7 Non–Abelian Gauge Theory

Perhaps the single most important class of quantum field theory to understand is non–Abelian Gauge Theory. The starting-point for any non–Abelian gauge theory is the statement that our world is not just a manifold M, but rather a *principal G-bundle*. In this section we'll begin by looking at these from a geometric and topological perspective, before going on to study particular

7.1 Principal bundles and vector bundles

$P \to M$. These words mean that P is a manifold that comes with a projection map $\pi : P \to M$, such that for any $x \in M$, $\pi^{-1}(x) \cong G$ for some Lie group G. The space $\pi^{-1}(x)$ is known as the *fibre* of P over x and is often denoted by P_x, while the space M is called the *base*. You should think of P as M with a copy of G attached at each point (see figure **??**). In physics, the Lie group G is known as the *gauge group*, while in maths it's often called the *structure group*. For example, electromagnetism is the case $G = U(1)$, while for reasons nobody really understands[36] the Standard Model has $G = SU(3) \times SU(2) \times U(1)$.

Principal bundles come with a natural (right) group action $G : P \to P$ that preserves the fibres. In other words, if $p \in P$ is a point in the fibre over $x \in M$ then acting with a group element $g \in G$ gives another point $pg \in P$, with the property that $\pi(pg) = \pi(p) = x$ so that pg and p both lie in the *same* copy of the fibre. Thus the group action allows you to move around within each copy of G, but does not move you around in M.

To get more of a handle on these abstract ideas, it's useful to describe the situation just in a small region. Given an open set $U \subset M$, a *local trivialization* is a choice of isomorphism

$$\Phi : \pi^{-1}(U) \to U \times G \tag{7.1}$$

and so gives a way to identify $P|_U$ with $U \times G$. Explicitly, if we're given a point $p \in \pi^{-1}(U) \subset P$, then we can always write $\Phi(p) = (\pi(p), \phi(p))$ where $\pi(p) \in U \subset M$ is just whatever point p projects to, and $\phi(p)$ is some group element. Exactly which group element we get will of course depend on exactly how we choose Φ, but we require that this is compatible with the action of G on the bundle itself. In other words, if $G : p \to pg$, then

$$\Phi(pg) = (\pi(pg), \phi(pg)) = (\pi(p), \phi(p)g)\,. \tag{7.2}$$

Notice that while it's true *locally* that any principal bundle looks like $U \times G$, this might not be true globally. The simplest example is to take $G = \mathbb{R}$ thought of as a one–dimensional Abelian group, and $M = S^1$. Then both the cylinder $S^1 \times \mathbb{R}$ and the Möbius strip are principal R-bundles — they both look locally like $U \times \mathbb{R}$ — but topologically they are different.

[36] As you'll learn if you're taking the Part III Standard Model course, the running of the coupling constants for each of the three semi-simple factors, together with the particular representations of $SU(3) \times SU(2) \times U(1)$ in which quarks and leptons lie, suggests that this group may just be a low-energy remnant of a larger 'grand unified gauge group'. This grand unified group is often thought to be $SU(5)$. Or $SU(5) \times U(1)$. Or perhaps $SO(10)$. Or maybe E_6. Like I said, no one really knows.

Now suppose that $\{U_\alpha\}$ are a collection of open sets in M. Given a local trivialization Φ_α of $\pi^{-1}(U_\alpha)$ on each region U_α, if the open sets overlap then we can ask how the trivializations are related on their common domain of definition $\pi^{-1}(U_\alpha \cap U_\beta)$. First, let's just consider what happens at one point $p \in U_\alpha \cap U_\beta$. By definition, both trivializations involve projecting p to the same point in M, but $\phi_\alpha(p)$ may be a different group element than $\phi_\beta(p)$. Nonetheless, since both $\phi_\alpha(p)$ and $\phi_\beta(p)$ are certainly both in G, we must be able to find a group element $t_{\alpha\beta} \in G$ such that

$$\phi_\beta(p) = \phi_\alpha(p) t_{\alpha\beta} \,. \tag{7.3}$$

This group element allows us to relate our two local trivializations at some point p. If we wish to compare our trivializations throughout $\pi^{-1}(U_\alpha \cap U_\beta)$ then we must allow $t_{\alpha\beta}$ to vary. Thus, given a pair of open sets U_α and U_β, we define a *transition function* to be a map

$$T_{\alpha\beta} : U_\alpha \cap U_\beta \to G \,. \tag{7.4}$$

which we can think of as a G-valued function $t_{\alpha\beta}(x)$, defined at each $x \in U_\alpha \cap U_\beta$. So that we can compare Φ_α with Φ_β as well as compare Φ_β with Φ_α, we ask that the transition functions are invertible: $t_{\beta\alpha}(x) = t^{-1}_{\alpha\beta}(x)$. We also impose the compatibility relation $T_\alpha\gamma = T_\alpha\beta \circ T_\beta\gamma$ on triple overlaps $U_\alpha \cap U_\beta \cap U_\gamma$, which says that the result of comparing Φ_α with Φ_γ is the same as first comparing Φ_α with Φ_β and then comparing Φ_β with Φ_γ in any region where all three are defined. Finally, it's natural to ask that all these transition functions vary smoothly over $U_\alpha \cap U_\beta$, so we will.

In physics, the most common case of all this is when $U_\alpha = U_\beta$ are actually the *same* set U, and we're just comparing two different ways of identifying $\pi^{-1}(U)$ with $U \times G$. In this case, the local trivializations are thought of as choices of *gauge*, while the transition function is usually called a *gauge transformation*. For example, a familiar case might be to choose $M \cong \mathbb{R}^{3,1}$ and $G \cong U(1)$, whereupon for each x we could write $t(x) = \mathrm{e}^{\mathrm{i}\lambda(x)} \in U(1)$ with $\lambda(x)$ a gauge parameter in electrodynamics. Another example that should be familiar from General Relativity is to take M to be (curved) space-time and G to be $GL(d, \mathbb{R})$. In this case, a local trivialization is a choice of *coordinate system* above an open patch U, whereas the transition functions $T_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(d, \mathbb{R})$ are general coordinate transformations. Of course, in any open region of M there could be many valid coordinate systems, and a large part of the Principle of Relativity is the statement that the actual *physics* doesn't depend on which coordinates (= local trivialization) we use. You probably also know that spaces with non-trivial topology (*e.g.* just a circle S^1 or sphere S^d) cannot be described by just a single set of coordinates, which is why all our constructions are defined only locally.

Let me also point out that the mathematical picture of principal bundles is also the inspiration for efforts to recover the Standard Model from higher dimensional theories, initiated in the 1930s by Kaluza and later by Klein. For example, the Lie group $SU(2)$ is isomorphic to S^2, so a principal $SU(2)$-bundle over 'our' space–time M can be thought of as a six dimensional space–time where every point of M comes with a copy of S^2. If the sphere is very small, then present technology won't allow us to concentrate enough energy

in a small region so as to excite spherical harmonics on the S^2, since these oscillations will involve some very high frequency set by the inverse radius. Thus, at low energies, this theory should involve just the constant modes on S^2 and so look like a theory on M. We might hope to find some remnant of the spherical extra dimensions by examining how low–energy fields transform under higher dimensional coordinate transformations. You can find out more about these ideas in the Extra Dimensions course in Part III.

7.1.1 Choosing a representation

As you learned in the Symmetries, Particles & Fields course, whenever we're given a Lie group it's often a good idea to look at the *representations* of this group. Recall that a representation ρ is a choice of vector space (usually either $\mathbb{R}^r$ or $\mathbb{C}^r$) together with map

$$\rho : G \to \mathrm{Mat}(r;\mathbb{R}) \qquad \text{or} \qquad \rho : G \to \mathrm{Mat}(r;\mathbb{C}) \tag{7.5}$$

to the space of $r \times r$ matrices (with real or complex values) that tells us how elements of G act on the vector space. This map should be compatible with the group structure in the sense that

$$\rho(gh) = \rho(g) \circ \rho(h)\,, \tag{7.6}$$

where on the left gh denotes multiplication in G, whereas the *rhs* denotes matrix multiplication.

For example, the fundamental representation of the rotation group $SO(3)$ represents elements $g \in SO(3)$ by 3×3 matrices that are orthogonal and have unit determinant; these matrices tell us how the components of a standard vector $\mathbf{v} \in \mathbb{R}^3$ change as we rotate. The same group also has a spinor representation where an element $g \in SO(3)$ is represented by a 2×2 unitary matrix U_g with unit determinant, often written in terms of the Pauli sigma matrices σ as $\mathrm{e}^{\mathrm{i}\alpha_g\cdot\sigma}$, where the parameters α_g depend on which group element we're considering. This representation tells us how the two complex components of a spinor (such as an electron wavefunction) change under rotations.

Now, if we have a principal bundle then we have not just a single copy of a Lie group G, but a whole family of copies, one at each point of M. If we pick a representation ρ, we thus get a whole family of vector spaces. This structure is known as a *vector bundle* $E \to M$. In a vector bundle the fibre $\pi^{-1}(x)$ at each point $x \in M$ is now a vector space — the one we got when we chose a G-representation. Just like above, a vector bundle has local trivializations $\Phi : \pi^{-1}(U) \cong U \times \mathbb{C}^r$ and transition functions (or gauge transformations) are maps $T_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Mat}(r;\mathbb{C})$. If the principal bundle we started with is a subgroup of $GL(r;\mathbb{C})$ then these transition functions will preserve some structure on the vector space. For example, if $G = U(r) \subset GL(r;\mathbb{C})$ and we choose the fundamental r-dimensional (complex) representation, then the transition functions will be *unitary* matrices preserving the inner product $\sum_{a=1}^{1} |z^a|^2$ on each fibre, while if $G = SU(r)$ then the transition functions will additionally have *unit determinant*, and so will also preserve the top holomorphic form[37]

$$\epsilon_{a_1\ldots a_r} \mathrm{d}z^{a_1} \wedge \mathrm{d}z^{a_2} \wedge \cdots \wedge \mathrm{d}z^{a_r}$$

[37] Don't worry if you don't know what this means.

on each fibre.

Vector bundles are of relevance to physics because a charged matter field is a *section* of E. This is a map

$$s : M \to E \tag{7.7}$$

that obeys $\pi \circ s = \mathrm{id}$. Given a choice of trivialization, we can think of this section as the assignment of a vector $s^a(x) \in \mathbb{C}^r$ (for a complex vector bundle) to each point x. If we change our local trivialization using a gauge transform, then the particular vector components we get will change according to

$$s_\beta(x) = s_\alpha(x)\, t_{\alpha\beta}(x) \tag{7.8}$$

where $t_{\alpha\beta}(x)$ are the transition functions. We'll sometimes write $\Omega^0_M(E)$ to denote the space of all smooth sections of $E \to M$.

As a simple example, consider a complex scalar field on M. Usually, we think of this as just a function $\phi : M \to \mathbb{C}$ where $\phi : x \mapsto \phi(x)$ is the value of the field at $x \in M$. However, if this scalar carries electric charge q, then it doesn't really have any preferred 'value' because, as you learnt in electrodynamics, we can change $\phi(x) \to \mathrm{e}^{\mathrm{i}q\lambda(x)}\phi(x)$ by a gauge transform. The correct interpretation of our charged scalar is that ϕ is really section of a vector bundle $E \to M$ associated to the principal $U(1)$ bundle of electromagnetism. One we pick a local trivialization — *i.e.* pick a gauge — then we can think of $\phi(x)$ as a (one component) vector, at least for some open region $U \subset M$. However, there's no preferred way to choose this gauge, and making different choices (changing gauge) will cause ϕ to transform as is familiar.

The electromagnetic example is special because the only irreducible representations of $U(1)$ are one–dimensional; they're just labelled by the *charge* q of the field, with $q = 0$ being the trivial representation. As you saw in Symmetries, Particles & Fields, non-Abelian groups typically have (infinitely) many different irreps. For example, **BLAH BLAH BLAH**

Notice that the statement that a section is a map $s : M \to E$ means that, once we've picked a gauge, the resulting field *depends only on* $x \in M$. This is where the notion of a gauge theory differs from Kaluza–Klein theory, where the fields are allowed to vary over the whole higher–dimensional space.

Above, we've constructed vector bundles from principal bundles by making a choice of representation. Finally, let me mention that we can also go the other way and construct a principal bundle starting from a vector bundle, at least in the case of a matrix Lie group (those that are subgroups of $GL(r)$). For given any rank r vector bundle $E \to M$, we define the *frame bundle* to be the principal $GL(r)$-bundle whose fibre $\pi^{-1}(x)$ is the collection of all basis vectors (= frames) in the fibre E_x. There is a natural action of $GL(r)$ on this frame bundle, relating any pair of frames. If we wish to construct a principal G-bundle for a subgroup of $GL(r)$ then as above we just require that our basis vectors are compatible with some extra structure. For example, in the real case we obtain a principal $O(r)$-bundle by asking that our basis vectors are orthogonal, and a principal $SO(r)$ bundle by asking that they also define a fixed volume element.

The most common Lie groups that arise in physics are indeed matrix Lie groups, so the two viewpoints are equivalent. However, in some exotic theories (especially string theory and some grand unified theories) exceptional Lie groups such as E_6 play an important role, so the fundamental picture is really that of principal bundles. With this caveat, I'll mostly use vector bundles from now on.

7.1.2 Connections and curvature

So far, we have described a vector bundle E as just a collection of vector spaces parametrized by a base space M. If we wish to write down the kinetic terms of any matter field, we will need derivatives and we'd usually write

$$v^\mu \partial_\mu \phi(x) \stackrel{?}{=} \lim_{\epsilon\to 0} \frac{\phi(x+\epsilon v) - \phi(x)}{\epsilon}\,, \tag{7.9}$$

where v^μ is a vector at x. However, for a charged field this expression is meaningless as it stands, because the two terms on the right live in completely different spaces: $\phi(x+\epsilon v)$ lives in the fibre $E_{x+\epsilon v}$ while $\phi(x)$ lives in a different copy E_x of the fibre.

To make sense of this, we need a way to compare vectors in different fibres, which is what a *connection*, or *covariant derivative* provides. This is a linear map

$$\nabla : \Omega^0_M(E) \to \Omega^1_M(E) \tag{7.10}$$

from the space of sections to the space of 1-forms (covectors) on M with values in E. The connection is defined by the properties of *linearity*:

$$\nabla(\alpha_1 s_1 + \alpha_2 s_2) = \alpha_1 \nabla(s_1) + \alpha_2 \nabla(s_2) \tag{7.11}$$

for any two sections s_1, s_2 and constants $\alpha_1\alpha_2$, and the *Leibniz rule*:

$$\nabla(fs) = df\, s + f\nabla(s)\,, \tag{7.12}$$

where $f \in \mathbb{C}^\infty(M)$ is a smooth function. More specifically, for every tangent vector v on M, the connection defines a derivative $v\cdot\nabla s = v^\mu \nabla_\mu(s)$, thought of as the derivative of our section $s(x)$ in the direction of v, and then the Leibniz rule says $v\cdot\nabla(fs) = fv^\mu\nabla_\mu s + (v^\mu\partial_\mu f)s$, where ∂_μ is the standard partial derivative of the function f. Notice that if ∇ and ∇' are any two connections, then the difference obeys

$$(\nabla - \nabla')(fs) = f(\nabla - \nabla')s\,. \tag{7.13}$$

Thus $(\nabla - \nabla')$ maps $\Omega^0_M(E) \to \Omega^1_M(E)$ in a way that is linear over *functions* $f \in C^\infty(M)$. Hence the difference between any two connections is an element of $\mathrm{Hom}(E, E\otimes T*M) \cong \mathrm{End}(E)\otimes T^*M$.

To understand what this means, let's again look in a small region. Suppose we have a trivialization $\Phi : E|_U \to U\times\mathbb{C}^r$. Then in this region, any section $s : U \to E$ can be thought of as a vector–valued function on U, *i.e.* given $s : M \to E$, we can write $\Phi\circ s : U \to U\times\mathbb{C}^r$ with $\Phi(s(x)) = (x, s_U(x))$ for some vector $s_U(x)$. Then, within U, we have

$$\Phi(\nabla s)(x) = (x, \nabla s_U) \qquad \text{where} \qquad (\nabla s)_U = ds_U + A_U s_U \tag{7.14}$$

where A_U is a section of $U \otimes \mathrm{End}(\mathbb{C}^r) \otimes T^*M|_U$ that is independent of the particular section s. The object A_U is thus a matrix–valued 1-form (or covector) known to mathematicians as a *connection 1-form*. In physics, A_U is the gauge field.

To make this more familiar, suppose U_α and U_β are each open sets in M with overlap $U_\alpha \cap U_\beta \neq$. Suppose we're given trivializations $\Phi_\alpha : E|_{U_\alpha} \to U_\alpha \times \mathbb{C}^r$ and $\Phi_\beta : E|_{U_\beta} \to U_\beta \times \mathbb{C}^r$, and let $g_{\beta\alpha} : U_\beta \cap U_\alpha \to \mathrm{End}(\mathbb{C}^r)$ denote the transition function. In physics, we'd say we're working in one gauge (given by Φ_α) in U_α, and in another gauge in U_β, and that $g_{\beta\alpha}$ is the gauge transformation that takes us between the two gauges. We have $s_\beta = g_{\beta\alpha} s_\alpha$ and similarly $(\nabla s)_\beta = g_{\beta\alpha}(\nabla s_\alpha)$, since these are sections of E and $E \otimes T^*M$, respectively. It now follows that the gauge fields A_α and A_β on the two open patches must be related by

$$A_\beta = -g_{\beta\alpha} dg^{-1}_{\beta\alpha} + g_{\beta\alpha} A_\alpha g^{-1}_{\beta\alpha} \,. \tag{7.15}$$

As a special case, U_β and U_α might both be the *same* region, in which case (7.15) tells us how A changes under a change of trivialization on a given region. Notice also that if $\mathrm{rk}(E) = r = 1$, then $g(x) \in \mathrm{End}(\mathbb{C}^r)$ is just a single function at each point of $x \in U_\beta \cap U_\alpha$ which we can write as $\mathrm{e}^{\mathrm{i}\lambda}$ whereupon (7.15) reduces to $A_\beta = A_\alpha - \mathrm{i} d\lambda$, which is the familiar behaviour under a gauge transformation of the vector potential in electromagnetism.

Given a connection $\nabla : \Omega^0_M(E) \to \Omega^1_M(E)$, we can extend its definition to sections of $E \otimes \wedge^p T^*M$. Such sections are to be thought of as p-forms[38] with values in sections of E. The space of such is written $\Omega^p_M(E)$. This extension is also called ∇, and is again defined by

$$\begin{array}{ll} \text{linearity} & \nabla(\mathfrak{s}_1 + \mathfrak{s}_2) = \nabla\mathfrak{s}_1 + \nabla\mathfrak{s}_2 \\ \text{Leibniz} & \nabla(s\omega) = \nabla(s) \wedge \omega + s\, d\omega \end{array} \tag{7.16}$$

where $\mathfrak{s}_{1,2} \in \Omega^p_M(E)$ are p-form sections, while $s \in \Omega^0_M(E)$ is a regular section as above and $\omega \in \Omega^p_M$ is a p-form on M.

The construction above shows that the connection ∇ behaves as a sort of exterior derivative, generalized the usual de Rham exterior derivative d to the case of sections of vector bundles. However, there is a crucial difference. While $d^2 = 0$ automatically, it is not in general true that $\nabla^2 = 0$. To see this, note that

$$\begin{aligned} \nabla(\nabla(s\omega)) &= \nabla((\nabla s) \wedge \omega + s\, d\omega) \\ &= (\nabla^2 s) \wedge \omega - (\nabla s) \wedge d\omega + (\nabla s) \wedge d\omega + s d^2\omega \\ &= (\nabla^2 s) \wedge \omega \end{aligned} \tag{7.17}$$

where the second and third terms have cancelled and the last term is identically zero by the nilpotency of the de Rham exterior derivative. This calculation shows that $\nabla^2 : \Omega^p_M(E) \to \Omega^{p+2}_M(E)$ is linear over multiplication of the section by an arbitrary form,

$$\nabla^2(\mathfrak{s} \wedge \omega) = (\nabla^2 \mathfrak{s}) \wedge \omega \,.$$

[38] A p-form on M can be thought of as a tensor with p contravariant indices, antisymmetrized on these indices. We write Ω^p_M for the space of such p-forms. Thus, if $\omega \in \Omega^p_M$ then in local coordinates on M we have $\omega = \omega_{\mu_1\mu_2\cdots\mu_p}(x)\, dx^{\mu_1} \wedge dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_p}$.

It therefore must correspond to multiplication by some section $F_\nabla \in \Omega^2_M(\mathrm{End}(E))$. In particular, if $s \in \Omega^0_M(E)$, we have

$$\nabla^2(s) = F_\nabla s\,. \tag{7.18}$$

The $\mathrm{End}(E)$–valued 2-form F_∇ is called the *curvature* of the connection.

To understand this more explicitly, let's again choose a local trivialization $\Phi : E|_U \to U \times \mathbb{C}^r$, with $(\nabla s)_U = ds_U + A_U s_U$. Then we have

$$\begin{aligned}(\nabla^2 s)_U &= \nabla(ds_U + A_U s_U)\\ &= d^2 s_U + d(A_U s_U) + A_U \wedge (ds_U + A_U) s_U\\ &= (dA_U + A_U \wedge A_U) s_U\end{aligned} \tag{7.19}$$

and indeed all the derivatives of s_U itself cancel out. Thus on $E|_U$ with the trivialization given by Φ, we can identify the curvature as

$$\begin{aligned}(F_\nabla)_U &= dA_U + A_U \wedge A_U = (\partial_\mu A_\nu + A_\mu A_\nu)\, dx^\mu \wedge dx^\nu\\ &= \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])\, dx^\mu \wedge dx^\nu =: \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu\end{aligned} \tag{7.20}$$

where $F_{\mu\nu}$ are the curvature components. Recall that A_U is a matrix–valued 1-form, so that the commutator term $[A_\mu, A_\nu]$ does not in general vanish.

Given that $\nabla^2(\mathfrak{s}) = F_\nabla \mathfrak{s}$ for any section $\mathfrak{s} \in \Omega^p_M(E)$, so that ∇^2 acts as a purely multiplicative operator, it's interesting to compute what happens when we act with ∇ for a third time. On the one hand, we have

$$\nabla^3(\mathfrak{s}) = \nabla(F_\nabla \mathfrak{s}) = \nabla(F_\nabla)\mathfrak{s} + F_\nabla \nabla(s) \tag{7.21}$$

while on the other hand,

$$\nabla^3(\mathfrak{s}) = \nabla^2(\nabla\mathfrak{s}) = F_\nabla \nabla\mathfrak{s}\,. \tag{7.22}$$

The two equations (7.21) & (7.33) are compatible iff

$$\nabla(F_\nabla) = 0\,, \tag{7.23}$$

which is known as the *Bianchi identity* for the curvature F_∇. In a local trivialization where $\nabla = d + A_U$ and $F_\nabla = dA_U + A_U \wedge A_U$ the Bianchi identity can be seen explicitly by calculating

$$\begin{aligned}\nabla(F_\nabla)|_U &= dF_\nabla + A_U \wedge F_\nabla - F_\nabla \wedge A_U\\ &= d(dA_U + A_U \wedge A_U) + A_U \wedge (dA_U + A_U \wedge A_U) - (dA_U + A_U \wedge A_U) \wedge A_U\\ &= dA_U \wedge A_U - A_U \wedge dA_U + A_U \wedge dA_U + A_U^3 - dA_U \wedge A_U - A_U^3\\ &= 0\,,\end{aligned} \tag{7.24}$$

where we've used the fact that, in a local trivialization, the covariant derivative acts as $\nabla\phi = d\phi + A \wedge \phi - (-)^p \phi \wedge A$ on any section $\phi \in \Omega^p_M(\mathrm{End}(E))$. This agrees with our global argument above.

7.1.3 Holonomy

Some other time.

7.2 Classical Yang–Mills theory

The first and most important example of a non–Abelian gauge theory was introduced to physics 1954 by Chen Ning Yang and Robert Mills, and then almost completely ignored for nearly a decade. From a phenomenological point of view, the importance of Yang–Mills theory arises because (as you're surely aware) the Standard Model — the most fundamental description of Nature we currently possess — is at it's heart a non–Abelian gauge theory based on $SU(3) \times SU(2) \times U(1)$, though it took much hard work and many further new ideas before this became apparent. From the perspective of a theoretical physicst, Yang–Mills is theory important also because it's the *only* QFT in $d = 4$ that might a continuum limit, as realized by Coleman & Gross in 1973. In mathematics, Yang–Mills theory is at the heart of Simon Donaldson's exploration of the wild world of four–manifolds. More recently, it's even been related to the (geometric) Langlands Program. In a deep sense, Yang–Mills theory is the right four–dimensional analogue of geodesics in $d = 1$ and harmonic maps in $d = 2$.

7.2.1 The Yang–Mills action

To describe Yang–Mills theory, we pick a d–dimensional (pseudo-)Riemannian manifold (M, g) complete with a choice of metric g. The Yang–Mills action is then defined to be

$$S_{\mathrm{YM}}[\nabla] = \frac{1}{2\mathrm{g}_{\mathrm{YM}}^2} \int_M \mathrm{tr}(F_\nabla \wedge *F_\nabla) = \frac{1}{4\mathrm{g}_{\mathrm{YM}}^2} \int_M g^{\mu\nu}\, g^{\rho\sigma}\, F^{\mathrm{a}}_{\mu\rho}\, F^{\mathrm{a}}_{\mu\sigma}\, \sqrt{g}\, \mathrm{d}^d x \tag{7.25}$$

where g_{YM} is a coupling constant. The Yang–Mills action is thus just the (square of) the L^2-norm of F_∇ with respect to the standard volume element on M supplied by the metric g. It's the natural generalization of the Maxwell action

$$S_{\mathrm{Max}}[\nabla] = \frac{1}{4e^2} \int F^{\mu\nu}\, F_{\mu\nu}\, \mathrm{d}^4 x = \frac{1}{4e^2} \int (\mathbf{E} \cdot \mathbf{E} - \mathbf{B} \cdot \mathbf{B})\, \mathrm{d}t\, \mathrm{d}^3 x \tag{7.26}$$

of electromagnetism, to which it reduces when $G = U(1)$ and $(M, g) = (\mathbb{R}^{3,1}, \delta)$. I've written the action as $S_{\mathrm{YM}}[\nabla]$ to emphasize that we should treat the action as a function of the connection, not of the curvature. Again, this is familiar to you from deriving Maxwell's (vacuum) equations as the Euler–Lagrange equations of (7.26).

From the point of view of physics, the most important difference between the Yang–Mills action for a non–Abelian group G and the Abelian (Maxwell) case is that, for non–Abelian G, *the Yang–Mills field interacts with itself.* We can see this by restricting to a local trivialization on $U \subset M$ where[39] $\nabla = d + A$ and $F = dA + A^2$. On this patch the

[39] Henceforth, I'll drop the subscript on A_U, as is common.

Yang–Mills action becomes

$$\begin{aligned} S_{\rm YM}[A] &= \frac{1}{2{\rm g}_{\rm YM}^2}\int_U {\rm tr}((dA+A^2)\wedge *(dA+A^2)) \\ &= \frac{1}{4{\rm g}_{\rm YM}^2}\int_U g^{\mu\nu}\,g^{\rho\sigma}\,\left(\partial_{[\mu}A^{\rm a}_{\rho]}+\frac{1}{2}f^{\rm a}_{\rm bc}A^{\rm b}_{[\mu}A^{\rm c}_{\rho]}\right)\left(\partial_{[\nu}A^{\rm a}_{\sigma]}+\frac{1}{2}f^{\rm a}_{\rm de}A^{\rm d}_{[\nu}A^{\rm e}_{\sigma]}\right)\sqrt{g}\,{\rm d}^dx \end{aligned} \tag{7.27}$$

Thus we see that the action contains a cubic vertex of the schematic form $\sim AAdA$, and a quartic vertex $\sim A^4$. Thus, even in the absence of any charged matter, Yang–Mills theory is a non–trivial interacting theory.

At the classical level, these self–interactions make themselves felt via the Euler–Lagrange equations that follow from (7.25). Varying the connection $\nabla \to \nabla + \delta a$ where $\delta a \in \Omega^1_M({\rm End}(E))$, to first order in δa we have that

$$F_{\nabla+\delta a} - F_\nabla = \nabla(\delta a), \tag{7.28}$$

or in other words, $\delta F_{\mu\nu} = \nabla_{[\mu}\delta a_{\nu]}$. Therefore, varying the action gives[40]

$$\delta S_{\rm YM} = \frac{1}{{\rm g}_{\rm YM}^2}\int_M {\rm tr}(\nabla\delta a\wedge *F_\nabla) = -\frac{1}{{\rm g}_{\rm YM}^2}\int_M {\rm tr}(\delta a\wedge\nabla * F_\nabla) + \text{boundary terms} \tag{7.29}$$

so that the Euler–Lagrange equations are

$$\nabla * F_\nabla = 0\,, \tag{7.30}$$

or equivalently

$$0 = \nabla^\mu F^{\rm a}_{\mu\nu} = \partial^\mu F^{\rm a}_{\mu\nu} + \frac{1}{2}f^{\rm a}_{\rm bc}A^{{\rm b}\mu}F^{\rm c}_{\mu\nu} \tag{7.31}$$

in terms of components in a local trivialization. However you wish to write them, these are known as the *Yang–Mills equations*. We recall from (7.23) that the curvature of any connection automatically obeys the Bianchi identity $\nabla F_\nabla = 0$, or

$$\nabla_\mu F_{\nu\lambda} + \nabla_\nu F_{\lambda\mu} + \nabla_\lambda F_{\mu\nu} = 0 \tag{7.32}$$

or

$$\partial_\mu F^{\rm a}_{\nu\lambda} + \partial_\nu F^{\rm a}_{\lambda\mu} + \partial_\lambda F^{\rm a}_{\mu\nu} + \frac{1}{2}f^{\rm a}_{\rm bc}\left(A^{\rm b}_\mu F^{\rm c}_{\nu\lambda} + A^{\rm b}_\nu F^{\rm c}_{\mu\lambda} + A^{\rm b}_\lambda F^{\rm c}_{\mu\nu}\right) = 0 \tag{7.33}$$

to be completely explicit. Mathematically, connections whose curvature obeys the Yang–Mills equation (7.30) are *critical points* of the function $S_{\rm YM}[\nabla]$ defined on the space of *all* connections on $P \to M$.

Unlike the vacuum Maxwell equations, the Yang–Mills equations and the Bianchi identity (7.33) are *non–linear* p.d.e.s. The non–linearities arise because of the presence of A in the covariant derivative ∇ and the non–linear terms $\sim A^2$ in the curvature itself. The reason this happened is that the curvature F transforms in the adjoint representation,

[40] As an exercise, you should go through this derivation for yourself, using the second line of (7.27) and checking you're happy with where all the indices go. Once you've done this and checked you are happy, I hope you'll begin to appreciate the usefulness of the form notation.

with $F \to gFg^{-1}$ under a gauge transformation $g(x)$. So any sort of differential equation obeyed by F will have to involve covariant derivatives, as these are the only derivatives that make geometric sense. Electrodynamics avoids this complication because the adjoint representation of an Abelian group is actually trivial, as we can see by the manipulation $gFg^{-1} = Fgg^{-1} = F$, which is allowed since all group elements commute. As a consequence of the non–linearity, unlike beams of light or radio waves in the Abelian case, we do not expect superposition of solutions, and propagating waves of Yang–Mills fields do not pass through one another freely.

Famously, General Relativity is also a geometric theory in which the field (Einstein) equations are a system of non–linear pdes. Indeed, there are many parallels between the two theories, most of which can be made apparent by treating General Relativity as a theory of connections on the tangent bundle $TM \to M$. Now I'm sure you can all write down several non–trivial solutions of the Einstein equations, probably including the Schwarzschild metric, various homogeneous cosmological models, the Kerr metric and perhaps a few others. Likewise, I certainly expect you've solved Maxwell's equations in the presence of any number of weird charge configurations, including electrical circuits, solenoids and all manner of other things. However, I doubt that many of you know any non–trivial solutions of the Yang–Mills equations at all.

Why not? The answer turns out to be revealing[41]. Turning the question around, the reason you *do* know lots of solutions of the Maxwell or Einstein equations is simply that the role these equations play in Nature has been understood for over a century. Their weak field, Newtonian approximations have been known considerably longer, and the phenomena they describe are apparent in everyday life. By contrast, our technology has only just reached the point where we can perform any experiment in which the classical Yang–Mills equations are relevant.

The reason this is so is an effect known as the *mass gap*. Skipping ahead of our story, the path integral for Yang–Mills will roughly take the form $\int DA\, \mathrm{e}^{-S_{\mathrm{YM}}[\nabla]/\hbar}$. Because the coupling constant appears only as an overall factor in the Yang–Mills action (7.25), it plays the same role as $\hbar$; the path integral depends on g_{YM} and $\hbar$ only through the combination $\mathrm{g}^2_{\mathrm{YM}}\hbar$ (at least in pure Yang–Mills). Thus we should expect that the quantum theory is well–approximated by the classical limit if $\mathrm{g}_{\mathrm{YM}} \to 0$ so that the theory is 'weakly coupled'. However, we've seen already that coupling constants can run depending on the scale at which we examine the physics. Below, we'll find that in a non-Abelian theory (with not too much matter), the gauge coupling actually *increases* as we view physics at lower and lower energy scales. In this region, Yang–Mills theory is an inherently *quantum* theory and can yield results that are very far from the classical story. For example, the energy scale at which QCD becomes strongly coupled is ~ 200 MeV. This is accessible to particle accelerators, but at everyday energy scales the role of Yang–Mills theory in Nature cannot be described without talking about the quantum theory.

In fact, when C.N. Yang first presented his work with Mills, he was strongly criticized

[41] No, it's not that the Yang–Mills equations are 'harder' to solve than the Einstein equations — quite the contrary.

by Pauli. Pauli had spotted that there is no term $\sim A^2$ in the Yang–Mills action (it would not be compatible with gauge invariance), so quanta of the Yang–Mills field will be massless. He concluded that, as in both electromagnetism and gravity, these particles will be responsible for some long range force[42]. Needless to say, in Nature no other such forces exist: this is why the Yang–Mills equations were only contemplated in the middle of the 20^{th} century. (Yes, strong and weak nuclear forces were known about earlier, but they're certainly not long range.) Pauli's conclusion was correct in the weak coupling approximation, but the classical Yang–Mills equations are a very poor guide to the low–energy physics.

7.2.2 Minimal coupling

We can also write down actions describing the coupling of Yang–Mills theory to charged matter. The simplest of these are the *minimally coupled* actions, which essentially says that you just take your favourite action for uncharged matter, and then replace all derivatives by gauge covariant derivatives.

For example, suppose we have a scalar field ϕ that lives in the adjoint representation of the gauge group so that $\phi \to h\phi h^{-1}$ under a gauge transformation $h(x)$. With minimal coupling, the kinetic terms for this scalar are

$$S_{\text{kin}}[\phi, \nabla] = \int_M \text{tr}(\nabla\phi \wedge *\nabla\phi) \tag{7.34}$$

and locally on $U \subset M$ this is

$$S_{\text{kin}}[\phi, A] = \frac{1}{2}\int_U g^{\mu\nu}\left(\partial_\mu\phi^{\text{a}} + \frac{1}{2}f^{\text{a}}_{\text{bc}}\,A^{\text{b}}_\mu\,\phi^{\text{c}}\right)\left(\partial_\nu\phi^{\text{a}} + \frac{1}{2}f^{\text{a}}_{\text{de}}\,A^{\text{d}}_\nu\,\phi^{\text{e}}\right)\sqrt{g}\,\text{d}^d x\,. \tag{7.35}$$

We can also construct potential terms for ϕ of the form

$$S_{\text{pot}}[\phi] = \int_U V(\phi)\,\sqrt{g}\,\text{d}^d x \tag{7.36}$$

where $V(\phi)$ is any gauge invariant polynomial in ϕ. For example, since ϕ is in the adjoint representation we can think of it as a matrix and then a simple choice would be

$$V(\phi) = \frac{m^2}{2}\text{tr}(\phi\phi) + \frac{\lambda}{4!}\text{tr}(\phi\phi\phi\phi) \tag{7.37}$$

where the traces ensure that $V(h\phi h^{-1}) = V(\phi)$.

Similarly, the minimally coupled action for a massive Dirac spinor ψ transforming the in fundamental representation of G is[43]

$$S_{\text{Dir}}[\psi, \nabla] = \int_{\mathbb{R}^d} \bar\psi(\text{i}\not\nabla + m)\psi\ \text{d}^d x \tag{7.38}$$

[42]Recall that in four dimensions, a particle of mass m gives rise to a potential $V(r) \sim \text{e}^{-mr}/r$. For any finite m the resulting force is negligible at distances $\gg 1/m$ from the centre of the potential, but when $m = 0$ the force can be felt right across the Universe.

[43]I'll write this just for $(M, g) = (\mathbb{R}^d, \delta)$, which will be sufficient for our purposes in this course. To do more we'd first need to discuss how to define spinors on a curved manifold.

where ∇ acts in the fundamental representation, while the conjugate spinor $\bar\psi$ transforms in the antifundamental as $\bar\psi \to \bar\psi h$. Explicitly, if $G = SU(N)$ and $i, j = 1, \dots, N$ label a basis of $\mathbb{C}^N$, then we think of the gauge field A_μ as an $N \times N$ matrix $(A_\mu)^j_i$ and the action is

$$S_{\text{Dir}}[\psi, \nabla] = \int_{\mathbb{R}^d} \bar\psi_i \, \delta^i{}_j (\mathrm{i}\gamma^\mu \partial_\mu + m)\psi^j + \mathrm{i}\bar\psi_i \gamma^\mu (A_\mu)^i{}_j \psi^i \; \mathrm{d}^d x \tag{7.39}$$

where γ^μ are the Dirac γ-matrices. Note that S_{Dir} reduces to the usual electron action in the Abelian case $G = U(1)$ for QED.

For simple operators such as $(\partial\phi)^2$ the replacement $(\partial\phi)^2 \to (\nabla\phi)^2$ is unambiguous; once we've declared in which representation ϕ transforms, there is a unique notion of the covariant derivative acting on that representation. However, you may feel uneasy about treating a more complicated operator such as $\partial^\mu\partial^\nu\phi\,\partial_\mu\partial_\nu\phi$. Clearly our replacement prescription should involve $\mathrm{tr}(\nabla^\mu\nabla^\nu\phi\,\nabla_\mu\nabla_\nu\phi)$, but since $[\partial_\mu, \partial_\nu] = 0$ whereas $[\nabla_\mu, \nabla_\nu] \neq 0$, how can we tell whether or not the operator we end up with should include the antisymmetric part of μ and ν? There is no unambiguous way to decide, but fortunately the issue is not very important: whether or not we include such terms in our initial action, if they are allowed by global symmetries then they will in any case be generated by quantum corrections to the effective action.

In the presence of charged matter, the Yang–Mills equations (7.30) are replaced by

$$\nabla^\mu F_{\mu\nu} = -\mathrm{g}^2_{\text{YM}} \, J_\nu \tag{7.40}$$

where $J^\nu(x) = \delta S_{\text{matter}}/\delta a_\nu(x)$ is the matter current. (The Bianchi identity still holds.) Notice that, since the matter action as a whole was invariant and a lives in the adjoint representation, the current J_μ also transforms in the adjoint. For example, in the case of our scalar above we have

$$J^{\text{a}}_\nu = \frac{1}{2} f^{\text{a}}_{\text{bc}} \phi^{\text{b}} (\nabla_\nu \phi)^{\text{c}} \tag{7.41}$$

whereas for the Dirac spinor

$$(J_\nu)^j_i = \mathrm{i}\bar\psi_i \gamma_\nu \psi^j \, . \tag{7.42}$$

In electromagnetism, the adjoint representation was trivial, so the electromagnetic current J^{em} satsified a standard conservation law $\partial^\mu J^{\text{em}}_\mu = 0$. In the non–Abelian case, it only makes sense to differentiate J covariantly using ∇ acting in the adjoint representation. Using the equation of motion (7.40) we have

$$-\mathrm{g}^2_{\text{YM}} \, (\nabla^\nu J_\nu)^{\text{a}} = (\nabla^\nu (\nabla^\mu F_{\mu\nu}))^{\text{a}} = \frac{1}{2}([\nabla^\nu, \nabla^\mu] F_{\mu\nu})^{\text{a}} = \frac{1}{4} f^{\text{a}}_{\text{bc}} F^{\text{b}\,\mu\nu} F^{\text{c}}_{\mu\nu} \tag{7.43}$$

where the second equality follows since $F_{\mu\nu}$ is antisymmetric in (μ, ν), the third equality uses the facts that the commutator of covariant derivatives is F and that these derivatives act on the original F in the adjoint representation. Finally, this expression vanishes by antisymmetry of the structure constants $f^{\text{a}}_{\text{bc}} = -f^{\text{a}}_{\text{cb}}$. Thus we have a *covariant conservation law*

$$\nabla^\mu J_\mu = 0 \tag{7.44}$$

for our current in a non–Abelian theory. The differences between this conservation law and the naive conservation law $\partial^\mu J^{\text{Noether}}_\mu$ we found for Noether currents in section **??** have some profound consequences: you'll explore these in Problem Set 3.

7.3 Quantum Yang–Mills theory

We're now ready to consider the quantum theory of Yang–Mills. In the first few sections, we'll treat the path integral formally as an integral over infinite dimensional spaces, without worrying about imposing cutoffs. We'll turn to questions about using renormalization to make sense of these formal integrals in section **??**.

To specify Yang–Mills theory, we had to pick a principal G bundle $P \to M$ together with a connection ∇ on P. So our first thought might be to try to define the Yang–Mills partition function as

$$Z_{\text{YM}}[(M,g),\mathrm{g}_{\text{YM}}] \stackrel{?}{=} \int_{\mathcal{A}} DA\ \mathrm{e}^{-S_{\text{YM}}[\nabla]} \tag{7.45}$$

where $\mathcal{A}$ is the space of all connections on P. To understand what this integral might mean, first note that, given any two connections ∇ and ∇', the 1-parameter family

$$\nabla^\tau := \tau\nabla + (1-\tau)\nabla' \tag{7.46}$$

is also a connection for all $\tau \in [0,1]$. For example, you can check that the *rhs* has the behaviour expected of a connection under any gauge transformation. Thus we can find a path in $\mathcal{A}$ between any two connections. Since $\nabla' - \nabla \in \Omega^1_M(\mathfrak{g})$, we conclude that $\mathcal{A}$ is an infinite dimensional affine space whose tangent space at any point is $\Omega^1_M(\mathfrak{g})$, the infinite dimensional space of all $\mathfrak{g}$–valued covectors on M. In fact, it's easy to write down a flat (L^2-)metric on $\mathcal{A}$ using the metric on M:

$$\mathrm{d}s^2_{\mathcal{A}} = \int_M \mathrm{tr}(\delta A \wedge *\,\delta A) = \frac{1}{2}\int_M g^{\mu\nu}\,\delta A^{\mathrm{a}}_\mu\,\delta A^{\mathrm{a}}_\nu\ \sqrt{g}\,\mathrm{d}^d x\ . \tag{7.47}$$

In other words, given any two tangent vectors $(a_1,a_2) \in \Omega^1_M(\mathfrak{g})$ at the point $\nabla \in \mathcal{A}$,

$$\mathrm{d}s^2_{\mathcal{A}}(a_1,a_2) = \int_M \mathrm{tr}(a_1 \wedge *a_2)\,, \tag{7.48}$$

independent of where in $\mathcal{A}$ we are. This is encouraging: $\mathcal{A}$ just looks like an infinite dimensional version of $\mathbb{R}^n$, with no preferred origin since there is no preferred connection on P.

We might now hope that the path integral (7.45) means formally that we should pick an arbitrary base–point $\nabla_0 \in \mathcal{A}$, then write any other connection $\nabla = \nabla_0 + A$, with the measure DA indicating that we integrate over all $A \in \Omega^1_M(\mathfrak{g})$ using the translationally invariant measure on $\mathcal{A}$ associated to the flat metric (7.47). (Such an infinite dimensional flat measure does not exist — we're delaying this worry for now.) For a connection $\nabla = \nabla_0 + A$, the action becomes

$$\begin{aligned} S_{\text{YM}}[\nabla] &= \frac{1}{2\mathrm{g}^2_{\text{YM}}}\int_M \mathrm{tr}(F_\nabla \wedge *F_\nabla) \\ &= \frac{1}{2\mathrm{g}^2_{\text{YM}}}\int_M \mathrm{tr}(F_{\nabla_0} \wedge *F_{\nabla_0}) \ +\ \frac{1}{2\mathrm{g}^2_{\text{YM}}}\int_M \mathrm{tr}(\nabla_0 A + A^2)\wedge *(\nabla_0 A + A^2)\,. \end{aligned} \tag{7.49}$$

For example, on a topologically trivial bundle a standard choice would be to pick the trivial connection $\nabla_0 = \partial$ as base–point. Then $F_{\nabla_0} = 0$ and the action takes the familiar form

$$S_{\rm YM}[\partial + A] = \frac{1}{2{\rm g}_{\rm YM}^2} \int_M {\rm tr}(dA + A^2) \wedge *(dA + A^2)\,. \tag{7.50}$$

The path integral (7.45) would be interpreted as an integral over all gauge fields A. However, in some circumstances we'll meet later (even when P is topologically trivial), it will be useful to choose a different base–point ∇_0 for which $F_{\nabla_0} \neq 0$, known as a *background field*. In this case, the first term on the *rhs* of (7.49) is the action for the background field and comes out of the path integral as an overall factor, while the remaining action for A involves the covariant derivative with respect to the background field.

Of course, there's a problem. By construction, the Yang–Mills action was invariant under gauge transformations, so the integrand in (7.45) is degenerate along gauge orbits. Consequently, the integral will inevitably diverge because we're vastly overcounting. You met this problem already in the case of QED during the Michaelmas QFT course. There, as here, the right thing to do is to integrate just over physically inequivalent connections — those that are *not* related by a gauge transform. In other words, the correct path integral for Yang–Mills should be of the form

$$Z_{\rm YM}[(M,g), {\rm g}_{\rm YM}] = \int_{\mathcal{A}/\mathcal{G}} D\mu\ {\rm e}^{-S_{\rm YM}[\nabla]} \tag{7.51}$$

where $\mathcal{G}$ is the space of all gauge transformations, so that $\mathcal{A}/\mathcal{G}$ denotes the space of all gauge equivalence classes of connections: we do not count as different two connections that are related by a gauge transformation. Note that this definition means *gauge 'symmetry' does not exist* in Nature! We've *quotiented* by gauge transformations in constructing the path integral, so the resulting object has no knowledge of any sort of gauge transformations. They were simply a *redundancy* in our construction. The same conclusion holds if we compute correlation functions of any gauge invariant quantities, whether they be local operators built from gauge invariant combinations of matter fields, or Wilson loops running around some curves in space.

However, we're not out of the woods. Whilst $\mathcal{A}$ itself was just an affine space, the space $\mathcal{A}/\mathcal{G}$ is much more complicated. For example, it has highly non–trivial topology investigated by Atiyah & Jones, and by Singer. Certainly $\mathcal{A}/\mathcal{G}$ is *not* affine, so we don't yet have any understanding of what the right measure $D\mu$ to use on this space is, even formally. In the case of electrodynamics, you were able to avoid this problem (at least in perturbation theory on $\mathbb{R}^4$) by picking a gauge, defining the photon propagator and just getting on with it. The non–linear structure of the non–Abelian theory means we'll have to consider this step in more detail.

7.3.1 A ghost story

The way to proceed was found by Feynman, de Witt, and by Faddeev & Popov. To understand what they did, let's warm up with a finite–dimensional example.

Figure 11: *The gauge slice C should be chosen to be transverse to the gauge orbits.*

Suppose we have a function $S : \mathbb{R}^2 \to \mathbb{R}$ defined at any point on the (x, y)-plane, and suppose further that this function is invariant under rotations of the plane around the origin. We think of $S(x, y)$ as playing the role of our 'action' for 'fields' (x, y), while rotations represent 'gauge transformations' leaving this action invariant. Of course, rotational invariance implies that $S(x, y) = h(r)$ in this example, where $h(r)$ is some function of the radius. We easily compute

$$\int_{\mathbb{R}^2} \mathrm{d}x\, \mathrm{d}y\, \mathrm{e}^{-S(x,y)} = 2\pi \int_0^\infty \mathrm{d}r\, r\, \mathrm{e}^{-h(r)} \tag{7.52}$$

which will make sense for sufficiently well–behaved $f(r)$. The factor of 2π appears here because the original integral was rotationally symmetric: it represents the redundancy in the expression on the left of (7.52).

In the case of Yang–Mills, if we integrated over the space $\mathcal{A}$ of all connections rather than over $\mathcal{A}/\mathcal{G}$, the redundancy would be infinite: while the volume $\mathrm{vol}(G)$ is finite for compact structure groups G, the volume $\mathrm{vol}(\mathcal{G})$ of the space of all gauge transformations is infinite — heuristically, you can think of this as a copy of $\mathrm{vol}(G)$ at each point of M. What we'd like to do is understand how to keep the analogue of $\int_0^\infty \mathrm{d}r\, r\, \mathrm{e}^{-h(r)}$ in the gauge theory case, without the redundancy factor. However, neither the right set of gauge invariant variables (analogous to r) nor the right measure on $\mathcal{A}/\mathcal{G}$ (generalizing $r\, \mathrm{d}r$) are obvious in the infinite dimensional case.

Returning to (7.52), suppose C is any curve traveling out from the origin that intersects every circle of constant radius exactly once. More specifically, let $f(\mathbf{x})$ be some function with the properties

- For any point $\mathbf{x} \in \mathbb{R}^2$ there exist a rotation $R : \mathbb{R}^2 \to \mathbb{R}^2$ such that $f(R\mathbf{x}) = 0$

- f is non–degenerate on the orbits; *i.e.*, $f(R\mathbf{x}) = f(\mathbf{x})$ iff R is the identity[44] in SO(2).

The curve $C = \{\mathbf{x} \in \mathbb{R}^2 : f(\mathbf{x}) = 0\}$ then intersects every orbit of the rotation group exactly once and so is isomorphic to $\mathbb{R}^2/\mathrm{SO}(2)$: we can think of the curve $C \subset \mathbb{R}^2$ as a way to embed the orbit space in the plane. Notice that the non–degeneracy property means that $f(\mathbf{x})$ itself is certainly not rotationally invariant. In anticipation of the application to Yang–Mills theory, we call $f(\mathbf{x})$ the *gauge fixing function* and the curve C it defines the *gauge slice*. (See figure 11.)

Now consider the integral

$$\int_{\mathbb{R}^2} \mathrm{d}x\, \mathrm{d}y\, \delta(f(\mathbf{x}))\, \mathrm{e}^{-S(x,y)} \tag{7.53}$$

over all of $\mathbb{R}^2$. Clearly, the δ-function restricts this integral to the gauge slice. However, the actual value we get depends on our choice of specific function $f(\mathbf{x})$; for example, even

[44] Technically, we should restrict to $\mathbb{R}^2 - \{0\}$ to ensure this condition holds. For smooth functions $S(x, y)$ this subtlety won't affect our results and I'll ignore it henceforth.

replacing $f \to cf$ for some constant c (an operation which preserves the curve C) reduces the integral by a factor of $1/|c|$. Thus we cannot regard (7.53) as an integral over the moduli space $\mathbb{R}^2/\mathrm{SO}(2)$ — it also depends on exactly how we embedded this moduli space inside $\mathbb{R}^2$. The problem arose because the δ-function changes as we change $f(\mathbf{x})$. To account for this, define

$$\Delta_f(\mathbf{x}) := \frac{\partial}{\partial\theta} f(R_\theta \mathbf{x})\bigg|_0 \tag{7.54}$$

where the right hand side means we compute the rate of change of f with respect to a rotation R_θ through angle θ, evaluated at the identity $\theta = 0$. Notice that we only need to know how an *infinitesimal* rotation acts in order to compute this.

It's clear that the new integral

$$\int_{\mathbb{R}^2} \mathrm{d}x\,\mathrm{d}y\,|\Delta_f(\mathbf{x})|\,\delta(f(\mathbf{x}))\,\mathrm{e}^{-S(x,y)} \tag{7.55}$$

involving the modulus of Δ_f doesn't change if we rescale f by a constant factor as above. Nor does it change if we rescale f by a non–zero, r-dependent factor $c(r)$, which means that (7.55) is completely independent of the choice of function used to define the gauge slice C. In fact, I claim that (7.55) is actually *independent of the particular gauge slice itself*. To see this, let f_1 and f_2 be any two different gauge–fixing functions. Since the curves $C_{1,2}$ they define each intersect every orbit of SO(2) uniquely, we can always rotate C_1 into C_2, provided we allow ourselves to rotate by different amounts at different values of the radius r. Thus we must have

$$f_2(\mathbf{x}) \propto f_1(R_{12}(r)\mathbf{x}) \tag{7.56}$$

for some r-dependent rotation $R_{12}(r)$ and where the proportionality factor depends at most on the radius. By rescaling invariance,

$$|\Delta_{f_2}(\mathbf{x})|\,\delta(f_2(\mathbf{x})) = |\Delta_{f_1}(\mathbf{x}')|\,\delta(f_1(\mathbf{x}')) \tag{7.57}$$

where we've defined $\mathbf{x}' := R_{12}\mathbf{x}$ for any point $\mathbf{x} \in \mathbb{R}^2$, whether it lies on our curves or not. Now, the statement that the action $S(\mathbf{x})$ is rotationally invariant means that it takes the same value all around every circle of constant radius, so $S(\mathbf{x}) = S(\mathbf{x}')$. Similarly,

$$\mathrm{d}x'\,\mathrm{d}y' = \mathrm{d}x\,\mathrm{d}y \tag{7.58}$$

because again this measure is rotationally invariant at every value of r.[45] Putting all this together, the integral in (7.55) is independent of the choice of gauge slice $\mathbb{R}^2/\mathrm{SO}(2) \hookrightarrow \mathbb{R}^2$, as we wished to show.

[45] Again, it's a good idea to check you're comfortable with this assertion by writing

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha(r) & \sin\alpha(r) \\ -\sin\alpha(r) & \cos\alpha(r) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and explicitly working out the transformation of the measure, allowing for the fact that the angle $\alpha(r) = \alpha(\sqrt{x^2+y^2})$ depends on the radius. You'll find the measure is nonetheless invariant.

As a concrete example, suppose we choose C to be the x-axis, defined by $f(\mathbf{x}) = y = 0$. With this choice, $f(R\mathbf{x}) = y\cos\theta - x\sin\theta$ where R represents anti-clockwise rotation through θ. Thus

$$\Delta_f(\mathbf{x}) = \frac{\partial}{\partial\theta}(y\cos\theta - x\sin\theta)\Big|_{\theta=0} = -x \tag{7.59}$$

and therefore our integral (7.55) becomes

$$\int_{\mathbb{R}^2} \mathrm{d}x\,\mathrm{d}y\,|\Delta_c(\mathbf{x})|\,\delta(f(\mathbf{x}))\,\mathrm{e}^{-S(x,y)} = \int_{\mathbb{R}^2} \mathrm{d}x\,\mathrm{d}y\,|x|\,\delta(y)\,\mathrm{e}^{-S(x,y)} = \int_{\mathbb{R}} \mathrm{d}x\,|x|\,\mathrm{e}^{-h(|x|)} \tag{7.60}$$

where in the last step we've used the fact that since $S(x,y)$ was rotationally invariant, along the line $y = 0$ it can only depend on $|x|$. Since $|x|$ is an even function of x, we have

$$\int_{\mathbb{R}} \mathrm{d}x\,|x|\,\mathrm{e}^{-h(|x|)} = 2\int_0^\infty \mathrm{d}r\,r\,\mathrm{e}^{-h(r)}\,, \tag{7.61}$$

which disagrees with the radial part of our original integral by a factor of 2. What's gone wrong is that circles of constant r intersect the x-axis twice — when $x > 0$ and when $x < 0$ — and our gauge fixing condition $y = 0$ failed to account for this; in other words, it slightly failed the non–degeneracy property. We'll see below that this glitch is actually a model of something that also happens in the case of Yang–Mills theory.

To recap, what we've achieved with all this is that, for any non–degenerate gauge–fixing function f, we can write the desired integral over the space of orbits $\mathbb{R}^2/\mathrm{SO}(2) \cong (0,\infty)$ as

$$\int_{\mathbb{R}^2/\mathrm{SO}(2)} \mathrm{d}r\,r\,\mathrm{e}^{-h(r)} = \int_{\mathbb{R}^2} \mathrm{d}x\,\mathrm{d}y\,|\Delta_f(\mathbf{x})|\,\delta(f(\mathbf{x}))\,\mathrm{e}^{-S(x,y)}\,. \tag{7.62}$$

The point is that the expression on the *rhs* refers only to functions and coordinates on the affine space $\mathbb{R}^2$, and uses only the standard measure $\mathrm{d}x\,\mathrm{d}y$ on $\mathbb{R}^2$. When the gauge orbits have dimension > 1 we must impose several gauge fixing conditions f^{a}, one for each transformation parameter θ^{a}. Then we take the integral to include a factor

$$|\Delta_f(\mathbf{x})| \prod_{\mathrm{a}} \delta(f^{\mathrm{a}}(\mathbf{x})) \tag{7.63}$$

where now Δ_f is the *Faddeev–Popov determinant*

$$\Delta_f(\mathbf{x}) := \det\left(\frac{\partial f^{\mathrm{a}}(R\mathbf{x})}{\partial\theta^{\mathrm{b}}}\right) \tag{7.64}$$

for a generic set of variables $\mathbf{x} \in \mathbb{R}^n$ where the action is invariant under some transformation $\mathbf{x} \to R\mathbf{x}$ (not necessarily a rotation). Again, this will allow us to write an integral over the space of orbits of these transformations in terms of an integral over the affine space $\mathbb{R}^n$. These are things we have access to in the gauge theory case[46] where the affine space in question is the space $\mathcal{A}$ of all gauge fields, and the transformation group is the space $\mathcal{G}$

[46] Modulo, as always, the problem that there is no Lebesgue measure on $\mathcal{A}$: this is what we'll treat with renormalization.

of all gauge transformations. Armed with these ideas, we now turn to the case of gauge theory.

In Yang–Mills theory, we can fix the gauge redundancy by picking a particular connection in each gauge equivalence class — in other words, by picking an embedding of $\mathcal{A}/\mathcal{G} \hookrightarrow \mathcal{A}$ specified by some gauge–fixing functional $f[A]$. The most common choices of gauge fixing functional are *local*, in the sense that $f[A]$ depends on the value of the gauge field just at a single point $x \in M$. Heuristically, we then restrict to $f[A(x)] = 0$ at every point $x \in M$ by inserting "$\delta[f] = \prod_{x\in M} \delta(f[A(x)])$" in the path integral. We'll consider how to interpret this infinite–dimensional δ-function below. The remaining ingredient we need is the Faddeev–Popov determinant

$$\Delta_f = \det \frac{\delta f^{\mathrm{a}}[A^\lambda(x)]}{\delta \lambda^{\mathrm{b}}(y)} \tag{7.65}$$

where $A^\lambda = A + \nabla\lambda$ denotes an infinitesimal gauge transformation of A with parameters $\lambda(x)$ valued in the adjoint. Like the δ-functional $\delta[f]$, this determinant is now of an infinite dimensional matrix; we'll consider what this determinant means momentarily. With these ingredients, our Yang–Mills path integral can be written as

$$\int_{\mathcal{A}/\mathcal{G}} D\mu \; \mathrm{e}^{-S_{\mathrm{YM}}[\nabla]} = \int_{\mathcal{A}} DA \; |\Delta_f| \, \delta[f] \, \mathrm{e}^{-S_{\mathrm{YM}}[\nabla]} \,, \tag{7.66}$$

where the factor of $|\Delta_f| \, \delta[f]$ restricts us to an arbitrary gauge slice, but leaves no dependence on any particular choice of slice, as above. Again, the advantage of the *rhs* is that it refers only to the 'naive' integral measure over all connections.

Now let's consider how to treat these $\delta[f]$ and Δ_f factors. Taking our lead from Fourier analysis, we introduce a new field h (sometimes called a *Nakanishi–Lautrup field*) and write

$$\delta[f] = \int Dh \, \mathrm{e}^{-S_{\mathrm{gf}}[h,\nabla]} \,, \tag{7.67}$$

where

$$S_{\mathrm{gf}}[h,\nabla] := \int_M \mathrm{tr}(h * f[A]) = \frac{1}{2}\int_M h^{\mathrm{a}} f^{\mathrm{a}}[A] \, \sqrt{g} \, \mathrm{d}^d x \tag{7.68}$$

is the *gauge–fixing action*. The idea is that h is a Lagrange multiplier — performing the path integral over h imposes $f[A(x)] = 0$ throughout M. Notice that since we needed one gauge–fixing condition for every gauge parameter, we take h to lie in the adjoint representation, $h \in \Omega^0_M(\mathfrak{g})$. This does *not* imply that (7.67) is gauge invariant: indeed it *cannot* be if we wish to use it to fix a gauge! For the Faddeev–Popov determinant Δ_f, recall that if M is an $n \times n$ matrix and $(c^i, \bar{c}_j)$ are n-component Grassmann variables, then $\det(M) = \int \mathrm{d}^n c \, \mathrm{d}^n \bar{c} \, \exp(\bar{c}_j M^j{}_i c^i)$. Applying the same idea here, we have[47]

$$\det \frac{\delta f^{\mathrm{a}}[A^\lambda(x)]}{\delta \lambda^{\mathrm{b}}(y)} = \int Dc \, D\bar{c} \; \mathrm{e}^{-S_{\mathrm{gh}}[\bar{c},c\nabla]} \tag{7.69}$$

[47] One can show that the determinant is positive–definite, at least in a neighbourhood of the trivial connection. Thus, for the purposes of perturbation theory around the trivial background, we can drop the modulus sign. Non–perturbatively we must be more careful.

where

$$S_{\mathrm{gh}}[\bar{c}, c, \nabla] := -\int_{M\times M} \bar{c}^{\mathrm{a}}(x) \frac{\delta f^{\mathrm{a}}[A^\lambda(x)]}{\delta \lambda^{\mathrm{b}}(y)} c^{\mathrm{b}}(y) \tag{7.70}$$

and the fields $(c^{\mathrm{a}}, \bar{c}^{\mathrm{a}})$ are *fermionic scalars*, again valued in the adjoint representation of G. They are known as *ghosts* (c) and *antighosts* ($\bar{c}$).

These ghosts certainly seem strange. In the Michaelmas QFT course you learned that in any unitary QFT, the spin–statistics theorem requires fermionic fields to have half–integral spin (at least for $d > 2$). How come we're now allowed fermionic scalars? In section 7.3.2 we'll understand that the space of states $\Psi[A, \bar{c}, c, h]$ including the ghosts is a complex vector space, but is not a Hilbert space: it's inner product fails to be positive definite. Thus the theory including ghosts is indeed non–unitary. However, we'll see that there *is* a positive–definite inner product on the space of *gauge invariant* states that are independent of the ghosts and Nakanishi–Lautrup fields. Thus, provided we only try to compute expectation values of gauge invariant operators that are independent of $(\bar{c}, c, h)$, we will have a unitary theory. The whole reason for introducing ghosts was just to remove unphysical gauge degrees of freedom in the naive path integral over the simple affine space $\mathcal{A}$ — the 'physically meaningful' integral was always supposed to be taken over $\mathcal{A}/\mathcal{G}$, where no ghosts arise. Operators in $\mathcal{A}/\mathcal{G}$ would certainly be written without ghosts.

Putting everything together, our Yang–Mills path integral can finally be written as

$$\int_{\mathcal{A}/\mathcal{G}} D\mu \; \mathrm{e}^{-S_{\mathrm{YM}}[\nabla]} = \int D[A, c, \bar{c}, h] \; \exp\left(-S_{\mathrm{YM}}[\nabla] - S_{\mathrm{gh}}[\bar{c}, c, \nabla] - S_{\mathrm{gf}}[h, \nabla]\right) \tag{7.71}$$

where the integral on the *rhs* is formally to be taken over the space of all fields $(\nabla, \bar{c}, c, h)$. Everything on the right now looks like some form of action, so we can hope to compute it perturbatively using Feynman rules.

Let's now make all this more concrete by seeing how it works in an example. An important, frequently occurring choice of gauge is *Lorenz*[48] *gauge* is to pick the trivial connection ∂ as a base-point by writing $\nabla = \partial + A$, and impose that A obeys

$$f^{\mathrm{a}}[A] = \partial^\mu A^{\mathrm{a}}_\mu(x) = 0 \tag{7.72}$$

for all $x \in M$ and for all $\mathrm{a} = 1, \ldots, \dim(G)$. An obvious reason to want to work in Lorenz gauge is that in the important case $(M, g) = (\mathbb{R}^d, \delta)$, it respects the $\mathrm{SO}(d)$ invariance of the flat Euclidean metric. Since $A^\lambda = A + \nabla\lambda$ (where $\nabla = \partial + A$ is the connection associated to A), in the case of Lorenz gauge we have

$$\begin{aligned} f^{\mathrm{a}}[A^\lambda] &= \partial^\mu (A_\mu + \nabla_\mu \lambda)^{\mathrm{a}} = \partial^\mu A^{\mathrm{a}}_\mu + \partial_\mu (\nabla_\mu \lambda)^{\mathrm{a}} \\ &= \partial^\mu A^{\mathrm{a}}_\mu + \partial^\mu (\partial_\mu \lambda^{\mathrm{a}} + \frac{1}{2} f^{\mathrm{a}}_{\mathrm{bc}} A^{\mathrm{b}}_\mu \lambda^{\mathrm{c}}) \,. \end{aligned} \tag{7.73}$$

Consequently, the matrix appearing in the Faddeev–Popov determinant is

$$\frac{\delta f^{\mathrm{a}}[A^\lambda(x)]}{\delta \lambda^b(y)} = (\partial^\mu \nabla_\mu)^{\mathrm{ab}}_x \, \delta^d(x - y) \tag{7.74}$$

[48] Poor Ludvig Lorenz. Eternally outshone by Hendrik Lorentz to the point of having his work misattributed.

where the subscript x denotes the variable on which the differential operator $\partial^\mu \nabla_\mu$ acts. The factor of $\delta^d(x-y)$ arose because our gauge–fixing condition was local: the object $\partial^\mu(\nabla_\mu \lambda)$ lives at one point x, so we get nothing if we vary it wrt to changes in λ at some other point. While it may look scary to have a differential operator acting on these δ-functions, using this result in the ghost action yields

$$\begin{aligned} S_{\rm gh}[\bar{c}, c, \nabla] &= -\int_{M\times M} \bar{c}^{\rm a}(x)\, (\partial^\mu \nabla_\mu)^{\rm ab}_x\, \delta^d(x-y)\, c^{\rm b}(y)\, {\rm d}^d y\, {\rm d}^d x \\ &= -\int_M \bar{c}^{\rm a} (\partial^\mu \nabla_\mu c)^{\rm a}\, {\rm d}^d x \\ &= \int_M (\partial^\mu \bar{c}^{\rm a})(\nabla_\mu c)^{\rm a}\, {\rm d}^d x \end{aligned} \tag{7.75}$$

where in the first step we integrated out y using the δ-function, recalling that the differential operators only care about x. Altogether, in Lorenz gauge we have the action

$$S[\nabla, c, \bar{c}, h] = \frac{1}{2{\rm g}^2_{\rm YM}} \int_M {\rm tr}\,(dA + A^2) \wedge *(dA + A^2) + \int_M {\rm tr}\,(d\bar{c} \wedge * \nabla c) + \int_M {\rm tr}\,(h\; d*A) \tag{7.76}$$

Except for the strange spin/statistics of the ghost fields and the mixture of covariant and normal derivatives, this is now a perfectly respectable, local action for scalar fields coupled to the gauge field. Notice that in the Abelian case where the adjoint representation is trivial, the Lorenz gauge ghost action would be $\int_M d\bar{c} \wedge *dc$ and in particular would be independent of the gauge field A. Thus the ghosts would have completely decoupled, which is why you didn't meet them last term.

The remainder of this section is really just for enthusiasts: if your brain's already swimming, it's better to skip on to section 7.3.2.

If you're still reading, then recall that in the finite dimensional case, we needed our gauge fixing function to obey two conditions: that we can indeed always find a gauge transformation such that $f[A] = 0$ holds, and that once we've found it, this gauge is unique so that in particular starting on the gauge slice and performing any gauge transformation takes us off the slice.

Let's start by considering whether we *can* always solve (7.72). In other words, let $\nabla \in \mathcal{A}$ be some arbitrary connection and let $\mathcal{G}(\nabla)$ denote the orbit of ∇ under $\mathcal{G}$. Then we must show that there is always *some* $\nabla' \in \mathcal{G}(\nabla)$ whose connection 1-form obeys (7.72). You've seen how to do this in the Abelian case of electrodynamics: you noted that under a gauge transform, $A \to A' = A - d\lambda$ for some λ. The condition that A be in Lorenz gauge says

$$0 = \partial^\mu A'_\mu = \partial^\mu A_\mu - \Delta\lambda \tag{7.77}$$

where Δ is the Laplacian on M. Regarding this as a condition on λ, we must solve $\Delta\lambda = u$ where $u(x) = \partial^\mu A_\mu(x)$ is essentially arbitrary. This can always be done provided u is orthogonal to the kernel of the adjoint of the Laplacian in the L^2 norm on (M, g). However, the Laplacian is self–adjoint ($\int_M \phi * \Delta\psi = \int_M (\Delta\phi) * \psi$) and $\ker \Delta$ consists of

constant functions, since if $u \in \ker \Delta$ then

$$0 = \int_M u * \Delta u = -\int_M du \wedge *du = -||du||^2 \tag{7.78}$$

whenever u has compact support. Hence $du = 0$ so u is constant. Thus, for any generic electromagnetic potential A_μ, we can find a gauge transform that puts it in Lorenz gauge. In the non–Abelian case, things are more complicated because the gauge transform of a connection is non–linear: $A \to A + \nabla_A \lambda$, whose value depends on which A we start with. It turns out that we can always solve (7.72), but the proof is considerably more complicated — one was found by Karen Uhlenbeck in 1982 (at least for some common choices of M), and an alternative proof was later found by Simon Donaldson.

We must still check that (7.72) singles out a *unique* representative, so that we count each gauge equivalence class only once. Encouragement comes from the fact that connections obeying (7.72) are orthogonal to connections that are pure gauge, with respect to the L^2-metric (7.47) on $\mathcal{A}$. For if a_1 is a tangent vector at the point $\nabla \in \mathcal{A}$ that obeys $\nabla * a_1 = 0$, while $a_2 = \nabla\lambda$ is also a tangent vector at ∇ that points in the direction of an infinitesimal gauge transform, then

$$ds^2_{\mathcal{A}}(a_1, a_2) = \int_M \mathrm{tr}(a_1 \wedge *a_2) = \int_M \mathrm{tr}(a_1 \wedge *\nabla\lambda) = -\int_M \mathrm{tr}((\nabla * a_1)\,\lambda) = 0 \tag{7.79}$$

using the Lorenz condition. Thus changing our connection in a way that obeys Lorenz gauge takes us in a direction that is *orthogonal* to the orbits of the gauge group. This certainly shows that starting from any base-point and integrating over all gauge fields along the slice incorporates only gauge inequivalent connections while we're *near* our base-point.

However, as in the finite dimensional example where the line $y = 0$ intersected each circle on constant radius twice, it doesn't guarantee that some other connection, far away along the gauge slice, isn't secretly gauge equivalent to one we've already accounted for. This troubling possibility is known as the *Gribov ambiguity*, after Vladimir Gribov who first pointed it out and showed it actually occurs in the case of Coulomb gauge $\partial^i A_i = 0$ (the indices just run over $\mathbb{R}^3 \subset \mathbb{R}^{3,1}$). Somewhat later, Iz Singer showed that the Gribov ambiguity is in fact *inevitable*: no matter which gauge condition you pick, the gauge orbit *always* intersects the gauge slice more than once (at least for most reasonable M). To show this, Singer noted that $\mathcal{A}$ is itself an infinite dimensional principle bundle over the space $\mathcal{B} := \mathcal{A}/\mathcal{G}$ where the group $\mathcal{G}$ of all gauge transformations plays the role of the *structure* group. A gauge slice amounts to a global section of this bundle — *i.e.*, the choice of a unique point in $\mathcal{A}$ for each point in $\mathcal{B}$. A result I won't prove states that a principal bundle only admits a global section when it's topologically trivial, so the existence of a global gauge slice would imply

$$\mathcal{A} \stackrel{?}{\cong} \mathcal{B} \times \mathcal{G}\,. \tag{7.80}$$

Since $\mathcal{A}$ is an affine space, clearly $\pi_k(\mathcal{A}) = 0$ for all $k > 0$ (*i.e.* $\mathcal{A}$ itself is topologically trivial and has no non–contractible cycles). However, Singer computed that $\pi_k(\mathcal{G}) \neq 0$ for at least some $k > 0$ which says that there *are* some non–contractible cycles in the space on the *rhs* of (7.80). Thus $\mathcal{A} \neq \mathcal{B} \times \mathcal{G}$, so $\mathcal{A}$ is non–trivial as a principal bundle

over $\mathcal{B}$, and no global gauge choice exists. In practice, we'll work perturbatively, meaning we never venture far enough from our chosen base-point connection to meet any Gribov copies. Non–perturbatively, we'd have to cover $\mathcal{A}/\mathcal{G}$ with different coordinate patches, pick different gauges in each one and then piece them together at the end. I'm not aware of anyone actually trying to do this.

7.3.2 BRST cohomology

We've seen that renormalization group flow generates an infinite series of interactions — every possible term that is not forbidden by symmetries of the original action and regularized measure — all but finitely many of which are strongly suppressed at low energies. The gauge–fixing and ghost terms in the previous section are *not* invariant under gauge transformations, so we cannot use gauge invariance as a criterion by which to restrict the possible terms that are generated by RG flow. But this is troubling: if gauge non–invariant terms are indeed allowed in the effective action, what is to stop (*e.g.*) a mass term[49] $\sim \mathrm{tr} A_\mu A^\mu$ from being generated automatically in the quantum theory?

Becchi, Rouet & Stora and independently Tyurin showed that the full gauge–fixed action (**??**) in fact does possess a remarkable symmetry that remembers the gauge invariance of the original Yang–Mills action. Consider the transformations

$$\begin{aligned} \delta A_\mu &= \epsilon \nabla_\mu c & \qquad \delta \bar{c} &= \mathrm{i}\epsilon h \\ \delta c &= -\frac{\epsilon}{2}[c,c] & \qquad \delta h &= 0 \end{aligned} \tag{7.81}$$

where ϵ is a constant, anticommuting parameter. Note that $[c,c]^{\mathrm{a}} = f^{\mathrm{a}}_{\mathrm{bc}} c^{\mathrm{b}} c^{\mathrm{c}}$ is not identically zero because the ghosts are Grassmann valued. Letting Ψ^i denote any of the fields $\{A^{\mathrm{a}}_\mu, c^{\mathrm{a}}, \bar{c}^{\mathrm{a}}, h^{\mathrm{a}}\}$, we will often write these transformations as $\delta\Psi = \epsilon\, \mathcal{Q}\Psi$ so that $\mathcal{Q}\Psi$ represents the *rhs* of (7.81) with the anticommuting parameter ϵ stripped away. Note that $\mathcal{Q}\Psi^i$ thus has opposite statistics to Ψ^i itself.

The expression for δA shows that, as far as the gauge field itself is concerned, these BRST transformations act just like a gauge transformation $A_\mu \to A_\mu + \nabla_\mu \lambda$ with gauge parameter $\lambda^{\mathrm{a}}(x) = \epsilon c^{\mathrm{a}}(x)$ given in terms of the ghost field. It follows that any gauge–invariant function of the connection alone, such as the original Yang–Mills action $S_{\mathrm{YM}}[\nabla]$, is invariant under the transformations (7.81). To see that the rest of the action is also invariant under (7.81), we'll first show that $[\delta_1, \delta_2] = 0$, where $\delta_{1,2}$ are transformations with parameters $\epsilon_{1,2}$. Since

$$[\delta_1, \delta_2]\Psi^i = \delta_1(\epsilon_2\, \mathcal{Q}\Psi^i) - \delta_2(\epsilon_1\, \mathcal{Q}\Psi^i) = -(\epsilon_1\epsilon_2 - \epsilon_2\epsilon_1)\mathcal{Q}^2\Psi^i = -2\epsilon_1\epsilon_2\, \mathcal{Q}^2\Psi^i \tag{7.82}$$

using the fact that the parameters are fermionic and so anticommute with $\mathcal{Q}$. Therefore the statement $[\delta_1, \delta_2] = 0$ amounts to the statement that the transformation $\Psi \mapsto \mathcal{Q}\Psi$ is nilpotent.

[49] Here I mean just a naive perturbative mass term for the gauge boson (or gluon), not the more sophisticated appearance of a mass gap in the non–perturbative theory, or phase transition to a massive theory by means of the Higgs mechanism.

For the Nakanishi–Lautrup field h, this assertion is trivial. Similarly, for the antighost $\bar{c}$ we have

$$[\delta_1, \delta_2]\bar{c} = \mathrm{i}\delta_1(\epsilon_2 h) - \mathrm{i}\delta_2(\epsilon_1 h) = 0 \tag{7.83}$$

since h itself is invariant. For the gauge field,

$$\begin{aligned}[\delta_1, \delta_2]A_\mu &= \delta_1(\epsilon_2 \nabla_\mu c) - \delta_2(\epsilon_1 \nabla_\mu c) \\ &= \epsilon_2 \nabla_\mu(\delta_1 c) + \epsilon_2[\delta_1 A_\mu, c] - (1 \leftrightarrow 2) \\ &= (\epsilon_2\epsilon_1 - \epsilon_1\epsilon_2)\left[-\frac{1}{2}\nabla_\mu([c,c]) + [\nabla_\mu c, c]\right] ,\end{aligned} \tag{7.84}$$

which vanishes since $\nabla_\mu[c,c] = [\nabla_\mu c, c] + [c, \nabla_\mu c] = 2[\nabla_\mu c, c]$ using antisymmetry of the Lie bracket $[\ ,\ ]$ together with the fact that the ghosts are Grassmann valued. Finally, for the ghost itself we have

$$\begin{aligned}[\delta_1, \delta_2]c &= -\frac{\epsilon_2}{2}([\delta_1 c, c] + [c, \delta_1 c]) + \frac{\epsilon_1}{2}([\delta_2 c, c] + [c, \delta_2 c]) \\ &= \frac{\epsilon_2}{4}([\epsilon_1[c,c], c] + [c, \epsilon_1[c,c]]) - (1 \leftrightarrow 2) \\ &= \frac{1}{2}(\epsilon_2\epsilon_1 - \epsilon_1\epsilon_2)[[c,c], c]^{\mathrm{a}} t_{\mathrm{a}} .\end{aligned} \tag{7.85}$$

Because the ghosts anticommute, the expression $[[c,c],c]^{\mathrm{a}} = f^{\mathrm{a}}_{\mathrm{bc}} f^{\mathrm{b}}_{\mathrm{de}} c^{\mathrm{d}} c^{\mathrm{e}} c^{\mathrm{c}}$ must be totally antisymmetric on $\{\mathrm{d}, \mathrm{e}, \mathrm{c}\}$ and hence it vanishes by the Jacobi identity. Thus $\mathcal{Q}^2\Psi^i = 0$ for any single field $\Psi^i \in \{A, c, \bar{c}, h\}$.

Now let's show that the BRST transformation is nilpotent even when acting on an arbitrary functional $\mathcal{O}(A, c, \bar{c}, h)$ of the fields. We compute

$$\mathcal{Q}^2\mathcal{O} = \mathcal{Q}\left((\mathcal{Q}\Psi^i)\,\frac{\delta\mathcal{O}}{\delta\Psi^i}\right) = \mathcal{Q}^2\Psi^i\,\frac{\delta\mathcal{O}}{\delta\Psi^i} - \mathcal{Q}\Psi^i\,\mathcal{Q}\Psi^j\frac{\delta^2\mathcal{O}}{\delta\Psi^j\,\delta\Psi^i} . \tag{7.86}$$

The first term vanishes by our calculations above. To see that the second term also vanishes, split the sums over all fields (labelled by i, j) into separate sums over bosonic fields $\Psi^i \in \{A_\mu, h\}$ and fermionic fields $\Psi^i \in \{c, \bar{c}\}$. In the case that i and j both refer either to bosonic or fermionic fields, the term cancels because $\mathcal{Q}\Psi$ has opposite statistics to Ψ^i itself, so that pre–factor is symmetric if the second derivatives are antisymmetric, and *vice–versa*. The mixed terms cancel among themselves.

We're now in position to see why the full, gauge–fixed action is BRST invariant. We have

$$\int \mathcal{Q}\,\mathrm{tr}(\bar{c}f[A])\,\mathrm{d}^d x = \int\left[\mathrm{i}\,\mathrm{tr}(hf[A]) - \mathrm{tr}\left(\bar{c}\frac{\delta f}{\delta\theta}\nabla_\mu c\right)\right]\mathrm{d}^d x = S_{\mathrm{gf}}[h, A] + S_{\mathrm{gh}}[A, c, \bar{c}] , \tag{7.87}$$

so the gauge–fixing and ghost terms in the action are the BRST transformation of $\mathrm{tr}(\bar{c}f[A])$. Since BRST transformations are nilpotent, it follows that these terms are BRST invariant for any gauge–fixing functional $f[A]$. Combined with the gauge invariance of the original Yang–Mills action, this shows that BRST transformations preserve the full Yang–Mills

gauge–fixed action. Provided we regularize the path integral measure in a way that preserves this (as will be true perturbatively in dimensional regularization), BRST symmetry will be a symmetry of the quantum theory, and all new terms that are generated by RG flow will also be constrained to be BRST invariant. In particular, terms that depend only on the original gauge field will be constrained to be gauge invariant, preventing the appearance of a mass term $\sim A^2$ even at the quantum level.

Furthermore, if we restrict ourselves to computing correlation functions of operators $\{\mathcal{O}_i\}$ that are BRST invariant, the result will be completely independent of the particular choice of gauge–fixing functional $f[A]$, just as for the partition function. This is because $f[A]$ appears only in the BRST exact term $\mathcal{Q}\,\mathrm{tr}(hf[A])$ in the action. Therefore, changing f leads to a change

$$\Delta\left\langle\prod_i \mathcal{O}_i\right\rangle = -\left\langle\int \mathcal{Q}\,\mathrm{tr}(h\,\Delta f)\,\mathrm{d}^d x\,\prod_i \mathcal{O}_i\right\rangle = -\int\left\langle\mathcal{Q}\left(\mathrm{tr}(h\,\Delta f)\prod_i \mathcal{O}_i\right)\right\rangle \mathrm{d}^d x \quad (7.88)$$

using the assumption $\mathcal{Q}\mathcal{O}_i = 0$ of BRST invariant operators. This vanishes because it is the integral of a BRST exact quantity, so is a total derivative on the space of fields.

It's revealing to view these results from a Hamiltonian / canonical quantization perspective. By Noether's theorem, the fact that the action is invariant under BRST transformations leads to a conserved charge, which you should check is given by[50]

$$\begin{aligned} Q &= \int_N \mathrm{tr}\left(\frac{1}{\mathrm{g}_{\mathrm{YM}}^2}\nabla c\wedge *F_\nabla + \mathrm{i}h * \nabla c + \frac{1}{2} * d\bar{c}\,[c,c]\right) \\ &= \int \mathrm{tr}\left(\frac{1}{\mathrm{g}_{\mathrm{YM}}^2}\nabla_i c\,F^{0i} + \mathrm{i}h\nabla_0 c + \frac{1}{2}\partial_0\bar{c}\,[c,c]\right)\mathrm{d}^{d-1}x \end{aligned} \quad (7.89)$$

in the case of the Lorenz gauge action (**??**), where the second line is valid in the standard case that the co-dimension 1 surface N used to define the charge is the plane $x^0 = \text{const.}$.

To be completed: BRST cohomology and relation to Hamiltonian approach.

7.4 Perturbative renormalization of Yang–Mills theory

Even though perturbation theory is of limited use at low energies, we'll begin our study of quantum Yang–Mills theory by trying to see what we can learn from it. As an incentive, we should expect that the perturbative description *will* be useful at high energies (where the renormalized coupling turns out to be small), so understanding perturbation theory will allow us to probe the UV behaviour of Yang–Mills theory.

7.4.1 Feynman rules in R_ξ gauges

If we write $\nabla = \partial + A$ to expand around the trivial connection, then $F = dA + A^2$, the Yang–Mills Lagrangian schematically contains terms

$$F^2 \sim (dA)^2 + A^2(dA) + A^4\,. \quad (7.90)$$

[50] In a more sophisticated treatment, I'd point out that under canonical quantization this charge is really the Chevalley–Eilenberg differential of the infinite–dimensional Lie group $\mathcal{G}$ of all gauge transformations, acting on the space of fields.

Thus, in addition to the purely kinetic term $(dA)^2$, we have a four–gluon interaction vertex A^4 and a three–gluon interaction vertex A^2dA. Let's start by taking a closer look at the propagators and interactions that result from this gauge–fixed action. For simplicity, we'll consider just pure Yang–Mills theory on $(M, g) = (\mathbb{R}^d, \delta)$, with the gauge–fixing functional taken to be $f[A] = \partial^\mu A_\mu$.

As it stands, the gauge–fixing term $\mathrm{i}\int \mathrm{tr}(h\,\partial^\mu A_\mu)$ is a little awkward to work with; the field h is non–dynamical, but integrating it out introduces a δ–function into the path integral that we don't know how to handle. For this reason, it will be convenient to add the BRST exact term

$$-\,\mathrm{i}\frac{\xi}{2}\int \mathcal{Q}\,\mathrm{tr}(\bar{c}h)\,\mathrm{d}^d x = \frac{\xi}{2}\int \mathrm{tr}(hh)\,\mathrm{d}^d x \tag{7.91}$$

to the action, where ξ is an arbitrary constant. Since this term is BRST exact, its presence doesn't affect the value of any correlation function of BRST invariant operators. However, integrating out h we now find the constraint $h = -\mathrm{i}\,\partial^\mu A_\mu/\xi$. Inserting this back into the action gives

$$\int \left[\mathrm{i}\,\mathrm{tr}(h\,\partial^\mu A_\mu) + \frac{\xi}{2}\mathrm{tr}(h^2)\right]\mathrm{d}^d x = \frac{1}{2\xi}\int \mathrm{tr}(\partial^\mu A_\mu\,\partial^\nu A_\nu)\,\mathrm{d}^d x \tag{7.92}$$

which can be seen as a modification of the kinetic term of the gauge field. Combining this with the kinetic part of the original Yang–Mills action, one finds that the momentum space gluon propagator is

$$D^{\mathrm{ab}}_{\mu\nu}(p) = -\frac{\delta^{\mathrm{ab}}}{p^2}\left[\delta_{\mu\nu} - (1-\xi)\frac{p_\mu p_\nu}{p^2}\right] \tag{7.93}$$

in momentum space. This propagator is often said to be in R_ξ gauge. Since it originally appeared in front of a BRST exact term, the value of ξ can be chosen freely; common choices are $\xi = 0$ (Landau's choice – this recovers the original Lorenz gauge as for electromagnetism) and $\xi = 1$ (Feynman and 't Hooft's choice).

The gluons can interact among themselves via both the A^2dA and A^4 interaction vertices in the original action. In detail, the flat space three–gluon vertex is

$$\Gamma^{\mathrm{abc}}_{\mu\nu\lambda}(k,p,q) = -\mathrm{g}_{\mathrm{YM}}\,f^{\mathrm{abc}}\left[(k-p)_\lambda\,\delta_{\mu\nu} + (p-q)_\mu\,\delta_{\nu\lambda} + (q-k)_\nu\,\delta_{\lambda\mu}\right] \tag{7.94}$$

in momentum space, while the four–gluon vertex is

$$\begin{aligned}\Gamma^{\mathrm{abcd}}_{\mu\nu\lambda\sigma} = &-\mathrm{g}^2_{\mathrm{YM}}\,f^{\mathrm{abe}}\,f^{\mathrm{cde}}\,(\delta_{\mu\lambda}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\lambda}) \;-\; \mathrm{g}^2_{\mathrm{YM}}\,f^{\mathrm{ace}}\,f^{\mathrm{bde}}\,(\delta_{\mu\nu}\delta_{\sigma\lambda} - \delta_{\mu\sigma}\delta_{\nu\lambda}) \\ &-\; \mathrm{g}^2_{\mathrm{YM}}\,f^{\mathrm{ade}}\,f^{\mathrm{bce}}\,(\delta_{\mu\nu}\delta_{\sigma\lambda} - \delta_{\mu\lambda}\delta_{\nu\sigma})\,,\end{aligned} \tag{7.95}$$

and is independent of the momenta since it is a local (non–derivative) interaction.

We must also consider ghost fields which can run around loops even if they do not appear externally. The gauge fixed action above yields a ghost propagator

$$C^{\mathrm{ab}}(p) = \frac{\delta^{\mathrm{ab}}}{p^2} \tag{7.96}$$

in momentum space. This is the standard form expected for massless scalars, except the fact that we get a + sign rather than a − sign reflects the fact that the ghosts are fermionic. Finally, in Lorenz gauge $\partial^\mu a_\mu = 0$ we have a $a_\mu bc$ vertex of the form

$$\Gamma^{\rm abc}(p) = -{\rm g}_{\rm YM} f^{\rm abc} p_\mu \tag{7.97}$$

where the component p_μ of the momentum of the antighost couples to the component a_μ of the gauge field. It's a really good exercise — and a standard exam question — to check you can derive all these terms from the action above.

The most likely feeling at this point is panic. It doesn't take much imagination to see that any attempt to using these vertices to construct Feynman diagrams will quickly run into a huge proliferation of terms. In fact, counting each term in the vertices (7.94) & (7.95) separately, even a fairly simple process like $2 \to 3$ gluon scattering receives contributions from $\sim$ 10,000 terms, already just at *tree* level! In theories with charged matter, such as QCD and the Standard Model, there are further interactions coming from the gluons in the covariant derivatives, and at loop level there are further contributions from the ghosts.

On the one hand, perhaps this is just the way it is. After all, Yang–Mills theory is a complicated, non–linear theory. If you come along and prod it in a more or less arbitrary way (*i.e.* do perturbation theory), you should expect that the consequences will indeed be messy and complicated. But another possible response to the above is a slight feeling of nausea. The whole point of our treatment of Yang–Mills theory in terms of bundles, connections and curvature was to show how tremendously natural this theory is from a geometric perspective. Yet this naturality is badly violated by our splitting of the Yang–Mills action into $(da)^2$, $a^2 da$ and a^4 pieces, none of which separately have any geometric meaning. Surely there must be a different way to treat this theory — one that is less ugly, and treats the underlying geometric structure with more respect?

Many physicists sympathize with this view (me included). In fact, over the years various different ways to think about Yang–Mills theory have been proposed, ranging from viewing Yang–Mills theory as a type of string theory, to writing it in twistor space instead of space-time, to putting it on a computer. Some of these approaches have been highly successful, others only partially so. We'll take a brief look at a few in later sections. For now though, we must soldier on and do our best to understand the theory perturbatively in the neighbourhood of the trivial connection. To do otherwise would be somewhat akin to trying to understand differential geometry without first knowing what a vector is.

7.4.2 Yang–Mills is perturbatively renormalizable!

7.4.3 The β-function and asymptotic freedom

7.5 Further aspects

7.5.1 A string theory in disguise?

't Hooft's picture of YM = string theory. Mention ribbon diagrams (cf matrix models). Show expansion around $N = \infty$ can be rearranged as a genus expansion of the ribbon graph. Vague allusion to AdS/CFT.