{\text{kin}}(x') = i\chi'^\dagger(x') \bar{\sigma}^\mu \partial_\mu \chi'(x')
\]

\[
= i\chi^\dagger(x) \sigma_2 \bar{\sigma}^\mu \sigma_2 \partial_\mu \chi^*(x)
\]

\[
= i\chi^\dagger(x) (\sigma^\mu)^\dagger (\partial_\mu \chi^\dagger(x))^\dagger
\]

\[
= -i\bar{\partial}_\mu \chi^\dagger(x) \sigma_\mu \chi(x)
\]

\[
= +i\chi^\dagger(x) \sigma^\mu \partial_\mu \chi(x) + \text{a total derivative}.
\]

Besides an irrelevant total derivative, this CP-transformed kinetic term differs from the original

\[
L_{\text{kin}}(x) = i\chi^\dagger(x) \bar{\sigma}^\mu \partial_\mu \chi(x)
\]

in two aspects: \( \sigma^\mu \) matrices instead of \( \bar{\sigma}^\mu \), and derivatives \( \partial_\mu \) wrt \( x'^\mu = (+t, -x) \) instead of \( \partial_\mu \) wrt \( x^\mu = (+t, +x) \). But these two differences cancel each other:

\[
\partial_\mu = (\partial_t, +\nabla), \quad \partial'_\mu = (\partial_t, -\nabla), \quad \bar{\sigma}^\mu = (1, -\sigma), \quad \sigma^\mu = (1, +\sigma), \quad \Rightarrow \quad \sigma^\mu \partial'_\mu = \partial_t - \sigma \cdot \nabla = \bar{\sigma}^\mu \partial_\mu,
\]

hence \( L'_{\text{kin}}(x') = L_{\text{kin}}(x) \).

Now consider the mass terms, and this time I do need the species indices \( j \) and \( k \). Under CP,

\[
\chi_j^\dagger(x) \sigma_2 \chi_k(x) \rightarrow \chi_j'^\dagger(x') \sigma_2 \chi_k'(x') = (\pm i\sigma_2 \chi_j^*(x)) \sigma_2 (\pm i\sigma_2 \chi_k^*(x))
\]

\[
= -\chi_j^\dagger(x) \sigma_2 \sigma_2 \chi_k(x)
\]

\[
= +\chi_j^\dagger(x) \sigma_2 \chi_k^*(x),
\]

and likewise

\[
\chi_j^\dagger(x) \sigma_2 \chi_k^*(x) \rightarrow \chi_j'^\dagger(x') \sigma_2 \chi_k^*(x') = \chi_j^\dagger(x) \sigma_2 \chi_k(x).
\]

The mass matrices \( M_{jk} \) and \( M_{jk}^* \) are fixed parameters of the theory, so they are not affected
by the CP symmetry (15). Consequently, the net mass term

\[ L_{\text{mass}} = \frac{1}{2} \sum_{j,k} M^{jk} \chi_j^\dagger \sigma_2 \chi_k + \frac{1}{2} \sum_{j,k} M^{*}_{jk} \chi_j^\dagger \sigma_2 \chi_k^* \]  

(S.112)

in the Lagrangian (14) transforms to

\[ L'_{\text{mass}} = \frac{1}{2} \sum_{j,k} M^{jk} \chi_j^\dagger \sigma_2 \chi_k^* + \frac{1}{2} \sum_{j,k} M^{*}_{jk} \chi_j^\dagger \sigma_2 \chi_k. \]  

(S.113)

Comparing these two formulae, we see that after CP, the coefficients of the \( \chi_j^\dagger \sigma_2 \chi_k \) terms change from \( M^{jk} \) to their conjugates \( M^{*}_{jk} \); likewise, the coefficients of the \( \chi_j^\dagger \sigma_2 \chi_k^* \) terms change from \( M^{*}_{jk} \) to \( M^{jk} \). Effectively, the CP transform (15) changes the mass matrix \( M^{jk} \) into its complex conjugate!

Altogether, the Lagrangian (14) is invariant under CP transform (15) if and only if the mass matrix \( M^{jk} \) is real (i.e., all its matrix elements are real). \( \Box.\mathcal{E}.\mathcal{D.} \)

Problem 3(*):

Consider a unitary transform of spinor fields \( \chi_j(x) \) into

\[ \tilde{\chi}_j(x) = \sum_k U^*_j k \chi_k(x), \]  

(S.114)

or in matrix notations \( \tilde{\chi}(x) = U \chi(x) \) for some constant unitary matrix \( U \). Note: In this part of the problem, the tildes have different meaning from parts (a–c) — here \( \chi_j(x) \) and \( \tilde{\chi}_j(x) \) are not independent fields but different bases for the same \( N \) fields. In matrix notations,

\[ L = i\chi^\dagger \bar{\sigma}^\mu \partial_\mu \chi + \frac{1}{2} \chi^\dagger \sigma_2 \chi + \frac{1}{2} \chi^\dagger M^{*} \sigma_2 \chi^* \]  

(S.115)

\[ = i\tilde{\chi}^\dagger \bar{\sigma}^\mu \partial_\mu \tilde{\chi} + \frac{1}{2} \tilde{\chi}^\dagger \tilde{\sigma}_2 \tilde{\chi} + \frac{1}{2} \tilde{\chi}^\dagger \tilde{M}^{*} \tilde{\sigma}_2 \tilde{\chi}^* \]

where

\[ \tilde{M} = U^{*} MU^\dagger. \]  

(S.116)

Indeed, since \( U \) commutes with \( \bar{\sigma}^\mu \partial_\mu \) and \( \sigma_2 \), and since \( U^\dagger U = U^{*} U^{*} = 1 \), we have

\[ i\tilde{\chi}^\dagger \bar{\sigma}^\mu \partial_\mu \tilde{\chi} = i\chi^\dagger U^\dagger \bar{\sigma}^\mu \partial_\mu U \chi = i\chi^\dagger \bar{\sigma}^\mu \partial_\mu \chi, \]

\[ \frac{1}{2} \tilde{\chi}^\dagger \tilde{\sigma}_2 \tilde{\chi} = \frac{1}{2} \chi^\dagger U^\dagger \sigma_2 MU^\dagger \sigma_2 U \chi = \frac{1}{2} \chi^\dagger \sigma_2 \chi, \]  

(S.117)

\[ \frac{1}{2} \tilde{\chi}^\dagger \tilde{M}^{*} \sigma_2 \tilde{\chi}^* = \frac{1}{2} \chi^\dagger U^\dagger U^{*} \sigma_2 U^{*} \chi^* = \frac{1}{2} \chi^\dagger M^{*} \sigma_2 \chi^*. \]

For any complex symmetric matrix \( M \) there is a unitary matrix \( U \) that would make \( \tilde{M} = U^{*} MU^\dagger \) real and diagonal; its diagonal elements are the physical masses of \( N \) species of fermions. For the present purposes, I do not care about diagonalization but from now on I assume that \( U \) is chosen such that the \( \tilde{M} \) matrix is real.
In part (e) we saw that for a real mass matrix, the \textbf{CP} transform defined according to eq. (15) is a good symmetry of the theory. Consequently, in the basis of $\tilde{\chi}_j$ fields, we have a \textbf{CP} symmetry that acts as

$$\textbf{CP} : \tilde{\chi}'(x') = i\sigma_2 \tilde{\chi}^*(x).$$ \hspace{1cm} (S.118)

In terms of the original $\chi_j$ fields, this symmetry acts as

$$\textbf{CP} : \chi'(x') = U^\dagger\tilde{\chi}'(x') = U^\dagger i\sigma_2 \tilde{\chi}^*(x) = i\sigma_2 U^\dagger (U\chi(x))^* = i\sigma_2 U^\dagger U^* \chi^*(x).$$ \hspace{1cm} (S.119)

In other words, it acts according to eq. (16) where the $C$ matrix is

$$C = iU^\dagger U^*. \hspace{1cm} (S.120)$$

It remains to show that $C^\dagger M C = -M^*$. Remember that we have chosen the $U$ matrix such that $\tilde{M}$ is real,

$$\tilde{M} = \tilde{M}^* \implies U^* M U^\dagger = (U^* MU^\dagger)^* = UM^* U^\dagger \implies U^\dagger U^* M U^\dagger U^* = M^*. \hspace{1cm} (S.121)$$

Consequently, for the $C$ matrix as in eq. (S.120),

$$C^\dagger M C = -(U^\dagger U^*)^\dagger M U^\dagger U^* = -U^\dagger U^* M U^\dagger U^* = -M^*. \hspace{1cm} (S.122)$$

Q.E.D.

Problem 3(f):

Note: in eq. (18–22) the $d_j(x)$ and $u_j(x)$ are LH Weyl spinor fields for the quarks; the $d_j$ are the quarks of charge $-\frac{1}{3}$, namely d, s and b; the $u_j$ are quarks of charge $+\frac{2}{3}$, namely u, c, and t. Only the flavor indices of quarks are shown in my notations; the colors and the Weyl spinor indices are suppressed.

The \textbf{CP} symmetry acts similarly on all quark flavors,

$$\textbf{CP} : d'_j(x') = i\sigma_2 d^\dagger(x), \quad u'_j(x') = i\sigma_2 u^\dagger(x), \quad d^{\dagger\dagger}(x') = -id_j(x), \quad u^{\dagger\dagger}(x') = -iu_j(x).$$ \hspace{1cm} (S.123)

Consequently,

$$u^i \bar{\sigma}^\mu d_j \to u^{\dagger\dagger} \bar{\sigma}^\mu d'_j = u^i \sigma_2 \bar{\sigma}^\mu \sigma_2 d^\dagger j^* = u^i (\sigma^\mu)^\dagger (d^{\dagger\dagger})^\dagger = -d^\dagger \sigma^\mu u_i \hspace{1cm} (S.124)$$
and likewise
\[ d^j \bar{\sigma}^\mu u_i \to d^j \bar{\sigma}^\mu u'_i = -u^i \sigma^\mu d_j. \]  
(S.125)

Plugging these formulae into eq. (22) for the weak interactions, we obtain
\[
\sum_{i,j} V^j_i u^i \bar{\sigma}^\mu d_j \to -\sum_{i,j} V^j_i d^j \sigma^\mu u_i \quad \text{and} \quad \sum_{i,j} V^{*j}_i d^j \bar{\sigma}^\mu u_i \to -\sum_{i,j} V^{*j}_i u^i \sigma^\mu d_j,
\]
or in matrix form,
\[
u^\dagger V \bar{\sigma}^\mu d \to d^\dagger V^\dagger (-\sigma^\mu) u \quad \text{and} \quad d^\dagger V^\dagger \bar{\sigma}^\mu u \to u^\dagger V^* (-\sigma^\mu) d. \]  
(S.126)

In terms of the weak interaction Lagrangian (22), this makes for 3 differences in the fermionic factors \( u^\dagger \cdots d \) and the \( d^\dagger \cdots u \) multiplying the weak fields \( W^\pm_\mu \):

1. The \( u^\dagger \cdots d \) and the \( d^\dagger \cdots u \) factors trade places. At the same time, the charge conjugation exchanges \( W^+_\mu \leftrightarrow W^-_\mu \), cf. eq. (23), so these two changes cancel each other.

2. The \( \bar{\sigma}^\mu = (+1, -\sigma^\mu) \) matrices turn into \( -\sigma^\mu = (-1, -\sigma^\mu) \); this changes the signs of fermionic components for \( \mu = 0 \) but not for \( \mu = 1, 2, 3 \). At the same time, the charge conjugation changes the overall signs of the vector fields while space reflection changes the signs of the 3-vector components only; the net effect is to change the signs of the \( W^\pm_\mu \) for \( \mu = 0 \) but not for \( \mu = 1, 2, 3 \). Again, the effects of \( \text{CP} \) on fermions and on the vector fields cancel each other.

3. Finally, while the \( u^\dagger V \bar{\sigma}^\mu d \) and \( d^\dagger V^\dagger \bar{\sigma}^\mu u \) fermionic factors trade places, the CKM matrix \( V \) becomes transposed instead hermitian-conjugated. The discrepancy amounts to complex-conjugating the \( V \) matrix, \( V \to V^* \).

Altogether, the net effect of \( \text{CP} \) on the weak interaction Lagrangian (22) amount to changing the Cabibbo–Kobayashi–Maskawa matrix \( V \) to its complex conjugate \( V^* \). Thus, if the CKM matrix happens to be real, the weak interactions are CP invariant, but if it’s complex, they the CP symmetry is broken.

In real life, the CKM matrix does have a complex phase that cannot be removed by a field redefinition, and that makes the weak interactions break the \( \text{CP} \) symmetry.