Gauge theories in particle physics(3rd, Aitchison and Hey)

Chapter 3, Electromagnetism as a gauge theory

This chapter introduces briefly the application of gauge theory to electromagnetism.

Sector 3.1. Introduction

Over a hundred years ago, Faraday, Maxwell and others developed the theory of electromagnetic interactions. Today, Maxwell's theory still stands – unlike Newton's "classical mechanics" which was shown by Einstein to require modification at relativistic speeds, approaching the speed of light.
Moreover, Maxwell's electromagnetism, when suitably married with quantum mechanics, gives us "quantum electrodynamics"(QED) which is remarkably agree with experiments. Taking QED as paradigmatic theory, EW and QCD were developed, and all these three theories are included in *Standard Model.*

From today's perspective, the crucial thing about electromagnetism is that it's a theory in which the *dynamics* (i.e. the behavior of the forces) is intimately related to a *symmetry* principle. On the other hand, though Newton's laws have two *invariances* also: "translation" and "rotation"; however, these invariances have no special connection with any particular force law(*dynamics*), instead, only constrain the form of the allowed laws, but by no means uniquely determine them(*dynamics*).

Global invariance: the same transformation is carried out at all space-time points.
Local invariance: different transformations are carried out at different individual space-time points.
In general, a theory that is globally invariant "won't" be invariant under locally varying transformations. However, by specific tricks, local invariance still be held.

Sector 3.2. The Maxwell equations : current conservation.

The *gauge theories* are characterized by a close interrelation between three conceptual elements: *symmetries*, *conservation laws* and *dynamics*. In previous sector, *symmetries* $\leftrightarrow$ *dynamics* was slightly touched, in this sector, *conservation laws* $\leftrightarrow$ *dynamics* will be touched here.

Maxwell noticed that taking the divergence of the Ampère's law for steady currents $\nabla \times \boldsymbol{B} = \boldsymbol{j}_{em}$ leads to conflict with the continuity equation for electric charge,

$$\frac{\partial \rho_{em}}{\partial t} + \nabla \cdot \boldsymbol{j}_{em} = 0 \tag{1}$$

so he modified to $\nabla \times \boldsymbol{B} = \boldsymbol{j}_{em} + \partial \boldsymbol{E}/\partial t$.

The extra term $\partial \boldsymbol{E}/\partial t$ owns its place in the dynamical equation to a local conservation requirement. How the local conservation requirement comes from?

Taking a look at equation 1 (with integrals act on left side), it states that *the rate of decrease of charge* in any arbitrary volume Ω is due precisely and only to *the flux of current* out of its surface; that is, no net charge can be created or destroyed in Ω. Since Ω could be made as small as possible, this means that *electric charge must be locally conserved.*

Here, it's no global invariance because the global form of charge conservation would necessitate the instantaneous propagation of signals and this conflicts with special relativity.

It's now widely believed that the only exact *quantum number conservation laws* are those which have an associated *gauge theory force field.*

For instance, the charge conservation law survived everywhere due to it has associated electromagnetic field. And baryon number is not absolutely conserved when unifying gauge theories of the strong, weak and electromagnetic interactions could be interpreted as without corresponding baryonic force field.

As a result, to know the particle **X**'s charge, one can using *charge conservation* or *dynamics*(by observing how particle **X** responded to known electromagnetic fields) to determine it. While to know particle **X**'s baryonic number, one could only use *dynamics.*

The four equations composed of Maxwell equations.

$$\nabla \cdot \boldsymbol{E} = \rho_{em} \tag{Gauss' law}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{Faraday-Lentz laws}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{no magnetic charges}$$

$$\nabla \times \boldsymbol{B} = \boldsymbol{j}_{em} + \frac{\partial \boldsymbol{E}}{\partial t} \tag{Maxwell modified Ampère's law}$$

Sector 3.3. Maxwell equations: Lorentz covariance and gauge invariance

At first, this sector introduced the Maxwell equations in the form of "Lorentz covariance" and "gauge invariance". Then, the authors provided an insight view to explain why Maxwell equation is "Lorentz covariance" and "gauge invariance". Finally, the key role of gauge theories in SM was metioned.

It's convenient to introduce the *vector potential*, $A_\mu(x)$, in place of the fields **E** and **B**.

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \qquad \boldsymbol{E} = -\nabla V - \frac{\partial \boldsymbol{A}}{\partial t}$$

which defines the 3-vector potential **A** and the scalar potential V.

The origin of gauge invariance in classical electromagnetism lies in the fact that the potentials **A** and V are not unique for given physical fields **E** and **B**.
The transformations that **A** and V may undergo while preserving **E** and **B**(and hence the Maxwell equation) unchanged are called *gauge transformations*.
And the associated invariance of the Maxwell equations is called *gauge invariance*.
The *gauge transformations* are,

$$\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \nabla\chi \qquad V \to V' = V - \frac{\partial \chi}{\partial t} \qquad A^\mu \to A'^\mu = A^\mu - \partial^\mu \chi$$

The Maxwell equation can also be written in a manifestly *Lorentz covariant* form using the 4-current j^μ_{em} given by $j^\mu_{em} = (\rho_{em}, \boldsymbol{j}_{em})$,

$$\partial_\mu F^{\mu\nu} = j^\nu_{em}$$

where, $F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$.
Because under the gauge transformation, $A^\mu \to A'^\mu = A^\mu - \partial^\mu\chi$, $F^{\mu\nu}$ remains unchanged: $F^{\mu\nu} \to F'^{\mu\nu} = F^{\mu\nu}$. So, $F^{\mu\nu}$ is gauge invariant and so are the Maxwell equations.

The "Lorentz-covariant and gauge-invariant" field equations satisfied by A^μ is,

$$\Box A^\nu - \partial^\nu(\partial_\mu A^\nu) = j^\nu_{em}$$

BTW, in above equation, "Lorentz-covariant" was described by covariant operator like ∂_μ, and "gauge-invariant" conveyed from the fact that $\partial_\mu F^{\mu\nu} = j^\nu_{em}$ is gauge invariant, and these two equations are equal.

The insight view of why Maxwell's equations are gauge invariant is very critical to understand how gauge theories developed in Standard Model.
Mathematically, $\partial_\mu F^{\mu\nu} = j^\nu_{em}$ is gauge variant because $F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$ is a four-dimensional kind of "curl", which is unchanged by a gauge transformation $A^\mu \to A'^\mu = A^\mu - \partial^\mu\chi$. The proof is as following.

$$
\begin{aligned}
&\partial^\mu A^\nu - \partial^\nu A^\mu \\
&\to \partial^\mu(A^\nu - \partial^\nu\chi) - \partial^\nu(A^\mu - \partial^\mu\chi) \qquad \text{(with } A^\mu \to A'^\mu = A^\mu - \partial^\mu\chi) \\
&= \partial^\mu A^\nu - \partial^\mu\partial^\nu\chi - \partial^\nu A^\mu + \partial^\nu\partial^\mu\chi \\
&= \partial^\mu A^\nu - \partial^\nu A^\mu
\end{aligned}
$$

But if the left side of this equation is gauge invariant, $\partial_\mu F^{\mu\nu} = j^\nu_{em}$, how the gauge invariant relates to right side(charge)?
Wigner related charge conservation to an invariance under transformation of the electrostatic potential with a constant. And changing the value of the electrostatic potential by a constant mount is an example of a *global transformation*(since the change of the potential is same everywhere).
Invariance under this global transformation is related to a conservation law: electronic charge. However this global invariance is not sufficient to generate the full Maxwellian dynamics(current $\boldsymbol{j}_{em}$ isn't included for instance)!
Remarked by 't Hooft, one can consider that the *local* change in the electrostatic potential V can be compensated – in the sense of leaving the Maxwell equations unchanged – by a corresponding local change in the magnetic vector potential $\boldsymbol{A}$. Thus, by including magnetic effects, the *global invariance* under a change of V by a constant can be extended to a *local invariance*
Hence, one might almost "derive" the complete Maxwell equations from the requirement that the theory be expressed in terms of potentials(like A^μ) in such a way as to be invariant under local(gauge) transformations on those potentials($A^\mu \to A'^\mu = A^\mu - \partial^\mu\chi$); of course, special relativity must play a role too.

Generally, when a certain global invariance is generalized to a local one, the existence of a new "compensating" field is entailed, interacting in a specified way.
The first example of dynamical theory "derived" from a local invariance requirement seems to be the theory of *Yang* and *Mills*.
It's a remarkable fact that the interactions in the Standard Model of particle physics are of precisely this type. The strong interaction between quarks and the weak interaction between quarks and leptons are also seem to be described by gauge theories. And Utiyama showed that even *special relativity* could be arrived at (gauge theories) by generalizing the coordinate transformations to local ones.

Sector 3.4. Gauge invariance(and covariance) in quantum mechanics

From previous sector, by adding magnetic field, Maxwell equations could keep invariance under gauge transform, or global invariance could change to local invariance under classical mechanics. And this kind of compensating field idea played a key role of particle physics.
In this chapter, the authors discussed how the Schrödinger equations keep gauge invariance(and covariance) under the context of *quantum mechanics*.

The Maxwell equation in "classical mechanics" keep invariance (mathematically because $F^{\mu\nu}$ itself is gauge invariant) under gauge transformation like this $A^{\mu} \rightarrow A'^{\mu} = A^{\mu} - \partial^{\mu}\chi$.
While, in "quantum mechanics", the Schrödinger equation under electromagnetic field is,

$$(\frac{1}{2m}(-i\nabla - q\boldsymbol{A})^2 + qV)\psi(\boldsymbol{x},t) = i\frac{\partial\psi(\boldsymbol{x},t)}{\partial t} \tag{2}$$

The solution of ψ in above equation should describe completely the state of the particle moving under potential A^{μ}. So, if A^{μ} is changed, the Schrödinger equation will change, accordingly, ψ will change to ψ'.
While, from Maxwell equation, given the A^{μ} is changed, the **E** and **B** remain the same.
The plausible explanation is ψ is not a directly observable quantity and, to allow quantum mechanics be consistent with Maxwell's equations, ψ has to change also : $\psi \rightarrow \psi'$.
Then, the problem is, once $\psi \rightarrow \psi'$, the two different ψ describe same physics or not?

$$(\frac{1}{2m}(-i\nabla - q\boldsymbol{A}')^2 + qV')\psi'(\boldsymbol{x},t) = i\frac{\partial\psi'(\boldsymbol{x},t)}{\partial t} \tag{3}$$

Because 2 and 3 have the exact same form, so that they will effectively ensure that both "describe the same physics". By substituting the gauge transformation of $\boldsymbol{A}'$ and V', the relationship of ψ and ψ' are:

$$\psi'(\boldsymbol{x},t) = e^{iq\chi(\boldsymbol{x},t)}\psi(\boldsymbol{x},t)$$

It's necessary to check ψ and ψ' are indeed equivalent. The probability of ψ and ψ' are obviously same since they're related by a phase transformaiton.
Other interesting observable usually involving ∇ or $\partial/\partial t$, like $\psi^*(\nabla\psi) - (\nabla\psi)^*\psi$. It's easy to check that this current is not invariant under $\psi'(\boldsymbol{x},t) = e^{iq\chi(\boldsymbol{x},t)}\psi(\boldsymbol{x},t)$, since the phase $\chi(\boldsymbol{x},t)$ is $\boldsymbol{x}$-depedent. So, to construct *gauge-invariant currents*, with the hint from equations (3.36) and (3.37), one has to replace ∇ by $\boldsymbol{D} \equiv \nabla - iq\boldsymbol{A}$, and replace $\partial/\partial t$ by $D^0 \equiv \partial/\partial t + iqV$. So the current change like this:

$$\psi^*(\nabla\psi) \rightarrow \psi^{*\prime}(\boldsymbol{D}'\psi') = \psi^* \cdot e^{-iq\chi} \cdot e^{iq\chi} \cdot (\boldsymbol{D}\psi) = \psi^*\boldsymbol{D}\psi$$

From above equation, it's obvious to see that the *equality* between the first and last terms is indeed a statement of *gauge invariance*, which means after the replacement($\nabla \rightarrow \boldsymbol{D}, \partial/\partial t \rightarrow D^0$), the current keeps *invariance*.
So, in the language of D^{μ}, ψ and ψ' indeed describe same physics. And it's therefore *gauge invariance*.

However, though with new notation D^μ, *gauge invariance* obtained/survived, the difference of the expressions like $\psi^*(\nabla\psi) \to \psi^*\boldsymbol{D}\psi$ undoubtedly indicates they're *covariance*, but not *invariance.* One could make same conclusion from the comparison of 3 and following equation.

$$\frac{1}{2m}(-iD')^2\psi' = iD^{0'}\psi'$$

The *gauge invariance* of Maxwell equations re-emerges as a *gauge covariance*(with replacement $\nabla \to \boldsymbol{D}, \partial/\partial t \to D^0$) in quantum mechanics provided we make the combined transformation on the potential and on the wave function.

$$\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \nabla\chi \qquad V \to V' = V - \frac{\partial\chi}{\partial t} \qquad \psi \to \psi' = e^{iq\chi}\psi$$

Moreover, besides above three transformation, if another additional replacement supplied,

$$\partial^\mu \to D^\mu \equiv \partial^\mu + iqA^\mu$$

the the *gauge invariance* of Maxwell equations could be survived!

In summary, as the title of this sector indicates, for Maxwell equations in quantum mechanics, it can be *gauge invariance* or *gauge covariance*, depends on which kinds of transformations utilized.

Sector 3.5. The argument reversed : The gauge principle

In sector 3.4, the authors took the Schrödinger equation as *known* for a charged particle in a electromagnetic field.

In this sector, however, the authors start by demanding that the theory is invariant under the *space-time-dependent phase transformation*, $\psi(\boldsymbol{x},t) \to \psi'(\boldsymbol{x},t) = e^{iq\chi(\boldsymbol{x},t)}\psi(\boldsymbol{x},t)$. Then, they demonstrate that such a phase invariance is not possible for a free theory, but rather requires an *interacting* theory involving a field whose interactions with the charged particle are precisely determined, and which undergoes the transformation

$$\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \nabla\chi \qquad V \to V' = V - \frac{\partial\chi}{\partial t}$$

when $\psi \to \psi'$.

In double-slit experiment, it's clear that if the individual phase δ_1 and δ_2 are each shifted by the same amount, there will be no observable consequences, since only the phase difference matters the superposition. Which is precisely what is happened by a symmetry or invariance principle in quantum mechanics.

Invariance under a constant change in phase is an example of a *global invariance*, like $\psi \to \psi' = e^{i\alpha}\psi(\alpha$ is a constant). However, it may seem an unnatural state of affairs, since *global invariance* requires that once a phase convention has been adopted at one space-time point, the same convention must be adopted at all others.

Rather, requiring invariance under *local* phase transformation: independent choices of phase convention at each space-time point sounds more "reasonable". The *local invariance*: $\psi(\boldsymbol{x},t) \to \psi'(\boldsymbol{x},t) = e^{i\alpha(\boldsymbol{x},t)}\psi(\boldsymbol{x},t)$. This *local invariance* is just the same as we demanded for electromagnetic *gauge invariance* $\psi(\boldsymbol{x},t) \to \psi'(\boldsymbol{x},t) = e^{iq\chi(\boldsymbol{x},t)}\psi(\boldsymbol{x},t)$ if we require $\alpha = q\chi$.

Accordingly, we must modify the Schrödinger equation from

$$\frac{1}{2m}(-i\nabla)^2\psi = i\frac{\partial\psi}{\partial t}$$

to

$$\frac{1}{2m}(-i\nabla - q\boldsymbol{A})^2\psi = (i\frac{\partial}{\partial t} - qV)\psi$$

(Current understanding on this change is, to introduce some kind of force field in which the particle moves like $\psi \to \psi' = e^{i\alpha}\psi$, then the related Schrödinger equation has to change its operator as : $-i\nabla \to -i\nabla - q\boldsymbol{A}$, $i\frac{\partial}{\partial t} \to i\frac{\partial}{\partial t} - qV$. $q\boldsymbol{A}\psi$ could be understood as the movement due to the existence of $qV\psi$. But given this is the truth, why the two fields "superpose" their Schrödinger equations like this ?)

The modified wave equation is precisely the Schrödinger equation describing the interaction of the charged particle with the electromagnetic field described by **A** and V.

A vector field such as A^μ, introduced to guarantee *local phase invariance*, is called a "gauge field". The principle that the interaction should be so dictated by the phase (or gauge) invariance is called the *gauge principle*: it allows us to write down the wave equation for the interaction directly from the free particle equation via the replacement.

$$\partial^\mu \to D^\mu \equiv \partial^\mu + iqA^\mu$$