Gauge Theory and Chemical Structure

JAMES MATTINGLY

Philosophy Department, Georgetown University, Washington, DC 20057, USA

ABSTRACT: The possibility of chemical structure in the context of quantized matter is examined by way of Richard Bader's *Atoms in Molecules*. I critically examine his notion of "electronic charge density"—showing that he cannot really mean "density of charge"—and I argue that the appropriate concept is expectation value of charge. This still allows him to define chemical structure, but it makes problematic his appeals to the explanatory power of structure. This is because, as Rosenfeld and Bohr showed, the expectation value of charge cannot be taken as the electronic field experienced by other charges. I suggest that we can recover the efficacy of structure by thinking of chemistry as a gauge theory. Current consensus in the study of gauge theories indicates that gauge potentials represent a new type of property; while no member of the family of functions comprising the gauge potential is real, the potential itself is causally potent. I illustrate this in the case of electrodynamics, where the vector potential can causally influence charges in the absence of electric or magnetic fields. I show how chemical structure can be considered to be a gauge field. Following Bader, I take it to be a family of geometric configurations, no one of which is possessed by a given molecule. I claim that current research in gauge theory licenses the attribution of causal potency to this notion of structure, despite its lack of reality. I thus begin the process of freeing the explanatory resources of gauge theory from physics alone.

KEYWORDS: gauge theory; chemical structure; Bader; scientific explanation

INTRODUCTION

Back in the good old days, we are told, electrodynamics was philosophically much simpler to understand than it is now. There were electric fields and magnetic fields, and when these were specified, so were all possible electrodynamical observables. Thus the picture of what was real for electrodynamics was just that: the electric and magnetic fields. Of course things hadn't always been so simple. Indeed in the good even older days, the whole idea of the field concept was itself new, and untried and, frankly, suspicious. Yet with the reception and widespread acceptance of Maxwell's formalization of Faraday's researches (and of course his own and others'), the field concept came more and more to dominate physical theorizing, until the whole practice of thinking about field action was enshrined by Einstein in the principle of local action. Now we are told that, rather than good old property-valued

fields, the world is populated by operator-valued distributions—and this is true for all quantum fields, including electrodynamics. If this were not enough of a strain on our metaphysical imaginative faculty, it turns out that even good old electricity and magnetism have suffered a reversal of fortune. To wit, there's an extra field there that has properties completely unlike those of earlier fields. This is all to say that our researches in quantum physics have led to a revision of how we ought to understand what was really going on in electrodynamics even at the classical level.[1]

A similar problem has arisen for students of chemical structure. It seemed, around the end of the 19th century, that the dream of Newton, Hooke, Boyle, and others that we resolve the action of chemicals into the behavior of very small bodies with hooks and other mechanical contrivances had finally succeeded, in a manner of speaking. There were indications that the molecular theory could explain the behavior of various chemicals by the interactions among them caused by their electric fields, and that what determined the details of these fields was the shape of the individual molecules. This shape was, moreover, taken to be a constant for a given species of molecule. So, roughly speaking and with hedging all around, the strong solvency properties of H_2O (as well as the fact that it expands at its freezing point) are to be understood as arising from the peculiar "V" shape of the water molecule. This shape gives water a strong dipole moment, and this in turn allows water to "pry" other species of molecule apart from each other. But with the advent of quantum mechanics, and thus quantum chemistry, the very notion of chemical structure has been again called into question. Since atoms have no particular locations in space, it is argued, then molecules can have no particular geometry. Somewhat more carefully, one would point out that, insofar as the relative momenta of the atoms in some molecule are well-defined, the relative positions of these atoms are not. But for the naive story I just told about the efficacy of molecular structure in explaining molecular behavior, we need the shape to be stable and so we need the relative momenta of the various atoms to vanish. Without getting caught up in the modern debate over chemical structure that has been going on more or less since the beginning of quantum mechanics, I do want to suggest that the principles of gauge theory may be able to make plausible the idea that chemical structure really is ill-defined, and yet at the same time it really is causally effective in chemical interactions.

A little context may be in order, and this context will, I hope, make sense of the connection between gauge theory and chemical theories. My plan for producing this context and illuminating the connection between chemical structure and gauge theory is as follows: I begin with a brief introduction to the notion of gauge theory. The main thrust of this section is the idea that, in current views of gauge theory, the properties of gauge fields are causally active despite our inability to say what kind of thing these properties are. The second part of the paper concerns the issue of chemical structure. Indeed it is quite narrowly focused on one particular proposal to recover the concept of chemical structure in the context of quantum mechanics—Richard Bader's *Atoms in Molecules*.[2] The book is an ambitious attempt to show that chemical structure is meaningful for quantum-mechanical molecules despite the fact that chemical shape is not. Moreover, Bader attempts to show how his notion of chemical structure can recover the explanatory power that chemical shape was supposed to have. A discussion of Bader's program occupies the second and third sections. Finally in the last section I propose that Bader's proposal, though philosophically unsatisfying at first blush, may be made more satisfying (or at least

brought into line with some popular views of proper explanatory structure) by recasting it in analogy with gauge theory.

GAUGE THEORY AND NOVEL PROPERTIES

To illustrate the notion of gauge theory, I'll appeal just to the example of electrodynamics. This will keep the discussion manageable, but will still allow its interesting and novel explanatory strategy to emerge. Naturally this discussion is too compressed for a full analysis of gauge theory. In particular, it is not uncontroversial to suggest that the appeal to novel properties is generic of gauge theories and not just those theories with Aharonov–Bohm type effects. I think it is correct, but I won't argue for it here.[3]

The vector potential in electrodynamics, which is the field that forced a change in our understanding of the ontology of electrodynamics, is not really a new field. What is new is the conviction, widely held by physicists and philosophers of physics, that we must take it seriously as a physically significant quantity. The potted history goes like this: from early in the development of electrodynamics as a mathematical field theory, it was known that the magnetic field could be used to define another field, the vector potential, via $\nabla \times \mathbf{A} = \mathbf{B}$, where $\mathbf{B}$ is the magnetic field. But this equation does not uniquely specify $\mathbf{A}$, for one can always add the gradient of an arbitrary scalar function to $\mathbf{A}$ to obtain a new vector potential $\mathbf{A}'$ that satisfies the defining condition. (A complementary point obtains with respect to the scalar potential ϕ. The electric field can be given as $\mathbf{E} = \nabla\phi$, but ϕ is not uniquely defined thereby. One can always add some constant to a given scalar potential to produce a different potential that is, nonetheless, physically equivalent.) This is what is called a gauge transformation, and $\mathbf{A}$ is called a gauge field or gauge potential. There is a significant number of different ways of expressing the mathematical concept of gauge, and some of these are explained in the literature I cite below. But for the purpose I have in mind, this description of $\mathbf{A}$ as a gauge potential should do nicely. The crucial feature I want to highlight is that no particular choice of $\mathbf{A}$ is *the* vector potential. The vector potential is in some sense all $\mathbf{A}$s and in some sense none. The equations of motion do not single out a unique $\mathbf{A}$, so we have no grounds for saying that one is privileged over the others. We are simply stuck with an ill-defined electromagnetic property.[4] Until the second half of the 20th century the ill-definedness of $\mathbf{A}$ could be safely ignored.

It was thought because of the gauge freedom in $\mathbf{A}$ that the vector potential had no metaphysical (or even physical for that matter) significance. Nothing distinguishes the "right" $\mathbf{A}$ from all the others, and $\mathbf{A}$ is dispensable in all of our electrodynamical explanations. Then rather than maintaining that despite all appearances to the contrary there really is just one real $\mathbf{A}$, why not simply abandon it? The appropriate stance seems to be to acknowledge the irreality of both $\mathbf{A}$ and ϕ, and to regard them as merely mathematical fictions. Assessments of the appropriateness of this stance changed with the publication of Aharonov and Bohm's famous paper in 1959.[5]

In that paper, Aharonov and Bohm proposed a thought experiment that would show the "reality" of both $\mathbf{A}$ and ϕ. For the former they suggested scattering a quantum-mechanical electron around a solenoid. Because the solenoid was so long, there would be no appreciable magnetic field in the region of motion of the electron even

when there was a current flowing in the solenoid and hence a magnetic field inside. And yet, when the current was flowing there would be a measurable effect: the interference pattern of the electron would be shifted from that produced when the current was zero. Thus, since **A** was non-zero in the region of travel, Aharonov and Bohm concluded that it was real. (There have been a number of experimental confirmations of this well-known effect. Indeed researchers now regularly use "Aharonov-Bohm flux rings" to probe quantum-mechanical mesoscopic effects.)

Before I say what lesson was drawn from this experiment, I want to explain why there is one lesson that certainly should not be drawn: that gauge fields are real only when they are quantum-mechanical. I make this explanation is more detail elsewhere,[6] but in its simplest form, the argument goes like this: the electromagnetic field used to observe the efficacy of **A** was itself—in the context of the derivation as well—a classical field. Thus it is the classical **A** that was observed, not a quantum-mechanical **A**. So whatever we decide about how to understand **A**, and gauge fields generally, will have to apply to our understanding of classical field theories as well.

To return to the Aharonov–Bohm effect: now that the reality of **A** has been established, there are a number of approaches one could take to its interpretation. The simplest, I suppose, would be simply to say "oops, I guess one of the fields *was* the real one," and then move on to more pressing issues. There are important considerations though that apparently disallow this option. Looming large among them are various combinations of worries about non-determinism, non-locality, and epistemological underdetermination. I will not consider these here in any detail.[7] Instead I will simply repeat the consensus view of those physicists and philosophers of physics who have written on this: one cannot choose the naive option. There, however, the consensus very nearly ends. For there are a number of conflicting proposals setting out the right way to understand **A**. But what all these approaches have in common is their willingness to accept the causal efficacy of **A**—while denying the efficacy of any one member of the family of functions that makes up **A**. For, it is argued, **A** is clearly what causes the shift in interference patterns: there is no **E** field there; there is no **B** field there; the shift is functionally related to **A**, which is non-zero there; so it has to be **A**. Thus in one way or another we have a new kind of entity that appears to be a family of vector fields—again, none of which exists—which itself does exist and exerts an influence on the paths of electrons. What could be stranger than that?[8] We conclude from this that on the current best view of the physics and philosophy of physics community that gauge fields are causally efficacious; that the "unreality" of a quantity is no argument against its role in causal explanations, for the sense of unreality may be no more than an effect of the need to choose a new kind of property.

The most important reason for considering gauge theory in this context is to see whether the techniques of physics can be imported into analyses of chemistry without begging the question of the autonomy (or dependency) of chemistry from (or on) physics. So while I think there is an important sense in which gauge theoretical techniques can be applied to chemistry—especially in the context of debates over chemical structure—this work is speculative. I will not be presenting a mathematical embedding of chemistry into the category of gauge theories. I will not even be showing the mathematical possibility of re-characterizing certain chemical properties as gauge fields. Instead I will be pursuing an analogy between chemical properties and gauge theories, and observing that the way theorists who reason about the connec-

tion between gauge theory and properties in physics has a resonance with some ways of reasoning about chemical properties. The conclusion I draw is that further work on this issue is warranted in order to see whether this analogy can be fruitfully supported with more detailed characterization of chemical properties as, say, constrained Hamiltonian systems.

MOLECULAR STRUCTURE: THE CHARGE DENSITY

As I pointed out above, concern with (and suspicion about) the notion of structure in chemistry is not new. But I can here offer neither a comprehensive overview of the history of structure debates in chemistry nor an analysis of the present consensus or lack thereof about what constitutes chemical structure.[10] Instead I will be focusing exclusively on Bader's account of how we ought to understand and define chemical structure. But note that I am not particularly concerned with whether Bader's proposal holds up to critical scrutiny. And I will not be offering much such scrutiny myself. I am, however, very interested in the pattern of explanation he employs, and the way this pattern connects with patterns of explanation employed by those who study gauge theories. It is this that I wish to explore here. So while the remarks I make will be specific to his account, I believe that the lessons I draw will be more general.

I begin this part of the paper with a brief account of Richard Bader's paradigm for attributing structure to molecules described by quantum-mechanical state functions. After a brief clarification in this section of one of his candidate physical properties, I suggest, in the next section, how this paradigm might be seen as an application of gauge theoretical ideas to the question of molecular structure. I will indicate below certain similarities between the resources to which Bader appeals in his account of structure and the general account of gauge theories given above. The key point is that appeals to underlying but unrealized features of the situation in chemistry resemble—both in a technical sense and in what might be called their philosophical flavor—the physicists' appeal to dynamical variables that never take on particular values.

Bader begins his account of chemical structure with a litany of the problems that cannot be solved by appeals to geometric analyses of structure: we are unable, he says

> …to assign a single geometric structure, average or otherwise, to rotation- or inversion-related isomers, to a molecule in an excited vibrational state with geometrical parameters very different from those for the same molecule in its ground state, or to a molecule in a "floppy" state wherein the nuclear excursions cover a wide range of geometrical parameters. (p. 54)

He identifies the cause of these inabilities as follows:

> In reality, these are shortcomings of attempts to impose the classical idea of geometry on a quantum system. The nuclei, like the electrons, cannot be localized in space and instead are described by a corresponding distribution function. (p. 54)

Before detailing Bader's solution to the conundrum of applying geometry to quantum systems, I need to back up a little and outline the most important resource he has developed for this purpose. This is what Bader calls "the charge density." Now this is a funny quantity. As he assures us in various places, "the charge density, $\rho(\mathbf{r})$, is a physical quantity which has a definite value at each point in space." (p. 13)

Since this quantity plays such a central role in his analysis, it is crucial that we see what precisely it is. The charge density, or "electronic" charge density as he sometimes calls it, is the probability density of the electron wave function multiplied by the unit of electronic charge. (chapt. 1, especially pp. 6–7) In what sense is this a charge density? In fact it simply isn't, except at best in some collapse theories of quantum mechanics where $\psi^*\psi$ is taken to be density of stuff. But Bader can't be using one of these versions since he needs his electrons to be fully non-localized. That is, he needs for what follows to define a charge distribution that is continuous and twice continuously differentiable (except at the location of the nuclei). But on collapse, the electron wave function localizes to a delta function distribution. Of course this localization spreads almost immediately to a distribution throughout space, but this doesn't help in Bader's quest for an in-principle notion of structure. More crucially Bader will need to be able to show that the properties of the electronic charge distribution attain local maxima only at the location of nuclei. But on any collapse model the electron's wave function will be peaked sharply about its localization position. Thus it is clear that Bader is using a no-collapse model of quantum mechanics. (Note that it won't do to suggest that there is some kind of continuous collapse going on here, induced by the coupling of the vibrational or rotational modes of the molecule, or by some other mechanism. Bader's argument for why we cannot use classical geometry is a *principled* objection. No matter how small the magnitude of the probability density far away from the classical nuclear location, that density is non-zero. Bader is under no illusions about this, but is not always explicit about what is entailed by the fact that the "nuclei…cannot be localized in space.")

I will not comment on the various complications introduced into Bader's account by his lack of attention to the quantum-mechanical measurement problem (which is what the above issue amounts to). Instead I will suggest that he would do just as well to consider what his charge density really is: the expectation value for the charge at each point of space. What's the difference? Well first, it isn't an average. For while the expectation value of the charge at some given point may be non-zero, on a no-collapse view of quantum mechanics there is no answer to the question "what's the average charge here?" To have an "average charge" at a point, one needs the charge at that point to be well-defined at each moment. But on a no-collapse view of quantum mechanics, there is no answer to the question "what's the charge here now?" Even on a collapse view, where these questions do have answers, the answer is not Bader's "charge density." It is perfectly consistent to suppose a non-zero expectation value for the charge at a point while also supposing that no charge is ever at that point. I submit that since Bader's "charge density" as defined is really just the expectation value of the charge, he should not obscure matters by calling it the average charge.[11] On the other hand, the expectation value won't do as a quantum-mechanical quantity which couples to, say, the nuclear charge of some atom. This is essentially the point to Rosenfeld and Bohr's 1933 paper, where they show that expectation values of the charge cannot, for experimental reasons, play the role of the electric field. Instead, they show, the electric field in quantum mechanics must itself be a true quantum-mechanical field.[12]

Now because Bader's account of the charge density is incorrect, we cannot say that the behavior of the molecule is specified by the action of this scalar field. So what do we have instead? I'll return to this question after outlining the rest of Bader's

proposal. Here I will merely suggest that the results he obtains employing the notion of equivalence classes of geometries suggest that we take seriously the idea that the topology of the "electronic charge density" is the real dynamical quantity of interest, and that we follow Bader in rejecting geometric configuration as primary. To this I now turn.

MOLECULAR STRUCTURE: EQUIVALENT GEOMETRIES

We have, as Bader shows us, a notion of structure given in the pre-quantum understanding of the atomic constituents of molecules. This is simply that molecular structure is the molecular geometry given by the arrangement of constituent atoms. The principle distinction between the quantum and pre-quantum cases is that in the latter we can make coherent the notion of a fixed geometry. But in the former case, recall, a fixed geometry of unsharply localized electronic and nuclear constituents is incoherent. To get around this problem, Bader will use instead the idea of equivalence classes of geometries as the structure of the molecule. An equivalence relation over possible nuclear configurations provides Bader with the notion of structure he requires to carry out his program of demonstrating "that the existence of atoms with definable properties and the associated concepts of the molecular structure hypothesis are a consequence of the quantum description of matter."(p. 2)

The new notion of structure is complicated in execution, but quite simple in idea. The rough and ready outline goes like this: pick, as exemplar, a classical geometry of the atoms involved in the molecule of interest. Consider, in "geometry space" the neighborhood of the exemplar geometry. For all geometric configurations in the region defined by the molecular bond structure of the exemplar molecule, we say that these geometric configurations have the same molecular structure.(p. 54ff) These then are equivalent geometries. But notice that we cannot mean by this the following hypothetical suggestion: "If the actual geometry of the molecule is in the equivalence class for most of some period of time, we assign it the structure associated to that class. If, on the other hand, it spends most of its time exemplifying the geometries of another class, we assign it the structure of that class." For the whole problem is that molecules just don't have a given geometry. On the other hand, if we allow that Bader has succeeded in his descriptive task, we *can* understand that the geometric possibilities of the molecule are constrained to lie inside a given equivalence class of geometries.

Bader takes it that a conception of structure developed along these lines will allow for a solution to the "central problem of molecular structure—to predict the discontinuous changes in structure that are caused by a continuous change in nuclear coordinates." (p. 88) Here Bader makes it clear that for him the issue is not primarily about whether we *can* attribute structure to molecules in the way outlined, since he takes himself to have succeeded already in that. Rather the issue is how, given that they do have structure, we can account for the observed evolution of molecular structures. As a philosophical issue, however, the connection between attribution of structure and its causal agency is not so clear.

Bader immediately goes on blandly to observe:

> It is important to remember that, while the structural aspects of the theory of atoms in molecules are described by the dynamical changes in the topology of the charge densi-

ty, the theory is rooted in quantum mechanics. It is the atom and its properties which are defined by quantum mechanics. The bond paths and the structure they define just mirror and summarize in a convenient way what the atoms are doing, performing the same role here as does the assignment of a set of bonds in the molecular structure hypothesis. (p. 89)

But the bond paths he mentions are summaries of the properties of the "charge density," as are the sets of bonds. For us it is important to remember this, for it is here that Bader's reasoning is most transparent. Molecular structure is irreal. It is a kind of "mirror" of the behavior of atoms. But the atoms don't have any particular behavior that is being mirrored. That is where we started this investigation. Instead the atoms have a connection to the expectation value of the charge distribution. But this is itself irreal—irreal in the sense that it cannot be taken to be the way various charges interact with electronic fields. As I said above, this is precisely the point that Rosenfeld and Bohr make. That said, the expectation value of charge is still well-defined. The problem is how to understand that it (or rather the structure it is used to define) has any efficacy. Assuming, along with Bader, that this notion of equivalence classes of geometry is sufficient to his purposes, however, seems to commit us to the idea that irreal properties are causally potent. For this reason, I suggest that the resources of gauge theory are sufficient to underwrite an explanation for the causal powers of chemical structure.

WHY CHEMICAL STRUCTURE IS A GAUGE PROPERTY

First to summarize the problem: The idea that there is some definite meaning that attaches to the topology of molecules but not to their geometries is a little puzzling. For again, as Bader tells us, the reason that molecules have no geometry is that the nuclei of their constituent atoms, being quantum-mechanical, have no definite location. (p. 54) And yet a similar problem obtains for the topological features of the molecules as well. There is no definite sense to the claim that a nucleus is on *this* side of some critical surface of structure change as opposed to *that* side. One could, I suppose, invoke some theory of quantum-mechanical measurement at this point and say that the nucleus is localized by the observer, or the environment, or whatever. But again, and more strongly, if there is any spread at all in the wave-function of the nucleus, then there is spread to arbitrary distances. And then there cannot be a sharp demarcation from one equivalence class of structure to another. So we must explain in a different way how we understand structure as well as how we understand change of structure.

It is here that I think the analogy between chemical theories and gauge theories may be fruitfully articulated. For we know that the attribution of structure is crucial to much of chemical practice. Some of the analogical features are the following: (1) The equivalence classes Bader defines may be seen as constraint manifolds, with change of structure given by the breaking of constraints by the introduction of out-side forces—a "lifting off of" the manifold; (2) Attributing a particular geometrical configuration to a molecule is like picking out a single point of the constraint mani-fold, and tracking the choice would be analogous to specifying a gauge trajectory—but this apparently violates determinism and is disallowed; we needn't (indeed we cannot) believe that a given structure is reflective of just one single configuration of

the molecule, just as a particular choice of gauge need not be associated to a given gauge potential—structure is crucial, but no given geometry is real.

Then to summarize the "solution": We have, in chemistry, a situation that is similar in certain respects to the situation in gauge theories. On the face of things it seems impossible to attribute structure to quantum-mechanical molecules. And yet we require structure to underwrite much of our explanatory practice in chemistry. Fortunately we can make sense of this explanatory practice without imposing an arbitrary notion of structure on chemistry, without, that is, appealing to extra-chemical physical resources, by noting how a similar practice in physics is under-written using gauge-theoretic techniques—techniques that are not themselves particular to physics but should be considered general principles of physical theorizing.

What remains to be done is clear, but not easy. One would need to show, at a minimum, that the analogy I have suggested between chemical structure and gauge theory can be extended in the following way: One would have to show that we can recast analyses of chemical structure (and here I refer specifically but not necessarily exclusively to Bader's account) into the mathematical language of gauge theory. There are a number of different candidates for that language, but the likeliest choice for use in chemistry is that of constrained Hamiltonian systems. The idea would be to represent the various different geometries in each equivalence class as points in a sub-manifold of configuration space, find an appropriate sense of dynamical evolution for geometric states, and then characterize the tendency of molecules to remain within in a single equivalence class as the operation of a weak constraint on the system. (We assign weak constraints because we know that molecules can, without outside interference, move between geometric equivalence classes.) We then would try to represent the interactions of molecules in terms of the constraints they jointly obey. It is too early to report much in the way of progress on this front, but the work is ongoing.

In advance of carrying out the project of the previous paragraph, I still think there is something that can be said in terms of the usefulness of the project for the philosophy of chemistry. What we see here is that possibilities exist for the peaceful co-existence of chemistry and physics as "fundamental" theories. How so? "If anything," a critic might say, "this program reduces even more chemistry to physics, by its importation of constrained Hamiltonian systems, principle fiber bundles—the whole panoply of physics tools." And yet that's all that's being imported: the relatively new mathematical tools associated with gauge theory. On the other hand, the application of these tools shows how to speak meaningfully about the autonomy of systems that are clearly built up out of other systems. If these tools are applied in the way I have suggested, then one need neither pose nor answer the question: "Is chemistry reducible to physics?" The irrelevance of *that* question will be manifest.

NOTES AND REFERENCES

1. The idea of giving a retrospective reevaluation of what *was* true for older theories is an extremely interesting topic in its own right, but it would take us too far afield to consider it now. Some of this is done in MATTINGLY, J. 2001. Singularities and scalar fields: matter theory and general relativity. Philos. Sci. **68:** S395–S406.
2. BADER, RICHARD F.W. 1990. Atoms in Molecules—A Quantum Theory. Clarendon Press. Oxford, UK.
3. MATTINGLY, JAMES. 2003. Why there are no gauge fields. Unpublished manuscript.

4. I speak all along in this paper as though I believe the story I'm telling about gauge theory. I don't. But the proper connection between what I do believe and the story I'm telling here will have to wait. What I express is the overwhelming consensus in present discussions of the nature of gauge theories.

5. AHARONOV, Y. & DAVID BOHM. 1959. Significance of potentials in the quantum theory. Phys. Rev. **115:** 485–491.

6. MATTINGLY, "Why there are no gauge fields."

7. I carry out a critical appraisal in the work cited in note 6.

8. These and other issues have received extensive treatment in the literature. See for instance: BELOT, G. 1998. Understanding electromagnetism. Br. J. Philos. Sci. **49:** 531–555; HEALEY, R. 1997. Nonlocality and the Aharonov–Bohm effect. Philos. Sci. **64:** 18–40; LEEDS, S. 1999. Gauges: Aharonov, Bohm, Yang, Healey. Philos. Sci. **66:** 606–627; and for an early and still crucial treatment, WU, T.T. & C.N. YANG. 1975. Concept of nonintegrable phase factors and global formulation of gauge fields. Phys. Rev. **D12:** 3845–3846.

9. See the first three references in note 4 for representative views.

10. However, these issues are addressed by other contributions to this volume, and references therein.

11. For those who are inclined to demand that "actually there" be given an operational definition, you are already opposed to saying that the expectation value is the average charge. Average charge is itself a nonoperational kind of quantity.

12. ROSENFELD, L. 1963. On quantization of fields. Nucl. Phys. **40:** 353–356.