Gauge Structure: Local Constructions and a Look Beyond

Frank Wilczek

Center for Theoretical Physics, MIT, Cambridge MA 02139 USA

February 25, 2013

Abstract

We discuss the mathematical structures associated with internal gauge symmetry. Specifically, we review of the standard local considerations, centered around the construction of covariant derivatives and curvature (= field strength). We first present those constructions in simple matrix form, then indicate the generalization to general Lie algebras and representations. We emphasize the fundamental nature of parallel transport ($\approx$ "nonintegrable frames"). That leads us into an introductory discussion of discrete gauge groups, gauge theory on discrete spaces, and gauge field topology.

Gauge symmetry, in general, involves the possibility of making ordinary symmetry transformations in a space-dependent way. The "space" involved need not be physical space or space-time in the conventional sense. Indeed, we will find that gauge symmetry arises in a natural way for spaces that parameterize families of quantum Hamiltonians, or geometric configurations of a molecule or of a cat, for example.

Our modern quantum theories of the strong and of the electroweak interactions – including, of course, as a most important special case electromagnetism (QED) – are based on postulated gauge symmetries. In those cases the symmetry transformations act within internal charge or color spaces. (The precise meaning of this jargon will be defined presently.) These symmetries, and their speculative extension in grand unified theories, might – or might not – have deeper origins in quantum theory and/or geometry, but so far it has been adequate, for a *very* wide range of practical purposes, to treat them as fundamental.

Topology, differential geometry, and general relativity bring in additional constructions, which we will consider in subsequent lectures.

In the basic formulation of gauge theories of fundamental interactions, local constructions mostly suffice. They center around the idea of gauge covariant derivatives.

Physicists usually first encounter gauge theory in its local form, which has a glorious history. Abelian gauge symmetry has its roots in the ideas of Faraday and Maxwell; came into its own in quantum theory, with seminal contributions from Hermann Weyl. We now recognize it as the foundational principle of quantum electrodynamics. The nonabelian generalization, introduced by Yang and Mills, guided us to today's Standard Model.

The fundamental idea of gauge symmetry is more general that the standard local constructions might immediately suggest. We get to the more general notion by considering parallel transport along paths that cannot be reduced to infinitesimals. This allows us to consider discrete gauge groups and, importantly, gauge theory on discrete spaces, such as lattices. The technique of lattice gauge theory, which is both fundamental and practical importance, is founded on these ideas. They are also the natural framework for considering gauge field topology. We will *introduce* those subjects in this chapter, too.

1 Covariant Derivatives

1.1 Matrix Realization

We start with "space" dependent fields $\psi(x)$ that support representations of some symmetry group G, according to

$$\psi(x) \rightarrow \Omega(g(x))\psi(x) \tag{1}$$

In different applications the "space" can be lots of different things. It can be curved, topologically nontrivial, even discrete. The $g(x)\epsilon G$ are position-dependent group elements and Ω is a group representation. We'll approach that full generality gently, though. In this subsection we assume that x lives within in a smooth patch – say a small ball – within a flat space (or space-time) R^d (R^{d+1}). We also assume that the group is one of $SO(n), SU(n), U(1)$ and that the representation is the basic vector representation in each case. For $U(1)$, we use $1{\times}1$ matrices, in the form $e^{i\theta}$.

For purposes of physics we often want to consider derivatives of our fields, but derivatives obey the ugly transformation law

$$\partial_\mu \psi \rightarrow \Omega \, \partial_\mu \psi + \partial_\mu \Omega \, \psi \tag{2}$$

which makes it difficult to form invariants using $\partial_\mu \psi$. On the other hand we typically need derivatives to form invariant Hamiltonians, or Lagrangians, that support nontrivial dynamics. The standard solution, ratified by Nature, is to add a correction piece, defining a *covariant derivative* by

$$\nabla_\mu \psi = \partial_\mu \psi + i A_\mu \psi \tag{3}$$

(The reason for the i will appear momentarily.) We suppose that A_μ, but not ∂_μ, transforms non-trivially. Then by demanding that $\nabla_\mu \psi$ transforms in the same way as ψ, we are led through

$$\begin{aligned} \nabla'_\mu(\Omega\psi) &= \Omega\,\partial_\mu\psi + \partial_\mu\Omega\,\psi + iA'_\mu\Omega\psi \\ &\overset{?}{=} \Omega(\partial_\mu + iA_\mu)\psi \end{aligned} \tag{4}$$

to require

$$A'_\mu = \Omega A_\mu \Omega^{-1} + i\partial_\mu\Omega\,\Omega^{-1} \tag{5}$$

For then indeed

$$(\nabla_\mu)'(\psi') = (\nabla_\mu\psi)' \tag{6}$$

or alternatively

$$\widetilde{\nabla_\mu\psi} = \widetilde{\nabla_\mu\psi} \tag{7}$$

What kind of object is A_μ? It must be consistent with the inhomogeneous transformation law Eqn. (5); that is, the additional term $i\partial_\mu\Omega\,\Omega^{-1}$ must be consistent with the assumed structure of A_μ. For the $U(1)$ case, the additional term is a real vector field:

$$i\partial_\mu e^{i\theta(x)}\,e^{-i\theta(x)} = -\partial_\mu\theta \tag{8}$$

In the $SU(n)$ case, we encounter the inhomogeneous term

$$i\partial_\mu\Omega\,\Omega^{-1} = i\partial_\mu\Omega\,\Omega^\dagger \tag{9}$$

From

$$0 = \partial_\mu(\Omega\Omega^\dagger) = \partial_\mu\Omega\,\Omega^\dagger + \Omega\,\partial_\mu\Omega^\dagger \tag{10}$$

we derive that the inhomogeneous term is Hermitean. Similarly, from the unit determinant condition $\det\Omega = 1$ we derive

$$\begin{aligned} 0 &= \partial_\mu(\det\Omega) \\ &= \partial_\mu e^{\operatorname{Tr}\,\log\Omega} \\ &= \partial_\mu\operatorname{Tr}\,\log\Omega\,e^{\operatorname{Tr}\,\log\Omega} \\ &= \operatorname{Tr}\,(\Omega^\dagger\,\partial_\mu\Omega)e^{\operatorname{Tr}\,\log\Omega} \end{aligned} \tag{11}$$

and conclude that the inhomogeneous term has zero trace. Similarly, in the $SO(n)$ case we find that the inhomogeneous term is an antisymmetric purely imaginary matrix. The factor i, which appears unnatural here, is the price of indulging the physicists' preference for Hermitean matrices.

In each case, then, the inhomogeneous term lives naturally in the standard, minimal representation of the infinitesimal version of the group, i.e. its Lie algebra. For example if we write $U \equiv 1 + i\epsilon M$, and demand U satisfies the unitarity and unit determinant conditions to order ϵ, we find precisely that M must be an Hermitean matrix with zero trace. So we can accommodate the inhomogeneous terms if A_μ is a vector field of members of the Lie algebra of G.

Conversely, it's not difficult to convince yourself that no smaller space of A_μ will serve. For at any given point x, we can transform $A_\mu(x)$ into any element of the Lie algebra, by an appropriate local symmetry operation.

1.2 More General Groups and Representations

Now let us move to define the covariant derivative for general Lie groups – actually, as we'll see, in effect for Lie algebras – and representations.

We let λ_a be a basis for the Lie algebra of G, and consider a representation defined by infinitesimal generators τ^a. We have the structure constants

$$[\lambda_a, \lambda_b] \equiv i f_{ab}^c \lambda_c \tag{12}$$

which also appear in the commutators of the τ^a – that is the defining property of a representation:

$$[\tau_a, \tau_b] \equiv i f_{ab}^c \tau_c \tag{13}$$

To see how this relates to invariance under the group, note that for $g\epsilon G$ in the neighborhood of the identity we have, on the group itself

$$g(\rho) = e^{i\lambda_a \rho^a} \approx 1 + i\lambda_a \rho^a \tag{14}$$

and for the representation Ω_τ corresponding to the τ_a

$$\Omega_\tau(g(\rho)) = e^{i\tau_a \rho^a} \approx 1 + i\tau_a \rho^a \tag{15}$$

(See the following paragraph for more on the mathematical niceties.) Let us spell out the group representation condition

$$\Omega_\tau(g(\rho_1)) \, \Omega_\tau(g(\rho_2)) = \Omega_\tau(g(\rho_1)g(\rho_2)) \tag{16}$$

4

to second order in ρ. We have

$$
\begin{aligned}
g(\rho_1)g(\rho_2) &\approx (1 + i\rho_1 \cdot \lambda - \frac{1}{2}(\rho_1 \cdot \lambda)^2)(1 + i\rho_2 \cdot \lambda - \frac{1}{2}(\rho_2 \cdot \lambda)^2) \\
&\approx 1 + i(\rho_1 + \rho_2) \cdot \lambda - \frac{1}{2}((\rho_1 + \rho_2) \cdot \lambda)^2 - \frac{1}{2}[\rho_1 \cdot \lambda, \rho_2 \cdot \lambda] \\
&= 1 + i(\rho_1 + \rho_2) \cdot \lambda - \frac{1}{2}((\rho_1 + \rho_2) \cdot \lambda)^2 - \frac{i}{2}\rho_1^a \rho_2^b f_{ab}^c \lambda_c \\
&\approx e^{i\lambda_c((\rho_1+\rho_2)^c - \frac{1}{2}f_{ab}^c \rho_1^a \rho_2^b)}
\end{aligned}
$$

$$(17)$$

and so

$$
\Omega_\tau(g(\rho_1)g(\rho_2)) \approx e^{i\tau_c((\rho_1+\rho_2)^c - \frac{1}{2}f_{ab}^c \rho_1^a \rho_2^b)} \tag{18}
$$

whereas

$$
\begin{aligned}
\Omega_\tau(g(\rho_1))\,\Omega_\tau(g(\rho_2)) &\approx (1 + i\tau \cdot \rho_1 - \frac{1}{2}(\tau \cdot \rho_1)^2)(1 + i\tau \cdot \rho_2 - \frac{1}{2}(\tau \cdot \rho_2)^2) \\
&\approx e^{i(\tau \cdot (\rho_1+\rho_2) - \frac{1}{2}[\tau\cdot\rho_1, \tau\cdot\rho_2])}
\end{aligned}
$$

$$(19)$$

The condition that these two expressions agree is precisely the Lie algebra representation condition Eqn. (13).

Mathematical niceties: We have not introduced the concepts of group manifold or Lie algebra in a mathematically precise or general way. There are many books that do that, and developing that material here would crowd out the further directions and applications I want to show you. As a compromise, I've shown you concretely, but without pretense of rigor, the origin of those concepts in the most important cases for physics (so far) – $SU(n), SO(n), U(1)$ – where the group manifolds are conveniently realized as spaces of matrices and the Lie algebra as providing a sort of linear approximation to the group near the origin.

If you'd like to go deeper into the foundations of Lie group theory, a good, clean but rather dry source is Nakahara *Geometry, Topology, and Physics*; a quirky but stimulating source, that I personally like very much, is Gilmore *Lie Groups, Lie Algebras, and Some of Their Applications*.

Here I'll adopt a compromise between simplicity and generality. We'll always consider that our group manifold G is realized by a set of matrices – that is, subsets of R^{n^2} or C^{n^2} subject to algebraic constraints, as in our standard examples, with the group operations implemented by matrix multiplication. The Lie algebra is then realized as the tangent space near the origin, which is also naturally a space of matrices – now, a linear one – and

the basic (Lie) product is simply matrix commutation. A major result of the general theory is that in the end surprisingly little is lost by making these assumptions. Specifically, in the case of compact connected Lie groups *nothing* is lost – they all can be faithfully represented as matrix groups. Stillwell *Naive Lie Theory* is a nice book (with a great title) that develops Lie theory systematically from this perspective.

After these preliminaries and understandings, we can develop a more flexible version of the covariant derivative, allowing both the Lie group G and the representation Ω to range beyond our previous, canonical examples.

Let us assume that we have $g(x_0) = 1$ at the point x_0 of interest. (That involves no real loss of generality, since we can trivially factor out a space-independent transformation if necessary.) We then expand in its neighborhood

$$
\begin{aligned}
g(x) &= 1 + i\lambda_a \rho^a \\
\Omega_\tau(g(x)) &= 1 + i\tau_a \rho^a
\end{aligned}
\tag{20}
$$

and impose our basic condition Eqn. (4) to linear order, starting from the *ansatz*

$$
\nabla_\mu = \partial_\mu + iA_\mu \cdot \tau
\tag{21}
$$

We compute

$$
\begin{aligned}
\nabla'_\mu \psi' &\approx (\partial_\mu + iA_\mu \cdot \tau + i(\delta A_\mu) \cdot \tau)(\psi + i\tau \cdot \rho\,\psi) \\
&\overset{?}{\approx} (1 + i\tau \cdot \rho)(\partial_\mu + iA_\mu \cdot \tau)\psi \\
&\approx (\nabla_\mu \psi)'
\end{aligned}
\tag{22}
$$

which leads us after some simple algebra to

$$
i\delta A_\mu \cdot \tau\,\psi = (-i\partial_\mu \rho \cdot \tau - [\tau \cdot \rho, A_\mu \cdot \tau])\psi
\tag{23}
$$

Crucially now, the representation condition Eqn. (13) allows us to rewrite the last term using

$$
[\tau \cdot \rho, A_\mu \cdot \tau] = if^c_{ab}\,\rho^a A^b_\mu \tau_c
\tag{24}
$$

and then to satisfy Eqn. (23) with

$$
\delta A^a_\mu = -\partial_\mu \rho^a - f^a_{bc}\,\rho^b A^c_\mu
\tag{25}
$$

A most notable feature of Eqn. (25) is that the dependence on τ has fallen away. As a consequence, the single gauge vector potential field A^a_μ suffices to build covariant derivatives for all representations.

6

1.3 Examples and Comments

1. For the group $U(1)$, the structure constants f^a_{bc} vanish, and we recover the gauge transformation law familiar from electrodynamics. The irreducible unitary representations of the group, parameterized as the unit circle with angle θ, are of the form

$$\Omega_n(\theta) = e^{in\theta} \tag{26}$$

with n an integer. The corresponding representation matrices are

$$\tau_n = n \tag{27}$$

and the covariant derivative is

$$\nabla_\mu = \partial_\mu + inA_\mu \tag{28}$$

For the additive group R^+ of real numbers, the structure constants likewise vanish. We get the same structure of irreducible unitary representations and covariant derivatives as for $U(1)$, with the sole difference that n can be an arbitrary real number, not necessarily an integer.

2. For the group $SO(3)$ of proper rotations in three dimensional space, the Lie algebra is the familiar algebra of quantum angular momentum:

$$[J_a, J_b] = i\epsilon^c_{\ ab} J_c \tag{29}$$

(The raised and lowered indices have no real significance in this case, but my fingers tremble and refuse to type out contractions over two upper or two lower indices ...). The gauge potential has three internal indices, and obeys the transformation law

$$\delta A^a_\mu = -\partial_\mu \rho^a - \epsilon^a_{\ bc}\, \rho^b A^c_\mu \tag{30}$$

for transformations near the identity. For spatially constant transformations, with $\partial_\mu \rho^a = 0$, we see that A^c transforms like a vector.

The irreducible unitary representations of $SO(3)$ are classified by their spin S, with S a non-negative integer. The representation is $2S + 1$ dimensional. I will not review this representation theory, which I hope is familiar to you, and is in any case well presented in many books. I particularly recommend the classic Edmonds *Angular Momentum in*

Quantum Mechanics, which is specifically devoted to the subject, reasonably short, and oriented toward physical applications. One elementary but fundamental point, related to our discussion in the preceding Chapter, is that the diagonal matrix representing rotation through θ around the $\hat{z}$ axis (in the standard implementation) is

$$\Omega_S(\hat{z}, \theta) = \text{diagonal}\,(e^{iS\theta}, e^{i(S-1)\theta}, ..., e^{-iS\theta}) \tag{31}$$

which we get by exponentiating J_3. and the requirement that it reduces to the identity for the rotation through 2π, which of course is the trivial (non-)rotation, constrains S to be an integer. Conversely, all such representations can be implemented globally, for example by using symmetric traceless S-index tensors.

The Lie algebra of the group $SU(2)$ is isomorphic to that for $SO(3)$ – that is, it has the same dimension and structure constants. Indeed, it is generated by the 2×2 zero trace Hermitean matrices, and the Pauli matrices $\sigma_1, \sigma_2, \sigma_3$, divided by two, satisfy Eqn. (29) with $J_j = \frac{\sigma_j}{2}$:

$$[\frac{\sigma_a}{2}, \frac{\sigma_b}{2}] = i\epsilon^c_{ab}\frac{\sigma_c}{2} \tag{32}$$

$e^{iJ_3\theta}$ is still leads to Eqn. (31), but back in the group manifold we do not have $e^{i\frac{\sigma_3}{2}2\pi} = 1$, but only $e^{i\frac{\sigma_3}{2}4\pi} = 1$. So we find that S is also allowed to be a half-integer. These representations can be implemented globally using symmetric $2S$-index multispinors. (For more on multispinors, which provide a very pretty, useful but underutilized approach to more advanced applications of rotation symmetry, see Appendix D.)

In all cases, the covariant derivative takes the form

$$\nabla_\mu = \partial_\mu + iJ_j^{(S)}A_\mu^j \tag{33}$$

where the $J_j^{(S)}$ are the angular momentum matrices for spin S.

3. To make contact between our special cases and the general case, we write

$$
\begin{aligned}
\Omega(x) &= e^{i\epsilon M(x)} \\
M(x) &= \rho^a\,\lambda_a \\
A_\mu(x) &= A_\mu^a(x)\,\lambda_a \\
[\lambda_a, \lambda_b] &= if^c_{ab}
\end{aligned}
\tag{34}
$$

and expand Eqn. (5) to first order in ϵ. The result is Eqn. (25).

4. The fact that the same A_μ appears in every covariant derivative has a most noteworthy physical consequence, known as *universality*. When we quantize the theory, so that A_μ creates and destroys gauge bosons (*e. g.* photons, W and Z bosons, or color gluons), we see that the coupling strength of those gauge bosons to matter fields is determined entirely group-theoretically, in terms of representation matrices, for which there only a discrete set of highly constrained possibilities. (R^+ is an exception, in that representations allow a continuous parameter. This is connected to the non-compactness of R^+, as I will discuss later.)

5. It might seem we have demonstrated too much, in that we seem to have left no room for coupling constants to appear. Where is the electron's charge, for example?

   In fact its deep origin lies elsewhere. We'll find it in the next Section.

2 Gauge Curvature, or Field Strength

Having seen how the general case proceeds, let us return to our standard examples, with $U(1), SU(n), SO(n)$ realized as numerical (phase), unimodular (that is, determinant unity) unitary, or proper (that is, determinant unity) orthogonal matrices. This allows us to use a transparent, concrete notation, while as we have discussed it entails no real loss of generality. Thus our transformations are of the type we originally contemplated in Eqn. (6), *viz.*

$$\psi'(x) \;=\; \Omega(x)\,\psi(x) \tag{35}$$

where $\psi(x)$ is appropriate vector – one-component complex, n-component complex, or n component real respectively.

It is useful, not to mention pretty, to recast our central definition/result for the covariant derivative,

$$
\begin{aligned}
\nabla'_\mu \psi' &= (\nabla_\mu \psi)' \\
\nabla'_\mu \Omega\psi &= \Omega\nabla_\mu\psi
\end{aligned}
\tag{36}
$$

into the form of an operator equation

$$\nabla'_\mu \;=\; \Omega\nabla_\mu\,\Omega^{-1} \tag{37}$$

Now we want to bring derivatives of A^a_μ into the dynamics. The straightforward derivative obeys a very messy transformation law, which makes it

9

difficult to see how to form invariant interactions. To finesse that difficulty, we observe that the commutator of covariant derivatives

$$[\nabla_\mu, \nabla_\nu] = i(\partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]) \equiv iF_{\mu\nu} \tag{38}$$

contains derivatives of A, but not of ψ. It will also, in view of Eqn. (37), obey the simple transformation law

$$F_{\mu\nu} \rightarrow \Omega F_{\mu\nu} \Omega^{-1} \tag{39}$$

$F_{\mu\nu}$ is what we want. It transforms simply, according to the so-called adjoint representation.

More precisely, the $F^a_{\mu\nu}$, defined through

$$
\begin{aligned}
F_{\mu\nu} &= \lambda_a F^a_{\mu\nu} \\
F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f^a_{bc} A^b_\mu A^c_\nu
\end{aligned}
\tag{40}
$$

are the real-valued fields. In this form, the equation defining $F^a_{\mu\nu}$ depends only on the structure constants, and is general.

2.1 Bianchi Identity

From the purely algebraic identity among operators ("Jacobi identity")

$$[\nabla_\alpha, [\nabla_\beta, \nabla_\gamma]] + [\nabla_\beta, [\nabla_\gamma, \nabla_\alpha]] + [\nabla_\gamma, [\nabla_\alpha, \nabla_\beta]] = 0 \tag{41}$$

we derive an important identity for the covariant derivatives of the field strength, known as the Bianchi identity:

$$\nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha} + \nabla_\gamma F_{\alpha\beta} = 0 \tag{42}$$

In electromagnetism, the Bianchi identity gives us two of the Maxwell equations, the magnetic Gauss law and Faraday's law of induction, as mathematical identities:

$$
\begin{aligned}
\vec{\partial} \cdot \vec{B} &= 0 \\
\vec{\partial} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0
\end{aligned}
\tag{43}
$$

This reduction of his induction law to hocus-pocus symbol manipulation might have come as a surprise to Faraday, who no doubt thought he was making an empirical discovery! From the present perspective, the significance of Faraday's law is that it shows that it is possible to represent physical electric and magnetic fields using potentials.

In general relativity, the Bianchi identity is closely connected to conservation of energy-momentum. It plays a crucial role in the mathematics of Chern classes, which describe topological aspects of gauge fields, as we'll be uncovering.

2.2 Yang-Mills Action and the Interpretation of Coupling Constants

To form local invariants, we can take any the trace of any product of $F_{\mu\nu}$s. Physically significant examples include

$$
\begin{aligned}
\mathcal{L}_{\mathrm{MYM}} &\propto \mathrm{Tr}\ F^{\mu\nu}F_{\mu\nu} \\
\mathcal{L}_{\mathrm{C1}} &\propto \mathrm{Tr}\ \epsilon^{\mu\nu}F_{\mu\nu} \\
\mathcal{L}_{\mathrm{C2}} &\propto \mathrm{Tr}\ \epsilon^{\alpha\beta\gamma\delta}F_{\alpha\beta}F_{\gamma\delta}
\end{aligned}
\tag{44}
$$

These are respectively (after appropriate normalization) the Maxwell-Yang-Mills, first Chern class, and second Chern class Lagrangian densities.

Let us first discuss the Maxwell-Yang-Mills action, which plays a basic role in the description of fundamental interactions.

The Maxwell-Yang-Mills density is relativistically invariant in any number of dimensions, and it is the object that enters into the Lagrangian density of the gauge theories of the standard model. It can occur with any coefficient. In the conventional formulation the coefficient is chosen so that the kinetic energy of the dynamical fields A^a are canonical; then it is a simple pure number. In fact, as may be familiar from electrodynamics, the number is $-\frac{1}{4}$. In electrodynamics, that leads us to

$$
-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}(\vec{E}^2 - \vec{B}^2)
\tag{45}
$$

We will have vector particles described by the A^a with the same field normalization for single quanta as we have for photons, if we choose λ_a matrices satisfying

$$
\mathrm{Tr}\lambda_a\lambda_b = \kappa\delta_{ab}
\tag{46}
$$

and use the Lagrangian density

$$
-\frac{1}{4\kappa}\mathrm{Tr}\ F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu}
\tag{47}
$$

In our test cases, it's easy to identify such λ matrices. For general Lie algebras, it is connected with the so-called "Killing form", which can be proven

to exist for so-called semisimple Lie algebras, including the Lie algebras corresponding to compact Lie groups. (In the literature, one commonly finds the choice $\kappa = \frac{1}{2}$.) That normalization can always be enforced, no matter what value the coefficient starts with, by rescaling A.

If instead we stick with the natural normalization of the covariant derivatives, then $\mathcal{L}_{\mathrm{MYM}}$ will occur with a coefficient proportional to $1/g^2$ (or $1/e^2$), the inverse coupling constant squared. Indeed, the canonically normalized A, $A_{\mathrm{canonical}}$, will be related to the natural A, A_{natural}, by

$$\frac{1}{g} A_{\mathrm{natural}} \;=\; A_{\mathrm{canonical}} \tag{48}$$

so that in covariant derivatives we have

$$\nabla_\mu \;=\; \partial_\mu + i A_{\mu\,\mathrm{natural}} \cdot \tau \;=\; \partial_\mu + i g A_{\mu\,\mathrm{canonical}} \cdot \tau \tag{49}$$

These considerations expose the true meaning of coupling constants: they measure the cost (in action, or ultimately in energy) of curvature in the gauge fields. $1/g^2$ is a measure of stiffness. Large $1/g^2$ – what we usually call *weak coupling* – means very stiff gauge fields, that are hard to excite. Small $1/g^2$ – *strong coupling* – means floppy gauge fields, that respond strongly to the motion of charges, and that are also, in the quantum theory, subject to vigorous quantum fluctuations. And they cast a clear light on universality, showing that there is just one coupling constant governing all the fields.

(Of course, before invoking this uniqueness result we must first break the group into simple factors, and the fields into irreducible representations. If we do not, there will be additional ways to form invariants, by taking sub-traces or by breaking up the representation matrices. In this regard, for example, let us note that if we have two independent local transformation laws, corresponding to groups G_1, G_2 and representations Ω_1, Ω_2 for the field ψ:

$$\begin{aligned}
\psi(x) &\;\rightarrow\; \Omega_1(x)\psi(x) \\
\psi(x) &\;\rightarrow\; \Omega_2(x)\psi(x)
\end{aligned} \tag{50}$$

then we can for a covariant derivative that works for both, simply by adding both $i A_{\mu\,1} \cdot \tau_1$ and $i A_{\mu\,2} \cdot \tau_2$ terms. Then when we change to the canonically normalized fields, two coupling constants will appear. I will not further belabor these important, but mathematically rather trivial, points.)

Because of the commutator terms, in nonabelian theories the A_μ^a appear as source terms of the (multicomponent) Maxwell-like equations for the A^a. Unlike the electrodynamic photon, nonabelian gluons obey nonlinear field equations, and are intrinsically self-interacting.

3 Parallel Transport as Primary Concept

3.1 From Potentials to Parallel Transporters

The covariant derivative, roughly speaking, gives us a way of comparing the values of a field at infinitesimally nearby points, that has a decent respect for the local symmetry of the field. A deeper consideration of that issue – comparing values of transformable fields at different points – both casts new light on the meaning of covariant derivatives, and motivates fruitful generalizations of the local gauge theory constructions we've considered thus far.

Previously, we considered the transformations $\Omega_\tau(g(x))$, in terms of which we phrase local or gauge symmetry, "actively", as implementing changes in fields. We can also consider them "passively", as changes in the way we view the fields – or, in proper mathematical language, as implementing changes of basis in the internal spaces where the $\psi(x)$ live. The passive perspective is more appropriate in one important respect – it emphasizes the fact that is we have several different kinds of fields $\psi^{(\kappa)}(x)$, perhaps transforming according to different representations, their transformations are all governed by the same $g(x)$. In fact, we can go a step further: The basis change really amounts to an x-dependent choice of elements of G; having chosen those elements, we implement the required basis changes, according to the appropriate group representation. (These ideas take us close to the mathematical language of fiber bundles.) From this new perspective, then, we think of gauge transformations as position-dependent changes of basis in the internal spaces, and gauge invariant quantities as quantities that are independent of such changes.

Along any given curve, we can choose bases that are aligned. To do this, we move our field in the straightest way possible, as measured by the covariant derivative. This defines parallel transport along the curve. Thus we impose

$$\nabla_\mu \psi = (\partial_\mu + iA_\mu)\psi = 0 \tag{51}$$

This equation can be "solved", formally, in terms of a path-ordered integral

$$\psi(x(\gamma(1))) = [P \exp -i \int_{x(\gamma(0))}^{x(\gamma(1))} A_\mu dx^\mu] \, \psi(x(\gamma(0))) \tag{52}$$

where $x(\gamma)$ indicates the path. The path-ordered integral is defined by the power series expansion of the exponential, where however in taking the

13

products of $A_\mu(x(\gamma))$ factors at different points, we always take the A with whose argument has the smaller value of γ to act first, i.e. on the right. Thus differentiating the exponentiated path-ordered integral brings down the value of the integrand at the upper limit, as a factor on the left.

The factor inside the brackets in Eqn. (52) defines the *parallel transporter* along γ. It can be defined for any curve γ. We'll introduce the condensed notation

$$\tau_\gamma(1 \leftarrow 0) \equiv P \exp -i \int_{x(\gamma(0))}^{x(\gamma(1))} A_\mu dx^\mu \tag{53}$$

The parallel transporters obey a composition rule: If $\gamma^{A \cdot B}$ is the path we get by first traversing B, then continuing with A, we have

$$\tau_{\gamma^{A \cdot B}}(2 \leftarrow 0) = \tau_{\gamma^A}(2 \leftarrow 1)\tau_{\gamma^B}(1 \leftarrow 0) \tag{54}$$

The parallel transporter $\tau_\gamma(1 \leftarrow 0)$ behaves very nicely under gauge transformations, as follows:

$$\tau_\gamma(1 \leftarrow 0)' = \Omega(1)\tau_\gamma(1 \leftarrow 0)\Omega(0)^{-1} \tag{55}$$

To prove this, we differentiate both sides with respect to $x(1)$, along γ. From the definition Eqn. (53) we have

$$\frac{dx^\mu}{d\xi} \partial_\mu^{(1)} \tau_\gamma(1 \leftarrow 0)' = -i\frac{dx^\mu}{d\xi}A'_\mu \tau_\gamma(1 \leftarrow 0)' \tag{56}$$

where ξ parameterizes the curve γ, while

$$\begin{aligned}
&\frac{dx^\mu}{d\xi} \partial_\mu^{(1)} \left(\Omega(1)\tau_\gamma(1 \leftarrow 0)\Omega(0)^{-1} \right) \\
=\; &\frac{dx^\mu}{d\xi} [(\partial_\mu^{(1)}\Omega(1))\tau_\gamma(1 \leftarrow 0)\Omega(0)^{-1} + \Omega(1)\left(\partial_\mu^{(1)}\tau_\gamma(1 \leftarrow 0)\right)\Omega(0)^{-1}] \\
=\; &\frac{dx^\mu}{d\xi} [(\partial_\mu^{(1)}\Omega(1)\,\Omega(1)^{-1})(\Omega(1)\tau_\gamma(1 \leftarrow 0)\Omega(0)^{-1}) \\
&-\; \Omega(1)\,iA_\mu\Omega(1)^{-1}(\Omega(1)\tau_\gamma(1 \leftarrow 0)\Omega(0)^{-1})] \\
=\; &\frac{dx^\mu}{d\xi} (\partial_\mu^{(1)}\Omega(1)\,\Omega(1)^{-1} - \Omega(1)\,iA_\mu\Omega(1)^{-1})\,(\Omega(1)\tau_\gamma(1 \leftarrow 0)\Omega(0)^{-1})
\end{aligned} \tag{57}$$

The gauge transformation law for A shows that these two derivatives are equal. (We could have argued the same thing more simply, though perhaps

less convincingly and instructively, from the transformation law Eqn. (37), noting that ∇ defines P through Eqn. (51).) Since both sides of Eqn. (55) are equal – indeed, both are equal to the identity – at the initial condition $x(1) = x(0)$, they are equal for all values.

In view of its transformation law Eqn. (55), the parallel transporter allows us to construct invariants involving the field ψ at different points:

$$\psi^\dagger(x_1)\tau_\gamma(1 \leftarrow 0)\psi(x_0) \;\; = \;\; \text{invariant} \tag{58}$$

In this precise sense, the parallel transporter allows us to correlate bases of the internal spaces at distant points, while respecting the freedom that gauge symmetry allows in the choice of such bases individually.

Note well that while the transformation law Eqn. (55) for $\tau_\gamma(1 \leftarrow 0)$ depends only on the endpoints, the value of $\tau_\gamma(1 \leftarrow 0)$ itself depends on the whole curve γ.

If the parallel transporters do not depend on the path, but only its endpoints, then we will have an unambiguous parallelization of the bases at all different points. In that case, the gauge structure is trivial. Indeed, if that is the case, then we can make a gauge transformation using

$$\Omega(x) \;\; = \;\; \tau_\gamma(x \leftarrow x_0)^{-1} \tag{59}$$

where x_0 is an arbitrarily chose reference point, to reduce the transformed parallel transporters all to the identity. This of course also reduces A to 0. In general, however, there is path dependence. Thus although we can parallelize unambiguously along any given path, we cannot do it over all space. C. N. Yang calls this phenomenon *nonintegrable* transport, and sees it as the primary concept of gauge theory.

In physics the transporters are often called Wilson lines. Parallel transporters around closed loops satisfy a particularly simple transformation law:

$$\tau_\gamma(1 \leftarrow 1) \;\; \rightarrow \;\; \Omega(1)\tau_\gamma(1 \leftarrow 1)\Omega(1)^{-1} \tag{60}$$

As a consequence, their trace – indeed, all their eigenvalues – are invariant under the gauge symmetry. In physics, closed-path transporters are often called Polyakov loops.

For infinitesimal closed loops we have a version of Stokes' theorem

$$\oint_\gamma iA_\mu dx^\mu \;\; \approx \;\; \int_\mathcal{S} iF_{\mu\nu}dx^\mu \wedge dx^\nu \tag{61}$$

where $\mathcal{S}$ is an approximately flat surface whose boundary is γ, and the corrections vanish faster than the area. I'll leave the demonstration as an

exercise. As a consequence, if we can smoothly deform one path into another (with the same endpoints) through a region where F vanishes, then transport along the two paths will give the same result.

3.2   From Transporters to Potentials, and More

Now we can, and should, change our point of view, as Yang advises: We take transporters as the primary objects; the potentials should be derived from them. One can't simply invert Eqn. (53) to recover A_μ from the transports, according to

$$A(2)_\mu \overset{??}{=} -\partial_\mu \tau_\gamma(2 \leftarrow 1)(\tau_\gamma(2 \leftarrow 1))^{-1} \tag{62}$$

because the answer can depend on the choice of path γ.

The root of this problem goes deep. To compute A, as an element of a fixed Lie algebra, we want to realize it as an infinitesimal transformation in a fixed space. But when we differentiate parallel transporters, we are implicitly comparing different spaces, i.e. the internal spaces at different values of x. Or, to put it starkly, we can't use gauge-invariant quantities to compute gauge-variant ones, without making some extra choices. To address this problem, we should refer everything to a standard space. This is a geometric version of *gauge fixing*.

To do this, let us fix a reference point x_0 and for every point x in a contractible neighborhood of x_0 a path $\gamma(x)$ from x_0 to x. We can do all that smoothly. One simple procedure – once we've realized our space as a manifold – is to take our region as a disk, and transport along the radii emanating from x_0. Given this choice of paths we can refer everything to x_0, and represent the infinitesimal transports – which should be governed by A_μ – in a fixed space. Concretely, we define

$$A_\mu(x) = -\lim_{\Delta x^\mu \to 0} \frac{1}{\Delta x^\mu} (P(x + \Delta x)^{-1} \tau_{\text{short}}(x + \Delta x \leftarrow x) P(x)) \tag{63}$$

where

$$P(x) \equiv \tau_{\gamma(x)}(x \leftarrow x_0) \tag{64}$$

and τ_{short} is a direct path between the close-by points. Here any ambiguity introduced by the choice of short path will be of higher order in Δx. Thus we have

$$A_\mu = -P^{-1} K_\mu P - \partial_\mu(P^{-1}) P \tag{65}$$

where

$$K_\mu \equiv \partial_\mu \tau_{\text{short}}(x) \tag{66}$$

The first term in Eqn. (65) represents the naive wave of comparing reference spaces, while the second term is a sort of kinematic correction. Later, we will identify this second term as the general, abstract form of the *geometric matrix* (or "Berry phase").

If we chose a different set of paths, we'll get a bridging function that is different from P; call it Q. The new A_μ is then

$$A'_\mu = -\partial_\mu(Q^{-1})Q - Q^{-1}K_\mu Q \tag{67}$$

The two potentials are related by

$$A'_\mu = Q^{-1}PA_\mu P^{-1}Q - \partial_\mu(Q^{-1}P)(P^{-1}Q) \tag{68}$$

i.e. through a gauge transformation with

$$\Omega = Q^{-1}P \tag{69}$$

– just right to interpolate between our changed bases.

Through our construction Eqn. (65) we can move from transporters to gauge potentials; conversely, given those potentials we can construct transporters, as we did earlier. The reconstructed transporters won't necessarily coincide with the ones we started with, but they will agree in all their gauge-invariant consequences, *within the small patch* where Eqn. (65) applies.

But while the notion of covariant derivative, and the appearance of A to implement that notion, is intrinsically local, the notion of parallel transport is more flexible. There are several physically important aspects of parallel transport that cannot be captured by smooth A fields, which I'll now briefly introduce. We'll go into these matters more deeply later, in the context of significant applications.

3.2.1 Discrete Gauge Symmetry

Consider a punctured plane, with the origin removed. For any path $\gamma(1 \leftarrow 0)$ in this space, running from $\vec{x}_0$ to $\vec{x}_1$, we can define an integer $w(\gamma)$ as follows. If neither $\vec{x}_0$ nor $\vec{x}_1$ lies on the positive $\hat{x}$ axis, then w is the number of times γ passes through the positive $\hat{x}$ axis going up, minus the number of times γ passes through the positive $\hat{x}$ axis going down. If either of $\vec{x}_0$ does lies on the positive $\hat{x}$ axis, displace that axis upward by an infinitesimal, and then count as before. For closed paths γ, w counts the number of times γ winds around the origin. Let $g\epsilon G$ be an element of a group G. Here G need not

be a Lie group – it could be the additive group Z of integers, for example, or the additive group Z_n of the integers modulo n, or (say) the permutation group S_3 on three objects.

With those understandings, we can define parallel transporters on the punctured plane by

$$\tau_\gamma^D(1 \leftarrow 0) \equiv \Omega(g)^{w(\gamma)} \tag{70}$$

It's not hard to see that these transporters satisfy the composition rule Eqn. (54).

For closed paths, $w(\gamma)$ counts the net number of times that the path γ winds around the origin. If we monitor the angle ϕ swept out by the path, as viewed from the origin, then $w = \frac{\phi}{2\pi}$. Any closed path that does not wind around the origin gives trivial parallel transport, so we'd be inclined to say – in the spirit of Eqn. (61) – that it enclosed zero field strength. Nevertheless, the transporter is definitely not trivial for all paths. We can say it is locally trivial, but globally nontrivial.

One can think of τ^D as representing the effect of a very singular A field, concentrated on a ray infinitesimally above the positive $\hat{x}$ axis.

Alternatively, we can use a patch construction, similar to what we did for the magnetic monopole on a sphere. Here we take two regions corresponding to angular sectors in the punctured plane, for example

$$
\begin{aligned}
I &: \quad \frac{-\pi}{4} < \theta < \frac{5\pi}{4} \\
II &: \quad \frac{3\pi}{4} < \theta < \frac{9\pi}{4}
\end{aligned}
\tag{71}
$$

and define wave functions ψ^I, ψ^{II} on these regions. With each region, the gauge structure is trivial. (Paths confined within either region give trivial, i.e. identity, parallel transport.) On the overlaps, we are required to have gauge transformations relating ψ^I and ψ^{II}. The overlap region consists of two disconnected pieces, namely the angular slices

$$
\begin{aligned}
\alpha &: \quad \frac{-\pi}{4} < \theta < \frac{\pi}{4} \\
\beta &: \quad \frac{3\pi}{4} < \theta < \frac{5\pi}{4}
\end{aligned}
\tag{72}
$$

Thus we can impose conditions

$$
\begin{aligned}
\psi^I(x) &= \Omega^{(\alpha)}\psi^{II}(x) \qquad x \epsilon \alpha \\
\psi^I(x) &= \Omega^{(\beta)}\psi^{II}(x) \qquad x \epsilon \beta
\end{aligned}
\tag{73}
$$

18

To maintain the trivial transport with the regions I, II – and in fact, for simple consistency – we want $\Omega^{(\alpha)}$ and $\Omega^{(\beta)}$ to be constants, as functions of x. If they are the same constant, then that constant transformation can be absorbed into a re-definition of ψ^{II}, and the gauge structure is completely trivial. If they are not the same constant, however, they cannot be so removed. Notionally following the motion of a point x circling around the origin, starting (say) on branch I, we see the basis of the internal space evolve as follows. At some time during the sojourn in region β we must switch to branch II, so that we can get past $\theta = \frac{5\pi}{4}$. To do that, we apply $(\Omega^{(\beta)})^{-1}$. Then at some time during the sojourn in region α we switch back to branch I, to complete the circuit. That requires applying $\Omega^{(\alpha)}$. So in total we have made the transformation $\Omega^{(\alpha)}(\Omega^{(\beta)})^{-1}$. Comparing with our previous approach, we have

$$\Omega^{(\alpha)}(\Omega^{(\beta)})^{-1} \ \leftrightarrow \Omega(g) \tag{74}$$

Discrete gauge groups can arise from continuous ones as effective theories at low energy, after gauged spontaneous symmetry breaking. In that context, we need have no true singularity in the gauge potentials. To be specific, let us supposed that the continuous group G breaks to a discrete subgroup H. Then in place of true singularities in fields, we will find failures of the reduced description, based solely on H, to be valid globally. The fields escape singularity, by wandering into a larger internal space.

3.2.2 Relation to Aharonov-Böhm

Closely related to these considerations is the Aharonov-Böhm effect. There one considers an inaccessible, long "secret solenoid" piercing the plane of motion near the origin (for example, it might be shielded by a large potential barrier). The $\vec{B}$ field vanishes outside the solenoid, but it induces a non-trivial parallel transport, of precisely the kind we've just analyzed. Here the underlying gauge group is continuous, but the phenomenon of "locally trivial, globally non-trivial" gauge structure associated with non-trivial spatial topology – specifically, punctures – is the common theme.

Again – more later!

3.2.3 Lattice Gauge Symmetry

A more radical step beyond the local construction, which is both profound and practically useful, abandons potentials and covariant derivatives altogether, in favor of parallel transporters. By taking that step, we can formu-

late gauge symmetry on *discrete spaces*, since we no longer need to consider derivatives!

To focus the discussion, let us consider the physically interesting application to formulating QCD, the $SU(3)$ gauge theory that governs the strong interaction, in a manner acceptable to a computer. As a matter of fact it has proved extremely fruitful to do that, since powerful computers have enabled us to calculate many consequences of the theory that would otherwise be far beyond our reach. For now we will consider only the pure gluon theory version of the theory.

I will take it as granted that the crux of the problem is to formulate a suitable action for gauge fields in four Euclidean dimensions, anticipating the use of path integrals in imaginary time. We'll introduce those ideas properly in due course.

To discretize the problem, we introduce a four-dimensional lattice of points labelled by integers (n_1, n_2, n_3, n_4). We can take integers modulo some finite L, and apply periodic boundary conditions, to keep things completely finite. On each link connecting nearest neighbors, we have an $SU(3)$ matrix $U_x(n_1, n_1 + 1; n_2; n_3; n_4), U_y(n_1; n_2, n_2 + 1; n_3; n_4), U_z(n_1; n_2; n_3, n_3 + 1; n_4)$, or $U_\kappa(n_1; n_2; n_3; n_4, n_4 + 1)$, where the subscripts indicate the direction of the link. We should think of these Us as parallel transporters along the links. Crucially, whatever we think, that is what they'll become in the continuum limit, as we'll discuss momentarily. The inverse matrices transport us in the opposite direction.

Our action is proportional to a sum over contributions from the minifaces, or "plaquettes", typified by these in the $x - y$ plane:

$$
\begin{aligned}
Pl_{xy}(n_1, n_2, n_3, n_4) &= \quad \text{Tr}\, U_{\text{right}}\, U_{\text{up}}\, U_{\text{left}}\, U_{\text{down}} \\
U_{\text{right}} &\equiv \quad U_x(n_1, n_1 + 1; n_2; n_3, n_4) \\
U_{\text{up}} &\equiv \quad U_y(n_1 + 1; n_2, n_2 + 1; n_3, n_4) \\
U_{\text{left}} &\equiv \quad U_x^\dagger(n_1, n_1 + 1; n_2 + 1; n_3; n_4) \\
U_{\text{down}} &\equiv \quad U_y^\dagger(n_1; n_2, n_2 + 1; n_3; n_4)
\end{aligned}
\tag{75}
$$

or more precisely its real part.

Where is the gauge symmetry? The desired interpretation of Us as transporters requires that we implement Eqn. (55). So the transformations $\Omega(n_1; n_2; n_3; n_4)$ will be associated with lattice sites, and we will have transformations typified by

$$
\begin{aligned}
&U_x(n_1, n_1 + 1; n_2; n_3; n_4) \;\rightarrow \\
&\Omega(n_1 + 1; n_2; n_3; n_4)\, U_x(n_1, n_1 + 1; n_2; n_3; n_4)\, \Omega(n_1; n_2; n_3; n_4)^\dagger
\end{aligned}
$$

for the Us.

The plaquette quantities Pl are invariant under these transformations. Note that we have an *enormous* symmetry group – $SU(3)^{L^4}$, since there is an independent $SU(3)$ for each lattice site. We take our action proportional to the sum over all the plaquettes of the corresponding Pls. This will be invariant under the lattice versions of translation and rotation symmetry, as well. A little more quantitively, we take

$$S \; = \; -c\frac{1}{g^2} \sum_{\text{plaquettes P}} \text{Re } Pl(P) \tag{77}$$

where c is a positive numerical constant. [It will be supplied in the final version of these notes.]

Now suppose that g is very small. Then there is a big penalty, in the action, for deviations of the Pl from the maximum. The Pl_{xy} are maximized – equal to 3 – when the product of the Us involved is the identity matrix. In that case, one can make a gauge transformations consistently, to bring each of those Us to the identity. So it should seem plausible that at weak coupling we can focus on configurations where the Us are all close to the identity, and only slowly varying with position. We then write then as $U = \exp iA$, defining slowly varying A fields. The leading invariant contribution, expanding in gradients, will then have the form of the continuum Maxwell-Yang-Mills action. In this way, we have constructed a discretized form of the continuum gauge theory, that retains an enormous exact symmetry group.

Appendix E develops these ideas in more detail, and takes them further.

4 Topological Gauge Actions

Besides the Maxwell-Yang-Mills term, there are a few more simple invariants we can form from low powers of the curvature, that can be considered as contributions to the action of a gauge invariant theory. These have a different flavor, in that they only exist in certain dimensions, and they have a topological character. They have played a significant role at the frontiers of research in recent years, and will be the focus of much of our attention later. Here, we make a first acquaintance.

4.1 1+1 Dimensions

For a $U(1)$ gauge theory in 1+1 dimensions, we have the term

$$\mathcal{L} \propto \epsilon^{\mu\nu} F_{\mu\nu} \tag{78}$$

Since

$$\epsilon^{\mu\nu} F_{\mu\nu} = 2\epsilon^{\mu\nu} \partial_\mu A_\nu = \partial_\mu (2\epsilon^{\mu\nu} A_\nu) \tag{79}$$

this term is a total derivative, and so it does not contribute to the classical equations of motion. On the other hand, we can fix the gauge $A_1 = 0$ and integrate over x to find

$$\int dt\, dx\, \epsilon^{\mu\nu} F_{\mu\nu} = -2 \int dt [A_0(\infty) - A_0(-\infty)] \tag{80}$$

This is the action associated with two particles of equal and opposite charge at plus and minus infinity. So we might expect that our term implements a capacitor plate, or in other words a constant background electric field. A more sophisticated treatment bears this out, as we'll see.

Note that the value of our term is associated with the large-scale, topological structure of the gauge field A. Applied in ordinary space, and integrated over a surface, it gives the total magnetic flux through that surface – a quantity whose topological significance we've already discussed, in connection with magnetic monopoles.

4.2 2+1 Dimensions

For a $U(1)$ theory in 2+1 dimensions, we have the so-called Chern-Simons term

$$\mathcal{L} \propto \epsilon^{\mu\nu\rho} A_\mu F_{\nu\rho} \tag{81}$$

Despite the explicit appearance of A_μ, this is "almost" gauge invariant, thanks to the Bianchi identity. Indeed, when we make the gauge transformation

$$A_\mu \rightarrow A_\mu - \partial_\mu \lambda \tag{82}$$

$F_{\mu\nu}$ is invariant, so

$$\begin{aligned}
\epsilon^{\mu\nu\rho} A_\mu F_{\nu\rho} &\rightarrow \epsilon^{\mu\nu\rho} A_\mu F_{\nu\rho} - \epsilon^{\mu\nu\rho} \partial_\mu \lambda F_{\nu\rho} \\
&= \epsilon^{\mu\nu\rho} A_\mu F_{\nu\rho} - \epsilon^{\mu\nu\rho} \partial_\mu \lambda F_{\nu\rho} \\
&= \epsilon^{\mu\nu\rho} A_\mu F_{\nu\rho} - \partial_\mu (\epsilon^{\mu\nu\rho} \lambda F_{\nu\rho})
\end{aligned} \tag{83}$$

So an action based on $\mathcal{L}$ will change by a total derivative, which at least doesn't change the classical equations of motion.

The Chern-Simons term is central to anyon physics and the theory of the quantum Hall effect, among other things.

There is also an important nonabelian generalization. It involves

$$\mathcal{L} \ \propto \ \mathrm{Tr} \ \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho - \frac{2}{3} A_\mu A_\nu A_\rho \tag{84}$$

which changes by a total derivative under the (infinitesimal) gauge transformation

$$\delta A_\mu \ = \ \partial_\alpha \Lambda - \Lambda\, A_\mu + A_\mu\, \Lambda \tag{85}$$

4.3 3+1 Dimensions

In 3+1 dimensions we have the term

$$\mathcal{L} \ \propto \ \mathrm{Tr} \ \epsilon^{\alpha\beta\gamma\delta} \, F_{\alpha\beta}\, F_{\gamma\delta} \tag{86}$$

In the $U(1)$ case it is manifestly equal to a total derivative:

$$
\begin{aligned}
\epsilon^{\alpha\beta\gamma\delta}\, F_{\alpha\beta}\, F_{\gamma\delta} \ &= \ 4\,\epsilon^{\alpha\beta\gamma\delta}\, \partial_\alpha\, A_\beta\, \partial_\gamma\, A_\delta \\
&= \ \partial_\alpha\, (\epsilon^{\alpha\beta\gamma\delta}\, A_\beta\, \partial_\gamma\, A_\delta)
\end{aligned}
\tag{87}
$$

In the nonabelian case it is also a total derivative, though slightly less obviously so.

This term, for QCD, leads to violations of parity P and time reversal T symmetry at empirically unacceptable levels, unless its coefficient is very small. This defines the so-called strong P, T problem, which motivates the consideration of Peccei-Quinn symmetry and axion physics as an attractive extension of the standard model. Recently the $U(1)$, electrodynamic version has appeared in the theory of topological insulators, realizing a vision I put forward in the 1980s.