

This is the last recitation, thanks for a great semester! I'll still have my usual office hours until the end of the semester, as I've detailed on Canvas. Since I won't be holding any more recitations, I won't be creating any more notes; I'm not sure how helpful y'all find these, but in the chance that they're helpful I'll drop a few of my handwritten Spring 2021 recitation notes into Canvas so you have something to help with the last few topics of the semester! Note that those notes are written up for a class with the $(+, -, -, -)$ metric: the concepts all still hold and they should still be informative, but you should double check any signs that enter the calculation.

Feynman rules for fermions

These rules are specifically for fermion fields. The Feynman rules are mostly the same, but they now have extra indices! The idea here is that each external line will contribute a "polarization" in spin space given by the free field spinors $u_s(p)$ or $v_s(p)$ (or their Dirac conjugates). You need to keep track of the s indices, and if you have n external legs you'll have n spin indices s_1, s_2, \dots, s_n . When you draw the Feynman diagrams for fermions, there are a few things to keep in mind:

- Treat the fermion diagrams like a complex field and draw charge arrows in the direction of fermion flow.
- Each external leg should be labeled with a unique spin $s \in \{1, 2\}$ (this is the polarization s on $u_s(p)$ or $v_s(p)$), as well as a momentum.
- External legs will contribute row / column vectors, and vertices will contribute matrices.

After you draw the diagrams, you can evaluate them like so:

1. Start at the tip of each arrow and work backwards. The reason we need this ordering is that now that we have Dirac spinors, ordering matters! External legs contribute nonzero factors which are either column vectors or row vectors.
2. **External legs** contribute $u_s(p)$ or $v_s(p)$ depending on if we're working with a particle or antiparticle. The first external leg you encounter (going from tip to tail) will give you the Dirac conjugate $\bar{u}_s(p)$ or $\bar{v}_s(p)$, while the next external leg will give you the regular column vector $u_s(p)$ or $v_s(p)$.
3. **Propagators** are denoted with solid lines of momentum p , and contribute:

$$\frac{1}{-i\not{k} + m - i\epsilon} = \frac{i\not{k} + m}{k^2 + m^2 - i\epsilon} \quad (1)$$

If you have multiple types of fields in a Lagrangian, each should be denoted with its own set of lines.

4. **Vertices** can now contribute matrix structures¹, i.e. the interaction $g\bar{\psi}\gamma^\mu\psi$ contributes a factor of $ig\gamma^\mu$. This is one of the reasons we need to be very careful about ordering, as matrices don't commute and so we need to keep track of their ordering!

Note that here, we can also instead view the interaction as $i\bar{\psi}_a(\gamma^\mu)_{ab}\psi_b$ if we write everything in components. This makes the analogy to the scalar field Feynman rules we had more clear: the " i times the coefficient in the Lagrangian" still works if we think about it like this, since the "coefficient in the Lagrangian" is $g(\gamma^\mu)_{ab}$.

¹They still contribute an additional factor of i times the coefficient in \mathcal{L} because of the e^{iS} in the path integral

5. Impose **momentum conservation** at each vertex.
6. Use momentum conservation to solve for all **internal momenta** in terms of a set of *independent* momenta. Integrate over each independent momentum k_i with a factor of

$$\int \frac{d^4 k}{(2\pi)^4}. \quad (2)$$

7. Divide by the symmetry factor.

The Feynman rules allow us to compute $i\mathcal{M}$, the invariant matrix element which we use for scattering. The matrix element squared, $|\mathcal{M}|^2$, is the quantity that we need to know for scattering computations to determine $d\sigma/d\Omega$ by integration over phase space, where $d\mu$ is the **Lorentz-invariant measure** which integrates over phase space:

$$d\mu = (2\pi)^4 \delta^{(4)} \left(\sum_i q_i - \sum_f p_f \right) \prod_{j=1}^f \frac{d^3 \vec{k}_j}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}_j}}. \quad (3)$$

Here q_i are the initial state momentum and $\{p_j\}_{j=1}^f$ are the final state momentum. Some special cases to consider are:

- $2 \rightarrow n$ scattering. In this case, the observable of interest is the **differential cross section**:

$$d\sigma = \frac{1}{4E_1 E_2 v} |\overline{\mathcal{M}}|^2 d\mu, \quad (4)$$

where E_1, E_2 are the energies of the incident particles in the center-of-mass frame, and v is their relative momentum in this frame. A useful relation is:

$$\vec{p} = p^0 v = E v. \quad (5)$$

Note that a lot of these expressions place explicit assumptions on the frame the calculation is done in; formulas will almost always be given in the center-of-mass frame, and if you want to use a different one you have to be very careful.

I'll explain the notation with $|\overline{\mathcal{M}}|^2$ in a bit; basically, it's the spin-averaged $|\mathcal{M}|^2$, which is different than just evaluating $|\mathcal{M}|^2$ in the case when we have polarizations in the scattering problem.

- $1 \rightarrow n$ scattering. In this case, we're interested in the **decay width** Γ and the **particle lifetime** τ , given by:

$$d\Gamma = \frac{1}{2E_1} |\overline{\mathcal{M}}|^2 d\mu \quad \tau = \frac{1}{\Gamma}. \quad (6)$$

Note here that the CoM frame is the one where the parent particle is at rest, so you can also use $E_1 = M$, where M is the mass of the parent.

A good resource for scattering equations that I like is Chapter 5 of Schwartz (he uses the notation $d\Pi_{\text{LIPS}}$ for the Lorentz-invariant measure $d\mu$).

To compute the cross section for physical scattering processes, we often have to consider some special cases.

- **Unpolarized scattering:** This is the case we'll typically consider in this class, because the unpolarized case allows us to simplify our expressions down. In this case, we consider the **spin-averaged matrix element squared**. The spin-averaged matrix element squared *averages over initial polarizations and sums over final polarizations*, and is:

$$\overline{|\mathcal{M}|^2} = \frac{1}{\# \text{ incoming pols}} \sum_{\text{all pols}} |\mathcal{M}|^2. \quad (7)$$

The reason we only average over the incoming number of polarizations is because the initial beam is unpolarized and we only have control over the incident polarizations that go into the beam, not the ones that come out. This summing is important because it allows us to convert things into spin traces, which makes a lot of things much easier to calculate. As a quick example: for unpolarized $e^-e^- \rightarrow e^-e^-$ scattering, $(\# \text{ incoming pols}) = 2 \cdot 2 = 4$, since there are two spin $\frac{1}{2}$ electrons in the initial state, and the sum on all polarizations sums on the polarizations of all 4 electrons.

- **Polarized scattering:** This case is less common (for this class). You polarize the incident beam with a specific spin s , and parameterize the spinors according to the spin that you input to the experiment. We likely won't consider this, but if you're interested there's a good example of it with muon scattering in Chapter 5.3 of Schwartz.

The 4-Fermi theory

- **History:** the 4-Fermi theory was the original theory of β decay that Fermi wrote down in the 1940s to describe $n^0 \rightarrow p^+e^-\bar{\nu}$. It turns out that it's an effective field theory of the electroweak sector of the Standard Model. In the Standard Model, β decay is induced through W boson exchange. However, the W is very heavy, $m_W \approx 80 \text{ MeV}$, and so at low energies $E \ll m_W$, the W boson is integrated out of the theory, and the decay is described through a 4-quark interaction.
- **Muon decay.** It turns out that the decay of the muon can also be described through the electroweak interactions, with almost exactly the same physics as for β decay. The Lagrangian we get is also a 4-Fermi Lagrangian, and we'll be using muon decay as our main example for the rest of this section (hadrons are a bit more complicated, so I'll stay away from talking about β decay for now). The corresponding μ^- decay that's described by a 4-Fermi Lagrangian is:

$$\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu \quad (8)$$

Now that we've done spinors, we have enough machinery to write down the interaction part of the 4-Fermi theory Lagrangian:

$$\mathcal{L}_{4F} \supset \frac{4G_F}{\sqrt{2}} (\bar{\nu}^{(\mu)} \gamma_\mu P_L \mu) (\bar{e} \gamma^\mu P_L \nu^{(e)}) \quad (9)$$

where the coupling is given by:

$$G_F = 1.166 \times 10^{-5} \text{ GeV}^{-2}. \quad (10)$$

Let's talk about all the other stuff in this Lagrangian. There are four fermion fields: the muon $\mu(x)$, muon-neutrino $\nu^{(\mu)}(x)$ (here (μ) is a label, not an index!), the electron $e(x)$, and the electron neutrino $\nu^{(e)}(x)$. Each fermion field is a four-component Dirac spinor. The γ are the γ -matrices, and P_L is the left-handed chiral projector $(1 + \gamma_5)/2$.

- **Feynman rules:** We have the usual propagators associated with each particle. For this, we'll be assuming that all particles other than the muon are massless, since $m_e, m_{\nu_k} \ll m_\mu$. The Feynman rule that we'll care about for this theory comes from a four-point vertex, since we have a four-fermion interaction. To do this, we label each external leg with a Dirac index, to avoid confusion. Note that in simpler theories, you typically don't have to explicitly label the Dirac indices: let's expand out the Lagrangian, and explicitly put in $P_L = (1 + \gamma_5)/2$:

$$\mathcal{L}_{4F} \supset \frac{G_F}{\sqrt{2}} \left(\bar{\nu}_\beta^{(\mu)} (\gamma^\mu (1 + \gamma_5))_{\beta\alpha} \mu_\alpha \right) \left(\bar{e}_\rho (\gamma_\mu (1 + \gamma_5))_{\rho\delta} \nu_\delta^{(e)} \right) \quad (11)$$

This lets us read off the Feynman rules for this theory. We have the vertex given as Diagram (1) above, and the rule is:

$$(\text{Diagram 1}) = \frac{iG_F}{\sqrt{2}} (\gamma_\mu (1 + \gamma_5))_{\beta\alpha} (\gamma^\mu (1 + \gamma_5))_{\rho\delta}. \quad (12)$$

The Dirac indices here helps us to remember the “tip to tail” mnemonic for fermion Feynman diagrams. The “tips” of the arrows in the diagrams are at $\bar{\nu}_\beta^{(\mu)}$ and \bar{e}_ρ , respectively. Since the first index on the Dirac matrices are a β and a ρ , this means we need to start at the “tip” of each fermion line and evaluate

- Why the factors of P_L ? This is just what we talked about last week coming into play, now for the full Standard Model! The electroweak force only couples together e_L and ν_L (as discussed last recitation, because electroweak theory is a chiral gauge theory). This means that the fermions in this interaction must be left-handed, and the anti-fermions must be right-handed.
- Computation of muon decay: There's only a single diagram (Diagram 3) that contributes to the muon decay, which is the four-point interaction. We can evaluate it as follows. Remember to keep track of the spin polarizations!

$$i\mathcal{M}_{s,r_1,r_2,r_3} = (\text{Diagram 3}) = \frac{iG_F}{\sqrt{2}} \left(\bar{u}_{r_1}(\vec{p}_1) (\gamma^\mu (1 + \gamma_5)) u_s(\vec{k}) \right) \left(\bar{u}_{r_2}(\vec{p}_2) (\gamma_\mu (1 + \gamma_5)) v_{r_3}(\vec{p}_3) \right) \quad (13)$$

If we want to compute the spin-averaged matrix element square, we average over the initial spin degrees of freedom and sum over the final ones. This is:

$$\overline{|\mathcal{M}_{s,r_1,r_2,r_3}|^2} = \frac{1}{2} \sum_{s,r_1,r_2,r_3} |\mathcal{M}_{s,r_1,r_2,r_3}|^2 \quad (14)$$

To compute the lifetime of the muon $\tau^{(\mu)}$, you need to use this spin-averaged cross section $\overline{|\mathcal{M}_{s,r_1,r_2,r_3}|^2}$. Since we haven't focused on actual computations of lifetimes yet, I'll abstain from doing the integration over phase space. Nevertheless, if you do this, you'll find when the smoke clears that (also setting $m_e, m_{\nu_e}, m_{\nu_\mu}$ equal to zero):

$$\tau^{(\mu)} = \frac{G_F^2 m_\mu^5}{192\pi^3} \approx 2 \mu s \quad (15)$$

Even though the calculation is done in an EFT and not the full Standard Model, it's surprisingly close to the actual value and shows how well an EFT can approximate a full calculation!

Gauge invariance and the covariant derivative

Unfortunately we don't have much time to go into gauge theory during recitation, since this is the last recitation and we had other things to touch on. Regardless, I wanted to add a little bit here about what the covariant derivative is, and the geometrical picture behind gauge invariance.

- For this section, we'll work in scalar QED, i.e. we'll have a scalar field $\phi(x)$ which transforms under a local (gauge) $U(1)$ symmetry² as

$$\phi(x) \longrightarrow e^{ie\lambda(x)}\phi(x). \quad (16)$$

The fact that λ is now a function $\lambda(x)$ makes all the difference, and is the reason that gauge theories are a difficult subject to deal with. The naïve scalar Lagrangian,

$$\mathcal{L}_0 = -(\partial_\mu\phi)^*(\partial^\mu\phi) - m^2\phi^*\phi \quad (17)$$

is invariant under the global rotation $\phi(x) \rightarrow e^{i\alpha}\phi(x)$, but **not under the local rotation** $\phi(x) \rightarrow e^{ie\lambda(x)}\phi(x)$. The reason for this is in the derivative term, since:

$$\partial_\mu\phi(x) \longrightarrow \partial_\mu(e^{ie\lambda(x)}\phi(x)) = \underbrace{e^{ie\lambda}\partial_\mu\phi}_{\text{"covariant"}} + \underbrace{ie\lambda e^{ie\lambda}\phi(x)}_{\text{not "covariant"}}. \quad (18)$$

In contrast, for the global symmetry case, the first part of that term vanished, since the $e^{i\alpha}$ passed through the derivative, which made it obvious that $|\partial_\mu\phi|^2$ was invariant under the symmetry.

- Why are we imposing gauge symmetry on our theories anyway? Hong likely has a better answer for this, but my answer is that they allow us to study novel forms of interactions. By saying that we want theory which is invariant under a gauge symmetry, we implicitly constrain the theory to have a very specific structure which is compatible gauge transformations. Furthermore, the Standard Model is a gauge theory, and it has invariance under local $SU(3) \times SU(2) \times U(1)$ invariance; studying gauge symmetry allows us to further understand the structure of these theories, and to see what physics we can extract from them.

In the more immediate case, gauge theories allow us to couple massless vector³ bosons to other fields. To study QED, that's exactly what we need! The photon is a massless vector boson, and the structure that we get from coupling it to another field is exactly the structure of a gauge theory.

- **Notation:** To keep this discussion a little more general, we'll abbreviate the gauge transformation as $U(x)$ (this also has the advantage that we don't have to write $e^{ie\lambda(x)}$ anymore), so:

$$U(x) = e^{ie\lambda(x)} \quad U^\dagger(x) = e^{-ie\lambda(x)} \quad \phi(x) \longrightarrow U(x)\phi(x). \quad (19)$$

- The problem in trying to impose a gauge symmetry in Eq. (17) lies in the derivative. Since we gauged the phase rotation, it no longer passes through the derivative, since we now have to use the product rule when we transform $\phi(x)$. Let's back up for a second, because there's a more fundamental reason that

²The e here is conventional, and can be absorbed into the $\lambda(x)$ factor if you would like. However, it does get important when we have multiple fields of different charges, as they need to be scaled by the strength of their coupling to the field.

³By "vector", we mean a 4-vector field $A_\mu(x)$.

gauge invariance and the derivative ∂_μ don't like one another. Consider two spacetime points $x \neq y$. If we look at $\phi(x)$ and $\phi(y)$, they **transform differently under gauge transformations**:

$$\phi(x) \longrightarrow U(x)\phi(x) \qquad \phi(y) \longrightarrow U(y)\phi(y). \quad (20)$$

as $U(x)$ and $U(y)$ are arbitrary numbers in the unit circle, and can be chosen to be different. What does that mean about taking a sum or difference of these fields? Well,

$$\phi(x) - \phi(y) \longrightarrow U(x)\phi(x) - U(y)\phi(y) = U(x) \underbrace{(\phi(x) - U^\dagger(x)U(y)\phi(y))}_{\text{not } \phi(x) - \phi(y)} \quad (21)$$

In other words, performing a gauge transformation on $\phi(x) - \phi(y)$ does not transform $\phi(x) - \phi(y)$ to a multiple of itself. When we have gauge symmetry, **we must be very careful about comparing fields at different points, as they don't transform under the same symmetry. We can't add or subtract fields at different points, because local symmetry means that this doesn't transform "covariantly" under the symmetry.**

This answers the question of why the derivative is so ill-behaved under gauge transformations. Its definition is:

$$\partial_\mu \phi(x) = \lim_{\epsilon \rightarrow 0} \frac{\phi(x + \epsilon \hat{\mu}) - \phi(x)}{\epsilon}. \quad (22)$$

This equation doesn't make any sense in the context of what we just talked about, because we see that in the definition of ∂_μ we're inherently comparing the value of ϕ at two different points, x and $x + \epsilon \hat{\mu}$! This means that it should be no surprise that there are issues when we use the derivative in conjunction with a gauge theory, because the objects in the definition of the derivative transform differently!

- There is a way to get around the fundamental issue in trying to compare the field at different points: it's called a **Wilson line**. Let's introduce an object $W(x, y)$ which connects every two points x, y in spacetime, and transforms in the following way:

$$W(x, y) \rightarrow U(x)W(x, y)U^\dagger(y). \quad (23)$$

Note also we want the Wilson line to leave identical points untouched,

$$W(x, x) = 1. \quad (24)$$

The transformation law of $W(x, y)$ makes it possible to compare the field at different spacetime points in a covariant fashion:

$$\phi(x) - W(x, y)\phi(y) \rightarrow U(x)\phi(x) - (U(x)W(x, y)\cancel{U^\dagger(y)})(\cancel{U(y)}\phi(y)) = U(x)(\phi(x) - W(x, y)\phi(y)). \quad (25)$$

The difference between these two fields now transforms correctly under our $U(1)$ gauge symmetry, because of the addition of the Wilson line⁴. We can also now fix our problem with the derivative. We define the **covariant derivative** as:

$$D_\mu \phi(x) \equiv \lim_{\epsilon \rightarrow 0} \frac{\phi(x + \epsilon \hat{\mu}) - W(x + \epsilon \hat{\mu}, x)\phi(x)}{\epsilon}, \quad (26)$$

which is now a well defined object, since $\phi(x + \epsilon \hat{\mu})$ and $W(x + \epsilon \hat{\mu}, x)\phi(x)$ transform identically.

⁴In geometrical terms, $W(x, y)$ performs parallel transport in the principal bundle; for more on this, see the next box.

- The **gauge field**: Let's study the Wilson line a bit more under an infinitesimal translation in the n^μ direction (note here the ∂'_μ means $\partial/\partial x'^\mu$):

$$\begin{aligned} W(x + \epsilon n, x) &= W(x, x) + \epsilon n^\mu \partial'_\mu W(x', x) \Big|_{x'=x} + \mathcal{O}(\epsilon^2) \\ &\equiv 1 + i\epsilon n^\mu A_\mu(x) \end{aligned} \quad (27)$$

where we have defined the **gauge potential** $A_\mu(x)$ as the derivative of the Wilson line, normalized by the electric charge e :

$$A_\mu(x) \equiv -\frac{i}{e} \underbrace{\partial'_\mu W(x', x) \Big|_{x'=x}}_{\text{just a field}}. \quad (28)$$

In this geometric picture, this is exactly where the gauge field comes from. This definition is all well and good, but so what? This seems like an arbitrary definition: can we reproduce the correct transformation law for the gauge field? It turns out that we can, by considering the transformation of $W(x + \epsilon n, x)$ under a gauge transformation. If we let $A_\mu(x) \rightarrow \tilde{A}_\mu(x)$ and $W(x, y) \rightarrow \tilde{W}(x, y)$, we can see how it transforms:

$$\begin{aligned} W(x + \epsilon n, x) &\rightarrow \tilde{W}(x + \epsilon n, x) = 1 + i\epsilon n^\mu \tilde{A}_\mu(x) + \mathcal{O}(\epsilon^2) \\ &= \underbrace{(U(x) + \epsilon n^\mu \partial_\mu U(x) + \dots)}_{U(x+\epsilon n)} \underbrace{(1 + i\epsilon n^\mu A_\mu(x) + \dots)}_{W(x+\epsilon n, x)} U^\dagger(x) \\ &= U(x)U^\dagger(x) + i\epsilon n^\mu \left(U(x)A_\mu(x)U^\dagger(x) - \frac{i}{e}(\partial_\mu U(x))U^\dagger(x) \right) \\ &= 1 + i\epsilon n^\mu \underbrace{\left(A_\mu - \frac{i}{e}(\partial_\mu U)U^\dagger \right)}_{\tilde{A}_\mu} \end{aligned} \quad (29)$$

Now if we plug in⁵ for $U(x) = e^{ie\lambda(x)}$

$$A_\mu(x) \rightarrow \tilde{A}_\mu(x) = A_\mu(x) + \partial_\mu \lambda(x) \quad (31)$$

we recover exactly the transformation law for the gauge field!

- **A formula for the covariant derivative**: How does the gauge field enter the covariant derivative? You can probably guess by now that it'll be exactly what we saw in class. Note that for this derivation,

⁵The reason I kept $U(x)$ general here is because up until the last step where we commuted the $U(x)$ and $U^\dagger(x)$ terms around, this derivation was general for any gauge group! In particular, even for non-commutative ones like $SU(3)$, which describes QCD. For something like that which is more complicated, $A_\mu(x)$ ends up being valued in the Lie algebra associated to the group (so is an $n \times n$ matrix), and has transformation law:

$$A_\mu(x) \rightarrow U(x)A_\mu(x)U^\dagger(x) + \frac{i}{g}(\partial_\mu U(x))U^\dagger(x). \quad (30)$$

we won't sum on μ . Let's expand the covariant derivative by Taylor expanding the Wilson line:

$$\begin{aligned} D_\mu \phi(x) &= \lim_{\epsilon \rightarrow 0} \frac{\phi(x + \epsilon \hat{\mu}) - W(x + \epsilon \hat{\mu}, x) \phi(x)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{\cancel{\phi(x)} + \epsilon \partial_\mu \phi(x) - (\mathbb{1} + ie\epsilon A_\mu(x)) \phi(x)}{\epsilon} \\ &= \partial_\mu \phi(x) + ie A_\mu(x) \phi(x). \end{aligned} \quad (32)$$

So, this reproduces the form of the covariant derivative **exactly** like it was presented in class:

$$D_\mu = \partial_\mu + ie A_\mu. \quad (33)$$

- A closed form for $W(x, y)$: The Wilson line can be evaluated in terms of the gauge field $A_\mu(x)$ by solving the differential equation which defines it, Eq. (28)⁶. The solution to that equation is just an exponential:

$$W(x, y) = \exp \left(-ie \int_C dz^\mu A_\mu(z) \right) \quad (34)$$

where C is a curve which connects y to x .

- **What is the field strength?** We can form a gauge-invariant object, called a **Wilson loop**⁷, defined in terms of a loop P starting and ending at x as:

$$W_P(x) \equiv \exp \left(-ie \oint_P dz^\mu A_\mu(z) \right). \quad (35)$$

The interesting part here is that we can use Stokes' theorem to rewrite this integral as a surface integral over the interior A of P (i.e. $P = \partial A$). The “derivative” of A that comes into play is nothing other than the field strength $F_{\mu\nu}$, and thus can be written as:

$$W_P(x) = \exp \left(-\frac{i}{2} e \int_A d\sigma^{\mu\nu} F_{\mu\nu} \right) \quad (36)$$

where $d\sigma^{\mu\nu}$ is the area element. So, we can see that **the field strength $F_{\mu\nu}$ measures the failure of W to be path-independent**. What do we know from geometry that measures the failure of parallel transport to be path-dependent? None other than the curvature of a space. This connection goes a lot deeper than this: **the field strength $F_{\mu\nu}$ is a curvature on gauge space**. To be more precise, $\mathcal{F} = \frac{1}{2} F_{\mu\nu} dx^\mu dx^\nu$ is the **curvature two-form** on the principal bundle which describes the gauge field, and $\mathcal{A} = A_\mu dx^\mu$ is the **connection one-form** on this principal bundle. Taking the exterior derivative $d\mathcal{A}$ exactly gives \mathcal{F} , which in coordinates is just the antisymmetrized derivative $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

- **Summary:** It's no coincidence that the covariant derivative has such a simple form. It's exactly enforced by the geometry of the problem and of the space of gauge field $U(x)$ (which as described in the next

⁶When the gauge group is non-abelian, things get more complicated and the differential equation must be solved in terms of a **path-ordered exponential**. Luckily for us though, $U(1)$ is an abelian group and we can simply commute its elements around.

⁷In lattice QCD, Wilson loops are called **plaquettes** and are one of the fundamental ways in which one studies gauge fields. The gauge action $S = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ can actually be rewritten on a lattice into a sum of all possible plaquettes.

box, looks like a principal $U(1)$ bundle over spacetime). Demanding that we can compare the field $\phi(x)$ to $\phi(y)$ at two separate points gives rise to a **Wilson line** (which you may know by another name: a *connection*). The Wilson line's job is to parallel transport a field from point x to point y in a gauge-covariant manner. Locally expanding the Wilson line gives us the **gauge field** $A_\mu(x)$, and the transformation properties satisfied by $W(x, y)$ propagate through to the desired transformation law for A_μ . The Wilson line also allows us to define a covariant derivative D_μ , and indeed $D_\mu = \partial_\mu + ieA_\mu$.

General relativity vs. gauge symmetry

Gauge symmetry is actually deeply related to curvature on manifolds, and in particular there are a lot of interesting connections between GR and gauge theories. Consider the fact that the derivative doesn't transform covariantly under gauge symmetry: have you seen this before? If you've taken GR, you know that this is a big problem with using the derivative on a manifold, since taking a derivative of a vector field V_ν , then $\partial_\mu V_\nu$ doesn't transform covariantly under coordinate transformations:

$$\partial_\mu V_\nu(x) \rightarrow \left(\frac{\partial x^\alpha}{\partial x'^\mu} \partial_\alpha \right) \left(\frac{\partial x'^\mu}{\partial x^\beta} V^\beta \right) = \underbrace{\frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x'^\mu}{\partial x^\beta} \partial_\alpha V^\beta}_{\text{covariant}} + \underbrace{\frac{\partial x^\alpha}{\partial x'^\mu} \left(\partial_\alpha \frac{\partial x'^\mu}{\partial x^\beta} \right) V^\beta}_{\text{not covariant}}. \quad (37)$$

The second term in this equation prevents the derivative ∂_μ from transforming covariantly: if it was a proper covector, it would have the transformation law given only by the first term! To ameliorate this, we introduce the **covariant derivative** ∇_μ :

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma_{\mu\sigma}^\nu V^\sigma \quad (38)$$

where $\Gamma_{\mu\sigma}^\nu$ are the **Christoffel symbols**. The Christoffel symbols basically transform in such a way that they cancel the bad part of the transformation law in Eq. (37), and make $\nabla_\mu V^\nu$ transform covariantly. Notice that ∇_μ **and the gauge covariant derivative** D_μ **have the same structure: each one is a derivative, plus an extra “correction” term which soaks up the offending piece and gives us the correct transformation law.**

From a mathematical perspective, these two objects are essentially the same: they both provide **connections** which allow us to move and compare vectors (or gauge group elements) in different tangent spaces. They appear in different contexts, because the setting is different when talking about vectors in GR or gauge group elements in QFT. However, they both have the same overarching structure: a **fiber bundle** over a manifold M . A fiber bundle over M is a space that locally looks like $M \times F$, where F is another space called the fiber. Intuitively, what this means is that a fiber bundle is a manifold with a copy of the space F glued on at every point, and the way that you connect together these different F spaces can be non-trivial and have interesting curvature. In the case of vectors, they live in a vector bundle^a: the **tangent bundle** TM . The picture of TM that you can have in your head is that of a manifold, but at each point in spacetime you attach a tangent space of vectors with the same dimension as the manifold. In the case of gauge group elements, they live in a **principal bundle**, where the fiber F is a **Lie group** (in our case, the group $U(1)$). The intuition for this is that at each point in spacetime, we have a copy of $U(1)$: the gauge field configuration $U(x)$ selects out an element of $U(1)$ at x for each point x in spacetime.

^aCalled a vector bundle because F is a vector space