Deriving the Dirac equation

Bill Celmaster

March 8, 2021

1 Previous episodes ...

In the last meeting

- I introduced the **Dirac equation** – a 4-component differential equation, linear in derivatives. **I pulled it out of a hat.**
- I then pulled out of a hat a set of transformations on the components and said these represented the Lorentz transformations.
- Then I showed that the Dirac equation doesn't change form when we apply those transformations. We say the Dirac equation is **Lorentz invariant**.
- Finally I showed you a Lagrangian from which the Dirac equation follows as an equation of motion.
- **In summary**, I pulled the Dirac Lagrangian out of a hat and showed it was invariant under transformations that I called Lorentz transformations.

Today, I'll derive the Dirac equation by directly considering how to make Lorentz invariant Lagrangians.
Why?

- It's nice to see that the Dirac equation wasn't pulled out of a hat.

  ...and I don't want to follow the historical approach which was very much like hacking down your lawn with a blunt machete.

- The method generalizes so that Lorentz invariant Lagrangians can be found for fields with more than 4 components (e.g. the gravitational field is given by the metric tensor $g^{\mu\nu}(x)$.)
- The method generalizes even further for other Lie groups like $SU(3)$. We find Lagrangians invariant under those groups.
- **This kind of process is called model-building and has dominated many of the theoretical discoveries of the past 60 years.**

2 Philosophical Interlude

- So far, the Dirac equation is a classical set of DE's for complex-valued functions.
- So far, the discussion about Lie groups and Lie algebras has also been classical. Nothing about quantum mechanics.
- So, why didn't mathematicians or physicists – before the birth of quantum mechanics – derive the Dirac equation as an interesting application of Lorentz symmetries?

The Dirac equation could have been classical??? Matthew has pointed out that it couldn't have been classical! The 2-D representation is **NOT** a Lorentz representation. In particular, a rotation by 2π becomes a multiplication by **-I** rather than **I**. Only in QM is that allowable, where a physical theory is regarded as Lorentz invariant if it's transformations are represented up to a phase. Or, said more mathematically, the dimension-2 representations are representations of the covering group of the Lorentz group. It's a deep mathematical fact that the rotation covering group is a factor of 2 larger than the original group, and this factor of 2 gives rise to the omnipresent factor of $\frac{1}{2}$ characteristic of dimension-2 (aka "spin 12") representations. So, as Matthew points out, the spin of the electron is half-integer, which has profound physical consequences, all because 2-D representations aren't proper representations of the Lorentz group.

In any case, even if one could ignore all these reasons for classicists having ignored the Dirac Lagrangians, it should be said that Lorentz symmetry was almost as new to physics as quantum mechanics. And besides, the classical Dirac Lagrangian has energies unbounded from below. That's bad, even in classical theory. How do we make the transition from classical to quantum?

- Remember your basic quantum mechanics. Physics is expressed in terms of states in a Hilbert space, operators on those states, and probability interpretations related to an inner product on the Hilbert space.
- **Importantly, the rules of QM prescribe a correspondence between specific physical observables and specific operators.**
- The most elegant way to obtain the correspondence is with a Lagrangian.
  - Start with the classical Lagrangian (which we've been dealing with).
  - Promote the fundamental variables (e.g. positions in particle theory, or fields in field theory) and their canonical momenta, to operators.
  - Operators are objects that don't commute with one another. **So the operator-promotion goes together with commutation rules amongst operators.**
  - Somewhat surprisingly, you don't need to separately prescribe the Hilbert space. Instead, you can derive a Hilbert space on which the operators obey the commutation rules!
  - **In summary: the variables of the Lagrangian are the observables of nature, and the commutation rules promote these to quantum operators.**
- So the quantization of a theory is simply the promotion of variables in the Lagrangian, to operators obeying certain commutation relations.

So far, the entire prescription of quantum mechanics is rule-based. The issue is that it's hard to picture why those are the rules. But maybe this is just a version of the question "Why is there something rather than nothing? (See the chapter in Robert Nozick Philosopical Explanations.)

The problem with a pure rule-based explanation of the universe, is that it makes it more likely that the rules contain a built-in paradox. Godel tells us this can always be the case, even with something as intuitive as arithmetic. In my opinion, the miracle of modern physics is that the complex rules of quantum and statistical physics haven't led to paradoxes ... **ALMOST**.

Today, there is an unresolved paradox that is the subject of lots of research – the **INFORMATION PARADOX**. It's regarded as a true paradox. Starting with the rules of physics, information should be preserved (in quantum mechanics, this is equivalent to saying that time-evolution is given by a unitary operator). But starting with the rule of physics, information should be lost. That's a paradox. Why is information lost? It's complicated, which is why people are still trying to find out where they've made a logical error (or possibly that the rules of physics must be modified to eliminate the paradox). The basic idea is this. We all know that stuff which falls into a black hole never again sees "the light of day". Information disappears from our sight, but it's hidden inside the black hole. But the laws of physics also say that black holes evaporate – without giving up any of the secret information hidden inside them. So far, so good. But eventually the black hole evaporates completely and now there's no remaining black hole to hide the information. Adacadabra – information is lost.

Anyway, all of this motivates me to attempt some kind of an "explanation" for quantum mechanics which feels (to me) more understandable than just a collection of rules. As it happens, I love the path integral approach. Roughly speaking, it tells us that every 'path' or 'configuration' of the universe has a corresponding phase, and that when you add up all the phases, you get an overall contribution which selects preferred configurations with some probability according to configuration. In macroscopic limits you get classical mechanics. On the microscopic scale you get quantum mechanics and quantum field theory.

There's a problem. All of this feels intuitively reasonable (to me) but it doesn't work for Dirac fields. We haven't discussed this yet. But as it turns out, the "rules" have to be changed for Dirac fields and the path integral approach fails dismally...UNLESS we invent a new meaning for variables and for integration of those variables. In my opinion, that ruins the beauty of the path integral approach. Or almost. One strange mathematical consequence of the Dirac-modified path integral approach, is that it can be converted to a regular path integral with regular variables that interact in a specific highly non-local manner. I once thought I'd investigate this for my Ph.D. thesis, but I've never been able to gain any useful insights from purusing that line of thinking. Oh well.

3 Solutions to exercises on Lorentz transformations

See separate set of notes.

4 Construction of a Lorentz invariant bilinear form with derivatives

- In the exercises, we had fields ν and ν' which transformed respectively as $(0, \frac{1}{2})$ and $(\frac{1}{2}, 0)$. By convention we say ν is "right" and ν' is "left".

- We showed that $\mathbf{S}' \equiv \nu^{*i}\nu'_i + \nu'^{*i}\nu_i$ is a Lorentz invariant bilinear form.

- The generic question is this: What kinds of matrices $\mathbf{M}$ preserve the value of $S = \sum_{ij} \psi_i^* M^{ij} \psi_j$ under a Lorentz transformation? In the above example, we showed that if you have a 4-spinor composed of the two-spinors ν and ν' – namely $\psi = \begin{pmatrix} \nu' \\ \nu \end{pmatrix}$, then

$$\psi^\dagger \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \psi$$

  is invariant under Lorentz transformations. So $\mathbf{M} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is such a matrix.

- This is described in group theory as the following question: "If a vector ψ transforms according to the representation $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ then what are the irreducible representations that can be formed from the tensor product terms of the form $\psi^a \psi^b$, and what are the coefficient matrices ($\mathbf{M}$) required to create those irreducible representations.

- Next, include derivatives. ∂_μ transforms as a Lorentz vector, which is the $(\frac{1}{2}, \frac{1}{2})$ representation. For example, a rotation around the z-axis, transforms $(\partial_0, \partial_1, \partial_2, \partial_3)\phi(x)$ to $(\partial_0, \partial_1 \cos\theta - \partial_2 \sin\theta, \partial_1 \sin\theta + \partial_2 \cos\theta, \partial_3)\phi(x')$, where ϕ is a generic field, and $\mathbf{x}' = \mathbf{R}(\theta)^{-1}\mathbf{x}$. Now the question we ask is "what coefficient matrix is required to obtain a scalar from the product of three tensors, $(0, \frac{1}{2}) \otimes (\frac{1}{2}, \frac{1}{2}) \otimes (\frac{1}{2}, 0)$?"

- In principle, this is a solvable mathematical problem.

- Suppose you solve that problem for a 2-spinor ψ, and you decide that the right scalar combination is $\psi(x)^\dagger \left(\sigma_0\partial_0 - \sigma_1\partial_1 - \sigma_2\partial_2 - \sigma_3\partial_3\right) \psi(x)$. Let's see if this form is preserved under a rotation by θ around the z-axis. Recall that $\psi(x) \rightarrow \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{+i\frac{\theta}{2}} \end{pmatrix} \psi(x')$. Similarly, $\psi^*(x) \rightarrow$

$\begin{pmatrix} e^{+i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix} \psi^*(x')$ where x' is the rotated x. After rotation, the derivative term is $\sigma_0\partial_0 - \sigma_1 (\partial_1 \cos\theta - \partial_2 \sin\theta) - \sigma_2 (\partial_1 \sin\theta + \partial_2 \cos\theta) - \sigma_3\partial_3$. Look at the middle two terms by expanding out the σ_1 and σ_2 matrices. These add up to

$$\begin{pmatrix} 0 & -(\partial_1 \cos\theta - \partial_2 \sin\theta) + i(\partial_1 \sin\theta + \partial_2 \cos\theta) \\ -(\partial_1 \cos\theta - \partial_2 \sin\theta) - i(\partial_1 \sin\theta + \partial_2 \cos\theta) & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & (-\partial_1 + i\partial_2)\, e^{-i\theta} \\ (-\partial_1 - i\partial_2)\, e^{i\theta} & \end{pmatrix}$$

Now put it all together (ignoring, for now, the ∂_0 and ∂_3 terms).

$$\begin{aligned} &\psi^*(x') \begin{pmatrix} e^{+i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix} \begin{pmatrix} 0 & (-\partial_1 + i\partial_2)\, e^{-i\theta} \\ (-\partial_1 - i\partial_2)\, e^{i\theta} & \end{pmatrix} \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{+i\frac{\theta}{2}} \end{pmatrix} \psi(x') \\ &= \psi^*(x') \begin{pmatrix} 0 & -\partial_1 + i\partial_2 \\ -\partial_1 - i\partial_2 & 0 \end{pmatrix} \psi(x') \\ &= \psi^*(x')\, (-\partial_1\sigma_1 - i\partial_2\sigma_2)\, \psi(x') \end{aligned} \tag{1}$$

Lo and behold, these two derivative terms transform to exactly the same form they had before transformation. Add in the remaining two derivative terms (which don't transform) and we've shown that – except for the transformation $x \to x'$, the bilinear form is unchanged by transformation. But, you might ask, what about the transformation $x \to x'$? Strictly speaking, since the Lagrangian is the integrand of the action ($\mathcal{S} = \int d^4x\mathcal{L}$), it's not the Lagrangian which needs to be invariant, but the action. By changing the integration coordinates from $x \to x'$ and noticing that the Jacobian factor (from changing coordinates) = 1, for rotations, the action is completely unchanged under rotations.

- You might object, quite rightly, that the derivative factor was pulled out of a hat. You had to take my word for it that there is a mathematical theory that tells us exactly how to take the tensor product down to a scalar. So be it. At least, that should give you a sense of how one proceeds in general.

- Finally, the pièce de resistance. What is the appropriate form of the tensor reduction for 4-spinors (above was just a 2-spinor)? The answer is

$$i\psi^\dagger\gamma^0\partial\!\!\!/\psi - m\psi^\dagger\gamma^0\psi$$

using the notation introduced in a previous lecture, where $\not{\partial} \equiv \gamma^\mu \partial_{mu}$. The factor of i is introduced to assure the Lagrangian is real.

The mathematical theory (of tensor products) tells us that Lorentz invariance γ^μ must satisfy the Dirac algebra anti-commutator condition $\{\gamma^\mu, \gamma^\nu\} = 2\delta^{\mu\nu}$.

There is one further notational simplification that is introduced, where $\bar{\psi} = \psi^* \gamma^0$. With that notation, the Lagrangian becomes

$$i\bar{\psi}\not{\partial}\psi - m\bar{\psi}$$

.