253a: QFT1
Fall 2010
Matthew Schwartz

Lecture 8: Spin 1 and Gauge Invariance

1 Introduction

Up until now, we have dealt with general features of quantum field theories. For example, we have seen how to calculate scattering amplitudes starting from a general Lagrangian. Now we will begin to explore what the Lagrangian of the real world could possibly be. Then, of course, we will discuss what it actually is, or at least what we have figured out about it so far.

A good way to start understanding the consistency requirements of the physical universe is with a discussion of spin. There is a deep connection between spin and Lorentz invariance that is totally obscure if you just work in non-relativistic quantum mechanics. For example, well before quantum field theory, it was known from atomic spectroscopy that the electron had two spin states. It was also known that light had two polarizations. The polarizations of light are easy to understand at the classical level, if light is a field, but how can an individual photon be polarized? For the electron, we can at least think of it as a spinning top, so there is a classical analogy, but light is massless, so what exactly is spinning? And can we actually predict from first principles the size of the electrons magnetic dipole moment? The answers to all these questions follow from an understanding of Lorentz invariance and the requirements of a consistent quantum field theory.

2 Unitary representations of the Poincare group

Our universe has a number of apparent symmetries that we would like our quantum field theory to respect. One is that no place in space time seems any different from any other place. Thus, our theory should be translation invariant: if we take all our fields $\psi(x)$ and replace them by $\psi(x+a)$ for any a, the observables should look the same. Another is Lorentz invariance: physics should look the same whether we point our measurement apparatus to the left or to the right, or put it on a train. The group of translations and Lorentz transformations is called the **Poincare group**.

We need to know how our states transform

$$|\psi\rangle \rightarrow \mathcal{P}|\psi\rangle \tag{1}$$

under a Poincare transformation $\mathcal{P}$. A set of fields ψ which mix under a transformation group is called a **representation** of the group. For example, scalar fields $\phi(x)$ at all points x form a representation of translations, since $\phi(x) \rightarrow \phi(x+a)$. Quite generally, there should be some kind of basis for our states $|\psi\rangle$, call it $\{|\psi_i\rangle\}$ where i is a discrete or continuous index, so that

$$|\psi_i\rangle \rightarrow \mathcal{P}_{ij}|\psi_j\rangle \tag{2}$$

When the basis closes into itself under action of the group, we have a representation. If no non-trivial subset of our basis closes into itself, we have an *irreducible representation* of the group. We will study representation theory a little more precisely when we get to spinors, but this is enough for now.

In addition, we want **unitary** representations. The reason for this is that the things we compute in field theory are matrix elements

$$\mathcal{M} = \langle\psi_1|\psi_2\rangle \tag{3}$$

which should be Poincare *invariant.* If $\mathcal{M}$ is Poincare invariant, and $|\psi_1\rangle$ and $|\psi_2\rangle$ transform covariantly under a Poincare transformation $\mathcal{P}$, we find

$$\mathcal{M} = \langle \psi_1 | \mathcal{P}^\dagger \mathcal{P} | \psi_2 \rangle \tag{4}$$

So we need $\mathcal{P}^\dagger \mathcal{P} = 1$, which is the definition of unitarity. There are many more representations of the Poincare group than the unitary ones. For example, as we will discuss, the 4-vector representation, A_μ is not unitary. But the unitary ones are the only ones we'll be able to compute Poincare invariant matrix elements from, so we have to understand how to find them.

As an aside, it is worth pointing out here that there is an even stronger requirement on physical theories: the S matrix must be unitary. Recall that the S-matrix is defined by

$$|f\rangle = S|i\rangle \tag{5}$$

Say we start with some normalized state $\langle i|i\rangle = 1$, meaning that the probability for finding anything at time $t = -\infty$ is $P = \langle i|i\rangle^2 = 1$. Then $\langle f|f\rangle = \langle i|S^\dagger S|i\rangle$. So we better have $S^\dagger S = 1$, or else we end up at $t = +\infty$ with less than we started! Thus unitarity of the S-matrix is equivalent to conservation of probability, which seems to be a property of our universe as far as anyone can tell. We'll see consequences of unitary of S later on, but for now let's stick to the discussion of spin.

One way to think of the allocation into irreducible representations is that our universe is clearly filled with different kinds of particles, and different spin states. By doing things, such as shoving an electron in a magnetic field, or sending a photon through a polarizer, we manipulate the spin states. Some states will mix with each other under these manipulations and some will not. We could treat all the particles in the universe together in one big fat representation. But since some states never mix with other states, this would be silly, and maybe overconstraining. So we look at the irreducible representations because those are the building blocks with which we can construct the most general description of nature.

The unitary representations of the Poincare group were classified by Eugene Wigner in 1939. As you might imagine, before that people didn't really know what the rules were for constructing physical theories, and by trial and error they were coming across all kinds of problems. Wigner showed that irreducible unitary representations are uniquely classified by mass m and spin J. m can take any value and spin is a half integer $J = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$. Moreover, Wigner showed that if $m > 0$ then there are $2J + 1$ basis states in the representation, and if $m = 0$ there are exactly 2, for any $J > 0$. These basis states correspond to linearly independent polarizations of particles with spin J. You can find the proof of Wigner's theorem in Weinberg's book. We're not going to reproduce the proof, but we will do some examples which will make the ingredients that go into the proof clear.

Knowing what the representations of the Poincare group is a great start, but we still have to figure out how to construct a unitary interacting theory of particles in these representations. To do that, we would like to embed the irreducible representations into objects with spacetime indices. That is we want to squeeze states of spin $0, \frac{1}{2}, 1, \frac{3}{2}, 2$ etc into scalars fields (ϕ), vectors (V_μ), tensors ($T_{\mu\nu}$), and spinors (ψ). That way we can write down simple looking Lagrangians and develop general methods for doing calculations. We see an immediate complication: tensors have $0, 4, 16, 64, \cdots, 4^n$ elements (forgetting half integer spin for now), but spin states have $0, 3, 5, 7, \cdots, 2j + 1$ physical degrees of freedom. The embedding of the $2j + 1$ states (for a unitary representation) in the 4^n dimensional tensors is tricky, and leads to things like gauge invariance as we will shortly see.

2.1 Unitarity and Lorentz invariance – the problem

We don't need fancy language to see the problem. First of all, let's recall how to thinking about different spin states. In non-relativistic quantum mechanics, you had an electron with spin up $|\uparrow\rangle$ or spin down $|\downarrow\rangle$. This is your basis, and you can have a state which is any linear combination of these two

$$|\psi\rangle = c_1|\uparrow\rangle + c_2|\downarrow\rangle \tag{6}$$

The norm of such a state is

$$\langle\psi|\psi\rangle = |c_1|^2 + |c_2|^2 > 0 \tag{7}$$

This norm is invariant under rotations, which send

$$|\uparrow\rangle \rightarrow \cos\theta|\uparrow\rangle + \sin\theta|\downarrow\rangle \qquad |\downarrow\rangle \rightarrow -\sin\theta|\uparrow\rangle + \cos\theta|\downarrow\rangle \tag{8}$$

(In fact, the norm is invariant under the whole group of 3D rotations, SO(3), which you can see using the Pauli matrices, but that's not important right now.)

Say we wanted to do the same thing with a basis or 4 states $|V_\mu\rangle$, which we would like to transform like a 4-vector. Then an arbitrary linear combination would be

$$|\psi\rangle = c_1|V_0\rangle + c_1|V_1\rangle + c_2|V_2\rangle + c_3|V_3\rangle \tag{9}$$

The norm of this state would be

$$\langle\psi|\psi\rangle = |c_1|^2 + |c_2|^2 + |c_3|^2 + |c_4|^2 > 0 \tag{10}$$

This is the norm for any basis – it is always positive, which is one of the postulates of quantum mechanics. However, the norm is not Lorentz invariant. For example, suppose we start with $|\psi\rangle = |V_0\rangle$ which has norm $\langle\psi|\psi\rangle = 1$. Then we boost in the 1 direction, so we get $|\psi'\rangle = \cosh\beta|V_0\rangle + \sinh\beta|V_1\rangle$. Now the norm is

$$\langle\psi'|\psi'\rangle = \cosh^2\beta + \sin^2\beta \neq 1 = \langle\psi|\psi\rangle \tag{11}$$

Thus the probability of finding the state $|\psi\rangle$ in the state $|\psi\rangle$ depends on what frame we're in! We see that norm is not invariant under the boost. In terms of matrices, the boost matrix

$$\Lambda = \begin{pmatrix} \cosh\beta & \sinh\beta \\ \sinh\beta & \cosh\beta \end{pmatrix} \tag{12}$$

is not unitarity: $\Lambda^\dagger \neq \Lambda^{-1}$.

Suppose instead, we try to modify the "postulates of quantum mechanics". After all, they were defined in the non-relativisitic limit, and the problem we are seeing has to do with boosts. So let's just define the norm to be

$$\langle\psi|\psi\rangle = |c_1|^2 - |c_2|^2 - |c_3|^2 - |c_4|^2 \tag{13}$$

This is Lorentz invariant, but not positive definite. That doesn't mean we will have negative probabilities – the probability $\mathcal{P} = |\langle\psi|\psi\rangle|^2 \geqslant 0$ for any state. However, the probabilities will no longer be $\leqslant 1$. For example, suppose $|\psi\rangle = |V_0\rangle$ so that $\langle\psi|\psi\rangle = 1$ as before. In a boosted frame, $|\psi'\rangle = \cosh\beta|V_0\rangle + \sinh\beta|V_1\rangle$. The norm is preserved under evolution, so $\langle\psi'|\psi'\rangle = 1$. However, the probability of finding $|\psi\rangle$ in the state $|V_0\rangle$ is $|\langle V_0|\psi\rangle|^2 = \cosh^2\beta$. For decent sized boosts, $\cosh\beta > 1$, so there is no way to interpret this projection as a probability. Thus, because Lorentz transformations can mix positive norm and negative norm states, the probabilities are not bounded. In fact, you can show that having a probability interpretation, with $0 \leqslant P \leqslant 1$, requires us to have positive norm states. So unitarity, with a positive definite norm, is critical to the having any physical interpretation of quantum mechanics.

Here, we are seeing the conflict between having a Hilbert space with a positive norm, which is a physical requirement leading to the $\delta^{\mu\nu}$ inner product preserved under Unitary transformations, and the requirement of Lorentz invariance, which expects the $\eta^{\mu\nu}$ inner product preserved under Lorentz transformations.

What do we do about it? Well, there are two things we need to fix. First of all, note that $V_\mu^2 = V_0^2 - V_1^2 - V_2^2 - V_3^2$ has 1 positive term and 3 negative terms. In fact, the vector representation of the Lorentz group V_μ has two spins in it: a spin 0 part, with 1 degree of freedom, and a spin 1 part, with 3 degrees of freedom. We need to separate them out so that the representation will always have the same sign norm. When a description only has a single spin in it, it is called an **irreducible** representation of the Lorentz group. We want to embed these irreducible representations into tensors like V_μ or $h_{\mu\nu}$ so that we can write local Lagrangians for them. But it turns out that it's impossible to put more than one spin in the same tensor field and have a universally positive norm.

The second thing is that while there are in fact no non-trivial finite-dimensional irreducible representations of the Lorentz group, there are some infinite dimensional ones. We will see that instead of constant basis vectors, like $(1,0,0,0)$, $(0,1,0,0)$ etc, we will need a basis $\epsilon_\mu(p)$ which depends on the momentum of the field. We will see this when we quantize the fields, which is when Unitarity comes in. So first we will see how to embed the right number of degrees of freedom for a particular mass and spin (irreducible representation of the Poincare group) into tensors like A_μ. Then we will quantize the theory and see how the infinite dimensionality of the representation plays in.

3 Embedding particles into fields

Let's now explore how we can construct Lagrangians for fields which contain only particles of single spins. We will begin with the classical theory. In this case, we don't have to worry about unitarity, but we can ask for the energy to be positive definite, which is basically equivalent. We'll be talking about representations of the Poincare group, which are characterized by mass m and spin J . The "Poincare" part is the mass, the important feature of which is whether $m=0$ or $m\neq 0$. For fixed m, we then consider irreducible representations of the Lorentz group, which are characterized by spin. For $m>0$ there are $2J+1$ states in the irreducible representation of spin J, and $m=0$ there will be two states in the irreducible representation for any J.

3.1 spin 0

For spin-0, the embedding is easy, we just put the 1 degree of freedom into a scalar ϕ. The Lagrangian is

$$\mathcal{L} = -\frac{1}{2}\phi\Box\phi - \frac{1}{2}m^2\phi \tag{14}$$

which is Lorentz invariant. The equations of motion are

$$(\Box + m^2)\phi = 0 \tag{15}$$

Which have solutions $\phi = e^{ipx}$ with $p^2 = m^2$. So this field has mass m. The Lagrangian is unique up to an overall constant. In Fourier space the Lagrangian (or Hamiltonian) is

$$\mathcal{L} = \frac{1}{2}\phi(p^2 - m^2)\phi = \frac{1}{2}\phi(E^2 - \vec{p}^{\,2} - m^2)\phi \tag{16}$$

Note that the energy term is positive definite since $E>0$. The positivity of the energy term determines the sign of the whole Lagrangian. We normalize to $\frac{1}{2}$ so that the equations of motion are $(\Box + m^2)\phi = 0$.

3.2 massive spin 1

For spin 1, there are three degrees of freedom if $m>0$ and two if $m=0$. This is a mathematical result, which we won't derive formally, but we will see how it works in practice. The smallest tensor field we could possible embed these 3 or 2 degrees of freedom in is a vector field A_μ which has 4-components. The challenge is to make it so that not all of the 4 degrees of freedom in A_μ are active.

Let's start with massive spin 1. A natural guess for the Lagrangian is

$$\mathcal{L} = -\frac{1}{2}A_\mu\Box A_\mu - \frac{1}{2}m^2 A_\mu^2 \tag{17}$$

then the equations of motion are

$$(\Box + m^2)A_\mu = 0 \tag{18}$$

which has four propagating modes. In fact, this Lagrangian is not the Lagrangian for a massive spin 1 field, but the Lagrangian for 4 massive scalar fields, A_0, A_1, A_2 and A_3. The energy terms are

$$\mathcal{L} = \frac{1}{2} A_\mu (E^2 + p^2 - m^2) A_\mu \tag{19}$$

which has a positive sign for the A_0 field and a negative sign for the A_i fields. Thus some of them have negative energy. So this Lagrangian will not produce a physical theory.

We could say that A_μ transforms like a vector or like 4 scalars. The Lagrangian is invariant under both. Why did we get 4 scalars when we wanted a vector? It's because **we don't get to pick how the fields transform, the Lagrangian does.** As a very general statement, we do not get to impose symmetries on a theory. We just pick the Lagrangian, then we let the theory go. If there are symmetries, and the Lagrangian is constructed correctly to preserve them, the symmetries will hold up in matrix elements in the full interacting theory. This is true even if we never figured out that the symmetries were there. For example, Maxwell's equations are Lorentz invariant. They work the same way if you have $\vec{E}$ and $\vec{B}$ instead of A_μ, but the Lorentz invariance is then obscure. But it still works. In fact, a very important tool in making progress in physics has been to observe symmetries in a physical result, such as a matrix element, then to go back and figure out why they were there at a deeper level, which leads to generalizations. That happened with Maxwell, for E&M, with Einstein with special and general relativity, with Glashow, Wienberg and Salam for the weak force, with Georgi for grand unification, and in many many other cases.

Back to massive spin 1. There is one more Lorentz invariant kinetic term we can write down. The most general Lagrangian we can write down for A_μ is

$$\mathcal{L} = \frac{a}{2} A_\mu \Box A_\mu + \frac{b}{2} A_\mu \partial_\mu \partial_\nu A_\nu + \frac{1}{2} m^2 A_\mu^2 \tag{20}$$

where a and b are numbers. As long as b is non-zero, the $\partial_\mu A_\mu$ contraction forces A_μ to transform like a 4-vector. If A_μ transforms like 4 scalars, it is not Lorentz invariant. But there still 4 degrees of freedom, which are three for spin 1 and one for spin 0. We want to get rid of the spin 0. To see the different irreducible sub-fields, note that some of these $A_\mu' s$ can be written as

$$A_\mu = \partial_\mu \phi \tag{21}$$

In this case, $A_\mu' s$ Lorentz transformation properties simply emerge from those of ∂_μ, but the actual degree of freedom is a spin-0 scalar ϕ. So we would like to isolate the *other* degrees of freedom.

The equations of motion are

$$a \Box A_\mu + b \partial_\mu \partial_\nu A_\nu + m^2 A_\mu = 0 \tag{22}$$

Taking ∂_μ of this equation gives

$$\left[(a+b)\Box + m^2 \right] (\partial_\mu A_\mu) \tag{23}$$

If $a = -b$ and $m \neq 0$, this reduces to $\partial_\mu A_\mu = 0$ which removes the scalar. With $\partial_\mu A_\mu = 0$, the Lagrangian reduces to $\mathcal{L} = A_\mu (\frac{a}{2} \Box + \frac{1}{2} m^2) A_\mu \sim -\frac{a}{2} A_\mu E^2 A_\mu + \cdots$. If we pick $a < 0$ we would have one mode A_0 whose energy term is positive, and three A_i with minus signs. Since we want three modes for spin 1, we're going to need the A_i terms to have a positive energy. Normalizing so that the equations of motion are $(\Box + m^2) A_\mu = 0$. We find $a = 1$ and $b = -1$. Then the Lagrangian is

$$\mathcal{L} = \frac{1}{2} A_\mu \Box A_\mu - \frac{1}{2} A_\mu \partial_\mu \partial_\nu A_\nu - \frac{1}{2} m^2 A_\mu^2 \tag{24}$$

$$= -\frac{1}{4} F_{\mu\nu}^2 - \frac{1}{2} m^2 A_\mu^2 \tag{25}$$

where the Maxwell tensor is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. We didn't say anything here about gauge invariance or electromagnetism, we just derived that $F_{\mu\nu}$ appears based on constructing a Lagrangian which propagates a massive spin 1 field.

Now let's find explicit solutions to the equations of motion. In Fourier space, we can write

$$A_\mu(x) = \int \frac{d^4p}{(2\pi)^4}\epsilon_\mu(p)e^{ipx} \tag{26}$$

where ϵ_μ is called a polarization. For any given physical momentum for the field, $p^2 = m^2$, we want to find all the independent solutions $\epsilon_\mu^i(p)$. In Fourier space, the $\partial_\mu A_\mu = 0$ equation of motion gives $p_\mu\epsilon_\mu = 0$. These polarization vectors ϵ_μ are conventionally normalized so that $\varepsilon_\mu^2 = -1$. We can also choose a canonical basis. Take the momentum of the field to be in the z-direction

$$p_\mu = (E, 0, 0, p_z), \quad E^2 - p_z^2 = m^2 \tag{27}$$

then two obvious vectors satisfying $p_\mu\varepsilon_\mu = 0$ and $\varepsilon_\mu^2 = -1$ are

$$\epsilon_\mu^1 = (0, 1, 0, 0), \quad \epsilon_\mu^2 = (0, 0, 1, 0) \tag{28}$$

the other one is

$$\epsilon_\mu^L = (\frac{p_z}{m}, 0, 0, \frac{E}{m}) \tag{29}$$

It's easy to check that $(\epsilon_\mu^L)^2 = -1$ and $p_\mu\epsilon_\mu^L = 0$. These three polarization vectors $\epsilon_\mu^i(p)$ form the irreducible representation. The basis vectors depend on p, and since there are an infinite number of $p's$, it is an infinite dimensional representation.

By the way, massive spin 1 fields are not a purely theoretical concept: they exist! There is one called the ρ-meson, that is lighter than the proton, but unstable, so we don't often see it. More importantly, there are really heavy ones, the W and Z bosons, which mediate the weak force and radioactivity. You will study them in great detail in a standard model class. But there's an important feature of these heavy bosons that is easy to see already. Note that at high energy $E \gg m$, the longitudinal polarization becomes

$$\epsilon_\mu^L \sim \frac{E}{m}(1, 0, 0, 1) \tag{30}$$

So if we scatter these modes, we might have a cross section that goes like $d\sigma \sim \varepsilon^2 \sim g^2\frac{E^2}{m^2}$, where g is the coupling constant. So, no matter how small g is, if we go to high enough energies, this cross section blows up. However, cross sections can't be arbitrarily big. After all, they are probabilities, which are bounded by 1. So at some scale, what we are calculating becomes not a very good representation of what is really going on. In other words, our perturbation theory is breaking down. We can see already that this happens at $E \sim \frac{m}{g}$. If $m \sim 100$ GeV and $g \sim 0.1$, we find $E \sim 1$ TeV. That's why we expect that our theory, with the W or Z bosons of $m_W \sim 100$ GeV, should act funny at the LHC. The truth is, we don't know exactly how to calculate things at that scale. We have some ideas, but can't prove any of them yet, which is why we are building the machine.

Also, the fact that there is a spin-1 particle in this Lagrangian follows completely from the Lagrangian itself. We never have to impose any spin-oney conditions in addition. We didn't even have to talk about spin, or Lorentz invariance at all – all the properties associated with that spin would just have fallen out when we tried to calculate physical quantities. That's the beauty of symmetries: they work even if you don't know about them! It would be fine to think of A_μ as 4 scalar fields that happen to conspire so that when you compute something in one frame certain ones contribute and when you compute in a different frame, other ones contribute, but the final answer is frame independent. Obviously it's a lot easier to do the calculation if we know this ahead of time, so we can choose a nice frame, but in no way is it required.

3.3 massless spin 1

The easiest way to come up with a theory of massless spin 1 is to simply take the $m \to 0$ limit of the massive spin 1 theory. Then the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 \tag{31}$$

which is the Lagrangian for electrodynamics. So this is the right answer.

But we also know that massless spin 1 should have only 2 polarizations. What happened to the third? Well, first of all, note that as $m\to 0$ the longitudinal polarizations blows up:

$$\epsilon_\mu^L = (\frac{p_z}{m}, 0, 0, \frac{E}{m}) \to \infty \tag{32}$$

Partly, this is due to normalization. Note that also the momentum becomes lightlike

$$p_\mu \to (E, 0, 0, E) \tag{33}$$

so $p_z \to E$ and $\epsilon_\mu^L \to p_\mu$ up to the divergent normalization. We'll come back to this. The other problem is that the constraint equation $m^2(\partial_\mu A_\mu) = 0$ is automatically satisfied for $m = 0$, so we no longer automatically have $\partial_\mu A_\mu = 0$. Thus taking the massless limit isn't so easy.

Instead, let's just postulate the Lagrangian and start over with analyzing the degrees of freedom. So we start with

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{34}$$

This has an important property that the massive Lagrangian didn't: gauge invariance. It is invariant under the transformation

$$A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x) \tag{35}$$

For any function $\alpha(x)$. Thus two fields A_μ which differ by the derivative of a scalar are physically equivalent.

The equations of motion following from the Lagrangian are

$$\Box A_\mu - \partial_\mu(\partial_\nu A_\nu) = 0 \tag{36}$$

This is really 4 equations and it is helpful to separate out the 0 and i components:

$$-\partial_j^2 A_0 - \partial_j A_j = 0 \tag{37}$$

$$\Box A_i + \partial_i(\partial_t A_0 - \partial_j A_j) = 0 \tag{38}$$

To count the physical degrees of freedom, let's choose a gauge. Instead of Lorentz gauge, we will use Coulomb gauge $\partial_j A_j = 0$. Then the A_0 equation of motion becomes

$$\partial_j^2 A_0 = 0 \tag{39}$$

which has no time derivative. Now under gauge transformations $\partial_i A_i \to \partial_i A_i + \partial_i^2 \alpha$, so Coulomb gauge is preserved under $A_\mu \to A_\mu + \partial_\mu \alpha$ for any α satisfying $\partial_i^2 \alpha = 0$. Since $A_0 \to A_0 + \partial_t \alpha$ and A_0 also satisfies $\partial_i^2 A_0 = 0$ we have exactly the residual symmetry we need to set $A_0 = 0$. Thus we have eliminated one degree of freedom from A_μ completely, and we are down to three. One more to go!

The other equations are

$$\Box A_i = 0 \tag{40}$$

which seem to propagate 3 modes. But don't forget that A_i is constrained by $\partial_i A_i = 0$. In Fourier space

$$A_\mu(x) = \int \frac{d^4p}{(2\pi)^4} \epsilon_\mu(p) e^{ipx} \tag{41}$$

And we have $p^2 = 0$ (equations of motion), $p_i \epsilon_i$ (gauge choice), and $\epsilon_0 = 0$ (gauge choice). Choosing a frame, we can write the momentum as $p_\mu = (E, 0, 0, E)$. Then these equations have two solutions

$$\epsilon_\mu^1 = (0, 1, 0, 0), \quad \epsilon_\mu^2 = (0, 0, 1, 0) \tag{42}$$

Often people use a different basis

$$\epsilon_\mu^L = (0, 1, i, 0), \quad \epsilon_\mu^R = (0, 1, -i, 0) \tag{43}$$

which are called helicity eigenstates, and correspond to circularly polarized light. But I have always found it easier to use the ϵ^1 and ϵ^2 basis. In any case we have found that the massless photon has two polarizations.

We could also have used Lorentz gauge, in which case we would have found that three vectors satisfy $p_\mu \epsilon_\mu = 0$

$$\epsilon_\mu^1 = (0, 1, 0, 0), \quad \epsilon_\mu^2 = (0, 0, 1, 0), \quad \epsilon_\mu^f = (1, 0, 0, 1) \tag{44}$$

However, the forward polarized photon ϵ_μ^f is not normalizable, so we have to throw it out. There is actually residual gauge symmetry in Lorentz gauge which can remove this polarization too. If $A_\mu = \partial_\mu \phi$ for some ϕ, then we can shift with $\alpha = \phi$ to remove this state. Again, this is the residual scalar degree of freedom. In the same way, if we had not imposed the second Coulomb gauge condition $\epsilon_0 = 0$, we would have found the other polarization satisfying $p_i \varepsilon_i = 0$ is $\epsilon^0 = (1, 0, 0, 0)$. But this also cannot be normalized so that $\epsilon_\mu^i \epsilon_\mu^j = -\delta^{ij}$, since ϵ^0 is timelike.

3.4 summary

To summarize, for massive spin 1, we needed to have the kinetic term $F_{\mu\nu}^2 + m^2 A_\mu^2$ in order to enforce $\partial_\mu A_\mu = 0$ which eliminated one degree of freedom from A_μ, leaving the three for massive spin 1. The Lagrangian for a massive spin 1 particle does not have gauge invariance, but we still need $F_{\mu\nu}^2$.

For the massless case, having $F_{\mu\nu}^2$ gives us gauge invariance. This allows us to remove an additional polarization, leaving two which is the correct number for a massless spin-1 representation of the Poincare group. Sometimes people say that one mode is eliminated because A_0 doesn't propagate, and the other mode is eliminated by a gauge transformation. Some people say the two modes are both eliminated by gauge invariance. I don't know if these words make a difference to anything, or even what exactly they mean. The only concrete statement is that the gauge invariant Lagrangian for a vector field propagates two massless spin 1 degrees of freedom.

For both massive and massless spin 1, we found a basis of polarization vectors $\epsilon_\mu^i(p)$, with $i = 1, 2, 3$ for $m > 0$ and $i = 1, 2$ for $m = 0$. The fact that the polarizations depend on p^μ make these infinite dimensional representations. The representation of the full Poincare group is **induced** by a representation of the subgroup of the Poincare group which holds p^μ fixed, called the **little group**. The little group has finite dimensional representations. For the massive case, the little group, holding for example, $p^\mu = (m, 0, 0, 0)$ fixed (or any other 4-vector of mass m) is just the group of 3D rotations, SO(3). SO(3) has finite-dimensional irreducible representations of spin J with $2J + 1$ degrees of freedom. For the massless case, the group which holds a massless 4-vector like $(E, 0, 0, E)$ fixed is the group ISO(2) which has representations of spin J with 2 degrees of freedom for each J.

4 Covariant derivatives

In order not to screw up our counting of degrees of freedom, the interactions in the Lagrangian should respect gauge invariance as well. For example, you might try to couple

$$\mathcal{L} = \cdots + A_\mu \phi \partial_\mu \phi \tag{45}$$

But this is not invariant.

$$A_\mu \phi \partial_\mu \phi \rightarrow A_\mu \phi \partial_\mu \phi + (\partial_\mu \alpha) \phi \partial_\mu \phi \tag{46}$$

In fact, it's impossible to couple A_μ to any field with only one degree of freedom, like the scalar field ϕ. We must be able to make ϕ transform to compensate for the gauge transformation of A_μ, to cancel the $\partial_\mu \alpha$ term. But if there is only one field ϕ, it has nothing to mix with so it can't transform.

Thus we need at least two $\phi' s$, ϕ_1 and ϕ_2. It's easiest to deal with such a doublet by putting them together into a complex field $\phi = \phi_1 + i\phi_2$, and then to work with ϕ and $\phi^\star$. Under a gauge transformation, ϕ can transform like

$$\phi \rightarrow e^{i\alpha(x)} \phi \tag{47}$$

Which makes

$$m^2 \phi^\star \phi \tag{48}$$

gauge invariant. But what about the derivatives? $|\partial_\mu\phi|^2$ is not invariant.

We can in fact make the kinetic term gauge invariant using covariant derivatives. If a field transforms as

$$\phi \rightarrow e^{i\alpha(x)}\phi \tag{49}$$

Then

$$(\partial_\mu + igA_\mu)\phi \rightarrow (\partial_\mu + igA_\mu - i\partial_\mu\alpha)e^{i\alpha(x)}\phi = e^{i\alpha(x)}(\partial_\mu + igA_\mu)\phi \tag{50}$$

So we define the *covariant* derivative by

$$D_\mu\phi \equiv (\partial_\mu + igA_\mu)\phi \rightarrow e^{i\alpha(x)}D_\mu\phi \tag{51}$$

which transforms just like the field does. Thus

$$\mathcal{L} = (D_\mu\phi)^\star(D_\mu\phi) + m^2\phi^\star\phi \tag{52}$$

is gauge invariant.

5 Conserved Currents

We can expand out the Lagrangian we just found (dropping the mass for simplicity):

$$\mathcal{L} = (D_\mu\phi)^\star(D_\mu\phi) = \partial_\mu\phi^\star\partial_\mu\phi + igA_\mu\phi^\star\partial_\mu\phi - igA_\mu(\partial_\mu\phi^\star)\phi + g^2A_\mu^2\phi^\star\phi \tag{53}$$

$$= \partial_\mu\phi^\star\partial_\mu\phi + A_\mu J_\mu + \mathcal{O}(g^2) \tag{54}$$

$$J_\mu = ig[(\partial_\mu\phi^\star)\phi - \phi^\star(\partial_\mu\phi)] \tag{55}$$

The equations of motion for the free field ϕ are simply $\Box\phi = 0$, which directly implies $\partial_\mu J_\mu = 0$. So by just demanding this symmetry hold for the interactions, we were automatically led to a conserved current. It turns out this follows from a general result.

Let's just write the thing that the gauge field couples to as J_μ:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + A_\mu J_\mu \tag{56}$$

then under a gauge transformation the action transforms as

$$\int d^4x\, A_\mu J_\mu \rightarrow \int d^4x(A_\mu + \partial_\mu\alpha(x))J_\mu \tag{57}$$

$$= \int d^4x[A_\mu J_\mu - \alpha(x)\partial_\mu J_\mu] \tag{58}$$

where we have integrated by parts in the last step. For the action to be invariant for any function $\alpha(x)$ we must have $\partial_\mu J_\mu = 0$, that is the current must be conserved.

We say "conserved" because the total charge

$$Q = \int d^3x\, J_0 \tag{59}$$

satisfies

$$\partial_t Q = \int d^3x \partial_t\, J_0 = \int d^3x \vec{\nabla}\vec{J} \tag{60}$$

which is Gauss's Law. If $\vec{J}$ vanishes at the boundary, meaning that no charge is leaving our experiment, then $\partial_t Q = 0$. Thus the total charge doesn't change with time, and is conserved.

5.1 Noether's theorem

The existence of conserved currents is quite a general consequence of symmetries in a Lagrangian. The direct connection is embodied in Noether's theorem. It's a beautiful theorem, and like many great theorems, is almost trivial to prove.

Say we have a Lagrangian $\mathcal{L}(\phi_i)$ which is invariant when a bunch of fields ϕ_i transform some way under a symmetry parameter α. We don't even need α to be a function of spacetime for this theorem, it can be a number. But it has to be a continuous symmetry (*i.e.* not $\phi \to -\phi$), so that we can take $d\phi_i$ arbitrarily small. Then

$$0 = d\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi_i}\frac{d\phi_i}{d\alpha} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\frac{d(\partial_\mu \phi_i)}{d\alpha} \tag{61}$$

Now if we use the equations of motion (the Euler Lagrange equations)

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} = \frac{\partial \mathcal{L}}{\partial \phi_i} \tag{62}$$

we find

$$0 = \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\right)\frac{d\phi_i}{d\alpha} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\frac{d(\partial_\mu \phi_i)}{d\alpha} \tag{63}$$

$$= \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\frac{d\phi_i}{d\alpha}\right) \tag{64}$$

Thus our conserved current is simply

$$J_\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\frac{d\phi_i}{d\alpha} \tag{65}$$

You can make this more general, by allowing for $\mathcal{L}$ not just to depend on ϕ and $\partial_\mu \phi$ but also on higher derivatives, and also by allowing that the symmetry only leaves the Lagrangian invariant up to a total derivative. But those cases are unenlightening so we'll stop here.

In summary

- Noether's Theorem: If the action has a continuous symmetry, there exists a current associated with that symmetry which is conserved on the equations of motion.

To emphasize

- the symmetry must be continuous
- it works for *global* symmetries, parametrized by numbers α not just *local* (gauge) symmetries parametrized by functions $\alpha(x)$
- the current is conserved *on-shell*, that is, when the equations of motion are satisfied
- the theorem itself has nothing to do with gauge fields. But gauge invariance require continuous symmetries, so we often use the theorem in the context of gauge fields.

5.2 example

The easiest way to understand this therm is with an example. Take a scalar field, with Lagrangian

$$\mathcal{L} = |\partial_\mu \phi|^2 - m^2|\phi|^2 \tag{66}$$

The Lagrangian is invariant under $\phi \to e^{i\alpha}\phi$, whose infinitesimal form is

$$\phi \to \phi + i\alpha\phi, \quad \phi^\star \to \phi^\star - i\alpha\phi^\star \tag{67}$$

That is

$$\frac{d\phi}{d\alpha} = i\phi, \quad \frac{d\phi^\star}{d\alpha} = -i\phi^\star \tag{68}$$

Then,

$$J_\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\frac{d\phi_i}{d\alpha} = i[(\partial_\mu \phi^\star)\phi - \phi^\star \partial_\mu \phi] \tag{69}$$

This is the same thing we saw was conserved before, because it's the same thing that couples to the gauge field when the symmetry is gauged. However, the current is conserved even if there is no gauge field. All we need is the *global* symmetry to have a conserved current.

It is instructive to imagine working backwards. Suppose we first observed the global symmetry and the conserved current. Then we could have written down $A_\mu J_\mu$ by hand to generate an interactions. We would then have seen that the Lagrangian could be interpreted as having as covariant derivatives. This approach is not so important for electromagnetism, but it would have been very handy if we didn't know what the covariant derivative should look like. That's exactly what happens with gravity, and the approach Feynman takes to deriving general relativity in his lectures on gravity (Pegasus, 1995).

5.3 summary

To summarize, we saw that to have a massless unitary representation of spin 1 we needed gauge invariance. This required a conserved current, which in turn required that charge be conserved. To couple the photon to matter, we needed more than one degree of freedom so we were led to ϕ and $\phi^\star$, which are related by symmetry transformations. So there are a lot of implications of demanding unitary representations!

6 Quantization

To quantize fields with multiple degrees of freedom, we simply need creation and annihilation operators for each degree separately. For example, if we have two spin 0 fields, we can write

$$\phi_1(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} (a_{p,1}^\dagger e^{ipx} + a_{p,1} e^{-ipx}) \tag{70}$$

$$\phi_2(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} (a_{p,2}^\dagger e^{ipx} + a_{p,2} e^{-ipx}) \tag{71}$$

Then the complex field is

$$\phi(x) = \phi_1(x) + i\phi_2(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left\{ [a_{p,1}^\dagger + i\,a_{p,2}^\dagger] e^{ipx} + [a_{p,1} - i\,a_{p,2}] a_{p,i} e^{-ipx} \right\} \tag{72}$$

$$= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \sum_{i=1}^{2} \left(\epsilon_i a_{p,i}^\dagger e^{ipx} + \epsilon_i a_{p,i} e^{-ipx} \right) \tag{73}$$

with $\epsilon_1 = 1$ and $\epsilon_2 = i$. In this notation you can think of ϵ as the polarization of the complex scalar field.

To quantize a spin 1 field, we just allow for the polarizations to be in a basis which has 4-components instead of one: $\epsilon \rightarrow \epsilon_\mu$ and can depend on momentum $\epsilon_i^\mu \rightarrow \epsilon_i^\mu(p)$.

6.1 massive spin 1

The quantum field operator for massive spin 1 is

$$A_\mu(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \sum_{i=1}^{3} (\epsilon_\mu^i(p) a_{p,i}^\dagger e^{ipx} + \epsilon_\mu^{i\star}(p) a_{p,i} e^{-ipx}) \tag{74}$$

There are separate creation and annihilation operators for each of the polarizations, and we sum over them. $\epsilon_\mu^i(p)$ represents a canonical set of basis vectors. You can think of A_μ as 4 separate fields which each create some linear combination of the three polarizations. If we are interested in a particular polarization $\hat{\epsilon}_\mu$, then we can project out the field which creates only this polarization by contracting $\hat{\epsilon}_\mu A_\mu$.

We know that the basis has to depend on p^μ because there are no finite dimensional unitary representations of the Lorentz group. To see it again, let's suppose instead that we tried to pick constants for our basis vectors. Say, $\epsilon_\mu^1 = (0, 1, 0, 0)$, $\epsilon_\mu^2 = (0, 0, 1, 0)$ and $\epsilon_\mu^3 = (0, 0, 0, 1)$. The immediate problem is that this basis is not complete, because under Lorentz transformations

$$\epsilon_\mu^i \rightarrow \Lambda_{\mu\nu} \epsilon_\nu^i \tag{75}$$

So that for boosts, these will mix with the timelike polarization (1, 0, 0, 0). We saw from the solving the classical equations of motion that we can choose a momentum dependent basis $\epsilon_\mu^1(p)$, $\epsilon_\mu^2(p)$ and $\epsilon_\mu^3(p)$. For example, for the massive case, for p_μ pointing in the z direction

$$p_\mu = (E, 0, 0, p_z) \tag{76}$$

we can use the basis

$$\epsilon_\mu^1(p) = (0, 1, 0, 0), \quad \epsilon_\mu^2(p) = (0, 0, 1, 0), \quad \epsilon_\mu^L(p) = (\frac{p_z}{m}, 0, 0, \frac{E}{m}) \tag{77}$$

Note, these all satisfy $\epsilon_\mu^2 = -1$ and $\epsilon_\mu p_\mu = 0$.

What happened to the fourth degree of freedom in the vector representation? The vector orthogonal to these is $\epsilon_\mu^S(p) = \frac{1}{m} p_\mu = (\frac{E}{m}, 0, 0, \frac{p}{m})$. In position space, this is $\epsilon_\mu^S = \partial_\mu \alpha(x)$ for some function $\alpha(x)$, which transforms as a scalar. So we don't want to include this guy in the sum. To see that the polarization based on the scalar doesn't mix with the other three is easy: if something is the divergence of a function $\alpha(x)$, under a Lorentz transformation it will still be the divergence of the same function, just in a different frame. So the polarization vectors in the spin-1 representation (the $\epsilon_\mu^i s$) don't mix with the vector in the spin-0 representation ϵ_μ^S.

Now, you may wonder, if we are redefining our basis with every boost, so that $\epsilon_\mu^1(p) \to \epsilon_\mu^1(p')$, $\epsilon_\mu^2(p) \to \epsilon_\mu^2(p')$, and $\epsilon_\mu^L(p) \to \epsilon_\mu^L(p')$, when do the polarization vectors ever mix? Have we gone too far and just made four separate one-dimensional representations? The answer is that there are Lorentz transformations which leave p_μ alone, and therefore leave our basis alone. For those Lorentz transformation, we need to check that our basis vectors must rotate into each other and form a complete representation. For example, suppose we go to the frame

$$q_\mu = (m, 0, 0, 0) \tag{78}$$

Then we can choose our polarization basis vectors as

$$\epsilon_\mu^1(q) = (0, 1, 0, 0), \quad \epsilon_\mu^2(q) = (0, 0, 1, 0), \quad \epsilon_\mu^3(q) = (0, 0, 0, 1) \tag{79}$$

and $\epsilon_\mu^S = (1, 0, 0, 0)$. Now, the group which preserves q_μ is simply the 3D rotation group SO(3). It is then easy to see that under 3D rotations, the 3 ϵ_μ^i polarizations will mix among each, and $\epsilon_\mu^S = (1, 0, 0, 0)$ stays fixed. If we boost it looks like the ϵ_μ^i will mix with ϵ_μ^S. However, we have to be careful, because the basis vectors will also change, for example to ϵ_μ^1, ϵ_μ^2 and ϵ_μ^L above. The group that fixes $p_\mu = (E, 0, 0, p_z)$ is also SO(3), although it is harder to see. And these SO(3) rotations will also only mix ϵ_μ^1, ϵ_μ^2 and ϵ_μ^L leaving ϵ_μ^S fixed. So everything works. The non-trivial effect of Lorentz transformations to mix up the polarization vectors at fixed p_μ. So the spin 1 representation is characterized by this smaller group, the Little group, which is the subgroup of the Lorentz group which stabilizes p_μ. This method of studying representations of the Lorentz group is called the method of induced representations.

Also note that this way of doing things makes it clear that we can have a positive definite probability density. In quantum mechanics, we can expand any polarization in this basis

$$|A_\mu(q)\rangle = c_i |\epsilon^i(q)\rangle \tag{80}$$

Then since the states all have $\langle \epsilon^i | \epsilon^i \rangle = 1$, we find

$$\langle A | A \rangle = |c_1|^2 + |c_2|^2 + |c_3|^2 \tag{81}$$

is positive definite and Lorentz invariant. For rotations, this is obvious, and for boosts it is trivial – the c_i's don't change because the ϵ_μ^i do.

In quantum field theory, we will be calculating matrix elements with the field A_μ. These matrix elements will depend on the polarization vector and must have the form

$$\mathcal{M} = \epsilon_\mu M_\mu \tag{82}$$

where M_μ transforms like a 4-vector. Here, ϵ_μ is the polarization vector, which can be any of the ϵ_μ^i or any linear combination of them. For example, say we start with ϵ_μ^1. Now change frames, so $M_\mu \to M'_\mu = \Lambda_{\mu\nu} M_\nu$. Then the matrix element is invariant

$$\mathcal{M} = \epsilon'_\mu M'_\mu \tag{83}$$

where $\epsilon'_\mu = \Lambda_{\mu\nu}\epsilon_\mu$. This new polarization ϵ'_μ is still a physical state in our Hilbert space, since the basis is closed under the Lorentz group. More simply, we can say that $\epsilon_\mu M_\mu$ is Lorentz invariant on the restricted space of $\epsilon_\mu = c_i \epsilon^i_\mu$.

6.2 massless spin 1

We quantize massless spin one exactly like massive spin 1, but summing over 2 polarizations instead of 3.

$$A_\mu(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \sum_{i=1}^{2} (\epsilon^i_\mu(p) a^\dagger_{p,i} e^{ipx} + \epsilon^i_\mu(p) a_{p,i} e^{-ipx}) \tag{84}$$

A sample basis is, for p_μ in the z direction:

$$p_\mu = (E, 0, 0, E) \tag{85}$$

$$\epsilon^1_\mu(p) = (0, 1, 0, 0), \quad \epsilon^2_\mu(p) = (0, 0, 1, 0) \tag{86}$$

these satisfy $\epsilon^2_\mu = 1$ and $\epsilon_\mu p_\mu = 0$. The bad polarizations are

$$\epsilon^f_\mu(p) = (1, 0, 0, 1), \quad \epsilon^b_\mu(p) = (1, 0, 0, -1) \tag{87}$$

where f and b stand for forward and backward.

But now we have a problem. Even though there is an irreducible unitary representation of massless spin 1 particles involving 2 polarizations, it is *impossible* to embed these polarizations in vector fields like ϵ_μ. To see the problem, recall that in the massive case, ϵ^1_μ and ϵ^2_μ mixed not only with each other under the little group SO(3), that is, Lorentz transformations preserving $p_\mu = (E, 0, 0, p_z)$, but they also mixed with the longitudinal mode $\epsilon^L_\mu(p) = (\frac{p_z}{m}, 0, 0, \frac{E}{m})$. We saw this because $|c_1|^2 + |c_2|^2 + |c_L|^2$ is invariant under this SO(3), but $|c_1|^2 + |c_2|^2$ itself would not be. There is nothing particularly discontinuous about the $m \to 0$ limit. The momentum goes to $p_\mu = (E, 0, 0, E)$ and the longitudinal mode becomes the same as our forward-polarized photon, up to normalization

$$\lim_{m\to 0} \epsilon^L_\mu(p) = \epsilon^f_\mu(p) = p_\mu \tag{88}$$

The little group goes to something called ISO(2). There are still little group members which mix ϵ^1_μ and ϵ^2_μ with the other polarization, $\epsilon^f_\mu(p) = p_\mu$. In general

$$\epsilon^1_\mu(p) \to c_{11}(\Lambda)\epsilon^1_\mu(p) + c_{12}(\Lambda)\epsilon^2_\mu(p) + c_{13}(\Lambda) p_\mu \tag{89}$$

$$\epsilon^2_\mu(p) \to c_{21}(\Lambda)\epsilon^1_\mu(p) + c_{22}(\Lambda)\epsilon^2_\mu(p) + c_{13}(\Lambda) p_\mu \tag{90}$$

where the $c_{ij}s$ are numbers.

To be really explicit, consider the Lorentz transformation

$$\Lambda = \begin{pmatrix} \frac{3}{2} & 1 & 0 & -\frac{1}{2} \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & 1 & 0 & \frac{1}{2} \end{pmatrix} \tag{91}$$

This satisfies $\Lambda^T \eta \Lambda = \eta$, so it is a Lorentz transformation. It also has $\Lambda_{\mu\nu} p_\nu = p_\mu$ so it preserves the momentum $p_\mu = (E, 0, 0, E)$ and is a member of the little group. However,

$$\Lambda_{\mu\nu}\epsilon^1_\nu = (1, 1, 0, 1) = \epsilon^1_\mu + \frac{1}{E} p_\mu \tag{92}$$

So it mixes the physical polarization with the momentum. This doesn't happen for the massive spin 1, since the basis vectors ϵ^i_μ in that case only mix with themselves.

Now, consider the kind of matrix element we would get from something involving the field A_μ. It would, just like the massive case, look like

$$\mathcal{M} = \epsilon_\mu M_\mu \tag{93}$$

Where now ϵ_μ is some linear combination of the two physical polarizations ϵ_μ^1 and ϵ_μ^2. Then under a Lorentz transformation

$$\mathcal{M} \to \epsilon_\mu' M_\mu' + c(\Lambda) p_\mu M_\mu' \tag{94}$$

where $M_\mu' = \Lambda_{\mu\nu} M_\nu$ and ϵ_μ' is a linear combination of ϵ_μ^1 and ϵ_μ^2, but p_μ is not. For example, under the explicit Lorentz transformation above,

$$\mathcal{M} = \epsilon_\mu^1 M_\mu \to (\epsilon_\mu^1 + \frac{1}{E} p_\mu) M_\mu' \tag{95}$$

So we have a problem. The state with polarization $\epsilon_\mu^1 + \frac{1}{E} p_\mu$ is not in our Hilbert space! Thus there is no physical polarization for which the matrix element is the same in the new frame as it was in the old frame. There is only one way out − if $p_\mu M_\mu = 0$. Then there is a physical polarization which gives the same matrix element and $\mathcal{M}$ is invariant. Thus, to have a Lorentz invariant theory with a massless spin 1 particle, we must have $p_\mu M_\mu = 0$.

This is **extremely important**, so I'll say it again. We have found that under Lorentz transformations the massless polarizations transform as

$$\epsilon_\mu \to c_1 \epsilon_\mu^1 + c_2 \epsilon_\mu^2 + c_3 p_\mu \tag{96}$$

Generally, this transformed polarization is not physical, and not in our Hilbert because of the p_μ term. The best we can do is transform it into $\epsilon_\mu' = c_1 \epsilon_\mu^1 + c_2 \epsilon_\mu^2$. When we calculate something in QED we will get matrix elements

$$\mathcal{M} = \epsilon_\mu M_\mu \tag{97}$$

for some M_μ transforming like a Lorentz vector. If we Lorentz transform this expression we would get

$$\mathcal{M} \to (c_1 \epsilon_\mu^1 + c_2 \epsilon_\mu^2 + p_\mu) M_\mu' \tag{98}$$

It is therefore only possible for $\mathcal{M}$ to be Lorentz invariant if $\mathcal{M} = \epsilon_\mu' M_\mu'$ which happens only if

$$\boxed{p_\mu M_\mu = 0} \tag{99}$$

This is known as a **Ward Identity**. It must hold by LORENTZ INVARIANCE and the fact that UNITARY representations for MASSLESS PHOTONS have TWO POLARIZATIONS.

We did *not* show that it holds, only that it *must* hold in a reasonable physical theory. That it holds in QED is complicated to show in perturbation theory, but you will do a couple of special cases on a problem set. We will eventually prove it non-perturbatively using path integrals. The Ward identity is closely related to gauge invariance. Since the Lagrangian is invariant under $A_\mu \to A_\mu + \partial_\mu \alpha$, in momentum space, this is $A_\mu \to A_\mu + p_\mu$. Since $A_\mu = \epsilon_\mu e^{ipx}$, this should directly imply that $\epsilon_\mu \to \epsilon_\mu + p_\mu$ is a symmetry of the theory, which is the Ward identity.

7 Photon propagator

In order to calculate anything with a photon, we are going to need to know its propagator

$$\Pi_A(x, y) = \langle 0 | T\{A_\mu(x) A_\nu(y)\} | 0 \rangle \tag{100}$$

The easiest way to derive it is to just look for the classical Green's function. Then turning it into a time ordered product is exactly the same as for a scalar.

Let us first try to calculate the photon propagator by using the equations of motion, without choosing a gauge. In the presence of a current, the equations of motion are

$$\partial_\mu F_{\mu\nu} = J_\nu \tag{101}$$

$$\Rightarrow \partial_\mu \partial_\mu A_\nu - \partial_\mu \partial_\nu A_\mu = J_\nu \tag{102}$$

$$\Rightarrow (p^2 \eta_{\mu\nu} - p_\nu p_\mu) A_\mu = J_\nu \tag{103}$$

We would like to write $A_\mu = \Pi_{\mu\nu} J_\nu$, so that $(p^2\eta_{\mu\nu} - p_\mu p_\nu)\Pi_{\nu\alpha} = \delta_{\mu\alpha}$. That is, we want to invert the kinetic term. The problem is that

$$\det(p^2\eta_{\mu\nu} - p_\alpha p_\mu) = 0 \tag{104}$$

We can see this because $p^2\eta_{\mu\nu} - p_\alpha p_\mu$ has a zero eigenvalue, with eigenvector p_μ. This simply represents the gauge invariance: A_μ is not uniquely determined by J_μ; different gauges will give different values for A_μ.

So what do we do? We could try to just choose a gauge, for example $\partial_\mu A_\mu = 0$. This would reduce the Lagrangian to

$$-\frac{1}{4}F_{\mu\nu} \to \frac{1}{2}A_\mu \Box A_\mu \tag{105}$$

However, now it looks like there are four propagating degrees of freedom in A_μ instead of two. In fact, you can do this, but you have to keep track of the gauge constraint $\partial_\mu A_\mu = 0$ all along. This is known as a second-class constraint, and it's a real pain to work with constrained systems – we would need more information than what is contained in the Lagrangian.

The solution is to add a new auxiliary (non-propagating) field which acts like a Lagrange multiplier to enforce the constraint through the equations of motion

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{\xi}(\partial_\mu A_\mu)^2 + J_\mu A_\mu \tag{106}$$

(Writing $\frac{1}{\xi}$ instead of ξ is just a convention). The equations of motion of ξ are just $\partial_\mu A_\mu = 0$ which was the Lorentz gauge constraint. In fact, we could just treat ξ as a number, and for very small ξ there is a tremendous pressure on the Lagrangian to have $\partial_\mu A_\mu = 0$ to stay near the minimum.

Now, the equations of motion for A_μ are

$$\left[p^2\eta_{\mu\nu} - (1 - \frac{1}{\xi})p_\mu p_\nu\right]A_\nu = J_\mu \tag{107}$$

Which can now be inverted to

$$\Pi_{\mu\nu}^{\text{classical}} = \frac{\eta_{\mu\nu} - (1-\xi)\frac{p_\mu p_\nu}{p^2}}{p^2} \tag{108}$$

(classical refers to this being the classical Green's function, as opposed to the Feynman propagator, which is the time-ordered two-point function). We can check:

$$(p^2\eta_{\mu\alpha} - (1 - \frac{1}{\xi})p_\mu p_\alpha)\Pi_{\alpha\nu}^{\text{classical}} = \left[(p^2\eta_{\mu\alpha} - (1-\frac{1}{\xi})p_\mu p_\alpha)\right]\left[p^2\eta_{\alpha\nu} - (1-\xi)p_\alpha p_\nu\right]\frac{1}{p^4} \tag{109}$$

$$= \eta_{\mu\nu} + \left[-(1-\frac{1}{\xi}) - (1-\xi) + (1-\frac{1}{\xi})(1-\xi)\right]\frac{p_\mu p_\nu}{p^2} \tag{110}$$

$$= \eta_{\mu\nu} \tag{111}$$

To work out the Feynman propagator, the steps are just like the scalar field, and the answer is

$$\langle 0|T\{A_\mu(x)A_\nu(y)\}|0\rangle = \int \frac{d^4p}{(2\pi)^4}\Pi_{\mu\nu} \tag{112}$$

$$\Pi_{\mu\nu} = \frac{-i}{p^2 + i\varepsilon}\left[\eta_{\mu\nu} - (1-\xi)\frac{p_\mu p_\nu}{p^2}\right] \tag{113}$$

Where the $i\varepsilon$ for time ordering, the i comes from the contour integral going from d^3p in A_μ to d^4p and the -1 indicates that it is the spacial A_i components which propagate like scalars, since

$$-i\eta_{\mu\nu} = \begin{pmatrix} -i & & & \\ & i & & \\ & & i & \\ & & & i \end{pmatrix} \tag{114}$$

and the scalar propagator was $\Pi_S = \frac{i}{p^2 + i\varepsilon}$.

7.1 gauges

For any value of ξ we get a different Lorentz-invariant gauge. Some useful gauges are:

- Feynman gauge $\xi = 1$

$$\Pi_{\mu\nu} = -i\frac{\eta_{\mu\nu}}{p^2 + i\varepsilon} \tag{115}$$

  This is the gauge we will use for most calculations

- Lorentz Gauge $\xi \to 0$

$$\Pi_{\mu\nu} = -i\frac{\eta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}}{p^2 + i\varepsilon} \tag{116}$$

  We saw that $\xi \to 0$ forces $\partial_\mu A_\mu = 0$. Note that we couldn't set $\xi = 0$ and then invert the kinetic term, but we can invert and then set $\xi = 0$. This gauge isn't the most useful for calculations.

- Unitary gauge $\xi = \infty$. This gauge is useless for QED, since the propagator blows up. But it is extremely useful for the gauge theory of the weak interactions.

7.2 gauge invariance

The final answer for any Lorentz invariant quantity had better be gauge invariant (independent of ξ). Since

$$\Pi_{\mu\nu} = \frac{-i}{p^2 + i\varepsilon}\left[\eta_{\mu\nu} - (1 - \xi)\frac{p_\mu p_\nu}{p^2}\right], \tag{117}$$

gauge invariance implies that whatever we contract $\Pi_{\mu\nu}$ with had better give 0 if $\Pi_{\mu\nu} = p_\mu p_\nu$. This is very similar to the requirement of the Ward identities, which say that the matrix elements vanish if the physical external polarization is replaced by $\epsilon_\mu \to p_\mu$. It turns out the proof of the two requirements are very similar. In fact, if we can prove gauge invariance (ξ-independence) it will imply the Ward identity, since if the p_μ part of the propagator always gives zero it will also give zero on an external state.

Now, I don't know of a general diagrammatic proof of gauge invariance (again, I just mean ξ-independence of Feynman diagrams by this – it's easy to prove that a Lagrangian is gauge invariant). We will do the proof for the special case of scalar QED (where the electron has no spin) and real QED (where the electron is a spinor). But if you have all kinds of crazy interactions like $(D_\mu\phi)^{10}$, I think you would have to do the proof all over again. That's a little unsatisfying. However, there is a way to show in general that if we start with 2 polarizations for the photon, we will never produce that unwanted third one (or the fourth), which was the whole point of gauge invariance to begin with. That proof uses something called BRST invariance, which we will introduce for QED once we get to path integrals.

8 Why gauge invariance?

The symmetry $A_\mu \to A_\mu + \partial_\mu\alpha(x)$ is known as a gauge symmetry, because it depends on a function of spacetime $\alpha(x)$. (The word *gauge* means size, and the original symmetry of this type was conceived by Hermann Weyl as an invariance under scale transformations, now known as Weyl invariance). Another word for *gauge* symmetry is *local* symmetry. The alternative is global symmetry, for which α is just a number. Any theory which has a gauge symmetry also has a global symmetry, since we can just take $\alpha(x) = \text{constant}$. And we saw that the global symmetry is all that's necessary to have a conserved current.

Note that under a global symmetry transformation, the gauge field A_μ does not change. So if we set $A_\mu = 0$, the rest of the Lagrangian will still have a global symmetry, since A_μ doesn't transform. Moreover terms like

$$m^2 A_\mu^2 \tag{118}$$

are invariant under the global symmetry, even though they violate gauge invariance.

Also note that even if we choose a gauge, such as $\partial_\mu A_\mu = 0$, the global symmetry is still present, but the gauge symmetry is broken (the gauge fixing $\partial_\mu A_\mu = 0$ is not invariant). But we know that we can always choose a gauge, and the physics is going to be exactly the same. So what does this mean? What's the point of having a local symmetry if we can just choose a gauge and the physics is the same? In fact, we *have to* choose a gauge to do any computations!

There are two answers to this question. First, it is fair to say that gauge symmetries are a total fake. They are just redundancies of description and really do have no physically observable consequences. In contrast, global symmetries are real features of nature with observable consequences. For example, global symmetries imply the existence of conserved charges, which we can test. So the first answer is that we technically don't need gauge symmetries at all.

The second answer is that local symmetries make it much easier to do computations. You might wonder why we even bother introducing this field A_μ which has this huge redundancy to it. Instead, why not just quantize the electric and magnetic fields, that is $F_{\mu\nu}$, itself? Well you could do that, but it turns out to be more of a pain than using A_μ. To see that, first note that $F_{\mu\nu}$ as a field does not propagate with the Lagrangian $\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2$. All the dynamics will be moved to the interactions. Moreover, if we include interactions, either with a simple current $A_\mu J_\mu$ or with a scalar field $\phi^\star A_\mu \partial_\mu \phi$ or with a fermion $\bar{\psi}\gamma_\mu A_\mu \psi$, we see that they naturally involve A_μ. If we want to write these in terms of $F_{\mu\nu}$ we have to solve for A_μ in terms of $F_{\mu\nu}$ and we will get some crazy non-local thing like $A_\nu = \frac{\partial_\nu}{\Box} F_{\mu\nu}$. Then we'd have to spend all our time showing that the theory is actually local and causal. It turns out to be much easier to deal with a little redundancy so that we don't have to check locality all the time.

Another reason is that all of the physics of the electromagnetic field is, in fact, not *entirely* contained in $F_{\mu\nu}$. There are global properties of A_μ that are not contained in $F_{\mu\nu}$ but have physical consequences. An example is the Aharanov-Bohm effect, that you might remember from quantum mechanics. Thus we are going to accept that using the field A_μ instead of $F_{\mu\nu}$ is a necessary complication.

So there's no physics in gauge invariance but it makes it a lot easier to do field theory. The physical content is what we saw in the previous section with the Lorentz transformation properties of spin 1 fields

Massless spin 1 fields have two polarizations

One implication of this is the Ward identity, which is related to gauge invariance. But I'm sorry to tell you gauge invariance doesn't seem to be necessary. Here are two examples to illustrate why gauge invariance isn't as powerful as it seems.

First, we saw that we could calculate the photon propagator from the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{\xi}(\partial_\mu A_\mu)^2 \tag{119}$$

This Lagrangian breaks gauge invariance, but the photon still has only 2 polarizations.

Second, consider the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + m^2(A_\mu + \partial_\mu \pi) \tag{120}$$

This has a gauge invariance under which $A_\mu(x) \to A_\mu(x) + \partial_\mu \alpha(x)$ and $\pi(x) \to \pi(x) - \alpha(x)$. However, I can use that symmetry to set $\pi = 0$ everywhere. Then the Lagrangian reduces to that of a massive gauge boson. So the physics is that of 3 polarizations of a spin 1 field, but there is an exact gauge invariance in the Lagrangian. I could do something even more crazy with this Lagrangian: integrate out π. By setting π equal to it's equations of motion and substituting back into the Lagrangian, it becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + F_{\mu\nu}\frac{m^2}{\Box}F_{\mu\nu} \tag{121}$$

This Lagrangian is also manifestly gauge invariant, but it is non-local. It looks like a massless spin 1 field, but we know that it's exactly the same as the massive spin-1 Lagrangian. It is not hard to see that any Lagrangian involving A_μ can be made gauge invariant by adding π in this way without changing the physical content of the theory. So gauge invariance it's tricky business indeed.

Now that I've convinced you that gauge invariance is misleading, let me unconvince you. It turns out that the gauge fixing in Eq. (119) is very special. We can only get away with it because the gauge symmetry of QED we are using is particularly simple (technically, because it is Abelian). When we do path integrals, and study non-Abelian gauge invariance, you will see that you can't just drop a gauge fixing term in the Lagrangian, but have to add it in a very controlled way. Doing it properly leaves a residual symmetry in the quantum theory called BRST invariance. In the classical theory you are fine, but in the quantum theory you have to be careful, or you violate unitarity (probabilities don't add up to 1). Ok, but then you can argue that it is not gauge invariance that's fundamental but BRST invariance. However, it turns out that sometimes you need to break BRST invariance. However, this has to be controlled, leading to something called the Bitalin-Vilkovisky formalism. Then you get Slavnov-Taylor identities, which are a type of generalized Ward identities. So it keeps getting more and more complicated, and I don't think anybody really understands what's going on.

The bottom line, and the point of all this, is that gauge invariance (and BRST invariance), is a feature of the way we compute things in quantum field theory. It is not a feature of our universe. The observable feature of our universe is that *massless spin 1 fields have two physical polarizations*.