U1 Symmetry and Electromagnetism

Michael J. Hurben

July 5, 2023

1 Introduction

Typically, the Schrödinger equation is written so as to include only a scalar potential $V(\mathbf{r})$:

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r},t). \quad (1)$$

This is amenable for solving many important problems, such as a particle in free space, a particle in a box, barrier tunneling, the simple harmonic oscillator, a Coulomb potential, and so on [1]. However, it cannot account for effects on a charged particle from a magnetic field, which arises from a vector potential. Care is needed if we are to introduce and use a new term, such as this vector potential, in the Schrödinger equation. Working through the process, moreover, is instructive and valuable, for it provides a clear demonstration of what local gauge symmetry means. Specifically, it will allow us to show that requiring a wavefunction to have local U(1) invariance leads to $\boldsymbol{E}$ and $\boldsymbol{B}$ fields [2].

The approach here will be as follows: we start with a Lagrangian for the classical motion of a charged particle experiencing both electric and magnetic fields, showing that it results in the Lorentz force law. Then we will construct the corresponding Hamiltonian and write out a revised Schrödinger equation. Finally we'll look at the wavefunction solutions, and we'll see that local U(1) symmetry leads to the same gauge terms which we know from classical electromagnetism.

2 The Electromagnetic Lagrangian

Consider a particle with mass m and charge q in the presence of $\boldsymbol{E}$ and $\boldsymbol{B}$ fields. What is the Lagrangian? There is no way to derive it from first principles; instead one uses the trusted Guess and Check Method. We assume some Lagrangian and run it through the Euler-Lagrange equation to see if the result accords with expectations. Here, it will need to produce the Lorentz force law.

We assert that the correct Lagrangian can be expressed in terms of the scalar and vector potentials as:

$$L = \frac{1}{2}m\dot{x}^2 - q\phi(\boldsymbol{x}) + \frac{q}{c}\dot{\boldsymbol{x}} \cdot \boldsymbol{A}(\boldsymbol{x}). \tag{2}$$

Let's verify that this works. (If you are already satisfied that this Lagrangian is correct, skip ahead, you will not miss anything interesting.) For simplicity we remove the explicit references to space dependence in the potential terms, and do the usual derivatives:

$$\frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + \frac{q}{c}A_i, \tag{3}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = m\ddot{x}_i + \frac{q}{c}\frac{d}{dt}A_i, \tag{4}$$

$$\frac{\partial L}{\partial x_i} = -q\frac{\partial}{\partial x_i}\phi + \frac{q}{c}\dot{\boldsymbol{x}} \cdot \frac{\partial}{\partial x_i}\boldsymbol{A}. \tag{5}$$

The Euler-Lagrange equation then gives:

$$m\ddot{x}_i + \frac{q}{c}\frac{d}{dt}A_i + q\frac{\partial}{\partial x_i}\phi - \frac{q}{c}\dot{\boldsymbol{x}} \cdot \frac{\partial}{\partial x_i}\boldsymbol{A} = 0. \tag{6}$$

We next focus on the terms that contain only $\boldsymbol{A}$ and expand them. Note that the time derivative needs the chain rule:

$$\begin{aligned} \frac{d}{dt}A_i - \dot{\boldsymbol{x}} \cdot \frac{\partial}{\partial x_i}\boldsymbol{A} = {} & [\frac{\partial}{\partial x_i}\frac{\partial x_i}{\partial t} + \frac{\partial}{\partial x_j}\frac{\partial x_j}{\partial t} + \frac{\partial}{\partial x_k}\frac{\partial x_k}{\partial t}]A_i \\ & -\dot{x}_i\frac{\partial}{\partial x_i}A_i - \dot{x}_j\frac{\partial}{\partial x_i}A_j - \dot{x}_k\frac{\partial}{\partial x_i}A_k. \end{aligned} \tag{7}$$

This becomes

$$\frac{d}{dt}A_i - \dot{\boldsymbol{x}} \cdot \frac{\partial}{\partial x_i}\boldsymbol{A} = \dot{x}_j[\frac{\partial}{\partial x_j}A_i - \frac{\partial}{\partial x_i}A_j] + \dot{x}_k[\frac{\partial}{\partial x_k}A_i - \frac{\partial}{\partial x_i}A_k]. \tag{8}$$

The differences in the spatial derivatives are components of curls, by inspection:

$$\frac{d}{dt}A_i - \dot{\boldsymbol{x}} \cdot \frac{\partial}{\partial x_i}\boldsymbol{A} = -\dot{x}_j[\boldsymbol{\nabla} \times \boldsymbol{A}]_k + \dot{x}_k[\boldsymbol{\nabla} \times \boldsymbol{A}]_j. \tag{9}$$

And this is just a cross product. With the definition $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$,

$$\frac{d}{dt}A_i - \dot{\boldsymbol{x}} \cdot \frac{\partial}{\partial x_i}\boldsymbol{A} = [\dot{\boldsymbol{x}} \times \boldsymbol{B}]_i. \tag{10}$$

The Euler-Lagrange equation is then

$$m\ddot{x}_i + q\frac{\partial}{\partial x_i}\phi - \frac{q}{c}[\dot{\boldsymbol{x}} \times \boldsymbol{B}]_i = 0. \tag{11}$$

Since $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$, we generalize from single components to full vector form and we finally have

$$\boldsymbol{F} = q\boldsymbol{E} + \frac{q}{c}\dot{\boldsymbol{x}} \times \boldsymbol{B}. \tag{12}$$

3 The Electromagnetic Hamiltonian

If you are already happy with the electromagnetic Hamiltonian, you can skip this too. Otherwise, we'll use the fact that the Hamiltonian comes from the Lagrangian according to

$$H = \sum_i \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} - L. \tag{13}$$

The canonical momentum is defined as $p_i = \frac{\partial L}{\partial \dot{x}_i}$, which we found in equation 3. Now we plug in L, and get rid of velocities by expressing them in terms of momentum:

$$H = \sum_i \frac{1}{m}\left[p_i - \frac{q}{c}A_i\right]p_i$$
$$- \sum_i \frac{1}{2}m\left[\frac{1}{m}\left[p_i - \frac{q}{c}A_i\right]\right]^2 + q\phi - \frac{q}{mc}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] \boldsymbol{\cdot} \boldsymbol{A}. \tag{14}$$

The sums are just dot products:

$$H = \frac{1}{m}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] \boldsymbol{\cdot} \boldsymbol{p} - \frac{1}{2m}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] \boldsymbol{\cdot} \left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] + q\phi - \frac{q}{mc}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] \boldsymbol{\cdot} \boldsymbol{A}. \tag{15}$$

Grouping the first and last terms:

$$H = \frac{1}{m}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] \boldsymbol{\cdot} \left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] - \frac{1}{2m}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] \boldsymbol{\cdot} \left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right] + q\phi. \tag{16}$$

We finally have a nice Hamiltonian which will reduce to the familiar one if we switch off the vector potential:

$$H = \frac{1}{2m}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right]^2 + q\phi. \tag{17}$$

4 The Schrödinger Equation Gets Magnetized

We make the substitution of the gradient for momentum:

$$\boldsymbol{p} \rightarrow -i\hbar\boldsymbol{\nabla}. \tag{18}$$

Then the Schrödinger equation is, after some factoring:

$$i\hbar\frac{\partial}{\partial t}\psi = \left[\frac{-\hbar^2}{2m}\left[\boldsymbol{\nabla} - \frac{iq}{\hbar c}\boldsymbol{A}\right] \boldsymbol{\cdot} \left[\boldsymbol{\nabla} - \frac{iq}{\hbar c}\boldsymbol{A}\right] + q\phi\right]\psi. \tag{19}$$

Note that we have dropped explicit reference to space and time in ψ just as we are not writing out that the potentials can vary with space, for brevity.

Equation (19) is a general result. We are not going to solve it for some particular potentials. Rather, what we will ask is: how might we alter ψ such that it still satisfies equation (19)? Our motivation for wanting to do such a thing comes from the understanding that it is only the squared magnitude of the wavefunction that gives us a measurable quantity; namely, the probability density. We know we can rotate the wavefunction according to

$$\psi'(\mathbf{r},t) = e^{i\theta}\psi(\mathbf{r},t), \tag{20}$$

for any constant θ, and that it will still be a solution to (19). Just as importantly, the probability density will be unaffected:

$$|\psi'(\mathbf{r},t)|^2 = |\psi(\mathbf{r},t)|^2 . \tag{21}$$

That $e^{i\theta}$ is a special and wonderful entity is already well-known; what is more, the set of all $e^{i\theta}$ terms also form a unitary group, $U(1)$. Groups must satisfy four requirements: (a) under multiplication, they must generate another element of the group (check: $e^{i\theta}\ e^{i\kappa} = e^{i(\theta+\kappa)}$); (b) they must have an identity element as a member (check: $e^0{=}1$); (c) multiplication must be associative (check: yes, obvious in this case); and (d) every element must have an inverse in the group (check: $e^{i\theta}$ has $e^{-i\theta}$).

The nomenclature $U(1)$ indicates dimension 1; $U(2)$ must consist of 2×2 unitary matrices with the same properties, but under matrix multiplication, and so on.

When we multiply by $e^{i\theta}$, we are effectively rotating the wavefunction in the complex plane. We are not surprised that such an operation would leave the physics unchanged, in the same way that nothing would change if we could rotate the universe by some arbitrary angle. In other words, ψ obeys global $U(1)$ symmetry.

But what about local $U(1)$ symmetry? What happens if we locally alter ψ, twisting it to point off in some new direction here, tweaking it, to point in another direction over there, and so on? At first glance that seems foolish. We work hard to solve the Schrödinger equation for wavefunction solutions, so why would we then take the solution and arbitrarily garble it up like that?

Let's try it anyway. Quantum Mechanics is already strange, and this new version of the Schrödinger equation looks even stranger. Maybe something interesting will happen.

So, to start, how do we do a local $U(1)$ transformation? Easy, just let θ be a function of space and time:

$$\theta \to \theta(\mathbf{r},t). \tag{22}$$

For brevity we again will drop explicit reference to space and time, and set to working out the impact of the $e^{i\theta}$ term in the modified Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}e^{i\theta}\psi = \left[\frac{-\hbar^2}{2m}\left[\boldsymbol{\nabla} - \frac{iq}{\hbar c}\boldsymbol{A}\right]\boldsymbol{.}\left[\boldsymbol{\nabla} - \frac{iq}{\hbar c}\boldsymbol{A}\right] + q\phi\right]e^{i\theta}\psi. \tag{23}$$

Let's look for low-hanging fruit first. The left hand side calls for a simple application of the product rule:

$$i\hbar\frac{\partial}{\partial t}e^{i\theta}\psi = -\hbar e^{i\theta}\frac{\partial\theta}{\partial t}\psi + i\hbar e^{i\theta}\frac{\partial\psi}{\partial t}. \tag{24}$$

We also note that the final term on the right hand side of equation (24), which involves the scalar potential, is trivial, as it does not involve differentiation. The part of equation (24) that will get ugly is obvious: we have a vector $\boldsymbol{A}$ which is itself a function of space, added to a gradient operator, with the result squared, and then applied on a product of functions. Attempting short-cuts here can lead to disaster. We must be diligent. For the sake of getting through the calculus, let's define $\boldsymbol{\alpha} = -\frac{iq}{\hbar c}\boldsymbol{A}$. We need to then expand as follows:

$$\begin{aligned}[\boldsymbol{\nabla} + \boldsymbol{\alpha}]\boldsymbol{.}[\boldsymbol{\nabla} + \boldsymbol{\alpha}]\psi' &= [\boldsymbol{\nabla} + \boldsymbol{\alpha}]\boldsymbol{.}[\boldsymbol{\nabla}\psi' + \psi'\boldsymbol{\alpha}] \\ &= \nabla^2\psi' + \boldsymbol{\nabla}\boldsymbol{.}(\psi'\boldsymbol{\alpha}) + \boldsymbol{\alpha}\boldsymbol{.}\boldsymbol{\nabla}\psi' + \psi'\alpha^2.\end{aligned} \tag{25}$$

The product rule for the gradient operator then gives

$$[\boldsymbol{\nabla} + \boldsymbol{\alpha}]\boldsymbol{.}[\boldsymbol{\nabla} + \boldsymbol{\alpha}]\psi' = \left[\nabla^2 + 2\boldsymbol{\alpha}\boldsymbol{.}\boldsymbol{\nabla} + (\boldsymbol{\nabla}\boldsymbol{.}\boldsymbol{\alpha}) + \alpha^2\right]\psi'. \tag{26}$$

Still not too bad. But we need to deal with the fact that $\psi' = e^{i\theta}\psi$. This will complicate the gradient term and the Laplacian term.

Doing the gradient first:

$$\boldsymbol{\nabla}\psi' = \boldsymbol{\nabla}e^{i\theta}\psi = ie^{i\theta}\psi\boldsymbol{\nabla}\theta + e^{i\theta}\boldsymbol{\nabla}\psi = e^{i\theta}\left[i\psi\boldsymbol{\nabla}\theta + \boldsymbol{\nabla}\psi\right]. \tag{27}$$

The Laplacian is the divergence of that gradient:

$$\nabla^2\psi' = \boldsymbol{\nabla}\boldsymbol{.}\boldsymbol{\nabla}e^{i\theta}\psi = \boldsymbol{\nabla}\boldsymbol{.}e^{i\theta}\left[i\psi\boldsymbol{\nabla}\theta + \boldsymbol{\nabla}\psi\right]. \tag{28}$$

Continuing to grind through this, we have

$$\boldsymbol{\nabla}\cdot e^{i\theta}\left[i\psi\boldsymbol{\nabla}\theta+\boldsymbol{\nabla}\psi\right]=ie^{i\theta}\boldsymbol{\nabla}\theta\cdot\left[i\psi\boldsymbol{\nabla}\theta+\boldsymbol{\nabla}\psi\right]+e^{i\theta}\boldsymbol{\nabla}\cdot\left[i\psi\boldsymbol{\nabla}\theta+\boldsymbol{\nabla}\psi\right]. \quad (29)$$

Expanding the right hand side and using the product rule again,

$$\nabla^2\psi'=-e^{i\theta}\psi\nabla^2\theta+ie^{i\theta}\boldsymbol{\nabla}\theta\cdot\boldsymbol{\nabla}\psi+ie^{i\theta}\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\theta+ie^{i\theta}\psi\nabla^2\theta+e^{i\theta}\nabla^2\psi. \quad (30)$$

That got ugly fast, but collecting terms, it simplifies to

$$\nabla^2\psi'=e^{i\theta}\left[\nabla^2-\nabla^2\theta+2i\boldsymbol{\nabla}\theta\cdot\boldsymbol{\nabla}+i\nabla^2\theta\right]\psi. \quad (31)$$

We can combine the Laplacian and gradient expansions from (27) and(31) in order to write (25) as

$$\begin{aligned}&\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}\right]\cdot\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}\right]\psi'=\\&e^{i\theta}\left[\nabla^2-\nabla^2\theta+2i\boldsymbol{\nabla}\theta\cdot\boldsymbol{\nabla}+i\nabla^2\theta+2i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\,\theta+2\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}+(\boldsymbol{\nabla}\cdot\boldsymbol{\alpha})+\alpha^2\right]\psi.\end{aligned} \quad (32)$$

Next, a clever hack. We go back to (25) and let $\boldsymbol{\alpha}\to\boldsymbol{\alpha}+\boldsymbol{\beta}$ in order to see what happens if the expression had been using two vectors:

$$\begin{aligned}&\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}+\boldsymbol{\beta}\right]\cdot\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}+\boldsymbol{\beta}\right]\psi'=\\&\left[\nabla^2+2\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}+2\beta\cdot\boldsymbol{\nabla}+(\boldsymbol{\nabla}\cdot\boldsymbol{\alpha})+(\boldsymbol{\nabla}\cdot\boldsymbol{\beta})+\boldsymbol{\alpha}^2+2\boldsymbol{\alpha}\cdot\beta+\beta^2\right]\psi'.\end{aligned} \quad (33)$$

These two equations, (32) and (33), are exactly the same if we set $\boldsymbol{\beta}=i\boldsymbol{\nabla}\theta$. In other words, we have shown that:

$$\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}\right]\cdot\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}\right]e^{i\theta}\psi=e^{i\theta}\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}+i\boldsymbol{\nabla}\theta\right]\cdot\left[\boldsymbol{\nabla}+\boldsymbol{\alpha}+i\boldsymbol{\nabla}\theta\right]\psi. \quad (34)$$

Perhaps there is an easier way to arrive at this result, but working through it was a worthwhile exercise and generates confidence in the expression. The Schrödinger equation then becomes

$$-\hbar e^{i\theta}\frac{\partial\theta}{\partial t}\psi+i\hbar e^{i\theta}\frac{\partial\psi}{\partial t}=e^{i\theta}\left[\frac{-\hbar^2}{2m}\left[\boldsymbol{\nabla}-\frac{iq}{\hbar c}\boldsymbol{A}+i\boldsymbol{\nabla}\theta\right]^2+q\phi\right]\psi. \quad (35)$$

We can discard the exponentials and massage various terms, getting it into a more familiar form:

$$i\hbar\frac{\partial\psi}{\partial t} = \left[\frac{-\hbar^2}{2m}\left[\boldsymbol{\nabla} - \frac{iq}{\hbar c}\boldsymbol{A} + i\boldsymbol{\nabla}\theta\right]^2 + q\phi + \hbar\frac{\partial\theta}{\partial t}\right]\psi. \tag{36}$$

We are almost within reach of a profound new expression... Let's do some factoring as follows:

$$i\hbar\frac{\partial\psi}{\partial t} = \left[\frac{-\hbar^2}{2m}\left[\boldsymbol{\nabla} - \frac{iq}{\hbar c}\left(\boldsymbol{A} - \boldsymbol{\nabla}\frac{\hbar c}{q}\theta\right)\right]^2 + q\left(\phi + \frac{1}{c}\frac{\partial}{\partial .t}\frac{\hbar c}{q}\theta\right)\right]\psi \tag{37}$$

If we define

$$\zeta = -\frac{\hbar c}{q}\theta, \tag{38}$$

then this becomes

$$i\hbar\frac{\partial\psi}{\partial t} = \left[\frac{-\hbar^2}{2m}\left[\boldsymbol{\nabla} - \frac{iq}{\hbar c}\left(\boldsymbol{A} + \boldsymbol{\nabla}\zeta\right)\right]^2 + q\left(\phi - \frac{1}{c}\frac{\partial\zeta}{\partial t}\right)\right]\psi. \tag{39}$$

You may weep joyous tears now. The vector and scalar potentials have just gotten their classic gauge transformation terms applied inside the Hamiltonian. The corresponding $\boldsymbol{E}$ and $\boldsymbol{B}$ fields have not changed. We are therefore free to twiddle ψ with local $e^{i\theta}$ rotations, and electromagnetism will make sure that it doesn't matter. Almost as if that weird redundancy of electromagnetic gauge was waiting out there in order to perfectly cancel out the weird redundancy of the wavefunction phase. It is deeply satisfying the more you think about it.

Now imagine what would have happened if we had not included that vector potential. Our attempts to include a local $e^{i\theta}$ transformation on the wavefunction would have generated $\boldsymbol{\nabla}\theta$ terms in the Schrödinger equation, but with no vector potential to absorb them and make them irrelevant.

We can interpret all this as follows: the state of a free, charged particle alone would not have local $U(1)$ symmetry; but if we demand such a symmetry for ψ, then there must be an electromagnetic field which satisfies the classical gauge conditions. In this sense, the symmetry requirement is equivalent to, or the source of, the interaction between a charged particle and the EM field. Hence the expression *local gauge symmetry*.

What we have established here is an illustration of one-third of the Standard Model's $SU(3) \times SU(2) \times U(1)$ symmetry. Electromagnetism embodies the simplest of these three symmetries; the strong force will go with the $SU(3)$ group, and the weak with the $SU(2)$ group.

What do those groups look like? $U(n)$ refers to unitary matrices of dimension n having a determinant $= 1$. Familiar examples of $SU(2)$ symmetry are the three Pauli matrices, although this is seldom pointed out when they are introduced in the context of spin.

There is a huge gulf between the way undergraduate quantum mechanics is discussed, even at the senior level, and the language employed in the Standard Model. The latter may seem unnecessarily remote and mysterious, with references to Yang-Mills theory, group theory, electroweak, and so on. It seems like a missed opportunity in undergraduate quantum mechanics to not work through such a relatively straightforward demonstration of how a local gauge symmetry leads to one of the fundamental forces.

References

[1] Shankar R, Principles of quantum mechanics (1980). Plenum Press. This classic text provides everything need to perform this demonstration of local gauge symmetry. It runs through a similar calculation in the context of the Aharanov-Bohm effect in chapter 18.

[2] Thomson M, Modern Particle Physics (2013) Cambridge University Press. A very brief introduction to the local gauge principle and the connection of electromagnetism to U(1) invariance is given in chapter 10. The treatment here comes from expanding upon that overview and working through it more explicitly.