International Journal of Theoretical Physics, Vol. 39, No. 2, 2000

U(1) Puzzle and the Strong CP Problem from a Holonomy Formulation Perspective

Hugo Fort[1] and R. Gambini[1]

Received July 2, 1999

We analyze the issue of a complete description of gauge field theories in terms of holonomies or nonlocal gauge invariants. In particular, we show that a formulation of QCD in terms of holonomies does not exhibit the strong CP problem, and at the same time solves the $U(1)_A$ puzzle from the very beginning.

1. INTRODUCTION

In physical terms, holonomies are related to parallel transport along a given path of a vector field. In the case of pure gauge theories, it is enough to consider closed paths or loops, while when matter is present one needs to include open paths connecting the matter fields. Since Yang [1] noticed in the seventies the important role of holonomies for a complete description of gauge theories, they have been increasingly used both in particle physics [2–5] and quantum gravity [6,7].[2] In the case of quantum gravity, loops give a natural geometrical description of the space-time at the Planck scale [8]. Therefore, a description of gauge theories in terms of holonomies, besides the general advantage of only involving gauge-invariant quantities (under small and large transformations), is appealing because it provides a common geometrical framework to treat gauge theories and quantum gravity.

The simplest case of electromagnetism is illustrative. At the classical level, a complete description of electromagnetism is given by the field strength tensor $F_{\mu\nu}$. Quantum mechanically, $F_{\mu\nu}$ underdescribes electromagnetism (Aharonov–Bohm effect), while the gauge potential A_μ overdescribes it. It turns out that the holonomy or loop-dependent phase factor $H(C) = \exp[ie$

[1] Instituto de Física, Facultad de Ciencias, Iguá 4225, 11400 Montevideo, Uruguay.
[2] See ref. 7 for a general review of the loop approach.

0020-7748/00/0200-0341$18.00/0

$\oint_C A_\mu\, dx_\mu$] contains the necessary and sufficient information about electromagnetism at the first quantization level. Pushing one step further and taking the holonomies instead of the potential as a complete description of Yang–Mills theory at the second quantized level poses general questions: Given the fact that loop invariants do not detect large gauge transformations, do holonomies provide a different physical picture of gauge theories than the one which emerges from the conventional formulation in terms of vector potentials or connections? Are holonomies flexible enough to capture all the observed physical phenomena? Is this formalism free from the known problems which lurk in the standard approach?

In particular, in this article we will show that a description in terms of holonomies provides an entirely different point of view on the nature of the $U(1)_A$ anomaly and the strong CP problem. Specifically, (i) there is no hidden θ_{QCD} parameter coming from the strong interaction sector, and (ii) the absence of a Goldstone boson for the broken $U(1)_A$ is straightforward without explicitly resorting to the instanton-based mechanism. Moreover, several recent calculations involving the inclusion of fermions in a loop formulation [9–11] support that picture for the two previously mentioned problems.

Let us briefly review the standard formulation of pure Yang–Mills theory. This theory exhibits a nontrivial topological structure which is manifest through the existence of *large* gauge transformations—characterized by a topological integer or *winding number* n—in addition to the ordinary *small* gauge transformations (generated by Gauss' law) with $n = 0$. This implies that there is an infinite set of degenerate vacuum states, each labeled by its index n. Instanton solutions provide a mechanism of "vacuum tunneling" between topological inequivalent n-vacua [12]. So the "true" vacuum—the so-called θ vacuum—is a linear superposition of n-vacua [13] and as a consequence the theory possesses a hidden parameter, the vacuum angle θ. When fermion fields are coupled to the gauge fields in quantum chromodynamics (QCD) the nontrivial topological structure of the vacuum has two remarkable effects. The first effect that was soon pointed out [12] is that it offers a solution of the "$U(1)_A$ problem." In a nutshell the $U(1)_A$ puzzle is: the approximate axial symmetry is known to be broken, so *where is the corresponding quasi-Nambu-Goldstone boson*? The answer to this question is a little bit involved and can be traced to the fact that $U(1)_A$ is broken, by virtue of the axial anomaly, to the discrete symmetry Z_{2N_f}, where N_f is the number of flavors. This incomplete breaking of the axial symmetry opens the possibility that this breaking is not accompanied by a Nambu–Goldstone boson [14]. This solution does not depend on the value of the θ parameter associated with the nonperturbative QCD vacuum. The second effect is that a nonzero value of θ implies a violation of CP invariance in strong interactions. This can be understood as follows: the rich structure of Yang–Mills vacuum corresponding to tunneling between states of different winding number gives rise to an effective Lagrangian term proportional to θ times the Chern–Pontryagin density $F \wedge F$, which violates P and CP conservation. Strong interaction processes conserve the CP symmetry, therefore the "strong CP" problem is: *why is* $\theta = 0$? The same θ vacuum that solves the $U(1)_A$ problem in QCD creates the strong CP problem. It is important to note that the existence of the θ vacua and the value of the θ parameter are different questions. Different solutions to avoid the CP problem have been considered, but all present some drawback. A first alternative is to solve the problem by postulating that one of the quark masses (presumably m_u) is equal to zero. However, a massless quark contradicts the current algebra calculations of the quark masses [15]. A second proposal, and by far the most popular approach, is the Peccei–Quinn (PQ) mechanism [16], which introduces a new additional chiral $U(1)$ symmetry which allows one to rotate the θ parameter to zero. Unfortunately, a by-product of this mechanism is the generation of a Nambu–Goldstone boson, the *axion*, which has eluded detection. A third proposed solution is simply to set $\theta = 0$ based on mathematical grounds [17]. However, this proposal relies on a formulation of QCD which is still not completely settled (see, e.g., ref.14).

This article is organized as follows. In Section 2 we point out the main differences resulting from a formulation in terms of holonomies: absence of a θ hidden parameter and breakdown of the axial symmetry from the very beginning. In Section 3 we illustrate the points with the simplest toy model which mimics QCD: the Schwinger model or $(1 + 1)$ QED. Section 4 is devoted to conclusions and final remarks.

2. THE VACUUM IN THE HOLONOMY REPRESENTATION AND THE AXIAL SYMMETRY BREAKDOWN

The standard definition of QCD is in terms of local fields, quarks and gluons, but the physical excitations are extended composites: mesons and baryons. Thus, a quantum formulation of gauge theories directly in terms of the nonlocal gauge invariants associated to the above extended physical excitations seems to be the natural one. It is worth noting that, obviously, in this formulation the distinction between large and small gauge transformations is meaningless: the states are invariant under both. In fact, the states may be considered as linear combinations of Wilson loops, and consequently, due to the cyclic property of the trace, they are invariant under small and large gauge transformations. Therefore, the generator of large gauge transformations is trivially equal to one, and the vacuum degeneration is absent. Thus *the vacuum is unique and no* θ_{QCD} *parameter is hidden*.

Before continuing with our analysis, it is worth mentioning that, even in the holonomy description, it is always possible to introduce a θ parameter by hand if one needs (for some particular reason) to do that. Let us recall that, from the canonical point of view, in the ordinary representation the generator of large gauge transformations has a nontrivial action on the wavefunctions of the pure gauge theory given by [18, 19]

$$\Omega_n \Psi[A] = \Psi[A_{g_n}] \tag{1}$$

where A_{g_n} is obtained by acting on A with a large gauge transformation with winding number n. The Gauss law does not enforce the invariance under this type of transformations. Now, as the Hamiltonian $\hat{H}$ and the unitary operator Ω_n commute, they are simultaneously diagonalizable

$$\Omega_n \Psi_\theta[A] = \exp[i\theta n]\Psi_\theta[A] \tag{2}$$

$$\hat{H}\Psi_\theta[A] = E_\theta \Psi_\theta[A] \tag{3}$$

For a fixed value of θ it is possible to introduce a change of variable of the corresponding Ψ_θ,

$$\Phi_\theta[A] = \exp[-i\theta S_{CS}[A]]\Psi_\theta[A] \tag{4}$$

such that Φ is invariant under both small and large gauge transformations and satisfies the following eigenvalue equation:

$$\hat{H}_\theta \Phi_\theta[A] = E_\theta \Phi_\theta[A] \tag{5}$$

Notice that now the Hamiltonian depends on θ. It may be obtained by following the usual canonical procedure from the action $S_\theta = S_{\mathrm{YM}} + \theta \int F \wedge F$, where the second term is the Chern–Pontryagin topological invariant. This term does not modify the field equations because it only adds a surface contribution. In the holonomy representation Ω_n is a trivial operator, proportional to the identity, and one only has a description of one of the gauge-invariant sectors with wavefunctions Φ characterized by a value of θ. While in the standard approach, if we start with an action with $\theta = 0$, the θ vacuum, associated to large transformations, still appears, in the holonomy description of QCD there is a CP violation term only if it is introduced by hand.

To analyze the solution to the $U(1)_A$ problem, the fermionic degrees of freedom must be included in the holonomy formulation. This was done some years ago by including gauge-invariant hadronic objects built on open paths, in addition to the closed ones or loops for the pure gauge theory, giving rise to the so-called *P-representation* [20].[3] The $U(1)_A$ puzzle is solved in this

[3] The P is for paths, which, in this case when matter fields are present, are in general open. In what follows, although strictly holonomies are defined for closed paths, we will use the term holonomy representation indistinguishably from P-representation.

description straightforwardly without resorting explicitly to topological instantons (the effects of the tunneling instantons are implicitly included in this description).

Let us review the main features of the P-representation. To introduce a gauge-invariant basis, one starts by considering an overcomplete (by virtue of Mandelstam's relations) set of gauge-invariant operators of the form $\psi_d^\dagger(x)H(P_x^y)\psi_u(y)$, where $H(P_x^y)$ is the holonomy associated with the open path P_x^y going from x to y, and ψ_d and ψ_u are respectively the up and down components of the spinor.

The above spatial decomposition of fermionic degrees of freedom is a standard procedure in the lattice approach, namely the staggered or Susskind fermions [21]. It turns out that the lattice is the natural arena to discuss the P-representation of 4-dimensional QCD. In fact, a mathematically rigorous 4-dimensional formulation only exists on the lattice and it is in this framework where a comparison with the standard approach in terms of fields makes sense. Furthermore, the lattice is the main approach for nonperturbative physics. Staggered fermions consist of one single-component fermion field $\phi(n)$ defined at each site n such that

$$\begin{aligned} \psi_u(n) &= \phi(n), \qquad n \text{ even} \\ \psi_d(n) &= \phi(n), \qquad n \text{ odd} \end{aligned} \tag{6}$$

This definition turns out to be equivalent to considering gauge-invariant operators starting at even sites and ending at odd sites. This set of operators defines the path basis $|P\rangle$. Notice that the choice of this basis automatically breaks the remnant discrete chiral invariance of the usual lattice staggered fermion formulation [21]. Now a lattice translation by odd integers of a basis vector is not a basis vector. Thus we see that the anomaly responsible of this breaking is *intrinsic* to the P-representation. This fact is not unexpected. As is well known [18], the anomaly occurs as a consequence of the incompatibility of two classical symmetries—gauge and chiral invariance—at the quantum level. It happens that the gauge symmetry may only be preserved at the price of sacrificing the chiral symmetry which become anomalously broken. The P-representation deals with gauge-invariant quantities and hence has no chance to implement the chiral symmetry. The same happens in the recently proposed Lagrangian counterpart of the P-representation, the *worldsheet formulation* with dynamical fermions [22]. The worldsheet partition function of lattice QED Z_P is a sum over the worldsheets of strings or paths of the P-representation. That is, surfaces such that (1) their borders (fermion worldlines) are self-avoiding polymer-like loops and (2) when intersected with a time $t =$ const plane they produce paths beginning at even sites and ending at odd ones. That is, the contributions to the partition function are not invariant under odd translations, so that there is no remnant of the chiral symmetry.

To end this section, we point out that in the conventional lattice formulation there is a twofold degeneracy connected with chiral symmetry, i.e., there are two vacuum states which transform into one another by interchanging odd and even sites [23]. In order to compute the hadron spectrum the procedure is to modify the Hamiltonian by adding an *irrelevant* term (i.e., an operator which has no effect in the continuum limit) such that it renders the vacuum well determined. On the other hand, the gauge-invariant P-representation selects one of the two possible chiral sectors from the outset. This is consistent with the fact that, in the continuum limit, both sectors are separated by an infinite gap for any value of the quark mass, so that the value of the physical observables should not be affected. Indeed this was confirmed for QCD [24] and for the Schwinger model [10, 11].

3. ILLUSTRATION: THE SCHWINGER MODEL

To illustrate the previous points, we shall resort to a simple model: $(1 + 1)$ QED with massless fermions or the Schwinger model. This model is rich enough to share with 4-dimensional QCD the issues we are concerned with, namely the topological structure which gives rise to the θ vacuum and the breaking of the chiral symmetry with an axial anomaly. The Schwinger model in the P-representation has been studied both in the continuum [9] and on the lattice [10, 11].

The analytical continuum study of ref. 9 showed that the divergence of the axial current is nonzero. Nevertheless, it is not clear if it is possible to cast it as the divergence of a Chern–Simons density.

A quantity which provides useful information is the chiral condensate $\langle\bar{\psi}\psi\rangle$, the order parameter of chiral symmetry. The lattice chiral condensate per-lattice site is defined as

$$\langle\bar{\chi}\chi\rangle = \frac{1}{2N_a} \sum_x (-1)^{x_1} \langle[\hat{\chi}^\dagger(x), \hat{\chi}(x)]\rangle$$

where N_s is the number of lattice sites. The corresponding operator is realized in the P-representation and thus we get for the chiral condensate [20]

$$\langle\bar{\chi}\chi\rangle = \frac{1}{2} - \frac{2\mathcal{N}_P}{N_s} \tag{7}$$

where $\mathcal{N}_P$ is the number of connected paths at a given time t. In ref. 10, using a Hamiltonian finite lattice analysis, a nonnull chiral condensate completely consistent with the known value in the continuum

$$\beta^{1/2}\langle\bar{\psi}\psi\rangle = \frac{e^{\gamma}}{2\pi^{3/2}} = 0.15995 \tag{8}$$

is found, where $\beta = 1/e^2$ (the inverse of the square of the coupling constant e) and γ is the Euler constant. This nonzero value of the chiral condensate is a manifestation of the axial anomaly.

Using the worldsheet formulation of ref. 22, we can express $Z_P^{\text{Schwinger}}$

$$Z_P^{\text{Schwinger}} = \sum_{S_{\mathcal{F}^c}} \exp\left\{-\frac{1}{2\beta}\sum_{p\in S_{\mathcal{F}^c}} n_p^2\right\} \tag{9}$$

where n_p is an integer variable attached to plaquettes (a 2-form) and $\mathcal{F}^c$ denote the fermionic worldline borders of the worldsheets. Figure 1 shows a configuration of worldsheets enclosed by self-avoiding fermionic loops $\mathcal{F}^c$ and Fig. 1b the paths of the P-representation we get from them at different time slices. The recipe is very simple: we get a link of $P(t)$ for every plaquette of the surface which connects the slice t with $t + 1$.

By virtue of Equation (7) it is easy to calculate directly the chiral condensate simply by counting the number of "electromesons" we have when we intersect their world sheets with each time slice t. A Monte Carlo simulation of this model [11] showed that the chiral symmetry is broken for the *strictly* massless case and it produces a chiral condensate which once more converges in the weak coupling (continuum) limit to its known exact value. Note that this is a clear difference to what happens in an ordinary simulation in terms of fields, for which in the massless case, given enough time, the system rotates through all the degenerate minima so that $\langle\bar{\psi}\psi\rangle = 0$. Therefore, one has to calculate this order parameter for several small nonzero masses m to get the sensible limit at $m \to 0$. On the contrary, as in the lattice holonomy representation there is not a discrete chiral symmetry remnant from the continuous one, it is enough to study the massless case.

All this evidence leads us to conclude that there is no $U(1)_A$ problem in the P-representation; in fact this representation does not bear $U(1)_A$ symmetry at the second quantized level.

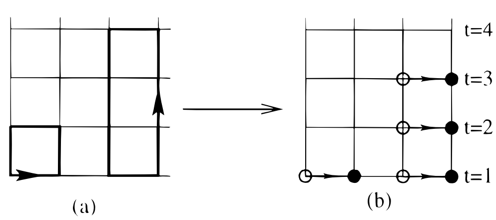

Fig. 1. (a) A possible configuration of self-avoiding paths on a 4 × 4 lattice. Fermionic loops are represented by heavy lines. (b) The corresponding stringlike excitations.

4. CONCLUSION AND FINAL REMARKS

Our general purpose is to study the consistency of the loop formulation as a general framework for particle physics and quantum gravity. By virtue of the growing interest in this formulation in both fields, we believe the issue deserves an analysis.

In the standard formalism, which takes the gauge potentials as dynamical variables, the resolution of the strong CP problem means finding a physical mechanism which constrains the θ parameter to be zero. This was not our goal. Basically, our statement is that holonomies can be taken as the real dynamical variables since all the nonperturbative known results can be reproduced. For the case of QCD, this description on one hand avoids the emergence of the strong CP problem and on the other hand, by virtue of the necessary staggering of the fermionic degrees of freedom, solves in a trivial way the $U(1)_A$ puzzle. We stress that we are not claiming that the $U(1)_A$ symmetry is broken for any theory with massless fermions, independently of its particle content or space-time dimensions, which is not correct. Our point concerns the consistency of QCD or QED when formulated in terms of holonomies in four dimensions.

Historically the physical interactions were formulated quantum mechanically in terms of gauge potentials. The success of the standard quantum field theory is closely tied to the impressive results collected by the subsequent applications of perturbation theory. However, later it was realized that gauge theories support a rich nonperturbative structure. It was observed that in order to capture the physics which is beyond perturbation theory, say by using the lattice approach, the natural formulation is the compact one, in terms of phase factors $U(x)_\mu = \exp[igA_\mu(x)]$. The ordinary lattice formalism works with gauge-invariant quantities, basically the phase factors attached to lattice plaquettes p, i.e., U_p, but it maintains the gauge redundancy in the measure (the integration is performed over the $A_\mu(x)$]. Hence, we propose to take an additional step and to avoid the gauge redundancy by integrating directly over gauge-invariant variables associated to the geometrical paths of the excitations, namely loops (pure gauge field excitations) and open ones (baryons and mesons).

It is interesting to speculate what would happen if from the beginning holonomies were used to describe the physical interactions instead of vector potentials. Probably we would not be discussing the strong CP problem. This would be simply considered as an artifact of an overdescription of nature, by means of gauge potentials, which is still necessary in order to compute quantities by using the powerful perturbative techniques. From this perspective, the strong CP problem is just a matter of how we describe nature rather than being a feature of nature itself.

To conclude, at first sight it might seem, according to the orthodox point of view, that a formalism in which the θ angle does not arise is a formalism not sufficiently flexible to describe all the physical possibilities. However, we have shown that a description in terms of holonomies allows one to recover the standard physical results and as a bonus it does away with the strong CP problem.

ACKNOWLEDGMENTS

We were supported in part by CONICYT, Projects No. 49 and No. 318. We are indebted to A. Ashtekar, G. Gonzalez-Sprinberg, R. Mendez, and M. Reisenberger for valuable discussions.

REFERENCES

1. C. N. Yang, *Phys. Rev. Lett.* **33** (1974) 445.
2. A. M. Polyakov, *Phys. Lett. B* **82** (1979) 249; *Nucl. Phys. B* **164** (1979) 171.
3. Yu. Makeenko and A. A. Migdal, *Phys. Lett. B* **88** (1979) 135.
4. R. Gambini and A. Trias, *Phys. Rev. D* **22** (1980) 1380; *Nucl. Phys. B* **278** (1986) 436.
5. A. M. Polyakov, *Nucl. Phys. B* **486** (1997) 23.
6. C. Rovelli and L. Smolin, *Phys. Rev. Lett.* **61** (1988) 1155; *Nucl. Phys. B* **331** (1990) 80.
7. R. Gambini and J. Pullin, *Loops, Knots, Gauge Theories and Quantum Gravity,* Cambridge University Press, Cambridge (1996).
8. A. Ashtekar, *Lectures on Non-Perturbative Canonical Gravity,* World Scientific, Singapore (1991).
9. R. Gambini, H. Morales-Técotl, L. Urrutia, and J. Vergara, *Phys. Rev. D* **57** (1998) 3711.
10. J. M. Aroca, H. Fort, and G. Alvarez, *Europhys. Lett.* **45** (1999) 565.
11. H. Fort, *Phys. Lett. B* **444** (1998) 174.
12. G. ’t Hooft, *Phys. Rev. Lett.* **37** (1976) 8.
13. C. Callan, R. Dashen, and D. Gross, *Phys. Lett.* **63B** (1976) 334; R. Jackiw and C. Rebbi, *Phys. Rev. Lett.* **37** (1976) 172.
14. R. E. Marshak, *Conceptual Foundations of Modern Particle Physics*, World Scientific, Singapore (1993).
15. H. Leutwyler, *Nucl. Phys. B* **337** (1990) 108.
16. R. Peccei and H. Quinn, *Phys. Rev. Lett.* **38** (1977) 1440.
17. T. Kugo and I. Ojima, *Prog. Theor. Phys. Suppl.* **66** (1979) 1.
18. R. Jackiw, In *Relativity Groups and Topology II*, North-Holland, Amsterdam (1984).
19. K. Huang, *Quarks, Leptons and Gauge Fields*, World Scientific, Singapore (1993).
20. H. Fort and R. Gambini, *Phys. Rev. D* **44** (1991) 1257.
21. L. Susskind, *Phys. Rev. D* **16** (1977) 3031.
22. J. M. Aroca, H. Fort, and R. Gambini, *Phys. Rev. D* **57** (1998) 3701.
23. T. Banks, D. R. T. Jones, J. Kogut, S. Raby, P. N. Scharbach, D. K. Sinclair, and L. Susskind, *Phys. Rev. D* **15** (1977) 1111.
24. R. Gambini and L. Setaro, *Nucl. Phys. B* **448** (1995) 67.