



Geometric foundations of classical Yang–Mills theory

Gabriel Catren

CREA, UMR 7656, École Polytechnique/CNRS, 1 rue Descartes, 75005 Paris, France

Received 27 July 2007; received in revised form 15 November 2007

Abstract

We analyze the geometric foundations of classical Yang–Mills theory by studying the relationships between *internal relativity*, *locality*, *global/local invariance*, and *background independence*. We argue that *internal relativity* and *background independence* are the two independent defining principles of Yang–Mills theory. We show that local gauge invariance—heuristically implemented by means of the *gauge argument*—is a direct consequence of internal relativity. Finally, we analyze the conceptual meaning of BRST symmetry in terms of the invariance of the gauge fixed theory under *general* local gauge transformations.

© 2008 Elsevier Ltd. All rights reserved.

Keywords: Yang–Mills theory; Gauge theories; BRST symmetry

When citing this paper, please use the full journal title *Studies in History and Philosophy of Modern Physics*

1. Introduction

In order to introduce Yang–Mills theory a formal procedure known as the *gauge argument* is commonly followed. This argument proceeds as follows. One begins with a free

E-mail address: gabrielcatren@yahoo.com.ar

fields theory that is invariant under a *global* symmetry of the form $\psi(x) \rightarrow e^{i\alpha}\psi(x)$. A “relativistic locality” is then invoked in order to justify the necessity of defining a theory with a *local* symmetry of the form $\psi(x) \rightarrow e^{i\alpha(x)}\psi(x)$, where the parameter α becomes dependant on space–time M .¹ Nevertheless, the precise physical meaning of the local gauge invariance requirement is still the object of an ongoing debate.² The common strategy is to justify this requirement by means of some sort of *locality principle*. However, far from being a univocally defined notion, there are many alternative definitions of locality. Even if the original Lagrangian $\mathcal{L}_0(\psi, \partial_\mu\psi)$ is not invariant under local gauge transformations, there exists a simple formal procedure which renders the theory locally invariant. One needs only to substitute the ordinary derivatives ∂_μ with covariant derivatives $D_\mu = \partial_\mu + A_\mu(x)$ in the Lagrangian $\mathcal{L}_0(\psi, \partial_\mu\psi)$ of the original non-interacting fields. If one assumes that the field A_μ transforms into $A'_\mu = g(x)^{-1}A_\mu g(x) + g(x)^{-1}\partial_\mu g(x)$ under a transformation defined by an element $g(x) = e^{i\alpha(x)}$ of a Lie group G , then the new Lagrangian $\mathcal{L}(\psi, D_\mu\psi)$ is invariant under these local gauge transformations. We can thus conclude that in order to guarantee local gauge invariance, it is necessary to introduce new fields, namely the so-called *gauge field* $A_\mu = A_\mu^a T_a \in \Omega^1(M) \otimes \mathfrak{g}$, where T_a is a basis of the Lie algebra \mathfrak{g} of G . The remarkable result is that the gauge field A_μ can be identified with the potential of a fundamental physical interaction (as for instance the electromagnetic interaction if $G = U(1)$). Under this identification, the new terms in the Lagrangian $\mathcal{L}(\psi, D_\mu\psi)$ represent the minimal coupling between the matter and the gauge fields. Hence, it is tempting to conclude that forcing local gauge invariance makes it necessary to introduce physical interactions. However, in order to find the complete Lagrangian of the matter field minimally coupled to a *dynamic* gauge field, it is still necessary to add a kinetic term for the gauge field of the form $-\frac{1}{4}\text{Tr}(F_{\mu\nu}F^{\mu\nu})$, with $F_{\mu\nu} = [D_\mu, D_\nu]$. Therefore, strictly speaking it is not true that a dynamic gauge field can be derived from the requisite of local gauge invariance. Following this and other arguments, some authors questioned the physical non-triviality of the gauge argument and the belief that it encompasses a fundamental principle of nature (see Martin, 2002).³ This criticism can be traced back to Kretschmann’s classic objection to the physical content of general covariance in general relativity (see Norton, 1992, 1993, 2003). The transition from what seems to be nothing more than an “epistemic constraint” (namely that the theory’s observational content has to be invariant under local transformations of the coordinate system) to the introduction of a fundamental physical interaction is not straightforward. Therefore, far from explaining the rational link between symmetries and interactions, the gauge argument’s heuristic value remains to be clarified.

In this paper, we will argue that *internal relativity* and *background independence* are the two independent defining principles of Yang–Mills theory. By means of the fiber bundle formalism, we will show that Yang–Mills theory depends on the ontological assumption

¹This is the kind of argument presented in the original seminal paper of Yang and Mills: “It seems that this [a *global symmetry*] is not consistent with the localized field concept that underlies usual physical theories.” (Yang & Mills, 1954, p. 192).

²See Martin (2002) for a criticism of the different justifications for the demand of local gauge invariance and O’Raifeartaigh (1995) for a historical account.

³The main arguments are the following: (1) the requisite of local gauge invariance is not clearly justified, (2) it is necessary to add a kinetic term for the gauge field by hand, (3) the resulting Lagrangian is not uniquely fixed by the gauge argument and (4) the final structure of the theory depends on other principles, as for example the requisite of renormalizability.

that geometric objects in the corresponding internal spaces (the fibers) are not *absolute geometric objects*, that is to say they cannot be intrinsically discerned. Since an active transformation of geometric objects in internal spaces can always be interpreted as a passive transformation of the coordinate system, the specification of the field's internal state at a certain position does not have a gauge invariant meaning. This property of internal spaces will be called *internal relativity*. Since internal states cannot be intrinsically discerned, different fibers cannot be canonically identified. This property of fiber bundles will be called *locality*. In such a geometric setting, the bases in the different fibers can be modified independently of one another. Therefore, the theory is naturally invariant under local gauge transformations. Local gauge invariance is thus the “epistemic” consequence (in the sense that it concerns our utilization of coordinate systems) of the non-trivial ontological assumption about the internal relativity of the matter fields' internal states.

Thanks to the fiber bundle formalism, we know that the gauge field A_μ describes a connection on a G -principal fiber bundle over space–time. However, the local theory that results from implementing the gauge argument depends on this arbitrarily fixed background connection (or *absolute object* in Anderson's words; see Anderson, 1967). In other words, the dynamics of matter fields is not relative to other degrees of freedom, but relative to a fixed background geometric structure. This means that a local theory is not necessarily a background independent theory. In order to render the theory background independent, it is necessary to include new degrees of freedom that describe the connection dynamics. This can be done by adding a kinetic term for the gauge fields to the Lagrangian $\mathcal{L}(\psi, D_\mu\psi)$. While general relativity succeeds in generalizing the background independence of external or spatiotemporal motion, Yang–Mills theory extends this requirement to the dynamics of matter fields' internal states.⁴

Since the quantum theory of gauge fields mixes quantum aspects with gauge aspects in a non-trivial way, in this work we will only treat classical Yang–Mills theory. It is true that Yang–Mills theory appeared for the first time in a quantum context (at least in the non-Abelian case). Nevertheless, the preliminary consideration of the classical theory allows us to distinguish between the gauge aspects coming from the classical theory (such as the *BRST symmetry*) and those features that are strictly of quantum nature (such as *unitary, renormalization and anomalies*).

In Section 2 we analyze the notions of *relative* and *absolute geometric objects*. In Section 3 we use the fiber bundle formulation of Yang–Mills theory in order to define *locality* and *internal relativity*. In Section 4 we consider the geometric conditions implicitly assumed in the framework of a theory with global gauge invariance. In Section 5 we analyze theories that are defined in a fiber bundle endowed with a connection. In Section 6 we analyze the physical significance of gauge freedom in Yang–Mills theory. In Section 7 we clarify the conceptual meaning of the BRST symmetry in Yang–Mills theory. In the final section we summarize the obtained results.

⁴It is sometimes stated that Einstein's theory of gravity failed to generalize the principle of relativity due to the irreducible absolute character of accelerated motion (see Friedman, 1983). Nevertheless, even these “absolute” motions are relative to a dynamic field of the theory, namely the gravitational field. In other words, “the inertial structure of the world”—which “is accepted as the cause for the dynamic inequivalence of motions”—is “*a real thing that not only exerts effects upon matter but in turn suffers such effects*. [...] The phenomena of gravitation thus divulge the flexibility of the field of inertia, or, as I prefer to call it, the ‘guiding field’, and its dependence on matter.” (Weyl, 1949, pp. 105–106). This way of considering general relativity was traditionally defended by Weyl (1949) and nowadays by Rovelli (2004).

2. Relative geometric objects

We will begin by analyzing the intrinsic nature of geometric objects. For the sake of simplicity, we will consider vectors belonging to a n -dimensional vector space V . A vector v —an “arrow”—does not depend on the particular basis s that we can arbitrarily choose in V . In a particular basis $\{\hat{e}_\mu\}_{\mu=1,\dots,n}$ in V , the vector v is expressed as $v = \hat{e}_\mu v^\mu$. We can also express v as $v = (\hat{e}_\nu A_\mu^\nu)(A_\rho^{-1\mu} v^\rho) = \hat{e}'_\mu v^{\mu'}$, with $A_\mu^\nu \in GL(n, \mathbb{R})$. The same geometric object v now has coordinates $v^{\mu'} = A_\rho^{-1\mu} v^\rho$ in the new basis s' defined by the vectors $\hat{e}'_\mu = \hat{e}_\nu A_\mu^\nu$. In this way, one can identify the intrinsic geometric object v with the equivalence class $[(s, v^s)]$ of pairs composed of a basis s and the components v^s of v in s . It is tempting to conclude that the demarcation line between the intrinsic geometric object v (the “arrow”) and its gauge-dependant coordinate representation has been successfully traced. Nevertheless, the standard distinction between the coordinate-independent—or intrinsic—mathematical object and the coordinate-dependent representation misses an important point which becomes relevant in the present context. Given a new set of coordinates $v^{\mu'} = A_\rho^{-1\mu} v^\rho$, it is *a priori* impossible to decide whether they are the components of the original vector v in the new basis $\hat{e}'_\mu = \hat{e}_\nu A_\mu^\nu$ (coordinate or passive transformation), or the components of a new vector v' in the old basis \hat{e}_μ (automorphism of V). In other words, given a coordinate transformation $s \rightsquigarrow s'$, one can always define an automorphism h such that the vector hv is given in the old coordinate system by the new coordinates of v . The automorphism h satisfies $s(hv) = s'(v)$, where $s: V \rightarrow \mathbb{R}^n$ with $s(v) = v^\mu$. Reciprocally, if h is an automorphism of V , one can always define an induced coordinate transformation $s \rightsquigarrow s'$ which assigns new coordinates to each vector v that are equal to the coordinates of hv in the original coordinate system. The new basis is defined by $s' = s \circ h$. We can thus conclude that an automorphism can always be trivialized in the form of a coordinate transformation.⁵ This well-known duality between coordinate transformations and automorphisms is a consequence of the fact that vectors in V lack intrinsic properties that could be used to discern them independently of the arbitrary election of a basis. This means that vectors in V , considered as intrinsic geometric objects, cannot be canonically individuated. It is worth stressing that this is an ontological statement about the intrinsic nature of vectors in V . The consequence of these considerations is that vectors in V cannot be interchanged in absolute terms, that is to say it is not possible to perform a “real” (non-trivializable) automorphism. In other words, vectors in V are unchangeable: an automorphism can always be reinterpreted as a relabelling of the vectors. This lack of intrinsic discernibility does not mean that geometric objects related by an automorphism should be identified. Even though elements of a vector space are intrinsically indiscernible, they are numerically distinct. If this were not so, we would lose the very notion of a vector space. This means that the leibnizian *identity of indiscernibles* is not valid here: two vectors are *different* and *indiscernible* geometric objects. Their difference is thus a primitive one, that is to say irreducible to properties. This means that in the present context the violation of the identity of indiscernibles is not due to the impossibility of an epistemic access to “metaphysical” properties. We could then say that “thisnesses are *metaphysically* primitive” (Adams, 1979, p. 10). In other words, a vector space is characterized by an

⁵The fundamental importance of the active interpretation of gauge transformations is a well-established fact in the case of general relativity (see the literature concerning the hole argument, as for example Earman & Norton, 1987).

irreducibility of “thisness” or “intuitive exhibition” to “suchness” or “conceptual determination”.⁶ This means that a vector space is a *primitive multiplicity* of indiscernible geometric objects that can only be individuated by means of a “primitive, nonqualitative thisness” (Adams, 1979, p. 10). Hence, we conclude that what we thought to be intrinsic geometric objects cannot be absolutely discerned in an ontological sense. In order to stress this intrinsic indiscernibility, we will call them *relative geometric objects*.⁷

On the other hand, if vectors in V were intrinsically discernible, then the coordinate transformations and the automorphisms could be distinguished by determining whether the basis or the vector had changed. Since an automorphism h would transform a vector v into another vector $v' = hv$ that can be intrinsically discerned from v , the automorphism h cannot be trivialized simply by being considered a mere coordinate transformation. We could imagine, for example, that each vector is colored differently. While a coordinate transformation changes the coordinates of each vector (but not its color), an automorphism assigns to each vector v another vector $v' = hv$ of a different color, that is to say intrinsically different. Now that the vectors differ in color, the identity of indiscernibles is saved. In other words, different vectors are different because they are intrinsically discernible, since their difference can be explained in terms of their properties. Intrinsically discernible geometric objects will be called *absolute geometric objects*.

These considerations show that the duality between passive and active transformations is the bridge allowing us to pass from what seems to be an epistemic consideration about coordinate systems to an ontological (or intrinsic) characterization of the relevant geometric objects, namely that they are numerically different—in the sense that there are many—and intrinsically indiscernible. The analysis of coordinate systems’ transformations only serves to demonstrate the intrinsic relativity of the geometry itself: any active automorphism can always be trivialized by interpreting it as a coordinate transformation. Therefore, the duality between passive and active transformations, far from being an *a priori* necessity, results from the intrinsic indiscernibility of vectors in V . Hence, it could be misleading to insist exclusively on the difference between the intrinsic (or coordinate-independent) geometry and its non-canonical coordinate representation. The important point here is that the intrinsic geometry itself is composed of relative geometry objects that lack intrinsic discernibility.

3. Locality

In Yang–Mills theory, matter fields are described as *sections* of a *vector bundle* $p : E \rightarrow M$ over space–time. At each point x of the base space M , a different copy V_x of a vector

⁶This terminology is borrowed from Adams (1979) and Weyl (1949). See also Auyang (1995) for a general discussion.

⁷In fact, choosing a basis s means choosing n linearly independent vectors of V . Hence, a vector v can only be defined *relatively* to other vectors. As it was clearly stated by Weyl, since the basis s is composed of vectors which, as any vector, lack of intrinsic discernibility (they cannot be “characterized conceptually”), they have to be exhibited by indexical expressions. In other words, the set of vectors chosen as a basis cannot be individuated by the coordinate system defined by themselves: “A conceptual fixation of points by labels [...] that would enable one to reconstruct any point when it has been lost, is here possible only in relation to a coordinate system, or frame of reference, that has to be exhibited by an individual demonstrative act. The objectification, by elimination of the ego and its immediate life of intuition, does not fully succeed, and the coordinate system remains as the necessary residue of the ego-extinction.” (Weyl, 1949, p. 75). This indexical resource is the “phenomenological residue” that makes it impossible to reduce *thisness* to *suchness*.

space V is attached. As we have seen in the previous section, a vector v in a fiber V_x is a relative geometric object described by an equivalence class of the form $[(s, v^s)]$, where v^s are the coordinates of v in the basis s . The equivalence relation is defined by $(s, v^s) \sim (s \cdot g, \rho(g^{-1})v^s)$, where ρ is a representation of a Lie group G (the so-called *structure group*) that defines the left action of G on V . A section of a vector bundle is a map $\psi : M \rightarrow E$ subject to the condition $p \circ \psi = id_M$. This means that ψ assigns a vector in the fiber V_x to each $x \in M$. The element $\psi(x) \in V_x$ will be called the *internal state* of the field ψ at x . Since a section ψ assigns a vector $v \in V_x$ to each $x \in M$ and since all fibers V_x are isomorphic to V , one might be tempted to consider ψ a function from M to the vector space V . Nevertheless, it should be stressed that the image of each $x \in M$ belongs to a different copy V_x of the vector space V . In fact, the distinction between *functions* and *sections* is one of the fundamental clarifications provided by the fiber bundle formalism. Even though all vector spaces V_x are isomorphic, there is no *canonical* isomorphism between them. This is a direct consequence of the fact that internal states in each fiber are *relative geometric objects* that cannot be intrinsically discerned. This property of the fibers will be named *internal relativity*. In other words, by internal relativity we mean that the internal state of a matter field ψ at a position $x \in M$ lacks a gauge invariant meaning (i.e. it is not like a colored vector). If one chooses a vector v in V_x , there is no canonical means of deciding which is the corresponding vector in another copy $V_{x'}$ of V at a different position. Copying an object whose points cannot be intrinsically discerned is an irreversible procedure: even when the copies are identical, one cannot identify them canonically. On the other hand, if vectors in V were intrinsically discernable (i.e. if they were like colored vectors), all copies V_x could be canonically identified. In this case, two vectors in different fibers will be identified if they are copies of the same intrinsically discernable vector in V (i.e. if they have the same color). We can thus conclude that the canonical identifiability of the different fibers depends on the intrinsic discernibility of geometric objects in each fiber. If vectors in each fiber V_x are intrinsically indiscernable (internal relativity), then the different fibers cannot be canonically identified. If internal states in different locations cannot be compared, then it is not possible to decide whether or not a section ψ changes with the position x . Since a matter field describes by definition the spatiotemporal dependence of local internal states, the very concept of a field is deprived of meaning if its spatiotemporal variation cannot be intrinsically defined.

To summarize, we can say that the space–time manifold M parameterizes identical copies of the vector space V that cannot be canonically identified. The intrinsic indiscernibility of vectors in V (internal relativity) results in the external unidentifiability of the different copies V_x .⁸ Since there is no canonical isomorphism between the different copies V_x of V , the procedure of making copies is not reversible. This means that their numerical difference cannot be canonically reduced. In Auyang’s terminology, we could say that the identical internal spaces V_x are “autonomous individuals” and that “[...] M , the set of all x , can be interpreted as the *identifying system for the individuals.*” (Auyang,

⁸It is worth noting that if the different fibers V_x are considered identical (as it is assumed here, as well as in Auyang, 1995; Weyl, 1949), then their spatiotemporal position cannot be interpreted as a property of the fibers. If spatiotemporal position were a “location property” of the fibers as it is proposed in Teller (1987), then the fibers would be different, since they would differ in their location property. On the contrary, we assume that the set of copies V_x of V is a primitive multiplicity parameterized by space–time of *numerically different* identical copies of V . In other words, we assume here that the copies V_x do not differ in any property. Hence, their difference is purely numerical.

1995, p. 132). This irreducible numerical difference between identical internal spaces V_x introduced by space–time is precisely what we shall understand by *locality*. In other words, locality means that the identical geometric structures that are defined over different space–time points cannot be canonically identified one to another. If this is the case, then space–time plays the role of a “[...] *principium individuationis*. It makes the existence of numerically different things possible which are equal in every respect.” (Weyl, 1949, p. 131). In sum, if internal states cannot be intrinsically discerned (internal relativity), then space–time localizes matter fields, that is to say the internal states in different locations cannot be compared. Hence, localized matter fields are just a collection, parameterized by space–time, of unrelated local internal states.

It is important to highlight that the spatiotemporal parametrization of internal spaces implies locality only if internal states in each fiber V_x cannot be intrinsically discerned. This means that locality is not strictly speaking a “principle”, but rather a consequence of the intrinsic relativity of internal states. If internal states were defined by means of absolute geometric objects (as in the example of the colored vectors), then internal states in different fibers could be compared. In this case, the fibers V_x could be identified by means of a canonical isomorphism, that is to say their numerical difference could be canonically reduced. The fibers V_x would no longer be “autonomous individuals”. Therefore, space–time would not localize the matter field. In other words, this definition of locality depends on an intrinsic (or ontological) feature of the geometric objects used for describing matter fields’ internal states. This intrinsic feature of internal states has “epistemic” consequences in what concerns our utilization of coordinate systems. If a matter field is localized, then the bases in the different fibers can be arbitrarily chosen independently one of another. This means that local gauge invariance is a natural consequence of internal relativity.

For a critical account of other definitions of locality that can be invoked for justifying the local gauge invariance requirement see Martin (2002). One possibility is to assert that “A ‘local’ field theory is [...] just one specified by a Lagrangian depending on the values of the field and a finite number of its derivatives at a *single* point.” (Martin, 2002, p. 226). Nevertheless, “[...] a field theory can be local in this sense and yet have a global gauge invariance.” (Martin, 2002, p. 226). Another possibility is to understand locality as a prohibition of instantaneous communication between space–time points. According to this interpretation, a globally invariant theory would not be compatible with special relativity. This notion of locality is closely related to the idea that “[...] every physical process is determined by what happens locally—that is to say, by what is happening at each space–time point where that process is going on.” (Healey, 1991, p. 406). Nevertheless, since a local gauge transformation can be consistently interpreted as a passive transformation of the local coordinate systems, “[...] there is nothing physical that gets changed under the transformations.” (Martin, 2002, p. 227). Moreover, as it is correctly remarked in Martin (2002), global gauge transformations are particular cases of local gauge transformations (where $g(x) = g \in G, \forall x \in M$). This means that by rendering the theory locally invariant, we are not excluding global gauge transformations. Another possibility would be to interpret locality in terms of Lorentz invariance. However, “[...] the requirement of Lorentz invariance does not exclude Lagrangians with global invariances: e.g., the Lagrangian for the free Dirac field.” (Martin, 2002, p. 226). In the next section, we will show that, according to our definition of locality, a theory invariant only under *global* gauge transformations is not local. Therefore, our definition of locality

allows us to trace a satisfactory demarcation line between globally invariant theories and locally invariant theories. Moreover, it allows us to ascertain the ontological assumption that is at the base of local gauge invariance, namely internal relativity.

4. Rigid fiber bundles

In this section we will analyze the ontological presuppositions implicitly assumed in the framework of a field theory with a global gauge invariance. To do so, we will begin by considering the corresponding G -principal fiber bundle $P \rightarrow M$. Instead of attaching a copy V_x of the vector space V to each point $x \in M$, we now attach the space R_x of all bases s in V_x . The structure group G rotates the bases in R_x by means of the free and transitive right action $R_x \times G \rightarrow R_x$, with $(s, g) \mapsto s' = s \cdot g$.⁹ We will now characterize the kind of objects that the spaces R_x are. If a basis $s_0(x)$ in a fiber R_x is arbitrarily chosen, an isomorphism $\varphi_{s_0(x)}$ between R_x and the structure group G is induced. The basis $s_0(x)$ is identified with the identity element e in G . Any other basis $s(x)$ in R_x is identified with the element g that defines the rotation between $s_0(x)$ and $s(x)$. In other words, the isomorphism $\varphi_{s_0(x)} : R_x \rightarrow G$ is such that $s_0(x) \mapsto e$ and $s(x) \mapsto g$ where $s(x) = s_0(x) \cdot g$. Technically speaking, each space R_x is a G -torsor, that is to say a set R_x with an operation $R_x \times R_x \rightarrow G$. As Baez says: “A torsor is a group that has forgotten its identity” (Baez, 2005, p. 3). The election of a local basis $s_0(x)$ in each fiber R_x defines a section $s_0 : M \rightarrow P$, which will be called a *moving frame of reference*.¹⁰

It is important to note that *even locally* a G -principal fiber bundle $P \rightarrow M$ is not, strictly speaking, a product space of the form $U_i \times G$. Fiber bundles are commonly characterized as globally twisted generalizations of the cartesian product of two spaces. According to this description, if a fiber bundle is not generally a cartesian product at the global level, it would nevertheless be so at the local level. The originality of the fiber bundle formalism would be that allows us to treat non-trivial global structures by gluing trivial local ones.¹¹ Even when this is the case, such a characterization misses a point of fundamental importance to the present context. In fact, the fiber bundle formalism introduces a novelty even at the local level (or, equivalently, in the case of trivial fiber bundles). From the previous analysis, it follows that a principal fiber bundle cannot be considered a product space, even at the local level. The fact that a fiber bundle can be globally trivialized does not mean that it is *per se* a cartesian product. This is because over each point $x \in M$ there is not group G , but rather a G -torsor R_x . The significant difference here is that the G -torsors R_x —like the vector spaces V_x —cannot be identified canonically. It is only after *arbitrarily* choosing a local moving frame of reference $s_0 : U_i \rightarrow P$, that the different G -torsors R_x over U_i are identified with the structure group G , and that the fiber bundle locally takes the trivial form $U_i \times G$. When a local section s_0 is chosen, the numerical or primitive heterogeneity between the different fibers is lost. Over each point $x \in M$ there is now a copy of the group G and the fiber bundle becomes a cartesian product.

⁹The vector bundle $E \rightarrow M$ can be defined as a bundle *associated* to the G -principal fiber bundle $P \rightarrow M$ by identifying E with $P \times_G V$, where the right action of G on $P \times V$ is given by $(s, v) \cdot g = (s \cdot g, \rho(g^{-1})v)$.

¹⁰In general this can only be done locally, that is to say on each open set $U_i \subset M$ of an open cover $\mathcal{U} = \{U_i\}_i$ of M . If the principal fiber bundle $P \rightarrow M$ is not trivial, it is not possible to define a *global* section $s : M \rightarrow P$.

¹¹For example, Lyre (2001, p. 376) states: “As long as we are concerned with trivial bundles, the notion of a fiber bundle is in a way superfluous (since we may simply use a direct product). For non-trivial bundles, however, the fiber bundle framework becomes indispensable.”

As we have already seen, the fact that different fibers cannot be canonically identified is not an *a priori* necessity.¹² Internal states in each fiber could conceivably be described by intrinsically discernable (or absolute) geometric objects (as in the example of the colored vectors). Since the different identical copies V_x of a vector space V composed of absolute vectors inherit the discernibility of these vectors, the different fibers become canonically identifiable and the theory is no longer local. This means that the numerical difference between the fibers V_x can be reduced by means of a canonical isomorphism. The intrinsic discernibility of elements of V privileges a certain class $[s_0]$ of moving frames of reference. A privileged moving frame of reference in this class is a section s_0 such that two vectors in different fibers $v \in V_x$ and $v' \in V_{x'}$ have the same coordinates in the local frames $s_0(x)$ and $s_0(x')$, respectively, if and only if they are copies of the same absolute vector in V (i.e. if they have the same color). The privileged sections in the class $[s_0]$ induce privileged isomorphisms between the frame spaces R_x and the structure group G . This means that there exists a class of privileged trivializations of the principal fiber bundle. Given a moving frame of reference s_0 in the privileged class $[s_0]$, another section in the same class can be obtained by rotating all the local frames of reference in s_0 in the same way. In other words, two sections in the privileged class can be related by means of a *global* transformation defined by a unique group element $g \in G$. In this way, vectors in different fibers that have the same coordinates in the moving frame of reference s_0 will have the same coordinates in the rotated moving frame $s' = s_0 \cdot g$. On the other hand, if the local frames in s_0 were rotated independently of one another (local gauge transformation), the identification defined by the privileged moving frame s_0 would be lost. In more common terms, a vector bundle with fibers V_x composed of absolute geometric objects is only invariant under *global* or *rigid* gauge transformations. By means of this analysis, we wish to identify the ontological presupposition of a theory invariant under global gauge transformations. The fundamental assumption of such a theory is that the fibers V_x are composed of absolute geometric objects that can be intrinsically discerned. In other words, such a theory assumes that internal states of a matter field at a certain position have an intrinsic physical meaning. Therefore, the duality between passive and active transformations is broken. While a passive global gauge transformation does not change absolute vectors, a global automorphism does change them. According to our definition of locality, a spatiotemporal parametrization of internal spaces composed of absolute geometric objects does not localize the matter field. In other words, the different internal spaces V_x can now be canonically identified. It is worth stressing that globally invariant theories do not presuppose that there is “one true gauge” (Maudlin, 1998, p. 367), but rather that there is a privileged class $[s_0]$ of moving frames of reference. This epistemic presupposition relies on the ontological assumption that internal states in different fibers are intrinsically discernable.

5. Connected fiber bundles

Apparently one has to choose between working with localized matter fields (composed of unrelated local internal states) on the one hand or working in a rigid fiber bundle where internal relativity and locality are lost on the other. The heuristic non-trivial content of the gauge argument is that it shows how to “connect” the different fibers without loosing

¹²See Norton (1992) for an analysis of an analogous situation in general relativity.

locality. Instead of supposing a fiber bundle with internal spaces composed of absolute geometric objects, the fiber bundle can be endowed with a fixed background connection A (see Appendix for the definition of connections on principal bundles). The different fibers are now “connected” by means of parallel transports $\pi_A^\gamma: R_x \rightarrow R_{x'}$ that depend in general on the path $\gamma(x, x')$ in M . What is significant here is that the “connection” between different fibers no longer relies on the supposition that internal states are intrinsically discernable. Since different fibers cannot be identified canonically, the connection does not break the spatiotemporal heterogeneity between them. Therefore, the resulting theory is local. It follows that local frames of reference in each fiber V_x can be modified independently of one another. Hence, the theory is naturally invariant under *local gauge transformations* generated by elements of the so-called *gauge group* \mathcal{G} . This group is the group of vertical automorphisms of the principal fiber bundle $P \rightarrow M$ (see Appendix for more details).

It is well known that theories such as Newtonian mechanics and special relativity can be reformulated in terms of generally covariant equations. This means that they can be given an intrinsic or coordinate independent formulation (Friedman, 1983; Misner et al., 1973). This result has been used to argue that covariance principles have no physical content (for a general discussion about the physical meaning of general covariance see Norton, 1992, 1993, 2003). In the case of a globally invariant theory of the kind that we are considering, an analogous reformulation can be performed by reinterpreting the original Lagrangian $\mathcal{L}_0(\psi, \partial_\mu \psi)$ as the *standard formulation* of a *locally* invariant theory defined in a fiber bundle endowed with a fixed flat connection A (see Appendix for the definition of flat connections). By “standard formulation” we mean the formulation of the theory in the coordinate system in which the Lagrangian takes its simplest form (Friedman, 1983). Since we can interpret $\mathcal{L}_0(\psi, \partial_\mu \psi)$ either as the Lagrangian of a globally invariant theory or as the Lagrangian of a locally invariant theory in a fiber bundle endowed with a fixed flat connection, local gauge invariance would seem to have no physical content. We will now argue that the ontological assumptions underlying these two alternative interpretations are completely different. Moreover, far from being a mere difference in the conceptual interpretation of the theory, the two formulations might have different physical content. This will be the case if space–time M is not simply connected, that is to say if the fundamental group $\pi_1(M)$ is non-trivial.

In order to analyze the situation in detail, let us consider the case of a fiber bundle $P \rightarrow M$ endowed with a flat connection $A \in \Omega^1(P) \otimes \mathfrak{g}$. Since the connection is flat, there always exists a class of moving frames of reference $[s]$ such that the components of the local form $A_s = s^*A \in \Omega^1(M) \otimes \mathfrak{g}$ are zero for $s \in [s]$.¹³ Different sections in the privileged class $[s]$ are related by means of global gauge transformations defined by elements of the structure group G . In these privileged moving frames of reference the covariant derivative $D_\mu = \partial_\mu + A_\mu(x)$ reduces to an ordinary derivative and the covariant Lagrangian $\mathcal{L}(\psi, D_\mu \psi)$ coincides with $\mathcal{L}_0(\psi, \partial_\mu \psi)$. In this way, the privileged class $[s]$ defines the standard formulation of a theory invariant under local gauge transformations. In a coordinate system $s' = s \cdot g(x)$ which does not belong to the privileged class, the

¹³The privileged class $[s]$ can be obtained by integrating the equivariant distribution H_A defined by the connection. This means that sections in the class $[s]$ are everywhere tangent to the horizontal subspaces defined by the distribution H_A . To show that $A_s = 0$, let us consider the action of A_s on any vector X on M :

$$A_s(X) = (s^*A)(X) = A(s_*X) = 0 \quad \forall X \in T_x M,$$

since s_*X belongs to the equivariant distribution H_A defined by the kernel of A .

components of the connection take the form $A_{s'} = g^{-1}(x) dg(x)$ and the connection is said to be “pure gauge”. From this description, it might seem that a flat connection A privileges the class of sections $[s]$ that defines the standard formulation of the theory. This class would establish a privileged identification between internal states in different fibers. Therefore, this tricky implementation of local gauge invariance does not seem to modify the ontological presupposition of rigid fiber bundles invariant under global gauge transformations, namely that internal states in each fiber are intrinsically discernable. Nevertheless, it is easy to see that the transition from a rigid fiber bundle to a fiber bundle endowed with a fixed flat connection does allow us to eliminate this ontological assumption. In other words, it is not true that the flat connection A privileges a particular class of sections. Given a local gauge transformation $s \rightsquigarrow s' = s \cdot g(x)$, one can always define a vertical automorphism $\Phi : P \rightarrow P$ such that the gauge equivalent connection $A' = \Phi^*(A)$ has zero local form in the coordinate system s' (i.e. $A'_{s'} = 0$).¹⁴ Whereas the connection A privileges the class $[s]$, the transformed connection $A' = \Phi^*(A)$ privileges the class $[s']$. Nevertheless, A and A' are gauge equivalent, that is to say they are different representatives of the same intrinsic connection $[A]$. This means that the class of sections that define the standard formulation depends on which representative of the orbit of gauge equivalent connections we are working with. Then, it is not true that the intrinsic connection $[A]$ privileges a certain class of sections. Hence, the different fibers cannot be canonically identified, that is to say they preserve their spatiotemporal irreducible heterogeneity (locality). This means that internal states in each fiber cannot be intrinsically discerned, i.e. they do not have a gauge invariant meaning (internal relativity). We have therefore shown that substituting a globally invariant theory (defined in a rigid fiber bundle) by a locally invariant theory (defined in a fiber bundle endowed with a fixed connection), allows us to eliminate the ontological assumption that the different fibers are composed of geometric objects that are intrinsically discernable. Therefore, by endowing the fiber bundle with a fixed background connection, it is possible to “connect” the different fibers without spoiling internal relativity and locality.

We will now argue that the covariant reformulation of a globally invariant theory is not a mere formal step, but rather requires, in the general case, the non-trivial election of a particular geometric background structure. In fact, if the fundamental group $\pi_1(M)$ of M is non-trivial (i.e. if M is not simply connected), there will be different inequivalent flat connections parameterized by the corresponding *moduli space* $\mathcal{M}(M, G)$ (see Appendix). This means that the covariant reformulation of a theory with a global gauge invariance makes it necessary to select a particular flat connection from this *moduli space*. In other words, the original theory described by the Lagrangian $\mathcal{L}_0(\psi, \partial_\mu \psi)$ does not univocally prescribe the flat connection of its covariant reformulation. Moreover, inequivalent flat connections can be distinguished by performing parallel transports around loops in M . Even if the curvature vanishes, parallel transport might be non-trivial if M is not simply connected. Hence, inequivalent flat connections will define different holonomy groups. Since the so-called *Wilson loops* are gauge invariant observables obtained from the holonomy defined by a connection, inequivalent flat connections will define different

¹⁴Let us show that it is possible to choose a vertical automorphism $\Phi : P \rightarrow P$ such that $A'_{s'} = 0$. To do so, let us write explicitly $A'_{s'}$:

$$A'_{s'} = s'^* A' = s'^*(\Phi^* A) = (\Phi(s'))^* A,$$

where we used that $(g \circ f)^* = f^* \circ g^*$. Let us choose Φ such that $\Phi(s') = s$. This condition can be satisfied by choosing $\Phi(s') = s' \cdot g(x)^{-1}$. In fact, $\Phi(s') = s' \cdot g(x)^{-1} = (s' \cdot g(x)) \cdot g(x)^{-1} = s$. Hence, $A'_{s'} = (\Phi(s'))^* A = s^* A = 0$.

physical observables (Gambini & Pullin, 1996). Hence, different elections of a background flat connection might have different physical consequences.

In sum, the introduction of a fixed background connection by means of the gauge argument has non-trivial consequences. Firstly, it allows us to get rid of the ontological assumption presupposed in the framework of globally invariant theories, namely that matter fields' internal states have a gauge invariant meaning. In this way, the gauge argument allows us to reestablish internal relativity and locality. Secondly, the introduction of a connection permits to define new physical observables associated to the parallel transport around loops of internal states. Nevertheless, the introduction of a fixed background connection by means of the gauge argument is not enough for reobtaining Yang–Mills theory. Since a kinetic term for A does not appear in $\mathcal{L}(\psi, D_\mu\psi)$, the connection is not described by new degrees of freedom, but rather is a fixed geometric background of the theory. A local field theory defined in a fiber bundle endowed with a fixed connection is a *background dependent theory*. In principle, there exist two strategies for trying to eliminate a geometric background structure. Firstly, one can try to establish a complete determination of the background structure by the material degrees of freedom. We will call this strategy *Machian programme*. In the context of space–time theories, this solution reflects Mach's idea that “[...] the local standards of inertial motion [...] are entirely fixed by the matter distribution throughout space and time [...]” (Huggett & Hoefer, 2006). According to the Machian programme, it should be possible to eliminate the “unnecessary” hypothesis of an absolute space by working only with relative quantities. In this way, it would be possible to define a background independent theory without extending the minimal ontology of material degrees of freedom. In the case that we are considering here, it is not the geometry of space–time, but rather the geometry of the G -principal fiber bundle $P \rightarrow M$ that should be completely fixed by the material degrees of freedom. The second possible strategy is to substitute the fixed geometric background with a new *physical* entity carrying its own degrees of freedom and coupled to matter fields, that is to say *partially* determined by them. We will call this strategy *Einsteinian programme*. In this case, the minimal ontology of matter fields is extended by introducing additional degrees of freedom that describe the dynamics of the physical connection.¹⁵ Since in Yang–Mills theory the degrees of freedom that describe the geometry of the fiber bundle are coupled to matter fields (instead of being completely determined by them), Yang–Mills theory can be understood to extend the Einsteinian programme to physical interactions other than gravity.¹⁶ It is worth noting that the background independence at stake in

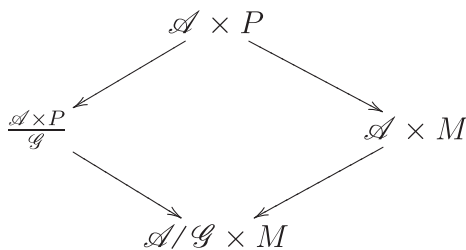
¹⁵In order to consider the spatiotemporal connection A as a dynamic entity, it is necessary to foliate space–time M by spacelike hypersurfaces. In other words, it is necessary to choose a diffeomorphism $\iota : M \rightarrow \mathcal{M} \times \mathbb{R}$, where \mathcal{M} is a smooth three-dimensional riemannian manifold. Using the diffeomorphism ι and the fact that \mathbb{R} is contractible, one can assume that the fiber bundle $P \rightarrow M$ over the space–time is the pullback, defined by means the projection $M \simeq \mathcal{M} \times \mathbb{R} \rightarrow \mathcal{M}$, of a fixed G -principal bundle $\mathcal{P} \rightarrow \mathcal{M}$. Each connection A on $P \rightarrow M$ then has a canonical decomposition $A = A(t) + A_0(t)dt$, where $A(t)$ denotes a time-evolving connection on $\mathcal{P} \rightarrow \mathcal{M}$ and $A_0(t)$ is a time-evolving section of $\text{ad}(P) = P \times_G \mathfrak{g}$.

¹⁶According to Einstein's point of view, general relativity satisfies what he calls *Mach's principle*. This principle is defined by Einstein as follows: “The G -field is *completely* determined by the masses of the bodies.” (quoted in Barbour, 1992, p. 149). Nevertheless, as Barbour says, this formulation “[...] unnaturally puts the entire responsibility for the determination of local inertial properties (represented by values of the metric tensor realized at and around a given space–time point) on the shoulders of matter alone, whereas the theory is actually about dynamic interaction and space and matter, which have their own independent dynamic degrees of freedom.” (Barbour, 1992, p. 150). This explains why we distinguish the Machian programme from the Einsteinian programme, in spite of Einstein's own formulation.

Yang–Mills theory exclusively concerns the geometry of the fiber bundle. The background independence concerning spatiotemporal geometry is accomplished by general relativity.

6. Physical meaning of gauge freedom

The configuration space associated to the connection’s degrees of freedom is an affine space modelled on $\Omega^1(M) \otimes \mathfrak{g}$. Since the action of the gauge group \mathcal{G} on this configuration space is not free, the quotient generally is not a manifold. This problem can be solved by using *framed connections* (Donaldson & Kronheimer, 1990). It can be shown that under certain conditions the \mathcal{G} -action induces the structure of a \mathcal{G} -principal fiber bundle $\pi : \mathcal{A} \rightarrow \mathcal{A}/\mathcal{G}$, where \mathcal{A} is the configuration space of framed connections. In principle, it is possible to directly work on the *orbit space* \mathcal{A}/\mathcal{G} , which only describes physical degrees of freedom (i.e. classes $[A]$ of gauge equivalent connections). Nevertheless, there are generally technical obstructions. The common solution is to work on \mathcal{A} by *fixing the gauge*, that is to say by selecting a particular representative $A \in \pi^{-1}([A])$ in the corresponding orbit for each class $[A] \in \mathcal{A}/\mathcal{G}$ of gauge equivalent connections. This amounts to choosing a gauge fixing section $\sigma : \mathcal{A}/\mathcal{G} \rightarrow \mathcal{A}$. In principle, we can thus choose between working in a space with non-physical degrees of freedom or working in the smaller space \mathcal{A}/\mathcal{G} where the gauge symmetry has been eliminated. More generally, it is well known that any ordinary Hamiltonian system can be *enlarged* to a gauge theory by adding non-physical degrees of freedom and constraints. Reciprocally, any gauge theory can be *reduced* to an unconstrained theory on the corresponding reduced phase space. Therefore, the formulation of a theory as a gauge theory could seem to be only a matter of technical “convenience” (mathematical tractability) and/or “perspicuity” (Belot, 2003, p. 220). Nevertheless, we will argue that the point of gauge freedom in Yang–Mills theory is independent of the choice between these two formulations of a physical theory. As we have explained above, local gauge invariance is a direct consequence of spatiotemporal locality, that is to say of the fact that fibers in different locations cannot be canonically identified. And locality is a consequence of internal relativity, that is to say of the fact that matter fields’ internal states do not have a gauge invariant meaning. Both locality and internal relativity are intrinsic features of Yang–Mills theory which do not depend on the chosen formulation. To clarify this point, let us consider the G -principal bundle $\mathcal{A} \times P \rightarrow \mathcal{A} \times M$. This bundle can be obtained as the pullback of $P \rightarrow M$ by means of the projection $\mathcal{A} \times M \rightarrow M$. The action of the gauge group \mathcal{G} on \mathcal{A} induces an action on $\mathcal{A} \times M$. This action is covered by the action of \mathcal{G} on $\mathcal{A} \times P$ induced by its action on both \mathcal{A} and P . Therefore, the G -principal bundle $\mathcal{A} \times P \rightarrow \mathcal{A} \times M$ induces the G -principal bundle $(\mathcal{A} \times P)/\mathcal{G} \rightarrow \mathcal{A}/\mathcal{G} \times M$ (Atiyah & Singer, 1984). The relationships between these bundles can be summarized by means of the following diagram:



where the arrows from left to right denote G -projections and the arrows from right to left denote \mathcal{G} -projections. The fiber bundle $\mathcal{A} \times P \rightarrow \mathcal{A} \times M$ can be interpreted as a universal family parameterized by \mathcal{A} of G -principal bundles $P \rightarrow M$ with tautological connections $A \in \mathcal{A}$. The G -principal bundle $(\mathcal{A} \times P)/\mathcal{G} \rightarrow \mathcal{A}/\mathcal{G} \times M$ is the intrinsic formulation of this universal family (in the sense that the gauge freedom was eliminated by taking the quotient by \mathcal{G}). If we fix an element $[A]$ in \mathcal{A}/\mathcal{G} we obtain a G -principal fiber bundle over M endowed with the intrinsic connection described by the class $[A]$. Therefore, even if we work in the orbit space \mathcal{A}/\mathcal{G} , we still have a G -principal fiber bundle over M endowed with a connection. In this principal bundle, internal states cannot be intrinsically discerned (internal relativity) and the different fibers cannot be canonically identified (locality). In this way, if we consider the intrinsic—or coordinate-independent—geometry (i.e. if we take the quotient by \mathcal{G}), local gauge invariance disappears. Nevertheless, the ontological assumptions underlying local gauge invariance (internal relativity and locality) are still present. In other words, the point of gauge freedom in Yang–Mills theory does not concern local gauge invariance itself, but the ontological assumptions that force the local gauge invariance of the theory. We could also argue as follows. The elimination of local gauge invariance (by passing from \mathcal{A} to \mathcal{A}/\mathcal{G}) does not erase the fact that the connection had to be introduced for guaranteeing internal relativity and locality (and therefore local gauge invariance). In other words, by taking the quotient by \mathcal{G} we eliminate the local gauge freedom in the expression of the connection, but we do not eliminate the connection that was introduced in order to guarantee the ontological presuppositions that underlie local gauge freedom. To summarize, we can say that the fiber bundle formalism strikes twice on Yang–Mills theory, namely through the \mathcal{G} -principal bundle $\mathcal{A} \rightarrow \mathcal{A}/\mathcal{G}$ and through the G -principal bundle $P \rightarrow M$ (or the associated bundle $E \rightarrow M$). Nevertheless, the essential point of gauge freedom in Yang–Mills theory does not concern the former, but rather the latter. The partition of the configuration space in gauge orbits is a common feature of any gauge theory (Henneaux & Teitelboim, 1994). On the contrary, the analysis of the fiber bundle structure over space–time let us ascertain the origin of gauge freedom in the particular context of Yang–Mills theory.

7. BRST symmetry

In the Lagrangian formulation of Yang–Mills theory, the gauge is fixed by means of a gauge fixing term in the corresponding Lagrangian. Commonly this term takes the form $\mathcal{L}_{gf} = \lambda F(A)$, where λ is a Lagrange multiplier whose variation imposes the gauge fixing condition $F(A) = 0$. In order to define a section $\sigma : \mathcal{A}/\mathcal{G} \rightarrow \mathcal{A}$, this condition has to be such that for each equivalence class $[A] \in \mathcal{A}/\mathcal{G}$ there is one and only one connection $A \in \mathcal{A}$ which satisfies the condition. In other words, a local gauge transformation applied to an element in the section σ must take it out of the section.¹⁷ This means that by definition an adequate gauge fixing term is not gauge invariant. Hence, the gauge fixed Lagrangian is not gauge invariant either. This fact is commonly presented in the following terms. It is held that the gauge fixing procedure breaks the gauge invariance of the theory. It is then argued that the gauge fixed theory is fortunately invariant under a mysterious

¹⁷Due to possible topological obstructions associated to the non-triviality of the bundle $\mathcal{A} \rightarrow \mathcal{A}/\mathcal{G}$, in some cases it is impossible to define a *global* gauge fixing section $\sigma : \mathcal{A}/\mathcal{G} \rightarrow \mathcal{A}$. This is the so-called *Gribov's obstruction* to the gauge fixing (Gribov, 1978; Singer, 1978).

new global—or rigid—symmetry, the so-called *BRST symmetry*. This symmetry would act as a substitute for the lost local gauge invariance. In particular, the BRST transformation of the gauge field A is given by the expression $A \rightsquigarrow A + D_A \eta$, where η is a new (anticommuting) field called *ghost field*. This transformation is similar to the usual infinitesimal local gauge transformation of A :

$$A \rightsquigarrow A + D_A \zeta(x), \quad (1)$$

where the role of the space–time function $\zeta(x)$ is played by the ghost field η . Nevertheless, the global nature of BRST symmetry seems to prevent us from mistaking it for the local gauge symmetry. As it is stated in Redhead (2003, pp. 135–136), the BRST symmetry would be “a kind of generalized gauge symmetry” of an “hybrid” nature: “It is in essence a non-linear rigid fermionic transformation, which contains within itself, so to speak, a local gauge transformation specified by a dynamical field, namely the ghost field.”

A rigorous conceptual and mathematical formulation of the BRST construction was performed in the context of the constrained Hamiltonian formalism (Henneaux & Teitelboim, 1994).¹⁸ The BRST formalism is a cohomological procedure for defining the observables of a gauge theory. In its Hamiltonian formulation, BRST transformations are generated by a nilpotent operator acting on an extended phase space. This space is obtained by introducing new conjugated canonical variables called *ghost fields* and *ghost momenta* (Henneaux & Teitelboim, 1994). The BRST transformation of a function f on this extended phase space is obtained by means of the expression $\delta_{\text{BRST}} f = \{f, \Omega\}$, where Ω is the so-called BRST charge and $\{\cdot, \cdot\}$ the Poisson bracket in the extended phase space.¹⁹ By using the Jacobi identity for the Poisson bracket, the nilpotency of the BRST transformation (that is to say the identity $\delta_{\text{BRST}}^2 = 0$) can be expressed as $\{\Omega, \Omega\} = 0$. The BRST charge Ω is constructed in such a way that the zero degree cohomology $H^0(\Omega)$ associated to the corresponding complex coincides with the theory’s set of observables (i.e. with the set $\mathcal{C}^\infty(\Sigma/\mathcal{G})$ of gauge invariant functions on the *reduced phase space* Σ/\mathcal{G} , where Σ is the constraint surface defined by the generalized Bianchi identities).²⁰ It is important to stress that the BRST formalism is a classical construction. Following the Faddeev–Popov method, it might seem that ghost fields are exclusively necessary for exponentiating the Faddeev–Popov determinant in the corresponding path integral (Faddeev & Popov, 1967; Faddeev & Slavnov, 1991). Certainly, BRST symmetry is a fundamental tool for quantifying gauge theories, and for guaranteeing the unitary and the renormalizability of the theory. Nevertheless, the Hamiltonian formulation of BRST

¹⁸The relationship between this formulation and the Lagrangian formulation that is usually used in quantum field theory (in terms of ghosts, antighosts and the Nakanishi–Lautrup auxiliary field) can be found in Henneaux and Teitelboim (1994, Chapter 11).

¹⁹In general, a function g on a phase space defines an induced classical operator v_g such that its action on a function f is given by the expression $v_g f = \{f, g\}$. The classical operator v_g , called *Hamiltonian vector field*, satisfies the expression $\omega(v_g) = dg$, where ω is the symplectic form of the phase space.

²⁰The BRST operator δ_{BRST} can be obtained in a two step procedure. Firstly, the restriction to functions on the constraint surface Σ makes it necessary to introduce a Koszul–Tate resolution of the algebra $\mathcal{C}^\infty(\Sigma)$. This resolution is generated by the ghost momenta and a homological differential δ . Secondly, the consideration of gauge invariant functions makes it necessary to introduce a vertical differential complex generated by the ghost fields and a vertical cohomological differential d . By using results from homological algebra, these two differentials can be combined in the BRST differential $\delta_{\text{BRST}} = \delta + d + \dots$, where the supplementary terms guarantee the nilpotence of δ_{BRST} (see Henneaux & Teitelboim, 1994, Chapters 8 & 9).

theory clearly shows that it is a construction associated to the geometry of the classical reduced phase space (Henneaux & Teitelboim, 1994).

We will now argue that the BRST symmetry should not be understood as a new rigid symmetry that replaces the lost local gauge symmetry. By definition, physical observables do not depend on the particular chosen frame of reference. Hence, we can fix the gauge without modifying the relevant physical content. By fixing the gauge, we only lose the explicit covariance of the motion equations. In a gauge fixed theory, motion equations are expressed in the chosen coordinate system. This is not a problem since the observable quantities are by definition gauge invariant. Nevertheless, it is not enough to test their invariance under a *particular* gauge transformation. In other words, the observables have to be invariant under *any* infinitesimal local gauge transformation. In order to test this *general* invariance, it is necessary to perform what we will call a *general* local gauge transformation. To clarify this notion, let us consider an analogous situation in differential geometry. The exterior derivative df of a function measures the general changes of f , in the sense that the particular direction of the transformation has not yet been specified. If this direction is specified by means of a vector v , the general expression df becomes $df(v) = v^\mu(\partial f / \partial x^\mu)$. This expression describes the change of the function f in the direction defined by v . As this example shows, defining these general transformations requires an exterior algebra of differential forms dual to the tangent space that defines the possible particular transformations. Analogously, we will now argue that a BRST transformation is a *general* infinitesimal local gauge transformation. In Yang–Mills theory, *particular* infinitesimal local gauge transformations are defined by elements $\xi(x)$ of the gauge group's Lie algebra $\mathcal{L}ie(\mathcal{G})$. Therefore, in order to define general gauge transformations it is necessary to introduce an exterior algebra dual to $\mathcal{L}ie(\mathcal{G})$. This explains why the definition of these general transformations requires to introduce the ghost field. In fact, in Bonora and Cotta-Ramusino (1983), it was shown that the ghost field can be identified with the Maurer–Cartan form of the gauge group \mathcal{G} , that is to say with the form $\eta \in \mathcal{L}ie(\mathcal{G})^* \otimes \mathcal{L}ie(\mathcal{G})$ that satisfies $\eta(\xi) = \xi, \forall \xi \in \mathcal{L}ie(\mathcal{G})$. This means that the ghost η is a $\mathcal{L}ie(\mathcal{G})$ -valued 1-form dual to $\mathcal{L}ie(\mathcal{G})$. The BRST transformation of the gauge field A is given by the expression

$$\delta_{\text{BRST}}A = D_A\eta. \quad (2)$$

This expression defines a *general* infinitesimal local gauge transformation. This explains why the Lie algebra element $\xi(x)$ (analogous to the vector v of the previous example) was substituted by the $\mathcal{L}ie(\mathcal{G})$ -valued differential form η (analogous to the form df of the previous example). A *particular* infinitesimal local gauge transformation of A can be obtained by contracting the ghost η appearing in the BRST transformation of A with an element $\xi(x) \in \mathcal{L}ie(\mathcal{G})$. As expected, the infinitesimally transformed gauge field is

$$A' = A + D_A\eta(\xi(x)) = A + D_A\xi(x), \quad (3)$$

where we used that the Maurer–Cartan form satisfies $\eta(\xi) = \xi, \forall \xi \in \mathcal{L}ie(\mathcal{G})$. As it is clearly stated in Bonora and Cotta-Ramusino (1983, p. 594): “Expression $A \rightarrow A + D_A\eta$ represents in a compact form all the possible infinitesimal gauge transformations, each of which is characterized by an element of the Lie algebra $\mathcal{L}ie(\mathcal{G})$ [. . .].”

It is worth stressing that the spatiotemporal dependence of a particular local gauge transformations is carried by the element $\xi(x) \in \mathcal{L}ie(\mathcal{G})$. Since a BRST transformation is a general local gauge transformation (where $\xi(x)$ was substituted by η), the spatiotemporal

dependence is not specified. In other words, the spatiotemporal dependence carried by the elements $\zeta(x)$ appear only when we pass from general (or BRST) transformations to particular local gauge transformations. Hence, a BRST transformation is *in a trivial sense* a “global” transformation with no spatiotemporal dependence. These arguments show that BRST symmetry should not be interpreted as a kind of generalized gauge symmetry different from the usual local gauge symmetry. In a gauge fixed theory, physical observables do not depend on the gauge, that is to say they are invariant under any local gauge transformation. Since BRST transformations define general local gauge transformations, physical observables have to be invariant under BRST transformations.

8. Conclusion

The analysis of classical Yang–Mills theory’s geometric foundations was performed by analyzing the relationships between *internal relativity*, *locality*, *global/local invariance*, and *background independence*.

Since elements of a vector space V are not intrinsically discernable, any automorphism of V can be reinterpreted as a transformation of the basis in V . This fact constitutes the well-known duality between passive and active transformations. Even though vectors in V are geometric objects independent of coordinate systems, they lack intrinsic geometric identity. The key point is not the relationship between the intrinsic geometry and the non-canonical election of a particular basis, but rather the relativity of the intrinsic geometry itself. According to the fiber bundle formulation of Yang–Mills theory, space–time parameterizes identical copies of V . As a consequence of the intrinsic indiscernibility of vectors in each fiber V (internal relativity), there is no canonical isomorphism between the different copies V_x of V (locality). Hence, locality is a direct consequence of internal relativity, that is to say of the fact that internal states lack a gauge invariant meaning. Since different fibers are not identified, the local frames of reference in each fiber can be locally modified. In other words, local gauge invariance is a direct consequence of internal relativity and locality. If the fiber bundle is not endowed with supplementary geometric structures, matter fields’ internal states $\psi(x) \in V_x$ in different locations $x \in M$ cannot be compared. We could say that local internal states in matter fields are mutually “disconnected”. Since a matter field is for the moment a set of local unrelated internal states, it is not strictly speaking, a “penetration of the This (here-now) and the Thus” (Weyl, 1949, p. 130).

One possible way of treating this dispersion of unrelated fibers is by supposing that internal spaces V_x are absolute spaces whose points can be intrinsically discerned (as in the example of the colored vectors). In this case, the theory is no longer characterized by internal relativity: the specification of internal states now has an intrinsic meaning. Making copies of a vector space V composed of intrinsically discernable elements is a reversible procedure and the different copies V_x can now be canonically identified. Hence, locality is lost. There exists now a privileged class $[s_0]$ of moving frames of reference. A section in this class assigns the same coordinates to vectors in different fibers that are intrinsically identical. Different sections in the privileged class $[s_0]$ are related by means of global gauge transformations. It is worth stressing that global invariance is the “epistemic” consequence (in the sense that it concerns our utilization of coordinate systems) of the intrinsic (or ontological) discernibility of internal states.

Another possibility is to endow the fiber bundle with a fixed connection. By means of parallel transports, internal states in different fibers can now be compared without supposing that they are intrinsically discernible. Since the connection does not establish any privileged identification between different fibers, locality is preserved. Hence, the theory is naturally invariant under local gauge transformations. In this way, local gauge invariance is the epistemic consequence of the intrinsic (or ontological) indiscernibility of internal states (internal relativity). The resulting locally invariant theory depends on a fixed geometric background, or what Anderson calls an absolute object (Anderson, 1967). However, by adding to the Lagrangian a kinetic term for the connection, the theory is rendered background independent. This means that the fixed geometric structure is replaced by new dynamic degrees of freedom. This solution to the problem of formulating a background independent theory was distinguished from the Machian programme, where the background structure is supposed to be *completely* determined by the material degrees of freedom.

In sum, we could say that in order to define a field theory characterized by *internal relativity* (and consequently by *locality* and *local gauge invariance*) and *background independence*, it is necessary to endow the fiber bundle with a dynamic connection. The requirement of internal relativity makes it necessary to endow the fiber bundle with a connection minimally coupled to matter fields. The requirement of background independence makes it necessary to consider the connection a dynamic physical entity. We can thus conclude that *internal relativity* and *background independence* are the two independent principles of Yang–Mills theory. Hence, in Yang–Mills theory the dynamics of internal states depends neither on an absolute definition of the latter (Yang–Mills theory’s internal relativity) nor on a fixed background geometric structure (Yang–Mills theory’s background independence). Just as spatiotemporal motion is relative to the local gravitational field in general relativity, the dynamic of matter fields’ internal states is relative to the dynamic gauge field A in Yang–Mills theory. Just as spatiotemporal localization does not have a gauge invariant meaning in general relativity, the internal states of a matter field in Yang–Mills theory does not have one either. In general relativity, the invariance under general diffeomorphisms removes any intrinsic physical meaning from naked points of space–time, that is to say the latter do not define absolute positions. Analogously, the invariance of Yang–Mills theory under local gauge transformations removes any intrinsic physical meaning from internal states, that is to say the latter do not define absolute internal states. In this way, Yang–Mills theory allows us to extend the criticism of *spacetime substantivalism* in general relativity to the case of internal states. According to Earman and Norton (1987), spacetime substantivalism can be defined by the denial of *Leibniz equivalence*, that is to say by the assumption that diffeomorphic models represent different physical situations. In Yang–Mills theory, field configurations related by means of local gauge transformations are interpreted as different representations of the same physical configuration. If we did not consider them physically equivalent, the resulting theory would be indeterministic.²¹ In other words, this “radical form of indeterminism” (Earman & Norton, 1987, p. 515) would be the price to pay for what we could call *internal substantivalism*. By interpreting the different solutions as different representations of a single physical configuration, it is possible to recover a deterministic

²¹It is a well known fact that in gauge theories the specification of a set of initial conditions for all the degrees of freedom determine infinitely many subsequent evolutions (Henneaux & Teitelboim, 1994).

theory. The gauge invariant quantities—the observables—thus have a deterministic evolution (Norton, 1992).

Finally, the BRST symmetry of the gauged fixed theory was interpreted in terms of *general* local gauge transformations. The most singular features of BRST symmetry, namely its action over an *extended* phase space and its *global* nature, follow naturally from this characterization.

Acknowledgment

This work was partially supported by the grants PICT 04-17687 (ANPCyT), PIP no. 1478/01 (CONICET), UBACyT no. X204 and the financial support of *Fondation Maison des Sciences de l'Homme* (Paris).

Appendix A

A.1. Connections on principal bundles

A connection on a G -principal fiber bundle $q : P \rightarrow M$ can be defined as an *equivariant distribution* H , that is to say a smooth field $H_p \subset TP_p$ (with $p \in P$) such that (1) there is a direct sum decomposition $TP_p = H_p \oplus \ker dq_p \forall p \in P$ and (2) the field is preserved by the induced action of G on TP ($H_{pg} = R_{g^*}H_p$, where R_{g^*} denotes the differential of the right translation by $g \in G$). We can assign a \mathfrak{g} -valued G -equivariant 1-form A on P to each connection, such that vectors in the distribution H belong to the kernel of A . In a local trivialization defined by local sections $s_i : U_i \rightarrow P$, the 1-form A induces \mathfrak{g} -valued 1-forms $A_{s_i} = s_i^*A$ on U_i (the gauge fields).

A.2. Vertical automorphisms of principal bundles

A vertical automorphism of $q : P \rightarrow M$ is a diffeomorphism $\Phi : P \rightarrow P$ that preserves the fibers ($q(\Phi(p)) = q(p)$) and that commutes with the action of the structure group G , i.e. $\Phi(p \cdot g) = \Phi(p) \cdot g, \forall p \in P, \forall g \in G$. If s_i is a local section over $U_i \subset M$, it is possible to define a map $g(x) : U_i \rightarrow G$ such that $\Phi(s_i(x)) = s_i(x) \cdot g(x)$. If the principal fiber bundle is trivial, it is possible to define a global section s . In this case, a vertical automorphism Φ can be described in terms of a map $g(x) : M \rightarrow G$ (that depends on the arbitrary section s).

A.3. Flat connections

A connection is flat if every $x \in M$ has an open neighborhood U such that the induced connection in $q^{-1}(U)$ is isomorphic with the canonical flat connection in $U \times G$. The equivariant distribution corresponding to this canonical flat connection is given by the tangent spaces to $U \times \{g\}$ at $u = (x, g) \in U \times G$. The corresponding 1-form is $f^*\theta$, where θ is the canonical 1-form on G and $f : U \times G \rightarrow G$ is the natural projection. The isomorphism is provided by the pullback of the canonical flat connection by means of the trivialization $\psi_s : q^{-1}(U) \rightarrow U \times G$ defined by a section s . This trivialization is defined as $\psi_s(p) = (q(p), g)$, where $g \in G$ is such that $p = s(q(p)) \cdot g$ (Kobayashi & Nomizu, 1963).

In general, the *moduli space* $\mathcal{M}(M, G)$ of flat connections in a G -principal fiber bundle $q : P \rightarrow M$ is given by $\mathcal{M} = \text{Hom}(\pi_1(M), G)/G$, where the quotient of G is defined by

conjugation (Birmingham, Blau, Rakowski, & Thompson, 1991). Roughly speaking, the holonomy $h_p(\gamma) \in G$ around a loop γ (where $p \in P$) defined by a flat connection A depends only on its homotopy class $[\gamma] \in \pi_1(M)$. Hence, it defines a map $\pi_1(M) \rightarrow G$, where the composition of loops is mapped into multiplication in G . If we start from $p \cdot g$ (with $g \in G$) the holonomy is $h_{p \cdot g}(\gamma) = g^{-1}h_p(\gamma)g$. This fact explains the necessity of taking the quotient by the conjugation action of G . In particular, if $\pi_1(M)$ is trivial, then there is only one possible flat connection.

References

- Adams, R. M. (1979). Primitive thisness and primitive identity. *The Journal of Philosophy*, LXXVI(1), 5–26.
- Anderson, J. L. (1967). *Principles of relativity physics*. New York: Academic Press Inc.
- Atiyah, M. F., & Singer, I. M. (1984). Dirac operators coupled to vector potentials. *Proceedings of the National Academy of Sciences of the USA*, 81, 2597–2600.
- Auyang, S. Y. (1995). *How is quantum field theory possible?*. Oxford: Oxford University Press.
- Baez, J. C. (2005). *Higher gauge theory, homotopy theory and n-categories*. Lectures at topics in homotopy theory, Graduate Summer School at the Pacific Institute for the Mathematical Sciences, University of Calgary. Available at: (<http://math.ucr.edu/home/baez/calgary/calgary.pdf>).
- Barbour, J. B. (1992). Einstein and Mach's principle. In J. Eisenstaedt, & A. J. Kox (Eds.), *Studies in the history of general relativity*. Boston: The Center for Einstein Studies, Boston University.
- Belot, G. (2003). Symmetry and gauge freedom. *Studies in History and Philosophy of Modern Physics*, 34, 189–225.
- Birmingham, D., Blau, M., Rakowski, M., & Thompson, G. (1991). Topological field theory. *Physics Reports*, 209.
- Bonora, L., & Cotta-Ramusino, P. (1983). Some remarks on BRS transformations, anomalies and the cohomology of the lie algebra of the group of gauge transformations. *Communications in Mathematical Physics*, 87, 589–603.
- Donaldson, S. K., & Kronheimer, P. B. (1990). *The geometry of four-manifolds*. Cambridge, MA: Oxford University Press.
- Earman, J., & Norton, J. (1987). What price spacetime substantialism? The hole story. *British Journal for the Philosophy of Science*, 38, 515–525.
- Faddeev, L., & Popov, V. N. (1967). Feynman diagrams for the Yang–Mills field. *Physics Letters*, 25B, 29–30.
- Faddeev, L., & Slavnov, A. (1991). *Gauge fields: An introduction to quantum theory* (2nd ed.). Frontiers in Physics, Perseus Books.
- Friedman, M. (1983). *Foundations of space–time theories. Relativistic physics and philosophy of science*. Princeton, NJ: Princeton University Press.
- Gambini, R., & Pullin, J. (1996). *Loops, knots, gauge theories and quantum gravity*. Cambridge, MA: Cambridge University Press.
- Gribov, V. (1978). Quantization of non-Abelian gauge theories. *Nuclear Physics B*, 139, 1.
- Healey, R. A. (1991). Holism and nonseparability. *The Journal of Philosophy*, 88(8), 393–421.
- Henneaux, M., & Teitelboim, C. (1994). *Quantization of gauge systems*. Princeton, NJ: Princeton University Press.
- Huggett, N., & Hoefer, C. (2006). Absolute and relational theories of space and motion. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Fall 2006 Edition) (<http://plato.stanford.edu/archives/fall2006/entries/spacetime-theories>).
- Kobayashi, S., & Nomizu, K. (1963). *Foundations of differential geometry* (vol. 1). New York: Wiley.
- Lyre, H. (2001). The principles of gauging. *Philosophy of Science*, 68(3), 371–381.
- Martin, C. A. (2002). Gauge principles, gauge arguments and the logic of nature. *Philosophy of Science*, 69, 221–234.
- Maudlin, T. (1998). Discussion: Healey on the Aharonov–Bohm effect. *Philosophy of Science*, 65, 361–368.
- Misner, C. W., Thorne, K. S., & Wheeler, J. A. (1973). *Gravitation*. W.H. Freeman.
- Norton, J. (1992). The physical content of general covariance. In J. Eisenstaedt, & A. J. Kox (Eds.), *Studies in the history of general relativity*. Boston: The Center for Einstein Studies, Boston University.
- Norton, J. (1993). General covariance and the foundations of general relativity: Eight decades of dispute. *Reports of Progress in Physics*, 56, 791–858.

- Norton, J. (2003). General covariance, gauge theories, and the Kretschmann objection. In K. Brading, & E. Castellani (Eds.), *Symmetries in physics: Philosophical reflections*. Cambridge: Cambridge University Press.
- O’Raifeartaigh, L. (1995). *The dawning of gauge theory*. Princeton, NJ: Princeton University Press.
- Redhead, M. (2003). The interpretation of gauge symmetry. In K. Brading, & E. Castellani (Eds.), *Symmetries in physics: Philosophical reflections*. Cambridge: Cambridge University Press.
- Rovelli, C. (2004). *Quantum gravity*. Cambridge, MA: Cambridge University Press.
- Singer, I. (1978). Some remarks on the Gribov ambiguity. *Communications in Mathematical Physics*, 60, 7–12.
- Teller, P. (1987). Space–time as a physical quantity. In R. Kargon, & P. Achinstein (Eds.), *Kelvin’s Baltimore lectures and modern theoretical physics* (pp. 425–447). Cambridge, MA: The MIT Press.
- Weyl, H. (1949). *Philosophy of mathematics and natural science*. Princeton: Princeton University Press.
- Yang, C. N., & Mills, R. L. (1954). Conservation of isotopic spin and isotopic gauge invariance. *Physics Review*, 96, 191–195.