The Witten $SU(2)$ Anomaly

Emily Nardoni[1]

[1]*Department of Physics, University of California at San Diego, La Jolla, CA 92093*

We explain why a quantum field theory with gauge group $SU(2)$ and an odd number of left-handed fermion doublets is inconsistent due to the so-called Witten anomaly. The origin of the problem is that there are physically inequivalent field configurations in the quantum theory that correspond to the same classical theory. We argue that any gauge group with nontrivial fourth homotopy group has the same type of inconsistency.

INTRODUCTION

An anomaly is a symmetry of the classical theory that fails at the quantum level. Here, we are in particular interested in global anomalies, which arise when the partition function fails to be invariant under part of the gauge group (the "large" part) which actually yields a different physical configuration.

Let's understand this statement. Suppose we have a disconnected gauge group G: it consists of a subgroup G_0 of transformations that are continuously connected to the identity–and thus can be written as infinitesimal gauge transformations–and a disconnected discrete subgroup which is not. A gauge transformation which is not continuously connected to the identity as such is called "large". It turns out that when this is the case, we cannot naively obtain our physical space of states by modding out the theory's Hilbert space H by the full gauge group G–doing so would group together states related by large gauge transformations that are physically distinct in the full quantum theory. Perhaps in this sense, it is misleading to really call "large gauge transformations" gauge transformations at all.

An intuitive way to understand these large gauge transformations is in the language of homotopy groups. Basically, two functions are said to be homotopic if they can be continuously deformed into each other. Homotopy groups are a way to classify how one can draw loops that can be continuously deformed into one another in a topological space.

In particular, we will be considering the gauge group $SU(2)$, and the relevant fact for our purposes is that the fourth homotopy group of $SU(2)$ is nontrivial:

$$\pi^4(SU(2)) = \mathbb{Z}_2. \tag{1}$$

That (1) is true means that $SU(2)$ contains two types of gauge transformations, those in the trivial homotopy class that are continuously connected to the identity (the subgroup I've called G_0 above), and those in the nontrivial homotopy class, call them U, which cannot be connected to the identity by $SU(2)$ transformations.

Witten was the first to understand a global anomaly, and in his classic paper [4] found that the inconsistency which had previously puzzled theorists in an $SU(2)$ gauge theory with an odd number of Weyl fermions was due to the above over-counting of the physical Hilbert space. In the following section, we show how this ambiguity arises. For the reader that requires a review of how to construct a Lagrangian for Weyl fermions, see an extended discussion in the appendix.

WITTEN'S ANOMALY

Consider the path integral for a gauge theory with $G = SU(2)$, and a single doublet of left handed fermions,

$$\begin{aligned} Z = &\int d\psi \, d\bar{\psi} \int dA_\mu \\ &\times \exp\left[\int d^4x \left(-\frac{1}{4}\mathrm{Tr}\ F_{\mu\nu}^2 + \bar{\psi} i \not{D} \psi\right)\right]. \end{aligned} \tag{2}$$

Above, D is the Dirac operator for the $SU(2)$ gauge theory as defined in (9). Performing the integral over the fermions [5], we find

$$\int d\psi \, d\bar{\psi} \ \exp\left[\bar{\psi} i \not{D} \psi\right] = \pm(\det\ i\not{D})^{\frac{1}{2}}. \tag{3}$$

All the trouble is in which sign to take for the square root. In particular, consider a gauge transformation in the non-trivial homotopy class of $SU(2)$, A_μ^U, related to a gauge transformation in the trivial sector, A_μ, by

$$A_\mu^U = U^{-1} A_\mu U - i U^{-1} \partial_\mu U. \tag{4}$$

One can show that these gauge transformations will necessarily lead to the sign difference

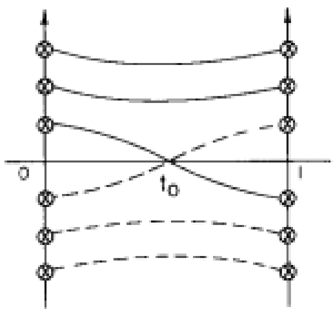
$$(\det\ i\not{D}[A_\mu])^{\frac{1}{2}} = -(\det\ i\not{D}[A_\mu^U])^{\frac{1}{2}}. \tag{5}$$

In other words, such a prescription for the sign of the square root of the Dirac operator is not gauge invariant. If you're happy to take this statement as a given, skip the box below; otherwise, read through for an overview of the technical details.

Important technical aside:

Consider the operator on the right hand side of (3), which is formally a product of half of all the eigenvalues of the Hermitian operator $i\not{D}$–half because of the square root. All of the eigenvalues are real, and for every eigenvalue λ there is an eigenvalue $-\lambda$ [6]. In choosing which half of the eigenvalues to include in (3), we may define $(\det i\not{D})^{1/2}$ to include either the plus or minus eigenvalue for each pair $(\lambda, -\lambda)$.

Just to start somewhere, let's define $(\det i\not{D})^{1/2}$ to be the product of the positive eigenvalues for the particular gauge field A_μ in the trivial sector of $SU(2)$. Now, imagine varying $A_\mu \to A_\mu^U$ continuously in field space.



The spectrum of $i\not{D}$ is precisely the same for both A_μ and A_μ^U, but it turns out that the individual eigenvalues themselves rearrange along such a variation. See the figure above for an illustration of such a rearrangement, where the sign for the square root of the Dirac operator is defined by the sign of the product of the eigenvalues indicated by the solid lines.

This is where we employ a useful piece of supermath, called the Atiyah-Singer index theorem, to argue that the number of eigenvalues of the operator in question that cross zero along such a smooth variation is always odd. Proving this is beyond the scope of this paper, but refer to [4] for more details, or [2] for a different take on the proof. The basic idea is that the index theorem tells you how to count zero modes (zero eigenvalues) of the Dirac operator, which one can then relate to the eigenvalue flow along the variation. This means that you will always pick up a minus sign going between the two sectors of gauge transformations, as in (5).

What our considerations above tell us is that $(\det\ i\not{D})^{1/2}$ is odd under the topologically non-trivial gauge transformation U. In general, for an odd number of left-handed fermion doublets n, the RHS of (3) becomes $(\det\ i\not{D})^{n/2}$, and the theory suffers from this sign inconsistency.

Now that we've established the inconsistency, we need to answer the question of why it's so bad. To answer this, consider the partition function for the $SU(2)$ theory with a single left-handed doublet by plugging (3) into (2),

$$Z = \int dA_\mu (\det\ i\not{D})^{1/2} \exp\left(-\frac{1}{4}\int d^4x\ \mathrm{Tr}\ F_{\mu\nu}^2\right). \quad (6)$$

This would vanish identically because the contribution of any gauge field A_μ is exactly cancelled by the equal and opposite contribution of A_μ^U. Clearly, such a theory with vanishing partition function is ill-defined.

One can then generalize this result to other gauge groups. As discussed in the introduction, the way to identify that $SU(2)$ in four dimensions has a subgroup of large gauge transformations is to note that its fourth homotopy group is nontrivial. Since $\pi^4(SU(N>2)) = 0$, but $\pi^4(Sp(N)) = \mathbb{Z}_2$ for all N, this anomaly holds as well for an $Sp(N)$ gauge theory with an odd number of left-handed Weyl fermions in the fundamental representation.

CONCLUSION

We've found that an $SU(2)$ gauge theory with an odd number of left-handed fermion doublets is inconsistent, which comes from the observation that the sign for the square root of the Dirac operator cannot be defined in a smooth and gauge invariant way.

One application of this result is that a global anomaly in a chiral gauge theory makes it impossible to formulate such a theory consistently on the lattice; the lattice analogue of the anomaly results in an inconsistency in defining the fermion measure in the path integral [1].

Appendix A: Weyl spinors

Since the whole content of this paper is the partition function for a quantum field theory of Weyl spinors, let's take some time to understand what a spinor is–just to make sure we're all on the same page.

An equation of motion is automatically Lorentz invariant if it follows from a Lagrangian that is a Lorentz scalar. Thus, in quantum field theory we wish to study the Lorentz transformation properties of quantum fields, such that we can put them together into a Lagrangian field theory. A useful way to label representations of the Lorentz algebra is to first note the fact that the Lie algebra of the Lorentz group splits into two copies of the $SU(2)$ Lie algebra. We know well how to classify irreducible representations of the $SU(2)$ algebra from quantum mechanics: they are labeled by $l = 0, \frac{1}{2}, 1, ...$, and each has dimension $2l+1 = 1, 2, 3, ...$. Thus each of the irreducible representations is classified by the numbers (s_+, s_-) where s_+ and s_- take values $0, \frac{1}{2}, 1, ...$, and the dimension of the irrep is $(2s_+ + 1)\cdot(2s_- + 1)$.

A spinor field, by definition, is a field that transforms as a spinor under Lorentz transformations. The representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ are of dimension 2, and act on spin-1/2 spinors.

I'll skip the details of the construction except to point out that it involves use of the Clifford algebra,

$$\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}, \tag{7}$$

which is why the matrices γ^μ show up in the spinor Lagrangian. The bottom line is that we can construct a dimension 2 representation of the Lorentz group corresponding to each of the representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$, where these representations act on 2-component vectors called Weyl spinors. These vectors are known as left-handed or right-handed chiral Weyl spinors respectively, commonly written ψ and $\bar{\chi}$, where I've suppressed the usual spinor indices.

One can go through the construction of a Lagrangian for these objects by insisting that it is composed of the fields ψ and $\bar{\chi}$, their Hermitian conjugates, and first derivatives of these, and that it is Hermitian and Lorentz invariant. Then, one promotes the theory to have an $SU(2)$ gauge symmetry, thus changing partial derivatives to covariant derivatives and introducing the vector field A_μ. The final Lagrangian for the massless left-handed Weyl fermions in an $SU(2)$ gauge theory is

$$\mathcal{L} = \bar{\psi} i \not{D} \psi \tag{8}$$

where D is the covariant derivative,

$$D_\mu = \partial_\mu - igA^i_\mu \frac{\sigma^i}{2}. \tag{9}$$

We use the Feynman slash notation, $\not{D} = \gamma^\mu D_\mu$. One can combine the 2-component Weyl spinors into a 4-component vector, called a Dirac spinor Ψ:

$$\Psi = \begin{pmatrix} \psi \\ \bar{\chi} \end{pmatrix}. \tag{10}$$

A single set of Dirac fermions in the fundamental representation of $SU(2)$ (a doublet of Dirac fermions) is exactly the same as 2 left-handed Weyl doublets–this why the integral (3) is evaluated to be the square root of the same integral done for Dirac fermions.

[1] O. Bär, I. Campos, "Global anomalies in chiral gauge theories on the lattice," Nucl. Phys. B **581**, 499 (2000).

[2] F.R. Klinkhamer, "Another look at the SU(2) anomaly," Phys. Lett. B **256**, 41 (1991).

[3] J. McGreevy, "Physics 239a: Where do quantum field theories come from?" UCSD Lecture Notes (2014).

[4] E. Witten, "An SU(2) Anomaly," Phys. Lett. B **117**, 324 (1982).

[5] The basics of Grassmann calculus are reviewed in p.67 of the lecture notes [3], and for further review consult Peskin and Schroeder.

[6] One can see this by noting that if $i\not{D}\psi = \lambda\psi$, then $i\not{D}(\gamma_5\psi) = -\lambda(\gamma_5\psi)$.