Gauge Theories and Holisms

Richard Healey

University of Arizona
(rhealey@U.Arizona.edu)

Abstract

Those looking for holism in contemporary physics have focused their attention primarily on quantum entanglement. But some gauge theories arguably also manifest the related phenomenon of nonseparability. While the argument is strong for the classical gauge theory describing electromagnetic interactions with quantum "particles", it fails in the case of general relativity even though that theory may also be formulated in terms of a connection on a principal fiber bundle. Anandan has highlighted the key difference in his analysis of a supposed gravitational analog to the Aharonov-Bohm effect. By contrast with electromagnetism in the original Aharonov-Bohm effect, gravitation is separable and exhibits no novel holism in this case. Whether the nonseparability of classical gauge theories of non-gravitational interactions is associated with holism depends on what counts as the relevant part-whole relation. Loop representations of quantized gauge theories of non-gravitational interactions suggest that these conclusions about holism and nonseparability may extend also to quantum theories of the associated fields.

Key words: holism, gauge, nonseparability, electromagnetism, gravitation, Yang-Mills

1. Introduction

Theories of contemporary physics have frequently been taken to offer holistic accounts of their subject matter, suggesting that in some sense we indeed live in an undivided universe (to echo the title of Bohm and Hiley(1993)). Elsewhere(1991, 2004) I have offered analyses of what physical holism might amount to, as well as of the distinct but related notion of nonseparability. Much of the contemporary interest in holism and nonseparability has stemmed from attempts to come to terms with physical phenomena associated with violations of Bell inequalities (Bell([1964], 1987), Cushing and McMullin(1989), etc.). Quantum mechanics predicted this violation by ascribing an entangled quantum state to pairs of "particles"–a state that was not determined by the quantum states of its component "particles". But the notions of holism and nonseparability are general enough to be applied more widely in physics–to classical as well as quantum systems, and to fields as well as particles. Contemporary physics describes all fundamental interactions by gauge field theories–field theories (classical or quantum) possessing local gauge symmetries. There are reasons to suspect that some such theories involve a rather different kind of nonseparability–one that may or may not be associated with some variety of holism. This paper aims to explain how the suspicions arise and to give a preliminary assessment of the extent to which they prove well-grounded.

Section 2 offers analyses of holism and nonseparability and contrasts these with alternative analyses with some currency. Section 3 briefly sketches the fiber bundle formalism that has been advocated as exhibiting the intrinsic structure of gauge theories of non-gravitational interactions, and shows how it may be extended to accommodate general relativity. But fiber bundles need not appear in an even more intrinsic characterization of a gauge theory, as section 4 explains; for the key notion of a holonomy may be defined in

terms of the group of loops, prior to the introduction of a fiber bundle representation. Section 5 argues in favor of an interpretation of classical electromagnetism according to which it offers a nonseparable account of an interaction with quantum "particles"in the Aharonov-Bohm effect and elsewhere. Section 6 describes a supposed gravitational analog to the Aharonov-Bohm effect in order to draw an important contrast with the separability of general relativity's treatment of the gravitational interaction. Section 7 considers the extent to which either electromagnetism or gravitation exhibit holism. The gauge theories used to describe nongravitational interactions in the standard model are not classical but quantized. Section 8 sketches procedures for quantizing these theories, with particular reference to loop representations. Section 9 makes a preliminary attempt to assess the implications of loop representations of quantized gauge theories for the issues of holism and nonseparability in physics. The paper's conclusions are summarized in the final section.

2. Holism and Nonseparability

Holism is the maxim that the whole is more than the sum of its parts. Like all maxims, this requires clarification before it can usefully be applied. Elsewhere(1991, 2004) I have explicated one variety of metaphysical holism as an assertion to the effect that the natures of some wholes are not determined by those of their parts. Here is a more precise statement.

Physical Property Holism: There is some set of physical objects from a domain *D*, not all of whose qualitative intrinsic physical properties and relations supervene on qualitative intrinsic physical properties and relations of their basic physical parts.

It is notoriously difficult to say, in general, what makes a feature qualitative, intrinsic or even physical. But these notions may be clear enough in the context of particular physical theories. Indeed, part of what is involved in interpreting a physical theory is to say what are the

qualitative intrinsic physical properties and relations of objects in its domain. What makes

something a basic physical part also needs to be determined in context, and we shall see that

there may be different ways of understanding that idea even after settling on an interpretation

of a theory. The modality implicit in the notion of supervenience involved here is not logical

or metaphysical, but broadly physical. The idea is that supervenience obtains if and only if

any models of the relevant physical theories that agree on the features of the parts also agree

on the features of the wholes they compose.

Teller(1987, p.214)) has introduced a similar notion as follows:

Relational Holism: There are non-supervening relations...that is, relations that do not

supervene on the nonrelational properties of the relata.

Within physics this specializes to a close relative of physical property holism, namely

Physical Relational Holism: There are physical relations between some physical

objects that do not supervene on their qualitative intrinsic physical properties.

Physical property holism entails physical relational holism, but not *vice versa*.[1] As we'll see,

this makes physical property holism the more radical, and perhaps more interesting, notion.

Not everyone has been content to equate holism in physics with failure of

supervenience. Jones(forthcoming) objects that mere appeal to an absence of supervenience

does not suffice to yield genuine understanding of the kinds of quantum correlations involved

_____

[1]Suppose that F is some qualitative intrinsic physical property or relation of one or more elements of D that fails to supervene on qualitative intrinsic physical properties and relations of their basic physical parts. We may define a (non-intrinsic) physical relation R_F to hold of the basic physical parts of elements of D if and only if F holds of these elements. Clearly R_F does not supervene on the qualitative intrinsic physical properties of these parts. So physical property holism entails physical relational holism. But the converse entailment fails. For let R_G be a physical relation that holds between the basic parts of some elements in D when and only when those elements are in the relation S_G. R_G may fail to supervene on the qualitative intrinsic physical properties of these basic parts, even though all qualitative intrinsic physical properties and relations of elements of D (including S_G) supervene on the qualitative intrinsic physical properties *and relations* of their basic parts.

in violations of Bell inequalities, and that this will require a "thicker" notion of holism in the quantum domain. Esfeld(2001) also takes holism, here and elsewhere, to involve not just a failure of supervenience. He maintains that a compound quantum system is holistic in that its subsystems themselves count as quantum systems only by virtue of their relations to other subsystems together with which they compose the whole. But while alternative explications of the notion of holism may be suited to different applications, I hope to show in this paper that a simple and relatively clear identification of holism with failure of supervenience still has much to teach us when applied in the philosophy of physics.

As a quick preliminary exercise before embarking on the main topic of gauge theories, consider Jones's objection against equating holism with failure of supervenience. He points out that while the quantum states of the members of a correlated EPR-Bohm pair do not determine the state of the pair, it is true also that the quantum states of two completely uncorrelated systems may not determine their joint quantum state. Suppose we have two uncorrelated pairs of spin ½ systems, with each pair in the spin singlet state. Call these pairs (1+2) and (3+4). Now consider the quantum states of systems 1 and 3. The spin state of each of 1 and 3 is mixed: it is represented by the identity operator I on its respective spin Hilbert space, as is the joint state of the compound system (1+3). But this joint mixed state does not supervene on the individual mixed states, since the individual state of each of 1 and 3 would also have been represented by I if (1+3) had instead been in the joint singlet spin state. Jones's point is that there is no reason to invoke talk of holism here to describe the behavior of (1+3), even though its state does not supervene on those of 1,3 separately.

But note that physical property holism denies supervenience not of states but of intrinsic properties and relations. And (assuming intrinsic spin properties and relations are here connected to states by the eigenvalue-eigenstate link) there is no intrinsic spin property

or relation possessed by (1+3) in the situation described that it would not also have possessed in the singlet state. Rather, the singlet state of (1+3) would have been one in which it had *additional* properties, such as having zero angular momentum. A careful application of physical property holism reveals no holism in the situation described: all the intrinsic spin properties and relations of (1+3) do in fact supervene on the intrinsic properties and relations of 1,3 separately. But it is noteworthy that to reach this conclusion it is necessary to deny that a "negative" property such as *not having zero angular momentum* counts as intrinsic.

Holism is closely related to nonseparability–a notion that may also be understood in terms of failure of supervenience, as follows (cf. Healey(1991, 2004)):

> *Nonseparability*: Some physical process occupying a region R of spacetime is not supervenient upon an assignment of qualitative intrinsic physical properties at spacetime points in R.[2]

Physical property holism typically leads to nonseparability. For suppose that a compound system undergoes a process occupying a region R of spacetime, during which its qualitative intrinsic physical properties are not supervenient upon those of its spatially separated basic physical parts, even together with the spatial or other qualitative intrinsic physical relations among these parts. Then that process will be nonseparable, since it involves states of the whole system that fail to supervene upon an assignment of qualitative intrinsic physical properties at spacetime points in R. But one can also imagine cases of nonseparability that do

---

[2]This raises the interesting issue of whether an assignment at a point of a "local" geometric object like velocity, or the spacetime metric or curvature, amounts to such an assignment. For the standard definition of such a quantity at point p is in terms of a limiting process that involves nearby points in a sequence of neighborhoods enclosing p. But having raised this issue, it will simplify the subsequent discussion to bracket it in what follows. The novel question posed by gauge theories is whether they demonstrate nonseparability even on the assumption that an assignment of "local" quantities at spacetime points *does* constitute an assignment of qualitative intrinsic physical properties at those points.

not involve any holism: even an object without proper parts could undergo a nonseparable process. Gauge theories may present a more interesting example of nonseparability without holism, as section 7 will explain.

3. Fiber Bundle Formulations of Gauge Theories

There is widespread agreement that the quantum field theories describing fundamental interactions of elementary particles in the so-called standard model are gauge theories. But what exactly is a gauge theory? Two general answers to this question have risen to prominence in recent years. While classifying theories of strong, weak and electromagnetic interactions as gauge theories, these answers do not always agree in affixing that label to other theories; and even when they do agree (e.g. in classifying general relativity as a gauge theory) they may do so for quite different reasons.

Following Henneaux and Teitelboim(1992), Earman(2002) and Belot(2003) employ an understanding according to which a gauge theory is one whose Hamiltonian formulation leads to constraints of a certain kind–identities satisfied by the canonical variables. That includes not only the contemporary Yang-Mills field theories of the standard model, but also certain classical particle theories, as well as general relativity as it is usually presented in terms of tensor fields on a differentiable manifold. This constrained Hamiltonian approach to gauge theories was pioneered by Dirac and others as a preliminary to quantization of field theories including general relativity. I will return to it in this context in section 8.

But I will begin by pursuing an alternative understanding of gauge theories that has often been taken to yield insights into the geometry and topology of Yang-Mills gauge fields like those that figure in the standard model of elementary particles. Trautman(1980) succinctly stated this approach, holding that it may be extended to encompass gravitation:

For me, a gauge theory is any physical theory of a dynamic variable which, at the

classical level, may be identified with a connection on a principal bundle. The

structure group *G* of the bundle *P* is the group of gauge transformations of the first

kind; the group Œ of gauge transformations of the second kind may be identified with

a subgroup of the group *Aut P* of all automorphisms of *P*. In this sense, gravitation is

a gauge theory. (p. 306)

Now while theories of electromagnetic, weak and strong interactions are often now

formulated in these terms, general relativity (currently our best theory of the gravitational

interaction) is not. But although it is not a Yang-Mills theory, general relativity may still be

formulated as a theory of the connection and curvature of a principal fiber bundle. However

unnatural, this formulation will further the goals of this paper by highlighting an illuminating

contrast between Yang-Mills theories and general relativity with respect to holism and

nonseparability.

One can take the structure group for a fiber bundle formulation of general relativity to

be a group of which the Poincaré group (including spacetime translations as well as

homogenous Lorentz transformations) is a subgroup. This may seem odd, since spacetime

translations are most naturally understood to be represented on the base manifold rather than

in the fibers (cf. Redhead(2002, p294)). But it is certainly possible to formulate general

relativity in this way, and doing so will help to reveal an interesting sense in which general

relativity is less holistic than electromagnetism or other Yang-Mills theories–a sense that will

be clarified in sections 6 and 7.

Classical electromagnetism will serve as a simple illustrative example of a gauge

theory in its fiber bundle formulation. Take the spacetime manifold as base space, and

assume for now this has the structure of Minkowski spacetime. The structure group is

U(1)–the group of rotations about the origin in the complex plane. The electromagnetic potential is represented by the bundle connection–a 1-form taking values in the Lie-algebra of the *U(1)* group. The connection is a geometric object that specifies the result of "lifting" a smooth curve on the base space "up" into a corresponding curve in the bundle–its so-called horizontal lift. It does this by mapping tangent vectors at a point to curves in the manifold onto a subspace of tangent vectors at each bundle point in the fiber "above" that point–the horizontal subspace. The bundle curvature represents the electromagnetic field strength: it is a 2-form that specifies how far away the horizontal lift of any closed curve is from closing. A bundle section consists of a smooth function constituting a selection of a single element from the fiber above each base point: provided there are no magnetic monopoles this is global, i.e. defined over the entire manifold. (Wu and Yang(1975) first introduced a fiber bundle formulation of Yang-Mills gauge theories in a paper that gave a singularity-free treatment of the effect of classical electromagnetism on quantum particles, even in the presence of hypothetical magnetic monopoles.) The usual 4-vector potential A_μ arises as the pull-back of the bundle connection onto the base manifold generated by a section: a change of section results in a gauge transformation of the form

$$A_\mu \ Y \ A_\mu \nmid= A_\mu \ + \ \partial_\mu \Lambda \qquad\qquad (1)$$

for an arbitrary but suitably differentiable function *Λ(x)*. The usual electromagnetic field tensor arises as the pull-back of the bundle curvature: it is represented on the spacetime manifold by the (gauge-invariant) object $F_{\mu\nu}(x)$.

The wave-function of particles of charge *e* is represented on an associated vector bundle whose fibers are complex numbers–a vector representation of *U(1)*. A choice of gauge corresponds to a simultaneous choice of section for the principal and associated vector bundles. The value of a wave-function at a spacetime point in that gauge is a complex

number corresponding to the fiber element (from the complex vector space $\div$) picked out by that section: a gauge transformation induces the transformation

$$\psi(x) \curlyvee \psi'(x) = \exp(/\ ie\Lambda(x)\ /\mathcal{S})\ \psi(x) \qquad (2)$$

All this generalizes to Yang-Mills theories with non-Abelian structure groups such as SU(2) and SU(3). After quantization, such theories have been used with great success to treat weak and strong interactions in the standard model of elementary particles. In each case a connection on the principal fiber bundle represents the gauge potential, while its curvature represents the corresponding field. And in each case, these objects may be projected down onto the spacetime base manifold in a particular section to represent the potential and field strength in a corresponding gauge. Gauge transformations of wave-functions and potentials induced by changes of section take a more complex form than for the case of electromagnetism; and field strengths are no longer gauge-invariant, but rather covariant, under such gauge transformations.

At first sight the fiber bundle formulation of electromagnetism appears to be more intrinsic than the older formulation in terms of 4-vector potentials, in so far as a gauge transformation corresponds to a mere change of section (analogous to a coordinate transformation), while the bundle connection and curvature are geometric features of the bundle that are independent of choice of section. But recall that Trautman took a gauge transformation (of the second kind) to correspond not to a change of section but to a bundle automorphism–a structure preserving mapping of the bundle onto itself. This alternative understanding will prove important in section 5, which will defend a more genuinely intrinsic representation in terms of the notion of holonomies introduced in section 4.

Consider now a formulation of general relativity in terms of a principal bundle *A(M)* of affine frames. Recall that a frame is a linearly independent set of basis vectors of the

tangent space at a point, and that an affine transformation is an invertible inhomogeneous linear transformation applied to this space, considered as an affine space. (Transformations in the Poincaré group are a subgroup of the group of affine transformations, namely those that preserve the Minkowski metric.) The gravitational gauge potential is now represented by an affine connection on this bundle. It is important not to confuse this with the usual linear connection ⌞ in general relativity, which is defined instead on the bundle of linear frames, *F(M)*. But though distinct, the two connections are intimately related.

For there is a natural homomorphism from *F(M)* onto *A(M)* which induces a mapping of the connection on *A(M)* onto a one-form on *F(M)* that decomposes into the sum of two parts. One part is simply the usual general relativistic connection ⌞. The other part is called the solder form on *F(M)*, since it "solders" *F(M)* to *M*. According to Trautman (*ibid*) "The most important difference between gravitation and other gauge theories is due to the soldering of the bundle of frames [*F(M)*] to the base manifold *M*". This occurs as follows. The solder form yields a unique way of "projecting" a vector in the tangent space to *F(M)* down onto a vector in the tangent space to *M*. Simply take the components of the original vector at point *e* in *F(M)* with respect to the basis defined by the frame at *e*, and regard these as the components of a new vector at the point *p* in *M* "below" *e* with respect to that same frame, now considered as furnishing a basis for the tangent space at *p*.

The solder form is important because it restricts the bundle automorphisms corresponding to gauge transformations of the second kind ("local" gauge transformations) to those that simply correspond to diffeomorphisms of the base space. They do not "rotate" the elements in the fibers, but just carry the horizontal lift of a curve in the base space into the horizontal lift of its image under such a diffeomorphism. Section 6 explains why this implies that general relativity, unlike Yang-Mills gauge theories, is separable.

4. Holonomies and the Group of Loops

It was assumed in the previous section that a gauge theory is naturally formulated in the fiber bundle framework, and that this formulation offers an intrinsic, gauge-invariant representation of its physical content. But a fiber bundle is a very abstract mathematical structure that is far removed from the physical world. The present section shows how to narrow the gap by explaining why a more basic physical structure may be represented by the connection on a principal fiber bundle. The structure concerns what is called the group of loops. Fiber bundles are carriers of representations of this structure in the following way. Any sufficiently smooth homomorphism from the group of loops into a suitable Lie algebra may be represented by the holonomies of a connection on a principal fiber bundle having that Lie algebra as its structure group. Let me explain.

A connection on a fiber bundle determines an associated set of holonomies. Consider a horizontal lift of a closed curve C beginning and ending at base point p. For each point e in the fiber above p, the horizontal lift through e will trace out a corresponding curve in the bundle that returns to a point f in the fiber above p. In this way the connection maps the fiber above p isomorphically onto itself. The map is called the holonomy map, and its action on a point e is an element of the bundle's structure group called the holonomy of C at e.

Conversely, the holonomies of a gauge theory (relative to an arbitrary base point p) determine the connection, but only up to a "vertical" bundle automorphism–that is, a smooth map that acts as an isomorphism from the fiber above each point onto itself.

In the case of electromagnetism, the holonomy of a closed curve C corresponds to the gauge-invariant Dirac phase factor

$$\exp[\,!\,(ie/S)\check{S}_C\, A_\mu(x^\mu)\, dx^\mu\,]$$

where e is the charge of particles subject to an interaction represented (in an arbitrary gauge)

by the 4-vector potential A_μ. In Yang-Mills theories with a non-Abelian structure group the holonomies themselves are not gauge invariant, but their traces are–these are known as Wilson loops. Various theorems (for references, see Gambini and Pullin(1996), §3.4.3) show that, in cases of interest, all gauge-invariant information contained in the holonomies can be reconstructed from Wilson loops. This is important, because the holonomies determine, up to a gauge transformation, the geometry of a connection on a principal fiber bundle, together with the topology of the bundle. So the Wilson loops already contain all the physical information represented in a fiber bundle representation of a Yang-Mills theory, with the possible exception of what is represented by a particular choice from within the class of gauge-equivalent connections. I shall argue in section 5 that this leaves nothing out.

A generalization of the notion of holonomy may be defined intrinsically without any reference to connections or fiber bundles. Holonomies can be viewed as homomorphisms from what is called the group of loops onto a Lie group ‹ . One can then recover the usual fiber bundle formulation from a sufficiently smooth homomorphism as a representation. While loops may not differ from closed curves at an intuitive level, closed curves do not form a group on the standard mathematical definition of a curve. To impose a group structure, it is therefore necessary to define loops as appropriate equivalence classes of closed curves[3]. This

---

[3]A closed curve C may be defined as a smooth (continuous and piece-wise differentiable) map from the unit circle into the differentiable manifold M. An equivalence class associated with a particular closed curve C includes all other curves that share its image and result from it by orientation-preserving reparametrizations. Such an equivalence class constitutes an unparametrized curve. Restrict attention now to unparametrized curves that begin and end at a fixed point o of M. These form a semi-group under the operation Β of composition of curves (which extends naturally to unparametrized curves), where the identity element is the curve whose image is simply o. But there is no inverse, since the image of a curve CΒCΝ that results from composing an arbitrary curve with a curve with the same image but opposite orientation is not just o. To arrive at a group structure, it is necessary to widen the equivalence classes so that CΒCΝand the curve whose image is simply o are both elements of the identity. This may be achieved by identifying unparametrized curves that differ only on a finite number of "trees"–curves that enclose no area. The resulting equivalence classes

defines the notion L_o of a group of loops beginning and ending at base point o of a manifold.

It is known (Gambini and Pullin(1996), §1.4) that any homomorphism $h{:}L_o\Upsilon\langle$ into a Lie group $\langle$ defines a holonomy map associated with a generalized connection, and that for a sufficiently smooth homomorphism there exists a principal fiber bundle whose connection yields precisely this holonomy map H. Since h is a homomorphism, we have $H(\alpha\mathbb{B}\beta) = H(\alpha)\mathbb{B}H(\beta)$ and $H(\alpha^{|1}) = (H(\alpha))^{|1}$ for loops α,β, where the holonomy map H has been uniquely extended in a natural way from curves to loops. While h thus defines a holonomy map H that is unique up to conjugation ($H\mathbb{N}(\alpha)=g\mathbb{B}H(\alpha)\mathbb{B}g^{|1}$ for a fixed element g), this holonomy map is compatible with more than one bundle connection: a vertical bundle automorphism preserves all holonomies (up to conjugation) while transforming the connection that gives rise to them. For Trautman this just counts as a local gauge transformation–one "of the second kind".

5. Classical Electromagnetism is Nonseparable in the Quantum Domain

In a wholly classical context, electromagnetism acts on charged particles only through the electromagnetic field that gives rise to the Lorentz force: the electromagnetic potential has no independent manifestations, and seems best regarded as an element of "surplus mathematical structure", in itself representing nothing physical. But the situation is different in the quantum domain. Phenomena such as the Aharonov-Bohm effect there provide vivid illustrations of the fact that there is more to classical electromagnetism than just the field. For the effects of electromagnetism on the phase (and subsequent observable behavior) of charged particles that pass through a region of space are not always wholly determined by the field there while they pass. In the quantum domain, it is not the field but the electromagnetic

---

are the *loops*. With these definitions, the loops with base point o constitute a group L_o under composition (uniquely defined by reference to composition of their constituent curves). L_o is called the *group of loops* at o.

potential itself that appears to give rise to these effects. Indeed, the vector potential appears explicitly and ineliminably in standard textbook explanations of Aharonov-Bohm type effects in terms of the action of classical electromagnetism on quantum particles.

From the point of view of the formulation of classical electromagnetism presented in section 3, such a vector potential arises merely as the representation in a particular gauge of a geometric structure–the connection on a principal fiber bundle. The gauge transformation (1) corresponds to a change from one to another equivalent representation of the same connection, induced by a change of bundle section. The connection itself is gauge-invariant, and offers a more intrinsic representation of the "extra" content of electromagnetism revealed in the quantum domain, beside the classical field strength. In this view, the electromagnetic potential is represented by the connection, while the electromagnetic field is represented by the curvature, of a principal fiber bundle with structure group $U(1)$. But there is a problem.

By performing a "vertical" bundle automorphism, one can transform one bundle connection into a distinct connection that is empirically equivalent to it, even in the quantum domain. If just one of these connections correctly represents the gauge potential, that potential remains epistemically inaccessible. A statement saying which is the true bundle connection may even lack cognitive significance. Such purportedly "active" gauge transformations are analogous to the manifold diffeomorphisms in general relativity that underlie the "hole" argument. One concludes in the case of the "hole" argument that diffeomorphic models of general relativity may be taken to represent the same physical situation. What is the appropriate conclusion in the case of a classical gauge theory of electromagnetism?

The Dirac phase factor offers a clue. It is a gauge-invariant object that suffices to account for the AB effect and all related phenomena involving the phases of charged particles

15

in the quantum domain. The Dirac phase factor generalizes to other gauge theories as the holonomy of a fiber bundle connection. In a classical Yang-Mills gauge theory like electromagnetism, distinct bundle connections which yield the same holonomies are empirically indistinguishable. All the evidence for a classical Yang-Mills gauge theory in the quantum domain is evidence not for any particular connection from a holonomy-equivalent family, but only for the holonomies they all generate. But, as it happens, electromagnetism is the only classical Yang-Mills gauge theory for which we have such evidence: the non-Abelian theories of weak and strong interactions are quantum field theories whose classical counterparts have no practical applications.

The status of the bundle connection in a classical Yang-Mills gauge theory is analogous to that of absolute space for Newton. According to his theories, there could be no evidence favoring any hypothesis about the absolute velocity of any inertially moving object. Those theories are therefore best interpreted as postulating no privileged state of rest in absolute space, but only an affine structure specifying states of inertial motion. A classical Yang-Mills gauge theory is best interpreted as postulating no privileged bundle connection, but only a holonomy structure specifying the effects of the associated interaction on the phase differences of a wave-function describing charged particles.[4]

How can this holonomy structure be understood? The previous section explained that a holonomy map of a gauge theory may be thought of as a homomorphism from the group of loops into a Lie group. Corresponding to each nontrivial loop γ is a closed, oriented one-dimensional spacetime region r_γ we may call a *ring*.[5] The holonomy of γ represents a property

---

[4] This conclusion is defended more fully in Healey(2001).

[5] Apart from the identity, each loop γ contains an unparametrized curve with no "trees" whose image defines an oriented "ring" r_γ in spacetime–an extended region represented by an oriented 1-dimensional embedded submanifold. The value of the holonomy

of r_γ (or an event located there)—a property that may manifest itself in its effects on matter that encounters that ring. The representation is in some ways analogous to the way a vector such as the electric field vector E represents properties. E manifests itself in its effects on charged matter that encounters a point where $E \neq 0$. But there are two significant differences.

1) The value of E may be given by a set of real numbers in some coordinate system. The value of a holonomy map is given more abstractly by an element of a Lie algebra in some gauge: only for an Abelian theory is this gauge-invariant (though Wilson loops are gauge-invariant).

2) While E represents properties at a point, a holonomy represents properties "at" a ring. By denying the physical significance of any connection or vector potential "underlying" the holonomy, one is committed to the nonseparable character of the classical gauge theory.

The first difference is relatively unimportant. A holonomy property at a ring acts by modifying the phase-change of particles' wave functions on traversing that ring. This phase-change is itself not gauge invariant in a non-Abelian theory, in so far as it depends on the initially arbitrary choice of phase at the base point to represent a particular physical state: it transforms by conjugation under a non-Abelian generalization of (2). And so the holonomy map cannot be expected to be gauge invariant either: it too is defined only up to conjugation in a non-Abelian theory. The choice of a particular holonomy map from this conjugacy class is then analogous to a choice of coordinate frame with respect to which the components of E are given at a point. Clearly, this affects the way the gauge potential and its effects are represented, not the underlying physical properties themselves.

It is the second difference that proves significant for the issue of holism. In so far as we have evidence supporting the application of a classical gauge theory like

homomorphism for γ is γ's holonomy.

electromagnetism in the quantum domain, we have reason to believe that it describes an interaction that is nonseparable. For it associates intrinsic physical gauge properties directly with rings in space (more generally spacetime) that fail to supervene on any intrinsic physical properties assigned at the points that lie on these rings (or even on their infinitesimal neighborhoods). Section 7 explains why this nonseparability of a classical gauge theory implies a certain kind of holism.

6. A Gravitational Analog to the Aharonov-Bohm Effect?

If classical electromagnetism is nonseparable in its action on quantum particles, and general relativity, our best classical theory of gravitation, may also be formulated in terms of the connection and curvature of a principal fiber bundle, then one might suspect that general relativity also acts nonseparably on quantum particles. Moreover, the striking quantum phenomena of electromagnetically induced phase-shifts in a field-free region used to illustrate the nonseparability of electromagnetism in the previous section have an apparent analog in general relativity. For there is a model of general relativity that represents a situation bearing a strong resemblance to the circumstances of the magnetic Aharonov-Bohm effect. But a closer examination of this model will show that the formal differences between general relativity and Yang-Mills gauge theories pointed out in section 3 give rise to significant physical disanalogies between the Aharonov-Bohm effect and its purported general relativistic analog. These will reinforce the conclusion of section 5 by highlighting an important contrast between general relativity and Yang-Mills gauge theories that bears importantly on the topic of this paper. Unlike Yang-Mills theories, general relativity is separable in the quantum domain.

The Aharonov-Bohm effect comes about because the effects of electromagnetism on

the phase (and subsequent behavior) of charged quantum particles that pass through a region of space are not always wholly determined by the electromagnetic field there while they pass. General relativity predicts an analogous effect in which the effects of gravity on the phase (and subsequent behavior) of quantum particles that pass through a region of space are not always wholly determined by the spacetime curvature in that region while they pass. The fiber bundle formulations described in section 3 make the analogy even clearer. In each case, the bundle curvature is zero everywhere in the region through which the particles pass. Nevertheless, the bundle connection defines parallel transport around a closed curve in that region, resulting in a phase change around any curve that encloses a separate central region in which the curvature is nonzero.

As Anandan(1993) notes, Marder(1959) described a static general relativistic spacetime with cylindrical symmetry which is empty of matter except for an infinite central cylinder. Everywhere outside the cylinder the Riemann curvature tensor is zero and the metric is Minkowski. The 2-dimensional spatial geometry of each normal section through the cylinder is the same, and may be represented in a hypothetical embedding space by the surface of a truncated cone with a smooth cap. If a tangent vector is parallel-transported along a closed curve once around the cylinder it is rotated through the angle of the cone, even though the curvature is zero everywhere along the curve.

In the Aharonov-Bohm effect, the interference pattern produced by a beam of charged particles passing through a region is not determined solely by the electromagnetic field there while they pass: it depends also on the field in regions from which they are excluded. In particular, the position of the intensity maxima produced by a beam after it has encircled an infinite solenoid while traversing a region throughout which the field is zero, depends on the current passing through the solenoid. The phase shift along a curve C is given by the

expression

$$F_C \; = \; \exp[\, ! \, (ie/\mathcal{S}) \, \check{\mathcal{S}}_C \; A_\mu(x^\mu) \, dx^\mu \,] \qquad\qquad (3)$$

In the gravitational analog, the interference pattern produced by a beam of spinning particles passing through a region is not determined solely by the spacetime curvature in that region: it depends also on the spacetime curvature in regions from which they are excluded. In particular, the position of the intensity maxima produced by a beam after it has encircled an infinite "string" while traversing a completely flat region, depends on the energy and momentum flow inside the "string". The phase shift along a curve C is given by the expression

$$F_C \; = \; \exp[\, ! \, (ie/\mathcal{S}) \, \check{\mathcal{S}}_C \; (e_\mu{}^a P_a + \tfrac{1}{2}\, \Gamma_\mu{}^{ab} M_{ab}) \, dx^\mu \,] \qquad\qquad (4)$$

where M_{ab} are (representations of) the generators of the Lorentz group and P_a of the translation group, $e_\mu{}^a$ is dual to the frame field $e_a{}^\mu$, $\Gamma_\mu{}^{ab}$ are the components of the connection $\llcorner$ with respect to $e_a{}^\mu$, and the integral is path-ordered. The second summand in equation (4) represents the phase-shift due to the particles' spin. These expressions for the phase change can both be seen as instances of the general equation

$$F_C \; = \; \exp[\, ! \, (ig/\mathcal{S}) \, \check{\mathcal{S}}_C \; A_\mu{}^k T_k \; dx^\mu \,] \qquad\qquad (5)$$

where the T_k generate the Lie algebra of the structure group of an arbitrary gauge field, and $A_\mu{}^k$ represents the corresponding gauge potential.

But despite these analogies between the effects of gravity on quantum particles in Marder spacetime and the Aharonov-Bohm effect, Anandan(1993) has also pointed out a significant disanalogy. While electromagnetic phase differences given by (3) are physically significant and measurable only around closed curves, gravitational phase differences given by (4) are physically significant, and potentially measurable, also along *open* curves. For

example, for spinless particles, the gravitational phase acquired by a locally plane wave along a classical trajectory C is given approximately by

$$\varphi = (1/\hbar) \int_C e_\mu{}^a p_a \, dx^\mu \qquad (6)$$

where p_a are the eigenvalues of the energy-momentum operator $\hat{p}_a$. Unlike the electromagnetically induced phase change in the Aharonov-Bohm effect, this gravitationally induced phase change is physically significant, and potentially measurable, along open curves C, as well as closed curves. It would be observable by the Josephson effect for a path across the Josephson junction, or by the strangeness oscillations in the neutral Kaon system for an open timelike path C along the Kaon beam.

The solder form described in section 3 is the formal reason for this physical difference between the gravitational and electromagnetic interactions. As one transports the phase of quantum particles along a curve C, the vector representing their phase lies in the tangent space to the vector bundle associated with the principal affine frame bundle $A(M)$. The solder form effectively "locks" the phase vector to a corresponding spacetime vector in the tangent space to M. It is the linear connection L that defines parallel transport of this latter vector along C; and, because of the solder form, it also therefore uniquely specifies how the phase changes from point to point along C. In the case of gravitation, it is the solder form that privileges a unique connection in the affine frame bundle $A(M)$. The class of gauge symmetries is restricted to those that preserve the solder form. This excludes "vertical" bundle automorphisms, and leaves only automorphisms of $A(M)$ that simply correspond to diffeomorphisms of M. The absence of the solder form in electromagnetism corresponds to the absence of any such privileged connection in the $U(1)$ bundle.

The existence of a uniquely privileged connection on the principal bundle $A(M)$ implies that classical gravitation is separable in its action on quantum particles. The intrinsic

geometric properties at (an infinitesimal neighborhood of) each point in any region of spacetime determine the effects of gravitation on quantum particles in that region, including effects on their spins. By contrast, the intrinsic electromagnetic properties pertaining to (even an infinitesimal neighborhood of) each point in an extended region of spacetime do not determine the effects of electromagnetism on charged particles in that region: those effects further depend on holonomy properties assigned "at" extended rings in the region. Since the classical gravitational field is separable even in the quantum domain, there is an interesting sense in which it is less holistic than the classical electromagnetic field, as the next section will show.

7. Do Electromagnetic or Gravitional Interactions Involve Holism?

As section 2 made clear, holism and nonseparability are distinct, but intimately related notions. If classical electromagnetism is nonseparable in its interaction with quantum particles, while general relativity is separable, does it follow that electromagnetism involves holism but gravity does not? The answer to this question depends on exactly what holism amounts to in this context. Section 2 understood holism as arising from a failure of features of wholes to be determined by those of their proper parts. The wholes with which we are here concerned are extended regions of spacetime (and/or events or processes that occupy them): the parts are regions (and/or their occupants) that in some sense compose these. Whether we take electromagnetism or gravitation to exhibit holism will depend not only on what we consider to be the relevant features of the parts, but also on how we understand the composition relation between parts and wholes.

Because it is nonseparable in the quantum domain, there is a sense in which classical electromagnetism exhibits holism. For, as section 5 argued, classical electromagnetism is

22

nonseparable in its interaction with quantum particles: its description of such processes ascribes qualitative intrinsic physical properties "at" a ring that do not supervene on qualitative intrinsic physical properties assigned at spacetime points on that ring (nor even on their infinitesimal neighborhoods). This is a case of physical property holism provided that points on the ring (or events that occur at them) count as basic physical parts of the ring (or of an event that occupies it). For the electromagnetic holonomy properties "at" a ring are not determined by electromagnetic or any other intrinsic properties of such constituents, even taken together with the spatiotemporal and other intrinsic relations among them.

Because it is separable in the quantum domain, there is no corresponding sense in which classical general relativity exhibits holism. For, as sections 3 and 6 made clear, classical general relativity is separable in its interaction with quantum particles since all the qualitative intrinsic physical properties it ascribes on a ring do supervene on qualitative intrinsic physical properties assigned on (infinitesimal neighborhoods of) spacetime points on that ring. By contrast with electromagnetism, gravity (as described by classical general relativity) does not exhibit this kind of physical property holism.

But there is another way of understanding the part-whole relation in these cases. Two rings may be considered to compose a third just in case the corresponding loops compose its loop under the product operation in the group of loops. With this alternative understanding of the composition relation, one may take any basis of arbitrarily small rings (or events occupying them) as basic parts of (an event occupying) any larger ring they make up. The holonomy properties of this composite object will indeed supervene on those of its basic parts, and so there will be no physical property holism in either classical electromagnetism or classical general relativity!

The first way of taking the part-whole relation seems more relevant if one is chiefly

interested in issues of locality, though the primary notion there is not holism but nonseparability. But I see no conclusive reasons for preferring one understanding of the part-whole relation over the other here. This does not show that it is wrong to identify holism with failure of supervenience. Rather, it illustrates the richness of that conception of what holism amounts to in physics.

Section 2 already brought out a further aspect of this richness in an abstract way, by noting two different supervenience claims that holism may be taken to deny, leading to the distinct, but related, notions of physical property holism and physical relational holism. This aspect may now be made more concrete by looking once more at the supposed gravitational analog to the Aharonov-Bohm effect discussed in the previous section. Even though this involves no nonseparability, intuitively there does seem to be something holistic about the behavior of systems in Marder spacetime[6]. No experiment that one can carry out in any finite "patch" of that spacetime that does not overlap the central cylinder will be able to distinguish that "patch" from a corresponding region of Minkowski spacetime, even though the geometry (in particular, the topology) of the union of these patches differs from that of Minkowski spacetime. Indeed, this difference manifests itself at the purely classical level. Two classical gyroscopes carefully transported around different sides of the central cylinder would point in different directions when they are brought together again, even though each was transported through perfectly flat spacetime. Intuitively, this does seem to be a situation in which the features of the whole are not determined by those of its parts.

Indeed this may be seen as a case of physical relational holism. The intrinsic physical (specifically, geometrical) properties of the "patches" do not determine their physical relations. There are similar examples in the case of unusual spacetime topologies, such as that

---

[6] Here I am indebted to correspondence with Paul Teller.

of Minkowski spacetime "rolled up" into a cylinder along a spatial dimension. Indeed, even though Marder spacetime has the standard Minkowski topology globally, it still presents such an example since the topology of the flat region is not R^4 but $S \times R^3$.

General relativity is a dynamical theory whose models can describe a variety of topologically distinct spacetime structures, and the intrinsic geometric properties of every point or coordinate patch in the manifold of a model do not determine the global topology of the manifold. This implies physical relational holism, since whether two points or coordinate patches are nearby or far apart is not determined by their intrinsic physical properties. But we do not need to look to exotic topological structures of dynamic spacetimes to find instances of physical relational holism. For geometry presents much simpler examples.

Even in the Euclidean plane, intrinsic spatial relations among n-tuples of points do not supervene on the local geometrical properties of those points. Indeed spatiotemporal relations have long been considered examples of external relations–relations that do not supervene on the intrinsic properties of their relata. This kind of failure of supervenience makes relational holism a less discriminating analytic tool in the philosophy of physics. The interest of Marder spacetime does not lie in the fact that it provides yet another example of relational holism. It is an advantage of the alternative explication of holism offered by physical property holism that it is able to mark the crucial distinction between the kind of holism revealed by the Aharonov-Bohm effect and the lack of this kind of holism in its supposed gravitational analog in Marder spacetime.

Whether or not general relativity itself involves an interestingly novel holism in physics, it does present a challenge to the conclusion that the original Aharonov-Bohm effect manifests the nonseparability (and associated holism) of electromagnetism in the quantum domain. Teller has pointed to difficulties in applying section 2's analysis of nonseparability in

the context of a dynamical spacetime such as that of general relativity. Theoretical accounts of the Aharonov-Bohm effect customarily proceed on the assumption that spacetime is nondynamical: the account offered in section 5 assumed a fixed Minkowski spacetime. While this is clearly an excellent approximation, it should be possible also to analyze the effect within a dynamic general relativistic spacetime. But this causes complications when one tries to apply the definition of nonseparability given in section 2.

Because the energy-momentum tensor in the spacetime region of the solenoid is not quite zero, the spacetime outside the solenoid in the magnetic Aharonov-Bohm effect will not be quite Minkowski, whether or not a current flows through the solenoid: indeed, its structure will depend (to a very small degree) on the magnitude of the current. The definition of nonseparability therefore requires a comparison of the intrinsic features of physical processes in distinct, slightly different, spacetimes. But there is no general way of uniquely identifying the same spacetime point in distinct spacetimes: indeed, on some views (such as Maudlin's metric essentialism) a given spacetime point will not exist in each of two sufficiently different possible spacetimes. At first sight, this may appear to trivialize the definition of nonseparability in a dynamic spacetime, so that general relativity renders classical electromagnetism trivially separable in the Aharonov-Bohm effect and elsewhere in the quantum domain.

There are two responses to this difficulty. One is to point out that, by judicious (hypothetical) addition of non-electromagnetic contributions to the energy-momentum tensor within the solenoid one can arrange that the spacetime geometry is the same everywhere in two hypothetical situations in which different currents flow through the solenoid. (Though note that this is possible only because the electromagnetic potential outside the solenoid does not contribute to the energy-momentum tensor there.) This does obviate the need to compare

features of individual points in spacetimes of different geometries. But it may require the acceptance as physically possible of situations corresponding to models of general relativity involving purely notional matter distributions inside a solenoid.

There is a more satisfying response. It is to note that the definition of nonseparability does not actually involve a comparison of the features of individual spacetime points in spacetimes with different geometries. All that is required is a comparison of the features of *the region outside the solenoid* in such different spacetimes. That region is readily identified in distinct spacetimes, whether or not it is possible also to identify a given point in that region in distinct spacetimes. The Aharonov-Bohm effect is nonseparable in a general relativistic spacetime because the holonomy properties of rings in the region outside the solenoid depend on the current flowing through the solenoid. This is true whether or not one can identify the same points, or the same rings, in distinct spacetimes in which different currents are flowing.

8. Quantized Gauge Theories and Their Loop Representations

The non-Abelian Yang-Mills gauge theories used so successfully in the standard model of elementary particles are quantum, not classical, field theories. The fundamental nature of these theories means that it is of great interest to enquire about the status of holism and nonseparability according to quantized gauge theories. Holism and nonseparability would be features of the world, not of our theories of it. But our best method of determining whether the physical world exhibits these features is to analyze the implications of our fundamental physical theories.

Since contemporary non-Abelian gauge theories are quantum field theories, this analysis involves coming to grips with major unresolved questions concerning the interpretation of a quantum field theory. And since holism concerns the relation between an

object and its parts, ontological questions will be prominent among these.[7] Some may therefore consider it premature to try to say whether the empirical success of contemporary gauge theories of the standard model gives us reason to believe that the world they describe is holistic or nonseparable. But even though it is certainly too early to render any definitive conclusion, this should not discourage one from attempting a preliminary sketch of relevant features of quantized gauge theories. The sketch provided in this and the following section already reveals features that suggest that quantized as well as classical gauge theories may be best understood as holistic by virtue of their nonseparability. These arise from what are called loop representations of quantized gauge theories.

While there is today no correspondingly successful quantum theory of gravity there is one approach to such a theory in which loop representations also play an important role. These representations have been employed to great effect in the theoretically progressive, but as yet empirically unsupported, loop quantum gravity program (see Gambini and Pullin(1996), Rovelli(2004) and references therein). It remains an open question how well loop quantum gravity describes our world. But it is not necessary to wait for an experimental answer before enquiring whether a world truly described by a loop quantum gravity theory would display holism or nonseparability of one sort or another. Unfortunately the technical preliminaries required to pursue this enquiry place it outside the scope of the present paper.

Quantizing a gauge theory involves complications that do not arise for other theories. There are alternative ways of handling these complications that result in different quantization techniques. One such technique proves of particular relevance to the present inquiry. It gives rise to loop representations of a quantized gauge theory. Unlike most standard techniques, this involves no arbitrary choice of gauge. It arises as a gauge-invariant

---

[7]The volume Kuhlmann *et.al.*(2002) provides a useful survey of such questions.

generalization of an approach to the quantization of a gauge theory due to Dirac. Since Dirac's approach may not be widely known I will sketch it first, and then explain how it generalizes to loop quantization.

Dirac(1964) pioneered a method for quantizing gauge theories which may be applied to a classical dynamical theory whose Hamiltonian formulation involves constraints–identities satisfied by the canonical variables. As section 3 explained, this includes classical Yang-Mills field theories, and also general relativity (though the application to the latter proves more problematic: for both applications, see Gambini and Pullin(1996)).

One starts with an algebra satisfied by basic variables of the theory–typically the Poisson algebra formed by the Poisson brackets of the classical canonical variables. One replaces these variables by operators, and the algebra by commutation relations of the operators. This abstract algebra is now represented by operators acting on a space of functionals $\Psi[\boldsymbol{\pi}(x)]$ to be thought of as candidate wave-functionals for the quantized theory. The space is restricted to those functionals that are annihilated by the constraint-functions of the operators, where a constraint in the classical theory sets such a function identically equal to zero. An inner product is defined on this space, making it a Hilbert space. Finally, the state of a system $\Psi[\boldsymbol{\pi}(x)]$ in the quantized theory is taken to satisfy a Schrödinger equation $\Psi = i\partial\Psi/\partial t$. (Units have been chosen throughout this section so that Planck's constant $\hbar$ equals 1.)

While it is the application of Dirac's method to non-Abelian Yang-Mills theories that is of greatest interest in the present context, it would be inappropriate here to go into the technical complications that arise for a non-Abelian theory. So consider instead the application to a classical electromagnetic field obeying Maxwell's equations in otherwise empty space.

29

The canonical configuration variables are the four components of A_μ: the electric field $\vec{E}$ gives three of the momentum variables as $-\vec{E}$, while the vanishing of the fourth momentum variable π^0 represents a constraint. A second constraint follows from Hamilton's equation of motion for π^0: it is just the Gauss law $\vec{\nabla}.\vec{E} = 0$. The method requires one to impose commutation relations on the canonical variables as follows:

$$
\begin{aligned}
[\hat{A}_i(\vec{x}, t), \hat{E}^j(\vec{x}', t)] &= -i\delta_{ij}^3(\vec{x} - \vec{x}') \\
\left[\hat{A}_0(\vec{x}, t), \hat{\pi}^0(\vec{x}', t)\right] &= \delta^3(\vec{x} - \vec{x}') \\
[\hat{A}_i(\vec{x}, t), \hat{A}_j(\vec{x}', t)] &= [\hat{E}^i(\vec{x}, t), \hat{E}^j(\vec{x}', t)] = 0
\end{aligned}
\tag{7}
$$

The first of these equations does not lead to an inconsistency, since the Gauss law is interpreted not as an operator equation, but as a restriction on the wave-functionals in the Hilbert space on which these operators act. The second equation looks strange in light of the classical constraint $\pi^0 = 0$, but again the corresponding operator equation is reinterpreted after quantization not as an identity, but as imposing a restriction on physical wave-functionals. These commutation relations may be realized by representing the action of the operators $\hat{A}_i, \hat{E}^j$ on the space of wave-functionals $\Psi[A]$ as follows

$$
\begin{aligned}
\hat{E}^j \Psi[A] &= -i\frac{\delta}{\delta A_j}\Psi[A] \\
\hat{A}_i \Psi[A] &= A_i \Psi[A]
\end{aligned}
\tag{8}
$$

The Gauss law is now represented by the equation

$$
\partial_j \hat{E}^j \Psi[A] \equiv -i\partial_j \frac{\delta}{\delta A_j}\Psi[A] = 0
\tag{9}
$$

in which the symbol $\frac{\delta}{\delta A_j}$ denotes a functional derivative. The Hamiltonian may be written in terms of the operators $\hat{A}_i, \hat{E}^j$, giving the Schrödinger equation $\hat{H}(\hat{A}_i, \hat{E}^j)\Psi[A] = i\partial\Psi[A]/\partial t$. This is called the connection representation because $\Psi[A]$ is a functional of the connection.

Wave-functionals and operators in the connection representation are gauge-dependent. One can arrive at a gauge-invariant quantization scheme by starting instead with gauge-invariant classical variables and applying Dirac's quantization scheme to the non-canonical algebra formed by taking their Poisson brackets. The result is the so-called *loop representation*, in which wave-function(al)s become functions of space-like loops. As Ashtekar and Rovelli(1992) stress, the free, quantized Maxwell field actually has many loop representations. I focus on one described by Gambini and Pullin(1996). They take as fundamental variables the electric field $E^i(x)$ and the Wilson loop $W_A(\gamma)$ where $\overrightarrow{A}$ is a magnetic vector potential, and γ is a loop in one of the foliation of 3D space-like hypersurfaces on which $\overrightarrow{E}$ and $\overrightarrow{A}$ are defined. In terms of these one defines new variables

$$
\begin{aligned}
T(\gamma) &= W_A(\gamma) \\
T^i(\gamma_x^x) &= E^i(x)W_A(\gamma)
\end{aligned}
\tag{10}
$$

The Poisson brackets of these new variables are

$$
\begin{aligned}
\{T(\gamma), T(\eta)\} &= 0 \\
\{T^j(\gamma_x^x), T(\eta)\} &= -iX^{jx}(\eta)T(\eta \circ \gamma) \\
\{T^j(\gamma_x^x), T^k(\eta_y^y)\} &= -iX^{jx}(\eta)T^k(\gamma \circ \eta)_y^y + iX^{ky}(\gamma)T^j(\eta \circ \gamma)_x^x
\end{aligned}
\tag{11}
$$

where γ_x^x is a loop beginning and ending at x and passing through the base point o, and $X^{ix}(\eta)$ is a tangent to η at x. The corresponding commutation relations are realized on a space of loop-dependent functions $\Psi(\gamma)$ by operators $\hat{T}(\gamma)$, $\hat{T}^i(\gamma_x^x)$ that satisfy

$$
\begin{aligned}
\hat{T}(\eta)\Psi(\gamma) &= \Psi(\eta^{-1} \circ \gamma) \\
\hat{T}^i(\eta_x^x)\Psi(\gamma) &= X^{ix}(\gamma)\Psi(\eta^{-1} \circ \gamma)
\end{aligned}
\tag{12}
$$

One can express operators for the electric field and the magnetic part of the field tensor in terms of the $T's$, with the result that

$$
\begin{aligned}
\hat{F}_{jk}(x)\Psi(\gamma) &= -i\Delta_{jk}(x)\Psi(\gamma) \\
\hat{E}^i(x)\Psi(\gamma) &= X^{ix}(\gamma)\Psi(\gamma)
\end{aligned}
\tag{13}
$$

Here $\Delta_{ij}(x)$ is a so-called loop derivative that specifies how a function of a loop varies under small extensions to that loop at point x. The loop representation is related to the connection representation by the so-called *loop transform*

$$
\Psi(\gamma) = \int DA . W_A^*(\gamma)\Psi[A]
\tag{14}
$$

where $W_A(\gamma) = \exp\left(i\int d^3z X^{jz}(\gamma)A_j(z)\right)$ is the Wilson loop for γ. This transform is analogous to the unitary transformation involved in a change of basis in the Hilbert space representing the state of a system of non-relativistic particles. The analogy is brought out by writing it in the form

$$
\Psi(\gamma) = <\gamma \mid \Psi> = \int DA <\gamma \mid A><A \mid \Psi>
\tag{15}
$$

which may be compared to

$$
\psi(p) = <p \mid \psi> = \int dx <p \mid x><x \mid \psi> = \int dx \exp(-ikx).\psi(x)
\tag{16}
$$

(The functional integral over the infinite-dimensional space of connections is a well-defined mathematical object here.)

One apparent disanalogy is that p is a dynamical variable, while γ is not. But, as Ashtekar and Rovelli(1992) point out, these are best thought of as *labelling* their respective wave-functions. A state of a hydrogen atom may be written as $\psi(n, l, m)$, where the numbers n, l, m indicate an eigenstate of energy, angular momentum, and z-component of angular momentum. In each case, the physical significance of the state is brought out only by interpreting the action on it of various operators.

32

9. Holism and Nonseparability in Quantized Yang-Mills Gauge Theories

It is striking that the search for a gauge-invariant quantization procedure for Yang-Mills gauge theories including electromagnetism leads to a representation in which wave-functions are defined not over spacetime points but over loops. This suggests, but does not establish, that such theories involve nonseparability and associated holism of much the same sort as their classical counterparts. Whether this conclusion is indeed justified will depend on how one interprets these quantum field theories

There are loop representations of the free, quantized Maxwell field that admit an interpretation in terms of photons, corresponding to a Fock space of states. Consider such an interpretation. Naively, a wave-function may appear to constitute a (complex) magnitude that takes values on space-like loops. If that were so, then we would have nonseparability analogous to the classical case, except the rings with non-supervenient properties would be spatial rather than more generally spatiotemporal. But this interpretation of the wave-function is surely too naive. Different loop representations yield different wave-functions, in addition to that of the connection representation: and in none of these is the wave-function an observable magnitude.

Moreover, Ashtekhar and Isham(1992) show that there are unitarily inequivalent loop representations of the free, quantized Maxwell field that *do not* admit a Fock space interpretation. They even argue that it is these representations that generalize more naturally to non-Abelian gauge theories. We are faced here with a proliferation of representations not only more extreme than that familiar from ordinary nonrelativistic quantum mechanics, but even more extreme than what occurs in the quantum field theory of a free scalar field. For the unitarily inequivalent representations described by Ashtekhar and Isham are based on different, though equally legitimate, operator algebras. In this context, any conclusion about

holism or nonseparability must be conditional on an assumption about the relative status of all these different representations.

One might, for example, attempt a Bohmian interpretation that takes as basic observables the holonomy magnitudes. These will be the analogs of position in the Bohmian interpretation of ordinary non-relativistic quantum mechanics. This interpretation privileges a particular loop representation and its corresponding wave-function, but acknowledges the unobservable character of the wave-function itself. Electric fields then have the same derived status as the momentum observable in ordinary Bohmian theory. Just like momentum in non-relativistic quantum mechanics, eigenstates of the electric field operator form an overcomplete, non-normalized basis of states in this loop representation. We then have nonseparability, since the basic holonomy magnitudes pertain to space-like rings, not points. This is associated with holism just as it was for the classical electromagnetic field when interacting with quantum particles.

Or one might defend a Copenhagen interpretation that treats all unitarily equivalent representations of a given operator algebra on a par. It will regard the wave-function in any of these as merely a repository of information concerning probabilistic dispositions to yield measurement results on gauge-invariant observables including holonomy properties as well as electric and magnetic fields. All these observables pertain to space-like rings, not points. Taking such dispositions as basic again yields a nonseparable ontology. The existence of inequivalent loop representations complicates matters. Perhaps complementarity can be extended to these also along the lines of Clifton and Halvorson(2001). But it is unclear what measurement procedures correspond to the different representations in this case.

10. Conclusion

We have reason to believe that, in a world of quantum particles, electromagnetism, described by a classical Yang-Mills gauge theory, is nonseparable. On one interpretation of the part-whole relation it therefore also exhibits a kind of holism–namely, physical property holism. But there is a different interpretation of the part-whole relation on which interactions described by a classical Yang-Mills gauge theory do not exhibit such holism.

General relativity may be viewed as a gauge theory, and even formulated in terms of a connection on a principal fiber bundle. But a general relativistic world does not exhibit nonseparability. Nor does it exhibit physical property holism. General relativity's admission of nonstandard spacetime topologies does provide vivid illustrations of relational holism. But more pedestrian examples of this occur in any world described by a spacetime theory that counts spatiotemporal relations as intrinsic (but external) physical relations.

Whether an interaction described by a quantized Yang-Mills gauge theory exhibits the kind of holism implied by nonseparability hinges on unresolved issues concerning the interpretation of quantum field theories. The existence of loop representations of quantized Yang-Mills gauge theories provides grounds for believing that these involve the same kind of holism associated with nonseparability as their classical counterparts, at least on some prominent interpretations. Exploring the question of whether we live in such a holistic world provides yet another incentive to those seeking an improved understanding of quantum field theory.

References

Anandan, J. (1993), "Remarks concerning the geometries of gravity and gauge fields", in B. Hu, M. Ryan and C. Vishveshwara (eds.), *Directions in General Relativity* (Cambridge: University Press).

Ashtekhar, A. and Isham, C. (1992), "Inequivalent observable algebras. Another ambiguity in field quantisation", *Phys. Lett. B 274*, 393-398.

Ashtekhar, A. and Rovelli, C. (1992), "A loop representation for the quantum Maxwell field", *Classical and Quantum Gravity 9*, 1121-1150.

Belot, G. (2003), "Symmetry and Gauge Freedom", *Studies in History and Philosophy of Modern Physics 34*: 189-225.

Bell, J.S. (1987), *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: University Press).

Bohm, D. and Hiley, B. (1993), *TheUndivided Universe* (London: Routledge).

Clifton, R. and Halvorson, H. (2001), "Are Rindler Quanta Real? Inequivalent Particle Concepts in Quantum Field Theory", *British Journal for the Philosophy of Science 52*, 417-470.

Cushing, J. and McMullin, E. (eds.) (1989), *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem* (South Bend, IN.:University of Notre Dame Press).

Dirac, P. (1964), *Lectures on Quantum Mechanics* (New York: Yeshiva University Press)

Earman, J. (2002), "Gauge Matters", in *PSA 2000: Proceedings of the Philosophy of Science Association Biennial Conference, Philosophy of Science 69*: S209-S220.

Esfeld, M. (2001), *Holism in Philosophy of Mind and Philosophy of Physics* (Dordrecht: Reidel).

Gambini, R. and Pullin, J. (1996), *Loops, Knots, Gauge Theories and Quantum Gravity* (Cambridge: University Press).

Healey, R. (1991), "Holism and Nonseparability", *Journal of Philosophy, LXXXVIII*, 393-421.

-------------(2001), "On the Reality of Gauge Potentials", *Philosophy of Science 68*, 432-455.

-------------(2004), "Holism and Nonseparability in Physics", entry in the *Stanford Electronic Encyclopedia of Philosophy*, http://plato.Stanford.edu/entries/physics-holism

Henneaux, M. and Teitelboim, C. (1992), *Quantization of Gauge Systems* (Princeton: University Press).

Jones, M. (forthcoming), "Holism and Nonsupervenience in Quantum Mechanics".

Kuhlmann, M., H. Lyre and A. Wayne (eds.) (2002), *Ontological Aspects of Quantum Field Theory* (Singapore: World Scientific).

Marder, L. (1959), "Flat Space-times with Gravitational Fields", *Proceedings of the Royal Society A 252*: 45-50.

Redhead, M. (2002), "The Interpretation of Gauge Symmetry", in Kuhlmann *et*. *al*. (2002).

Rovelli, C. (2004), *Quantum Gravity* (Cambridge: Cambridge University Press).

Teller, P. (1987), "Relational Holism and Quantum Mechanics", *British Journal for the Philosophy of Science 37*: 71-81.

Trautman, A. (1980), "Fiber Bundles, Gauge Fields and Gravitation", in A. Held (ed.), *General Relativity and Gravitation* (New York: Plenum Press).

Wu, T.T. and C.N. Yang (1975), "Concept of Nonintegrable Phase Factors and Global Formulation of Gauge Theories", *Physical Review D12*: 3845-3857.