Richard Healey, ***Gauging What's Real: The Conceptual Foundations of Contemporary Gauge Theories***. Oxford: Oxford University Press (2007), 240 pp., $99.00 (cloth).

This well-researched and authoritative book provides much material for philosophers of physics to discuss and debate. It is significant in particular for the attention it gives to non-Abelian Yang-Mills gauge theories (the theories that appear in the Standard Model), and the structural differences between them and the gauge theories that have typically appeared in the philosophical literature (i.e., Abelian electromagnetism and general relativity [GR]). Healey's specific goal is to defend an interpretation of Yang-Mills theories under which

(*a*) they refer to nonlocal properties encoded in holonomies; and
(*b*) the local gauge symmetries that characterize them are purely formal and have no direct empirical consequences.

A holonomy is a quantity associated with a closed curve. Under one representation, it is determined by the line integral along the curve of a gauge potential field. After a presentation of classical Yang-Mills theories in Chapter 1, Healey devotes Chapter 2 to the Aharonov-Bohm (AB) effect in semiclassical electromagnetism. This is a major motivation for Healey's holonomy interpretation and much detail is given to its exposition. Briefly, the AB effect demonstrates that, in nonsimply connected regions of space-time, two distinct electromagnetic potential gauge fields can give rise to the same electromagnetic gauge field. This suggests three interpretive options, which are extended in Chapter 4 to encompass classical gauge theories in general:

(1) *No gauge potential properties*. Classical gauge fields encode local gauge-invariant properties. Classical gauge potential fields have no physical content.
(2) *Localized gauge potential properties*. Classical gauge potential fields encode local properties.
(3) *Nonlocalized gauge potential properties*. Holonomies encode nonlocal properties.

Against Option 1, Healey notes that in non-Abelian Yang-Mills theories, two distinct gauge potential fields may give rise to the same gauge

Philosophy of Science, 75 (October 2008) pp. 479–485. 0031-8248/2008/7504-0008$10.00

field, even in simply connected regions of space-time. Healey cites an example due to Wu and Yang (1975) in which the gauge potential fields differ by a source term, and rightfully indicates that this provides a "powerful reason to accept that a classical non-Abelian gauge potential indeed represents qualitative intrinsic gauge properties over and above those represented by its associated gauge field" (Healey 2007, 84). He suggests this is due to the fact that in a non-Abelian gauge field, "a vector field representing the non-Abelian gauge potential appears independently and ineliminably" in the inhomogeneous Yang-Mills equations, and thus "alterations in the potential that preserve the associated field may consequently be balanced by corresponding alterations in the sources" (2007, 84). It's a bit unclear how this works, since the gauge potential field also appears in the inhomogeneous Abelian Yang-Mills equations (in the definition of the gauge field). Moreover, Deser and Wilczek (1976, 392) prove a necessary condition for the underdetermination of a non-Abelian gauge field by a gauge potential, and it depends primarily on the structure of the gauge field, and not on the presence of source terms. This suggests that source terms have nothing essential to do with this underdetermination. On the other hand, Healey might argue that the cases of physical interest are those in which the potentials differ by source terms. Note finally that similar results hold for general relativity, hence Option 1 appears inappropriate in this context, too.

Healey's analysis of Option 2 is a bit more nuanced. Ultimately he is willing to entertain Option 2 for GR, but not for Yang-Mills theories. This is due to a specific difference in one way of formulating both types of theories. In particular, Healey notes that, in fiber bundle formulations of GR, gauge transformations in the bundle space are soldered to the base space by a soldering form. Such a structure does not appear in fiber bundle formulations of Yang-Mills theories. Healey draws two interpretive conclusions from this formal difference:

The first is that GR is "separable" in the quantum domain, whereas Yang-Mills theories are not (78–81). The argument involves comparing the generalized AB effect in the latter with a gravitational analog. Briefly, glossing over Healey's intricate notions of separability and locality, Yang-Mills AB phase differences are measurable only along closed paths, whereas gravitational AB phase differences are measurable along open paths, and this is due explicitly to the presence in the latter of a soldering form (80). This suggests that a local interpretation of GR along the lines of Option 2 is possible, whereas it may be problematic for Yang-Mills theories.

The second conclusion Healey draws from his consideration of soldering forms is that, while Option 2 is viable for GR, it is not for Yang-Mills theories. He suggests that, under Option 2, a Yang-Mills theory must claim that there are localized gauge potential properties, but it cannot say what they are (96). In particular, under Option 2, a Yang-Mills theory is faced with the problem of identifying the real gauge potential from non-physical imposters related to it by gauge transformations. This suggests that a gauge-invariant interpretation is to be preferred for Yang-Mills theories. However, Healey does not think this is a problem for Option 2 in the context of GR, due to the presence, in the fiber bundle formulation of GR, of a soldering form. Before reconstructing his reasoning, note that arguments against Option 2 (in both the Yang-Mills and GR contexts) typically employ the threat of indeterminism: since a potential gauge field is only determined up to a gauge transformation by the Yang-Mills equations, an interpretation that awards it ontological status risks being indeterministic (and similarly in the GR context with the appropriate substitutions). Healey's critique of Option 2 is notable for its *lack* of reference to indeterminism, and this cries out for explanation.

For Healey, the presence of a soldering form in fiber bundle formulations of GR, and its absence in fiber bundle formulations of Yang-Mills theories, suggests that "there is no analog to Leibniz equivalence in the case of other [i.e., nongravitational] interactions" (98). Recall that 'Leibniz equivalence' in the context of GR holds between diffeomorphically related solutions of the Einstein equations that are identified as representing the same possible world. Assumedly, its analog in Yang-Mills theories would hold between solutions to the Yang-Mills equations related by a gauge transformation and deemed to represent the same possible world. Healey maintains that the presence of a soldering form allows one to claim that diffeomorphically related solutions in GR represent the same possible world, given that diffeomorphisms affect all parts of a fiber bundle model of GR (in particular the base space as well as the bundle space), whereas gauge-related Yang-Mills solutions do not represent the same possible world, given that gauge transformations in this context do not affect all parts of the model (they only affect the bundle space). Now such a claim is not forced upon us simply by the mathematical formalism of fiber bundles. In particular, it is unclear how the presence or lack of a soldering form in a formulation of a theory (i.e., a formal property of a theory) impinges on so general an interpretive claim as Leibniz equivalence. What seems to be needed is an interpretation of the soldering form itself that then justifies the subsequent attitude towards Leibniz equivalence. Note that the existence of a soldering form in fiber bundle formulations of GR is a way to encode the empirical fact that the gravitational force is universal; i.e., that all material objects experience it in the same way regardless of their gravitational charge (i.e., gravitational mass). Mathematically, the soldering form is the explicit mechanism by which gravity is geometrized in fiber bundle formulations of GR. The other known interactions (EM, weak, strong) can in principle also be geometrized; however, since their coupling to matter is dependent on their associated charges, this would require specifying one family of 'force-free' geodesics for every physically possible value of the associated charge (more precisely, the ratio of charge to inertial mass). One might wonder how the fact that gravity can be geometrized in a simple manner, while the other interactions cannot, has anything essential to do with Leibniz equivalence, and relatedly, indeterminism. Note that one specific consequence of the lack of a soldering form in Yang-Mills theories may be that in the latter, manifold substantivalism may not entail indeterminism. But this specific claim is a bit different from the general moral that 'Leibniz equivalence' doesn't apply to Yang-Mills theories; nor does it immunize all versions of Option 2 from the threat of indeterminism.

Under Healey's preferred interpretation Option 3, holonomies encode real properties that are predicated of closed curves. Abelian Yang-Mills holonomies are gauge-invariant, thus they escape Healey's critique of Option 2. However, in non-Abelian Yang-Mills theories, holonomies are not gauge-invariant; rather, they are invariant only under 'pointed' gauge transformations. These latter assign the identity element of the gauge group to an arbitrary point of the base space. Given that holonomies are uniquely determined by the gauge potential field, Healey's critique of Option 1 discussed above does not apply. On the other hand, one might wonder how non-Abelian holonomies escape the critique of Option 2. However, as Healey notes, the move to pointed gauge transformations is conceptually harmless. Healey demonstrates that a semiclassical non-Abelian Yang-Mills system is invariant under a pointed gauge transformation of both the (quantized) matter field and an associated (classical) holonomy; and this preserves the gauge-invariant nature of Option 3 (109). Another way of addressing this concern might be by extending the notion of a gauge group to a gauge groupoid. This would allow the holonomy interpretation to jettison all reference to base space points.

After a review of quantized Yang-Mills theories in Chapter 5, Healey turns in Chapter 6 to a discussion of the empirical import of gauge symmetry. He addresses a number of concerns that local gauge symmetries may have empirical significance, arguing that in each case no such conclusion is warranted. One of these cases involves an interesting discussion of the θ-vacuum that arises in certain solutions to the Yang-Mills equations. The space of this family of solutions is not simply connected, and this results in a degeneracy of the vacuum, labeled by the parameter θ. Healey's worry is that awarding physical content to degenerate θ-vacua, as is the custom among physicists, would entail awarding empirical import to local gauge symmetries, given that (a subset of) the latter transform θ-vacua into each other. Ultimately he argues against such a physical interpretation, and the details are informative and intricate. However, one might reach the same conclusion in a much quicker fashion by simply questioning the physical relevance of θ-vacuum solutions to begin with. To see how this might play out, note first that *not* all Yang-Mills theories suffer from θ-vacuum degeneracy. θ-vacua only arise in the context of finite-action solutions to the Yang-Mills equations in Euclidean 4-space, E^4 (which are generally referred to as instanton solutions). A theorem due to Uhlenbech (see, e.g., Ward and Wells 1990, 272) demonstrates that every such solution arises from a solution on the 4-sphere, the topologically nontrivial compactification of E^4. It is this *restricted* family of instanton Yang-Mills gauge fields that can be characterized by topological Chern numbers and degenerate θ-vacua. The question then is whether such solutions are physically relevant, given that the Yang-Mills gauge fields that describe the fundamental interactions are defined, not on E^4 but on Minkowski space-time.

To address this concern, note that in 4-dimensions, the source-free Yang-Mills equations $D{*}\mathbf{F} = 0$ are automatically satisfied by fields $\mathbf{F}$ for which ${*}\mathbf{F} = \lambda\mathbf{F}$, for some λ (where $*$ is the Hodge dual operator). As Nash and Sen (1983, 258) point out, in Minkowski space-time, this condition reduces to ${*}\mathbf{F} = \pm i\mathbf{F}$, and this entails that the associated gauge group must be *noncompact* (for instance, SL(n, C) or GL(n, C)). This rules out all the *compact* gauge groups associated with the fundamental interactions. In contrast, in E^4, the condition reduces to ${*}\mathbf{F} = \pm\mathbf{F}$, and this allows physically relevant compact gauge groups (in particular, SU(n)). The upshot is that in Minkowski space-time, there are no *physically relevant* Yang-Mills gauge fields that satisfy the simplifying condition (such fields are referred to as self-dual, or anti-self-dual, depending on the sign). On the other hand, in *physically nonrelevant* E^4, there *are* physically relevant (i.e., compact) (anti-)self-dual Yang-Mills gauge fields, namely, instantons with their θ-vacua. Now the physics literature suggests that we should take such Euclidean instantons seriously: they 'solve' one problem with QCD (the U(1) problem), while generating another (the strong CP problem; Kaku 1993, 559–565). One might attempt to justify taking Euclidean instantons seriously by considering E^4 as merely a formal device that simplifies calculations: We do our calculations in E^4, and then analytically continue the results back into Minkowski space-time (sentiments of this sort are expressed by Healey in 178, note 8). But this begs the question of what such analytically continued results represent. Can definite mathematical objects be identified in Minkowski space-time that correspond to Euclidean instantons, and if so, can they be realistically interpreted? One way to address this might be to seek out an alternative formalism in which the structure common to both Euclidean instantons and their Minkowskian correlates is made explicit. An example of this is the twistor formalism:

there are twistor constructions for both (anti-)self-dual Yang-Mills gauge fields in E^4 and in Minkowski space-time (for the SU(n) case in E^4, see Ward and Wells 1990, 390; for the cases of GL(n, C) and SL(n, C) in Minkowski space-time, see Ward and Wells 1990, 374, 386). This suggests that one way of taking Euclidean instantons seriously is to become a realist with respect to twistors (although what *that* entails is best left to another essay).

In Chapter 8 Healey considers the extent to which the holonomy interpretation of classical gauge theories can be applied to quantized Yang-Mills theories. An initial concern involves the extent to which the holonomy interpretation admits a concept of particle. Healey follows the standard line in the literature on this subject by maintaining that "a minimal requirement for any kind of particle ontology for a quantum field is the existence of a Fock representation of its ETCRs [equal time canonical commutation relations]" (206). Healey's concern now focuses on results due to Ashtekar and Isham (1992), who indicate that some Weyl algebras generated by loop observables in a non-Abelian Yang-Mills theory do not admit Fock space representations. Thus if one is concerned with providing non-Abelian Yang-Mills theories with particle interpretations, and one accepts the received view on particles, one might be hesitant in adopting the holonomy interpretation. Healey ultimately argues that such hesitance can be assuaged, but the initial motivation for doing so is a bit unclear. The received view assumes that particles possess certain properties that can only be represented by mathematical objects that occur in Fock space representations. One of these properties is a notion of localizability which gets encoded in Fock space local number operators. The question then is, is this notion of localizability compatible with Healey's nonlocal holonomy interpretation? If not, then a holonomy realist need not accommodate the received view.

As should now be evident, there is enough material in Healey's book to engage philosophers of physics for some time to come. It is a thought-provoking work and a significant contribution to the literature on gauge theories and philosophy of quantum field theory.

JONATHAN BAIN, POLYTECHNIC INSTITUTE OF NEW YORK UNIVERSITY

REFERENCES

Ashtekar, A., and C. Isham (1992), "Inequivalent Observable Algebras: Another Ambiguity in Field Quantisation", *Physics Letters B* 274: 393–398.

Deser, S., and F. Wilczek (1976), "Non-uniqueness of Gauge-Field Potentials", *Physics Letters B* 65: 391–393.

Kaku, M. (1993), *Quantum Field Theory.* Oxford: Oxford University Press.

Nash, C., and S. Sen (1983), *Topology and Geometry for Physicists.* London: Academic Press.

Ward, R., and R. Wells (1990), *Twistor Geometry and Field Theory.* Cambridge: Cambridge University Press.

Wu, T., and C. Yang (1975), "Some Remarks about Unquantized Non-Abelian Gauge Fields", *Physical Review D* 12: 3843–3844.