Chapter 13

Local Symmetry

So far, we have discussed symmetries of the quantum mechanical states. A state is a global (non-local) object describing an amplitude everywhere in space.

In relativistic physics, global objects are awkward because the finite velocity with which effects can propagate is expressed naturally in terms of local objects. For this reason high energy physics is expressed in terms of a field theory.

Consider the proton field $\psi_p(\vec{x},t)$, which is defined throughout space-time. We also have a neutron field $\psi_n(\vec{x},t)$. We know that physics is (sort of) invariant under an isospin rotation

$$\Psi(\vec{x},t) := \begin{pmatrix}\psi_p(\vec{x},t)\\ \psi_n(\vec{x},t)\end{pmatrix} \to e^{-i\vec{\omega}\cdot\vec{\tau}/2}\begin{pmatrix}\psi_p(\vec{x},t)\\ \psi_n(\vec{x},t)\end{pmatrix}$$

where the same rotation is applied to the field at each point in space-time. This is called a **global symmetry** transformation.

But if physics is local, shouldn't it be possible to choose our symmetry transformations differently at different points? Suppose we consider a transformation

$$\Psi'(\vec{x},t) = e^{-i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\Psi(\vec{x},t)$$

where the isorotation depends on position.

To consider whether this can be an invariance, it is good to consider the kind of terms which enter the Lagrangian or Hamiltonian.

The Hamiltonian for a noninteracting fermion[1] is

$$H = \int d^3x\,\mathcal{H}_1, \quad \text{with} \quad \mathcal{H}_1 = \sum_j -i\psi_j^\dagger\vec{\alpha}\cdot\vec{\nabla}\psi_j + m\psi_j^\dagger\beta\psi_j$$

where $j = p$ or n. In doublet form,

$$\mathcal{H}_1 = -i\Psi^\dagger\vec{\alpha}\cdot\vec{\nabla}\Psi + m\Psi^\dagger\beta\Psi.$$

Under the symmetry transformation,

$$\Psi'^\dagger(\vec{x},t) = \Psi^\dagger(\vec{x},t)e^{+i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2},$$

so[2]

$$m\Psi^\dagger\beta\Psi \to m\Psi^\dagger e^{+i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\beta e^{-i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\Psi = m\Psi^\dagger\beta\Psi$$

is unchanged, but

$$\begin{aligned}-i\Psi^\dagger\alpha_j\frac{\partial}{\partial x_j}\Psi \quad\to\quad & -i\Psi^\dagger e^{+i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\alpha_j\frac{\partial}{\partial x_j}e^{-i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\Psi\\ & = -i\Psi^\dagger e^{+i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\alpha_j\left[\frac{\partial}{\partial x_j}\left(e^{-i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\right)\right]\Psi\\ & \qquad -i\Psi^\dagger\alpha_j\frac{\partial}{\partial x_j}\Psi.\end{aligned}$$

Thus $\mathcal{H}_1$ is not unchanged — in fact the variation under a rotation $\vec{\omega}(\vec{x},t)$ adds a piece

$$\delta\mathcal{H}_1 = -i\Psi^\dagger\alpha_j\left(e^{+i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\frac{\partial}{\partial x_j}e^{-i\vec{\omega}(\vec{x},t)\cdot\vec{\tau}/2}\right)\Psi.$$

---

[1] The Dirac equation is first order in $\vec{\nabla}\psi$, unlike the quadratic $\left(\vec{\nabla}\phi\right)^2$ we had for the string or would have for scalar fields. But the important point is just that $\psi^\dagger\nabla\psi$ is involved, and symmetry would require $\nabla\psi$ to transform the same way ψ does. This is true for global transformations $\psi \to e^{i\vec{\omega}\cdot\vec{\tau}/2}\psi$ for constant $\vec{\omega}$, but not if $\vec{\omega}$ is a function of $\vec{r}$. Here α and β are constant matrices and do not disrupt the symmetry.

[2] Note that Ψ has both a two-dimensional isospin index and a four-dimensional spinor index. The τ is the two-dimensional isospin representation, acting only on the first index, and the $\beta = \gamma^0$ is a 4×4 matrix acting only on the spinor index, so τ and β commute.

This is a feature of any local symmetry. The momentum must appear in the Hamiltonian, and the momentum operator is a derivative. That is, the momentum operator compares the fields at different points. But how can we attach meaning to the difference of two values if we are allowed to rotate at each point independently? In order to subtract two vectors, we need to be able to express them in common coordinate systems.

The solution is to introduce a new field which specifies how the coordinates change as you move them in space. This is called the connection or gauge field. We need a field for each direction of motion, $\mathcal{A}_i$, and one for time too (that we need $\mathcal{A}_0$ would be more obvious if we talked about a Lagrangian formulation). These are elements of the Lie algebra, so let us generalize from SU(2) to some arbitrary Lie algebra. The gauge transformation then becomes

$$\Psi'(\vec{x},t) = e^{-i\omega(\vec{x},t)}\Psi(\vec{x},t),$$

where ω is the representation (suitable for Ψ) of an arbitrary Lie algebra element.

We need to introduce a term into the Hamiltonian whose variation will cancel what we got from $\vec{\nabla}\Psi$.

Try adding an interaction $\mathcal{H}_2 - -e\Psi^\dagger\vec{\alpha}\cdot\vec{\mathcal{A}}(\vec{x},t)\Psi$,, where each vector component of $\vec{\mathcal{A}}$ is a representative of a Lie algebra element, a 2×2 matrix in the case of isospin on an isodoublet. Then

$$\delta\mathcal{H}_2 = -e\Psi^\dagger\left(e^{i\omega(\vec{x},t)}\vec{\alpha}\cdot\vec{\mathcal{A}}'(\vec{x},t)e^{-i\omega(\vec{x},t)} - \vec{\alpha}\cdot\vec{\mathcal{A}}(\vec{x},t)\right)\Psi,$$

which will cancel $\delta\mathcal{H}_1$ if

$$e^{i\omega(\vec{x},t)}\mathcal{A}'_j(\vec{x},t)e^{-i\omega(\vec{x},t)} = \mathcal{A}_j - \frac{i}{e}e^{i\omega(\vec{x},t)}\frac{\partial}{\partial x_j}e^{-i\omega(\vec{x},t)},$$

or

$$\mathcal{A}'_j = e^{-i\omega(\vec{x},t)}\mathcal{A}_j e^{i\omega(\vec{x},t)} - \frac{i}{e}\left(\frac{\partial}{\partial x_j}e^{-i\omega(\vec{x},t)}\right)e^{i\omega(\vec{x},t)}.$$

This is actually familiar in another application. Instead of SU(2), consider ψ_1 and ψ_2 as the real and imaginary parts of a wave function, and $\omega\to\theta\tau_2$. Then the group transformations are $U(1)$, corresponding to a local change in phase of the field Ψ. As the group is Abelian, we have $\mathcal{A}'_j = \mathcal{A}_j - \frac{1}{e}\frac{\partial}{\partial x_j}\omega$, which is just the gauge transformation of electromagnetism.

It is important to note that, as a price for requiring local invariance under SU(2) rotations for the Ψ, we were forced to introduce a new dynamical field $\vec{\mathcal{A}}$. So this is not analogous to what we did earlier, in using a symmetry of the states to extract information. Here we have actually expanded the theory in a very elegant fashion to obtain a locally gauge invariant theory. This involved replacing the derivative $\vec{\nabla}$ by $\vec{\nabla} - ig\vec{\mathcal{A}}$, the **covariant derivative**.

Our new field $\vec{\mathcal{A}}$ needs to have dynamics of its own. So it needs a kinetic energy term. It needs to be of a special form, for $\sum_j\nabla^2\mathcal{A}_j$ would give additional pieces when we performed a symmetry transformation, which would ruin the invariance under this transformation. For electromagnetism, we know the answer is given in terms of $F_{\mu\nu}$, the field strengths.

This reminds us that we are interested in relativistic physics, (that is why we felt compelled to consider *local* invariance) and thus far we have done things nonrelativistically. Relativistic formulations of field theory generally start with the Lagrangian rather than the Hamiltonian. For a fermion field[3]

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi.$$

This $\mathcal{L}$ plays its role as part of the action

$$S = \int d^3x\,dt\,\mathcal{L}\left(\psi(\vec{x},t)\right),$$

and the dynamics takes place in spacetime, not just space. So a local gauge transformation (*i.e.* a local symmetry transformation) requires invariance under

$$\psi(\vec{x},t)\to\psi'(\vec{x},t) = e^{-i\omega(\vec{x},t)}\psi(\vec{x},t),$$

where the ω parameters of the group transformation are varying in space and in time as well. Invariance requires an A^0 field as well, now, because we have a $\partial_0\propto\frac{\partial}{\partial t}$ in the lagrangian.

There is much interest these days in what is called "Lattice gauge field theory". This involves approximating the continuous ψ fields by fields defined only on the sites of a (usually hypercubic) lattice. This is what we did nonrelativistically in one dimension with our string. The chief advantage of this approach is that it constitutes a way of defining the regulated field theory independent of perturbation theory. Field theories expressed perturbatively have **ultraviolet divergences** which must be cut off in some fashion, which come from interactions between field at points which approach each other in space-time. In perturbative theory one defines renormalized quantities which

[3]For a free massive scalar field $\mathcal{L} = \frac{1}{2}\left(\partial_\mu\phi\right)^\dagger\partial^\mu\phi - \frac{1}{2}m^2\phi^\dagger\phi$.

remain finite as the cutoff is removed (*i.e.* $p_{\text{cutoff}} \to \infty$). In the same way, but without recourse to perturbation theory, a lattice can be used to impose the cutoff (one automatically has $\Delta x \geq a$, the lattice spacing), and one can then talk about quantities which remain finite as $a \to 0$.

This is more than an analytic tool. Computer simulation has become an important approach in investigating QCD, and of course, a continuous field must be approximated by a grid of values on the computer.

Finally, the lattice has an advantage in discussing local gauge invariance, in that the gauge group plays a more natural role on the lattice, and the kind of structure we found in the continuous theory has a more natural motivation, and is more easily understood, as a connection on each link between nearest neighbors on the lattice.

So let us consider a theory defined on a rectangular lattice in space-time, with some fields ψ_n, which we will later call *matter fields*, defined on the sites $n = (n_0, n_1, n_2, n_3) \in (\mathbb{Z})^4$, corresponding to the space-time point $x^\mu = a^\mu n_\mu$ (no sum), where a^μ is the lattice spacing in the μ direction. We are allowing for unequal spacings, in particular so we may consider the continuous time, lattice space limit later.

The kinetic energy terms, derivatives in a continuum theory, become couplings between nearest neighbors, which arise from the terms $\psi_{n+\mu} - \psi_n$ either squared (for scalars) or contracted with $\psi^\dagger$ (for fermions). These are supposed to reflect the change in the field from one point to the next, but if the fields have local symmetry transformations, it may not be clear what it means for the field not to change. For example, the components of a constant vector field expressed in radial coordinates have components which vary from point to point, so moving a vector field while keeping it constant, called parallel transport, does not correspond to keeping the coordinates constant. If $U_{n,n+\mu}$ is the operator (a representation of some group element) that parallel transports ψ_n to lattice site $n+\mu$, the nearest neighbor coupling should come from $\psi_{n+\mu} - U_{n,n+\mu}\psi_n$ and would involve $\bar\psi_{n+\mu} U_{n,n+\mu}\psi_n$, as well as the local terms $\bar\psi_n\psi_n$. There may also be other local terms. These local terms will be symmetric under a group symmetry for which they transform with a unitary representation, but the nearest neighbor terms will need to have the parallel transport operator U transform suitably. Under a symmetry transformation

$$\bar\psi_{n+\mu} U_{n,n+\mu}\psi_n \to \bar\psi_{n+\mu} e^{i\omega_{n+\mu}} U'_{n,n+\mu} e^{-i\omega_n}\psi_n,$$

so we see that invariance requires the link fields to transform as

$$U'_{n,n+\mu} = e^{-i\omega_{n+\mu}} U_{n,n+\mu} e^{i\omega_n}.$$

The U needs to be multipliable on either side by a group element — why not make it a group element itself, or more accurately, the representation of the group element in the representation under which ψ transforms.

Thus we have a dynamics which consists of a field ψ living in a vector space transforming as a representation η of the group, and defined at each lattice site, and a group element g on each link between nearest neighbors.

Before we proceed with our lattice dynamics, let us ask how this relates to our continuum description. To get close to a derivative, add in a local interaction $\bar\psi_{n+\mu}\psi_{n+\mu}$ (as we have not yet specified our local interaction, except that it is group invariant, we can always add and subtract this) so that the link interaction is

$$-\bar\psi_{n+\mu}\left(U_{n,n+\mu}\psi_n - \psi_{n+\mu}\right).$$

In the continuum limit, one expects ψ varies slowly, $\psi_{n+\mu} \sim \psi_n + a^\mu\partial_\mu\psi_n$. We also expect U, which gives the rotations necessary to get from analogous states at n to states at $n+\mu$, should differ from the identity by a small amount, of order a, so we may write $U_{n,n+\mu} = e^{ia^\mu\omega_{n,n+\mu}} \sim \mathbb{1} + ia^\mu\omega_{n,n+\mu}$, so to first order in a^μ, we find the interaction to be $a^\mu\bar\psi_n\left(\partial_\mu - i\omega_{n,n+\mu}\right)\psi_n$. This is what we expect for the covariant derivative, with[4] $\omega_{n,n+\mu} = gA_\mu$, $U = e^{iga^\mu A_\mu}$.

Now consider a local symmetry transformation

$$\begin{aligned} U \to U' &= e^{-i\omega_{n+\mu}} U e^{i\omega_n} \\ &= \left(e^{-i\omega_n} + a^\mu\partial_\mu e^{-i\omega_n}\right) e^{iga^\mu A_\mu} e^{i\omega_n} = e^{iga^\mu A'_\mu}. \end{aligned}$$

Expand to first order in a

$$iga^\mu A'_\mu = e^{-i\omega_n} iga^\mu A_\mu e^{i\omega_n} + a^\mu\left(\partial_\mu e^{-i\omega_n}\right) e^{i\omega_n}$$

or

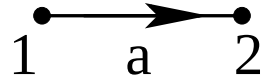
$$A'_\mu = e^{-i\omega_n} A_\mu e^{i\omega_n} - \frac{i}{g}\left(\partial_\mu e^{-i\omega_n}\right) e^{i\omega_n},$$

[4] Here g is the coupling constant, the generalization of the electric charge e. There are two normalizations of the gauge field commonly used. Here we are using the standard one, in which the covariant derivative has the gauge field multiplied by the charge before adding to the derivative operator. Sometimes the field is defined to include the charge.

which agrees with what we found from the continuous derivation.

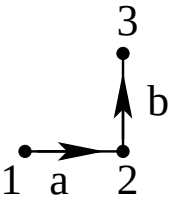
Now let us go back to the general case with finite a. We have a dynamical variable U on each link, but we do not yet have a kinetic energy term, or any term depending solely on U. How can one make a gauge invariant (*i.e.* local symmetry invariant) object of the U's to put in the $\mathcal{L}$?

Each U transforms with an $G_n^{-1} = e^{i\omega_n}$ on the right from its starting point and an $G_{n+\mu} = e^{-i\omega_n}$ on the left from its ending. These G's are representations (in the mathematician's sense) of group elements in the representation appropriate for the ψ field. The dynamical variable for the link a from n to $n+\mu$ is also a group element represented in the same representation, so lets call $U = G_a$. Clearly G_a is not invariant, and there is no way to make a nontrivial invariant function from just one G_a, as applying the symmetry G_a to its tail makes it the identity.

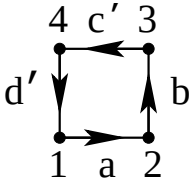


$$G_a \to G_2 G_a G_1^{-1}$$

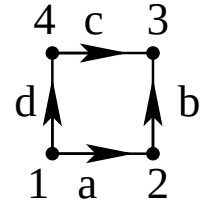
The only way to combine these U's to tranform reasonably is to multiply them so the beginning of the second is at the end of the first, Thus we can combine the G's for two links with a common point and the product becomes invariant under gauge transformations G_2 at the joint, $G_b G_a \to G_3 G_b G_a G_1^{-1}$.



We can continue this to a string of G's, but we still have dependence on the first and last point. This suggests closing the loop by going around one square of the lattice, called a **placquette**. So we consider having a piece of our Lagrangian depend on $G_p := G_{d'} G_{c'} G_b G_a$. (We will define our link variables so they point to the right or up, links c and d are "backwards" here, so I have primed them.) But even G_p is not invariant, for $G_p \underset{\text{gauge}}{\longrightarrow} G_1 G_p G_1^{-1}$. Nonetheless this is an improvement, because now there is an invariant we could make, for in the trace we can move the G_1 to the end and it cancels.



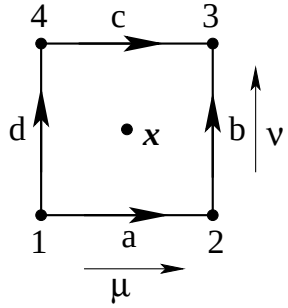
What is the relation of $G_{c'}$ to the right-pointing variable G_c? We are adding a term $\bar\psi_{n+\mu} U_{n,n+\mu} \psi_n$ to the Lagrangian, with U pointing to the right from n to $n+\mu$, but the Lagrangian must be hermitean, so we also need the complex conjugate $\bar\psi_n U^\dagger_{n,n+\mu} \psi_{n+\mu}$. Thus $U^\dagger_{n,n+\mu}$ is the left-pointing link $U_{c'}$, and is the inverse of U_c, for (1) our representations are unitary, and (2) the transformation which takes the basis at $n+\mu$ to that at n is the inverse of the transformation which takes the basis at n to that at $n+\mu$. Thus



$$G_p = U_d^{-1} U_c^{-1} U_b U_a.$$

To understand what this means and just how to add G_p dependence to the Lagrangian, let's consider the continuum limit. In approaching this limit, the ψ's are assumed to differ only of order the lattice spacings a^μ between neighboring points, and the group elements for each loop can be expanded in a power series in a^μ. $G_b = e^{iga^\nu A_\nu} \sim \mathbb{1} + ia^\nu g A_\nu - \frac{1}{2} a^{\nu\,2} g^2 A_\nu^2 + \ldots$.

So let us carefully evaluate G_p for a placquette in the $\mu\nu$ plane with x at its center. We may use the expansion above, but we also need to Taylor series expand the dependence on position. Let Δ^μ be a four vector with 1 in the μ direction and zero in the others. For example,



$$G_b = e^{iga^\nu A_\nu(x + a^\mu \Delta_\mu/2)}.$$

Expanding each link to order $\mathcal{O}(a^2)$

$$\begin{aligned}
G_b &\approx 1 + ia^\nu g \mathcal{A}_\nu(\mathbf{x} + \frac{1}{2} a^\mu \Delta_\mu) - \frac{1}{2} a^{\nu\,2} g^2 \mathcal{A}_\nu^2(\mathbf{x} + \frac{1}{2} a^\mu \Delta_\mu) \\
&\approx 1 + ia^\nu g \mathcal{A}_\nu(\mathbf{x}) + \frac{1}{2} ia^\mu a^\nu g \partial_\mu \mathcal{A}_\nu(\mathbf{x}) - \frac{1}{2} a^{\nu\,2} g^2 \mathcal{A}_\nu^2(\mathbf{x}) \\
G_a &\approx 1 + ia^\mu g \mathcal{A}_\mu(\mathbf{x} - \frac{1}{2} a^\nu \Delta_\nu) - \frac{1}{2} a^{\mu\,2} g^2 \mathcal{A}_\mu^2(\mathbf{x} - \frac{1}{2} a^\nu \Delta_\nu) \\
&\approx 1 + iag \mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2} ia^{\nu\,2} g \partial_\nu \mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2} a^{\mu\,2} g^2 \mathcal{A}_\mu^2(\mathbf{x})
\end{aligned}$$

$$\begin{aligned}
G_c^{-1} &\approx 1 - ia^\mu g \mathcal{A}_\mu(\mathbf{x} + \frac{1}{2} a^\nu \Delta_\nu) - \frac{1}{2} a^{\mu\,2} g^2 \mathcal{A}_\mu^2(\mathbf{x} + \frac{1}{2} a^\nu \Delta_\nu) \\
&\approx 1 - ia^\mu g \mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2} ia^\mu a^\nu g \partial_\nu \mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2} a^{\mu\,2} g^2 \mathcal{A}_\mu^2(\mathbf{x}),
\end{aligned}$$

so, to order a^2,

$$\begin{aligned} G_P &= \left(1 - ia^\nu g\mathcal{A}_\nu(\mathbf{x}) + \frac{1}{2}ia^\mu a^\nu g\partial_\mu\mathcal{A}_\nu(\mathbf{x}) - \frac{1}{2}a^{\nu\,2}g^2\mathcal{A}_\nu^2(\mathbf{x})\right) \\ &\quad \left(1 - ia^\mu g\mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2}ia^\mu a^\nu g\partial_\nu\mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2}a^{\mu\,2}g^2\mathcal{A}_\mu^2(\mathbf{x})\right) \\ &\quad \left(1 + ia^\nu g\mathcal{A}_\nu(\mathbf{x}) + \frac{1}{2}ia^\mu a^\nu g\partial_\mu\mathcal{A}_\nu(\mathbf{x}) - \frac{1}{2}a^{\nu\,2}g^2\mathcal{A}_\nu^2(\mathbf{x})\right) \\ &\quad \left(1 + ia^\mu g\mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2}ia^\mu a^\nu g\partial_\nu\mathcal{A}_\mu(\mathbf{x}) - \frac{1}{2}a^{\mu\,2}g^2\mathcal{A}_\mu^2(\mathbf{x})\right) \\ &= 1 + a^\mu a^\nu g\left\{g\left[\mathcal{A}_\mu(\mathbf{x}),\mathcal{A}_\nu(\mathbf{x})\right] + i\partial_\mu\mathcal{A}_\nu(\mathbf{x}) - i\partial_\nu\mathcal{A}_\mu(\mathbf{x})\right\} \end{aligned}$$

As we might expect, G_p differs from the identity by order $a^\mu a^\nu$. If either a vanished, two links would become identities and the other two cancel. Also, as G_p is unitary, its lowest order difference from $\mathbb{1}$ is antihermitean, so let us define the **field strength tensor**

$$\mathcal{F}_{\mu\nu}(\mathbf{x}) = \partial_\mu\mathcal{A}_\nu(\mathbf{x}) - \partial_\nu\mathcal{A}_\mu(\mathbf{x}) - ig\left[\mathcal{A}_\mu(\mathbf{x}),\mathcal{A}_\nu(\mathbf{x})\right], \qquad G_P \approx 1 + iga^\mu a^\nu\mathcal{F}_{\mu\nu}.$$

Note that $\mathcal{F}_{\mu\nu}$ is

- a Lie-algebra valued field, $\mathcal{F}_{\mu\nu}(\mathbf{x}) = \sum_b F_{\mu\nu}^{(b)}(\mathbf{x})L_b$.
- An antisymmetric tensor, $\mathcal{F}_{\mu\nu}(\mathbf{x}) = -\mathcal{F}_{\nu\mu}(\mathbf{x})$.
- As $\mathcal{F}$ is a Lie algebra element, we can take the group element for the placquette to be $U_p = e^{iga^\mu a^\nu\mathcal{F}_{\mu\nu}}$ away from the continuum limit. $\mathcal{F}$ is hermitean and U_p is unitary.
- Because the Lie algebra is defined in terms of the structure constants, ${c_{ab}}^d$ by
$$[L_a, L_b] = i{c_{ab}}^d L_d,$$
the field-strength tensor may also be written
$$F_{\mu\nu}^{(d)} = \partial_\mu A_\nu^{(d)} - \partial_\nu A_\mu^{(d)} + g{c_{ab}}^d A_\mu^{(a)} A_\nu^{(b)}.$$

The way we have defined U_p, its gauge transformation depends on the gauge at its lower left corner, but in the continuum limit this can be described

as the gauge at the point $(\vec{x},t)$, so

$$\begin{aligned} G_p(\vec{x},t) &\to e^{-i\omega(\vec{x},t)}G_p(\vec{x},t)e^{i\omega(\vec{x},t)} \\ \text{and}\quad \mathcal{F}_{\mu\nu}(\vec{x},t) &\to e^{-i\omega(\vec{x},t)}\mathcal{F}_{\mu\nu}(\vec{x},t)e^{i\omega(\vec{x},t)}. \end{aligned}$$

Note that the field strength tensor transforms homogeneously, unlike the gauge field $\mathcal{A}$, which picks up a piece independent of the initial value of the field.

How should the gauge field and field strength tensor appear in the Lagrangian? We hit upon the placquette and the field strength tensor in trying to find a gauge invariant expression to include in $\mathcal{L}$. As we mentioned earlier, $\mathcal{F}$ is not invariant, but its trace is. So we could add a piece $\mathrm{Tr}(U_p - \mathbb{1})$, except that is not hermitean, so we need to include a piece proportional to $\mathrm{Tr}(U_p + U_p^\dagger - 2\mathbb{1}) = 2\,\mathrm{Tr}\left[\cos\left(ga^\mu a^\nu\mathcal{F}_{\mu\nu}\right) - \mathbb{1}\right]$. The lowest order in lattice spacing is $\propto \mathrm{Tr}\,\mathcal{F}_{\mu\nu}^2 = \sum_d {F_{\mu\nu}^{(d)}}^2$. Of course the correct way to sum over directions of the placquettes means in the continuum limit we need to get $-\frac{1}{4}\sum_d F_{\mu\nu}^{(d)}F^{(d)\,\mu\nu}$. For electromagnetism, there is only one value of (d), and this is the familiar term for the electromagnetic fields.

We are going to turn to a Hamiltonian description, which will give us an interesting entrée into the invariant metric on a group. But before we do that, I want to say a few words about the covariant derivative, which in space-time we may write as

$$D_\mu = \partial_\mu - ig\mathcal{A}_\mu = \partial_\mu - ig\sum_b A_\mu^{(b)}L_b.$$

Notice that it involves both a derivative in space-time and an infinitesimal generator in the group, which is a derivative in group space. If we take the commutator of two of these operators at the same space-time point but in different directions,

$$\begin{aligned} [D_\mu, D_\nu] &= [\partial_\mu - ig\mathcal{A}_\mu, \partial_\nu - ig\mathcal{A}_\nu] \\ &= -ig\partial_\mu\mathcal{A}_\nu - g^2\mathcal{A}_\mu\mathcal{A}_\nu - (\mu\leftrightarrow\nu) \\ &= -g^2\left[\mathcal{A}_\mu,\mathcal{A}_\nu\right] - ig\partial_\mu\mathcal{A}_\nu + ig\partial_\nu\mathcal{A}_\mu \\ &= -ig\mathcal{F}_{\mu\nu}. \end{aligned}$$

Notice that although the covariant derivative is in part a differential operator, the commutator has neither first nor second derivatives left over to act on whatever appears to the right. It does need to be interpreted, however, as specifying a representation matrix that will act on whatever is to the right.

13.1 Gauge Invariance

We have seen that, by looking for local symmetries, we have come across a new class of theories with gauge invariance. These theories require a quite different attitude than ordinary theories with global symmetries. In ordinary theories, a symmetry transformation acting on a physical state gives a new physical state, distinct from the original state, though it has the same energy and other properties are related.

But local symmetries, or gauge transformations, are not like that. To understand the physical significance of these gauge transformations, consider Maxwell theory, with which we are well familiar. We noted that here the group is U(1) phase shifts, $\psi \to e^{i\theta}\psi$, and $\mathcal{A}_\nu$ has only one direction in group space, $\mathcal{A}_\nu \to A_\nu$. What are the physical degrees of freedom of the electromagnetic field? The physical fields are $\vec{E}$ and $\vec{B}$, defined at each point of space and time. That would appear to be six independent degrees of freedom at each point. But these are not independent degrees of freedom. In the absence of charged particles, $\vec{\nabla}\cdot\vec{E} = \vec{\nabla}\cdot\vec{B} = 0$, which are not equations of motion, for they have no derivatives in the time direction. We conclude that these fields are constrained and that the theory can be better described by four unconstrained fields $A_\mu(\vec{x},t) = \Big(\Phi(\vec{x},t), \vec{A}(\vec{x},t)\Big)$, with $\vec{E} = -\vec{\nabla}\Phi - \frac{1}{c}\vec{A}$ and $\vec{B} = \vec{\nabla}\times\vec{A}$, which can be put together in four-dimensional language into $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Clearly $\vec{E}$ and $\vec{B}$ are determined by $A_\mu(\vec{x},t)$, so now it appears we have four degrees of freedom at each point in space-time, but this isn't right either. Maxwell's equations are satisfied by A_μ when

$$\partial_\nu F_\mu{}^\nu = \frac{4\pi}{c} j_\mu \to 0$$

in the absence of charged particles, and these do not, in fact, determine the evolution of A_μ in time, as equations of motion for degrees of freedom ought to do. For if A_μ satisfies this equation, and if $\Lambda(\vec{x},t)$ is an arbitrary twice-differentiable function of space and time,

$$A'_\mu = A_\mu + \partial_\mu\Lambda$$

gives

$$F'_{\mu\nu} = \partial_\mu A_\nu + \partial_\mu\partial_\nu\Lambda - (\partial_\nu A_\mu + \partial_\nu\partial_\mu\Lambda) = F_{\mu\nu}.$$

So A' also satisfies all the equations, and has exactly the same physical fields as A, so

- these are not distinct states of the system
- as we could choose $\Lambda(\vec{x},0) = 0$ and $\dot{\Lambda}(\vec{x},0) = 0$, so as to leave the gauge field state unchanged at $t = 0$, we can still make $\Lambda(\vec{x},t{>}0)$ anything we want, changing the future development of A_μ.

So we see that $A_\mu(\vec{x},t)$ is not fully determined by its initial conditions.

One thing that is clear is that the ambiguity of Λ means that one of the degrees of freedom in A_μ has no physical meaning, so we might think there are three degrees of freedom at each point. We might eliminate the spurious degree of freedom by imposing a constraint, a **gauge condition**, such as the **Lorenz gauge**, $\partial^\mu A_\mu = 0$. As we can find an equivalent gauge field A' from any initial gauge field A_μ, with

$$\partial^\mu A'_\mu = \partial^\mu A_\mu + \partial^\mu\partial_\mu\Lambda,$$

we can take any gauge field A_μ and replace it with one that satisfies the Lorenz gauge $\partial^\mu A'_\mu = 0$ if we can solve the inhomogeneous wave equation $\partial^\mu\partial_\mu\Lambda(\vec{x},t) = -\partial_\nu A^\nu(\vec{x},t)$ for $\Lambda(\vec{x},t)$. But this equation is always solvable, and this determines the evolution of Λ in time, so we have a deterministic system. The equations are now

$$\partial^\mu F_{\mu\nu} = 0 = \partial^\mu\partial_\mu A_\nu - \partial^\mu\partial^\nu A_\mu = \partial^\mu\partial_\mu A_\nu,$$

together with the Lorenz gauge condition $\partial^\mu A_\mu = 0$.

The solutions to the equations of motion in the absence of charges and in the Lorenz gauge are

$$A_\mu(\vec{x},t) = A_\mu^{(\vec{k})} e^{ik_\mu x^\mu} \qquad \text{with} \quad k_\mu k^\mu = 0,$$

together with the Lorenz contraint $k^\mu A_\mu^{(\vec{k})} = 0$. So each three-momentum $\vec{k}$ has three degrees of freedom, but there is still and ambiguity due to gauge, because $\Lambda(\vec{x},t) = \lambda^{(\vec{k})} e^{ik_\mu x^\mu}$ can be added to Λ without changing its solution to the Lorenz gauge condition, as $\partial^\mu\partial_\mu\Lambda = -k^2\lambda^{(\vec{k})}e^{ik_\mu x^\mu} = 0$. So

$$A_\mu^{(\vec{k})} \to A_\mu^{(\vec{k})} + \lambda^{(\vec{k})} k_\mu$$

is a residual gauge invariance that does not affect the physics. So each 3-momentum has only two physical degrees of freedom, which is why there are two polarizations for a photon.

Now what does gauge invariance mean quantum mechanically? In quantum mechanics, the transition amplitude from an initial state at one time to a final state at a later time is given by the integral of $e^{iS/\hbar}$ over all "paths", possible values of the degrees of freedom at intermediate times. With a Lagrangian depending on fields and their derivatives, the fields are subject to arbitrary variation but the time derivatives are evaluated from the fields as functions of time. In our lattice discussion the time derivatives reside in the U's on links in the time direction. We will need to examine how this effects the momenta in the problem.