Basic examples on physical gauge theories

Fernando Chamizo

July 2016

A motivating example

Let us start with a very illustrative and simple example. It is taken from [1] (in an abbreviated form) and, as far as we know, it is not a historical example, but looking into retrospective it has some connections with the ideas introduced in 1926 by V. Fock [7] shortly after E. Schrödinger stated his equation. We are going to use the kind of *first quantization* arguments that a pioneer of quantum mechanics could have employed at that time.

After the classical contribution of J.C. Maxwell and H. Lorentz, it was known that the *Lorentz force* $q(\vec{E}+\vec{v}\times\vec{B})$ on a charge q derives from a classical Hamiltonian

$$(1) \qquad \mathcal{H} = \frac{1}{2m}\left(\vec{p} - q\vec{A}\right)^2 + q\varphi$$

where $(A_\alpha) = (\varphi, -\vec{A})$ is what we call today the 4-potential. We have

$$(2) \qquad \vec{E} = -\nabla\varphi - \frac{\partial\vec{A}}{\partial t} \qquad \text{and} \qquad \vec{B} = \nabla \times \vec{A}.$$

These equations, that embody two of the *Maxwell equations*, the Lorentz force and consequently the dynamics of the system, are invariant by the *gauge transformation*

$$(3) \qquad \varphi \longmapsto \varphi + \frac{\partial\chi}{\partial t}, \qquad \vec{A} \longmapsto \vec{A} - \nabla\chi.$$

On the other hand, the (first) quantization of this electromagnetic problem via *Schrödinger's equation* is (in natural units $\hbar = c = 1$)

$$(4) \qquad \left(\frac{1}{2m}\left(-i\nabla - q\vec{A}\right)^2 + q\varphi\right)\Psi = i\frac{\partial\Psi}{\partial t}.$$

There is something strange in this equation: It seems that the transformations (3) change it substantially and then the choice of different gauges could give, in principle,

1

different physical conclusions. But χ in (3) is a mathematical artifact related to the non uniqueness of the solution of the partial differential equations $\nabla \cdot \vec{F} = f_0$ and $\nabla \times \vec{F} = \vec{F_0}$, and there is not a clear rule to privilege a particular solution. At least in the classic setting the value of the 4-potential at a point has not physical significance. At a fixed point, we can only measure the electric and magnetic fields $\vec{E}$ and $\vec{B}$.

It turns out that the gauge change (3) that apparently modifies completely (4) does not act dramatically on the solutions. It can be checked that if Ψ is a solution of (4) for certain $(\varphi, -\vec{A})$ then $e^{-iq\chi}\Psi$ is a solution after the gauge change (3) (shortly we shall see the calculation in detail in another example). With this information at hand, one can argue that the "probability density" $|\Psi|^2$ is invariant under phase changes and take it as an explanation to save the gauge invariance coming from the well-settled Maxwell equations. But this explanation has a flaw, if we allow a possibly time dependent phase in the wave function, then the probability interpretation and its conservation, is in danger. Note that $\Psi^* \nabla \Psi - \Psi \nabla \Psi^*$ is no longer a current, as in the case of a free particle. The important point to be noted here is that (4) becomes the free particle Schrödinger's equation if

$$(5) \qquad \nabla - iq\vec{A} \longmapsto \nabla \qquad \text{and} \qquad \frac{\partial}{\partial t} + iq\chi \longmapsto \frac{\partial}{\partial t}.$$

This sounds relativistic and it is noteworthy in our context in which relativity was not explicitly considered. The important point is that it gives a big clue about the right conserved current and, in general, an answer about why different gauges lead to the same physics: The operator

$$(6) \qquad D = \left(\frac{\partial}{\partial t} + iq\varphi, \nabla - iq\vec{A} \right)$$

is in some sense gauge invariant, meaning that if D' is the operator in other gauge, and $\Psi' = e^{-iq\chi}\Psi$ is the solution of (4) in that gauge, then

$$(7) \qquad D'\Psi' = e^{-iq\chi}D\Psi.$$

Let us insist on the same point reviewing with care the computations in a relativistic example to avoid any paradox coming from the combination of Newton's dynamics and electromagnetism. Imagine that, as Schrödinger tried in first place [4], we want to study the quantum relativistic corrections for a charged particle. The natural quantization for the uncharged free particle in natural units ($c = \hbar = 1$) is

$$(8) \qquad E^2 + \vec{p}^2 = m^2, \qquad E \leftrightarrow i\frac{\partial}{\partial t}, \qquad \vec{p} \leftrightarrow -i\nabla,$$

that leads to the Klein-Gordon equation

$$\partial_\alpha \partial^\alpha \Psi + m^2 \Psi = 0. \tag{9}$$

If we now switch the electromagnetic field on, one should add the corresponding potential energy to the 4-moment, i.e. one should replace $i\partial_\alpha$ by $i\partial_\alpha + qA_\alpha$. Let us write $q = -e$ having in mind the electron. Consequently, (9) becomes

$$\left(\partial_\alpha + ieA_\alpha\right)\left(\partial^\alpha + ieA^\alpha\right)\Psi + m^2 \Psi = 0. \tag{10}$$

Again, if Ψ solves (10) then $e^{-ie\chi(x)}\Psi$ solves (10) after the gauge change (3). The key point is

$$\begin{cases} \left(\partial^\alpha + ie(A^\alpha + \partial^\alpha \chi)\right)\left(e^{-ie\chi(x)}\Psi\right) = e^{-ie\chi(x)}\left(\partial_\alpha + ieA^\alpha\right)\Psi, \\ \left(\partial_\alpha + ie(A_\alpha + \partial_\alpha \chi)\right)\left(e^{-ie\chi(x)}\Psi\right) = e^{-ie\chi(x)}\left(\partial_\alpha + ieA_\alpha\right)\Psi. \end{cases} \tag{11}$$

There is nothing deep in these relations, we simply compensate the extra term in the derivative of a product using $\left(\partial_\alpha - \partial_\alpha f\right)e^f = 0$, the defining property of the exponential.

We can express the situation in a very succinct way introducing the *covariant derivative D* and the *gauge transformation G*

$$D_\mu = \partial_\mu + ieA_\mu \qquad \text{and} \qquad G\Psi = e^{-ie\chi(x)}\Psi. \tag{12}$$

Then our observation is that under $\Psi \longmapsto G\Psi$, we have

$$D_\mu \Psi \longmapsto GD_\mu \Psi \qquad \text{and} \qquad A_\mu \longmapsto GA_\mu G^{-1} + ie^{-1}(\partial_\mu G)G^{-1}. \tag{13}$$

The second equation is just verbosity meaning simply $A_\mu \mapsto A_\mu + \partial_\mu \chi$.

Extending these ideas to the right context of spin $1/2$ particles (to include the electron) leads to write the QED Lagrangian as

$$\mathcal{L} = \bar{\Psi}(i\slashed{D} - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \qquad \text{with} \quad \slashed{D} = \gamma^\mu D_\mu. \tag{14}$$

It is clearly invariant by (13). Note that the interaction Lagrangian $-j^\mu A_\mu = -e\bar{\Psi}\gamma^\mu \Psi A_\mu$ is obtained by *minimal coupling* changing usual derivatives ∂_μ by covariant derivatives D_μ. In other words, from the *Dirac equation* for the free particle[1].

---

[1]In the case of the electromagnetic field, this is not so spectacular. In the famous paper [2] in which P.A.M. Dirac introduces his equation (published in 1928), we can read "we adopt the usual procedure of substituting $p_0 + e/cA_0$ for p_0 and $\mathbf{p} + e/c\mathbf{A}$ for $\mathbf{p}$ in the Hamiltonian for no field."

A non-Abelian example

Let us consider now an example with less physical significance but closer to the ideas appearing in the Standard Model.

Say that we have N real scalar fields ϕ_1, ϕ_2, $\ldots \phi_N$ behaving as harmonic oscillators with equal mass and no interaction. The Lagrangian is

$$(15) \qquad \mathcal{L} = \frac{1}{2} \sum_{k=1}^{N} \partial_\mu \phi_k \partial^\mu \phi_k - \frac{1}{2} m^2 \sum_{k=1}^{N} \phi_k^2.$$

We can define artificially a column vector field Φ in $\mathbb{R}^N$ having ϕ_k as its k-th coordinate, and write

$$(16) \qquad \mathcal{L} = \frac{1}{2} (\partial_\mu \Phi)^\dagger (\partial^\mu \Phi) - \frac{1}{2} m^2 \Phi^\dagger \Phi$$

(of course, the dagger can be changed by transposition).

There is no a priori physical reason to consider the different fields as components of the same object. In the same way, we could say that neutrinos and electrons are different animals and coupling both together in the Standard Model Lagrangian [13] is, in principle, an economic usage of the mathematical notation like in (16).

The group $SO(N)$ acts naturally on Φ preserving (16). In fact, $O(N)$ also leaves it invariant, but by technical reasons, we only consider the connected component containing the identity. Each transformation of $SO(N)$ mixes the real fields ϕ_k together in the same way as $SU(2)$ mixes neutrinos and electrons. Physics shows up when we assume a kind of *gauge principle*, claiming a minimal coupling rule. Shortly we shall see it in a broader and more formal context. Now we are going to tackle the problem through our example.

In the electrodynamic example, the formalism allows us to interpret the interacting field as a term to be added to the derivative in such a way that the result behaves well under the symmetry transformations, even if they vary from a point to another. Then we consider as in (12)

$$(17) \qquad D_\mu = \partial_\mu + i g A_\mu \qquad \text{and} \qquad G\Psi = S(x)\Psi.$$

where $S(x) \in SO(N)$ for each x. Now A_μ, the *gauge field* is at each point an $N \times N$ matrix, and g is just a constant introduced to recall the physical situations ($g = e$ in the previous case).

If we wish the covariant derivative D_μ to be really covariant, $D_\mu S(x) = S(x) D_\mu$, then A_μ must transform in the particular way stated in (13) with $e = g$. Both formulas in (13) become equivalent. To keep a complete analogy one would like to

write $S(x)$ as an exponential $G\Psi = e^{-igX(x)}\Psi$ where $X(x)$ is a purely imaginary anti-symmetric matrix to assure $e^{-igX(x)} \in SO(N)$. We can identify then the exponents with elements of the Lie algebra $\mathfrak{so}(N)$. This simply reflects that derivatives in a Lie group give rise to elements in its Lie algebra. Expanding this comment, since ∂_μ and A_μ stand with the same role in (17), it is natural to impose $A_\mu \in \mathfrak{so}(N)$. Hence we could express each A_μ as a linear combination of the $N(N-1)/2$ elements of a fixed basis of $\mathfrak{so}(N)$. On the other hand, μ varies in $\{0,1,2,3\}$, then we can think the gauge field as a superposition of $N(N-1)/2$ vector fields (with matrix coefficients). In the Standard Model, the strong force corresponds to the group $SU(3)$ (there are three colors), then there are $\dim \mathfrak{su}(3) = 3^2 - 1$ basic gauge (gluon) fields represented by the Gell-Mann matrices.

Under the "spell" [12] of the gauge principle, we can guess that in our example the Lagrangian including the interaction term is

$$(18) \qquad \mathcal{L} = \frac{1}{2}(D_\mu\Phi)^\dagger(D^\mu\Phi) - \frac{1}{2}m^2\Phi^\dagger\Phi.$$

Note that, unwrapping the notation, it gives an interaction Lagrangian that is not trivial to guess

$$(19) \qquad \mathcal{L}_{\text{int}} = \frac{1}{2}ig(A_\mu\Phi)^\dagger\partial^\mu\Phi + \frac{1}{2}ig(\partial^\mu\Phi)^\dagger A_\mu\Phi - \frac{1}{2}g^2(A_\mu\Phi)^\dagger A^\mu\Phi.$$

If we compare it to the electromagnetic example, this is not the whole story. In (14) we had the term $F_{\mu\nu}F^{\mu\nu}$ giving the Maxwell equations when taking variations [5]. Its proxy in the non-Abelian setting is the Yang-Mills term.

Some historical geometric aspects and Yang-Mills theories

The inaugural lecture "On the hypothesis which lie at the foundations of geometry" given by B. Riemann in 1854 is considered a landmark in mathematics (see [11] for an English translation with detailed explanations). It is known that C.F. Gauss praised it greatly[2]. But it was not published until the year in which Riemann passed away and the first impression for a reader is that of a vague outreach paper with very few formulas. This style reflects the hope of Riemann to be understood by the most of the audience attending the lecture.

A main point in Riemann's communication is that one can define intrinsically geometric objects by local metric properties, without any reference to an outer space. It took many years and many authors to develop this idea, that was extremely important in general relativity. One of the pioneer contributors to this new view

---

[2]He forced the topic breaking the tradition of admitting the candidate first choice that was Riemann's Habilitationsschrift, a masterpiece on Fourier series.

of geometry was T. Levi-Civita who defined an *absolute derivative* [8] introducing a method to displace a vector to a nearby point, this was a way to "connect" points and the origin of the term *connection*. Without entering into the mathematical definitions, the idea is very simple: If there is not a privileged global orthonormal frame, to study the variation (the derivative) of a vector field V, we have to take into account the variation of the coordinates plus the variation of the reference frame, this gives *covariant derivatives* or *connections* (in a strict sense)

$$(20) \qquad D_i V = \left(\partial_i V^k + \Gamma_{ij}^k V^j \right) \partial_k.$$

The functions Γ_{ij}^k must satisfy certain consistency conditions to avoid contradictions when changing coordinates. It turns out that there is only a possible choice of the Γ_{ij}^k if one wishes to keep some compatibility with the metric structure [11]. This is the so-called *Levi-Civita connection* in which Γ_{ij}^k are the Christoffel symbols.

It is apparent the similarity between (20) and the covariant derivative in the context of gauge fields. This similarity becomes identity when the concept of *vector bundle* or, more in general, that of *fiber bundle* is taken from mathematics. Surprisingly, it seems that, after the breakthrough in the physical gauge theory by C.N. Yang and R. Mills in 1954 [15], it took many years to note this coincidence. The following opinion was expressed by Yang in a recent colloquium [14]:

> It came as a great shock [...] when it became clear in the 1970s that the mathematics of gauge theory, both Abelian and non-Abelian, is *exactly* the same as that of fiber bundle theory. [...] it served to *bring back* the close relationship between the two disciplines [mathematics and physics] which had been interrupted through the increasingly abstract nature of mathematics since the middle of the 20th century.

Without entering into details, we are going to give a formulation of what can be called a *gauge principle* with a mathematical flavor (but not very technical). Let G be a generic element of a matrix Lie group included in $GL(N)$ and consider the covariant derivative

$$(21) \qquad D_\mu \Psi = \partial_\mu \Psi + ig A_\mu \Psi \qquad \text{with } A_\mu \text{ in the Lie algebra.}$$

Assume for simplicity that $\Psi : \mathcal{U} \longrightarrow \mathbb{R}^N$ with $\mathcal{U} \subset \mathbb{R}^n$ (more formally, Ψ should be a *section* of the bundle). Under

$$(22) \qquad \Psi \longmapsto G\Psi \qquad \text{and} \qquad A_\mu \longmapsto G A_\mu G^{-1} + ig^{-1} (\partial_\mu G) G^{-1}$$

we have that $D_\mu \Psi \mapsto G D_\mu \Psi$, i.e. the derivative is actually covariant. Here g is a *coupling constant* that is separated from A_μ for convenience in the physical interpretation. In our first examples was the charge of the electron.

Let us write G_0 to indicate a generic element of the Lie group constant in $\mathcal{U}$. Then we can deduce that for any Lagrangian $\mathcal{L}$,

(23)
$$\mathcal{L}(\partial_\mu \Psi, \Psi) \text{ invariant under } \Psi \mapsto G_0 \Psi \quad \Rightarrow \quad \mathcal{L}(D_\mu \Psi, \Psi) \text{ invariant under (22)}.$$

As in the examples above, (23) can be used to create involved Lagrangians once one guesses the *gauge group*. This idea has been specially successfully in the creation and development of the Standard Model. We can rephrase (23) in physical terms saying that the invariance of the Lagrangian with respect to local transformations determines the interaction of Ψ with the gauge field A_μ. But we still need a term, like $F_{\mu\nu}F^{\mu\nu}$ in (14), giving the Lagrangian of the field itself.

This new term must depend only on A_μ and must be gauge invariant. In Riemannian geometry an important construction associated to the Levi-Civita connection and already appearing in Riemann's inaugural talk is the *curvature tensor*. It corresponds to the difference between cross partial covariant derivatives. Define $F_{\mu\nu}$ such that for any Ψ

(24)
$$\left[D_\mu, D_\nu\right]\Psi = igF_{\mu\nu}\Psi.$$

Note the analogy with Faraday's tensor. It can be proved that it is well defined and behaves as a covariant tensor[3] in the indexes μ and ν. On the other hand, under gauge transformations on A_μ, as in the second part of (22), it changes as $F_{\mu\nu} \mapsto GF_{\mu\nu}G^{-1}$ (this is simpler than it seems, just the covariance of the derivatives). Then the natural "scalar" $F_{\mu\nu}F^{\mu\nu}$, actually a matrix, obeys the same rule and we can do it invariant under the gauge transformations taking traces. Summing up, a plausible term to add in the Lagrangian is the *Yang-Mills Lagrangian*

(25)
$$\mathcal{L}_{\text{YM}} = -\frac{1}{2}\text{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right).$$

The Euler-Lagrange equations corresponding to (25) are the *Yang-Mills equations* [9]

(26)
$$\partial_\mu F_{\mu\nu} + ig[A_\mu, F_{\mu\nu}] = 0,$$

that in the case of the electromagnetic field are the Maxwell equations.

Ironically, this cumbersome geometrization of gauge theories in which the potential corresponds to the connection and the field strength to the curvature [6] that has been crucial in the modern understanding of the Standard Model, is very close

---

[3]The usual introduction of the curvature tensor in Riemannian geometry is as a tensor field covariant in three indexes and contravariant in one. To recover the analogy one has to note that in our setting, for each μ and ν, $F_{\mu\nu}$ is a matrix, a tensor of type $(1,1)$, at each point.

to the initial introduction and development in the period 1918–1929 of the concept of gauge done by H. Weyl. His aim was the unification of the electromagnetism and gravitation (general relativity), where the curvature plays an important role in the theory. In his works the term "gauge", introduced in 1918 by himself (in German), is more natural than nowadays because it corresponded to a change of scale in the metric [10]. The modern language of differential forms initiated by É. Cartan is specially useful to present these ideas in a compact form. We have avoided it because it requires a stronger background.

A final comment is that the electro-weak sector in the Standard Model presents an important variation with respect the scheme explained here. It is a Yang-Mills theory with group $U(1) \times SU(2)$ giving $1 + \dim \mathfrak{su}(2) = 4$ fields, associated to the photon and the bosons $W^{\pm}$ and Z. The problem is that the massive nature of the particles $W^{\pm}$ and Z (related to the short range of the weak interaction) does not fit the pure Yang-Mills scheme. A new term was added to the Lagrangian to solve this problem breaking the symmetry imposed by the gauge group (see [3, §9.3] for a simple and convincing mathematical explanation). It is associated to the *Higgs particle* that was recently detected at the LHC.

References

[1] I.J.R. Aitchison and A.J.G. Hey. *Gauge Theories in Particle Physics: Volume I: From Relativistic Quantum Mechanics to QED, Third Edition.* Graduate Student Series in Physics. CRC Press, 2002.

[2] P. A. M. Dirac. The quantum theory of the electron. *Proc. Royal Soc. London A*, 117(778):610–624, 1928.

[3] G. B. Folland. *Quantum field theory*, volume 149 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008. A tourist guide for mathematicians.

[4] A. Galindo and P. Pascual. *Quantum mechanics. I.* Texts and Monographs in Physics. Springer-Verlag, Berlin, 1990. Translated from the Spanish by J. D. García and L. Alvarez-Gaumé.

[5] D. V. Gal′tsov, Yu. V. Grats, and V. Ch. Zhukovskiĭ. *Campos clásicos*. Editorial URSS, 2005.

[6] M. Göckeler and T. Schücker. *Differential geometry, gauge theories, and gravity.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 1987.

[7] J. D. Jackson and L. B. Okun. Historical roots of gauge invariance. *Rev. Modern Phys.*, 73(3):663–680, 2001.

[8] T. Levi-Civita. *The absolute differential calculus*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2005. Reprint of the 1926 translation.

[9] S. P. Novikov and I. A. Taimanov. *Modern geometric structures and fields*, volume 71 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2006.

[10] L. O'Raifeartaigh and N. Straumann. Gauge theory: historical origins and some modern developments. *Rev. Modern Phys.*, 72(1):1–23, 2000.

[11] M. Spivak. *A comprehensive introduction to differential geometry. Vol. II.* Publish or Perish, Inc., Wilmington, Del., second edition, 1979.

[12] G. 't Hooft. *Under the spell of the gauge principle*, volume 19 of *Advanced Series in Mathematical Physics*. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[13] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[14] C. N. Yang. Conceptual origin of Maxwell equations and of gauge theory. ASIAA/CCMS/IAMS/LeCosPA/NTU-Phys Joint Colloquia. `https://www.youtube.com/watch?v=SG343kojbnU`, 2015.

[15] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Physical Rev. (2)*, 96:191–195, 1954.