Draft. For PSA Y2K Symposium on the Concept of Gauge in Modern Physics. Not for quotation

Gauge Matters

John Earman
Dept. of History and Philosophy of Science
University of Pittsburgh

1 Introduction

Among the concepts of modern physics that have been largely neglected by philosophers of science, that of "gauge" is arguably the most important. A PSA 1998 symposium on gauge theories made a start on redressing the neglect,[1] but much remains to be done. The first order of business is demystification.

The physics literature contains many examples of the so-called "gauge argument" and "gauge principle." In barest outline, the argument runs thus. One starts with a free field which obeys a conservation law. By Noether's (first) theorem there is an associated "global symmetry." Appeals to relativity theory and "locality" are made to motivate passing to a "local symmetry." The gain of such a passage is startling: it turns out that in order to implement the local symmetry requirement it is necessary to introduce a new field which interacts with the original field (and, possibly, itself) in a specified way. Thus, the "gauge principle" can be glossed as saying that the interactions of the field are dictated by the requirement that the global symmetry be made local. At first sight the success of this gauge argument in capturing some of the fundamental physical interactions is nothing short of amazing. But a closer look reveals that there is both more and less than initially meets the eye. Chris Martin's contribution to this symposium cautions against a literal reading of the "gauge argument" and reveals a competing logic of nature in which the "gauge principle" is the output rather than the input.

But an even more fundamental kind of demystification is needed. Here I want to be deliberately provocative and intentionally unkind to my colleagues in accusing them of being blinded by the glitz. A prime example of what I have in mind is being blinded by the glare of the fibre bundle formalism. This formalism is very powerful and it is arguably essential to understanding some aspects of gauge theories. But all too often the discussions of gauge matters are so smothered in this formalism that they are virtually impenetrable.[2] And in any case, a satisfactory explication of the concept of gauge should explain how the relevant fibre bundle structure emerges. I will recommend an approach to getting a grip on the key issues–the constrained Hamiltonian formalism–that neither uses nor mentions fibre bundles–though, not surprisingly, fibre bundles do pop out in a natural way, at least in well-behaved cases. In these cases the fibres are the gauge orbits, which are identified in an independent manner. A related fallacy derives from the notion that the best way to understand the concept of gauge is to draw examples from those theories of contemporary physics that are self-consciously labeled gauge theories. An instance of an error to which this fallacy leads is the false conclusion that a gauge theory is a Yang-Mills theory. But although Maxwell's theory of electromagnetism and the non-abelian Yang-Mills generalizations of this theory form an important class of gauge theories, they do not exhaust the domain. For example, the standard formulation of the general theory of relativity (GTR) arguably embodies a big gauge group even though it is not a Yang-Mills theory.

After the demystification program is carried through, a host of foundational issues beg for attention. I will concentrate on one, some of whose many facets will be treated in this symposium. It concerns the concept of "observable." If a theory is formulated in a way that admits a non-trivial gauge freedom, there arises the challenge of reformulating the theory so that it uses only genuine observables or gauge invariant quantities. (This challenge becomes especially important when one is trying to quantize (see below), but even when one is not, meeting the challenge is an important part of interpreting the theory in the sense of saying what the world would have to be like for the theory to be true.) In some instances there is an obvious solution. For example, in classical electromagnetism formulated in terms of potentials one has the freedom to perform gauge transformations on the vector and scalar potentials; the genuine observables are the electric and magnetic fields or, more properly, the electromagnetic field tensor. The case of GTR is more challenging and perplexing, at least if one takes the point of view that the diffeomorphism group is the gauge group of the theory as it is formulated in standard textbook form. The construction of the gauge invariant quantities for GTR is an unsolved problem. But even prior to a solution, one knows in advance that most of the quantities ordinarily taken to be observables will not appear on the final list. And in fact, the list is so niggardly that a modern version of McTaggartism, according to which there is not even any "B-series" change, threatens (see section 3).

Issues about observables become especially acute when one tries to quantize a theory with gauge freedom since, presumably, the quantities that get promoted to quantum observables (in the sense of self-adjoint operators on an appropriate Hilbert space) are the gauge invariant quantities. One route to quantization is to move from the original phase space to a reduced phase space whose points are the gauge orbits in the original phase space, and then to quantize the true degrees of freedom, captured in the reduced phase space, in the normal way one quantizes any unconstrained Hamiltonian system.[3] In simple and well-behaved cases this program of reduced phase space quantization proceeds smoothly. But in more complicated and physically interesting cases it runs into various technical obstructions. Moreover, even in cases where the program of reduced phase space quantization can be carried through, it can lead to results that are physically at variance with Dirac quantization which proceeds at the level of the unreduced phase space and which implements gauge invariance by requiring that the physical Hilbert space consist of vectors that are annihilated by the quantum counterparts of the classical gauge constraints. In cases where there is a variance one can ask: Which is the "correct" quantization? And how is one to settle this question? Gordon Belot will comment on these issues in his contribution to this symposium.

2 Gauge and gauge transformations: the constrained Hamiltonian formalism

Before embarking on my approach let me acknowledge that it is certainly not the only one, nor do I claim that it is the best (indeed, I'm not sure there is any such thing as *the* best approach to gauge). But the approach I am going to describe does have the virtues of clarifying a number of issues and of linking the gauge concept to a number of interesting foundations problems in physics. In any case, it is an approach worth describing because although it is mothers' milk to a segment of the physics community, it is relatively unknown to philosophers of science.

The literature on gauge theories is filled with talk about "global" and "local" symmetries, talk which is annoying both because it is often unaccompanied by any attempt to make it precise and because it is potentially very misleading (e.g. a global mapping of the spacetime onto itself can count as a "local" symmetry in the relevant sense of corresponding to a gauge transformation). One way to get a grip on the global vs. local distinction is to use Noether's two theorems.[4] Both theorems apply to theories whose equations of motion or field equations are derivable from an action principle and, thus, are in the form of (generalized) Euler-Lagrange equations (ELEs). And both concern *variational symmetries*, that is, a group $\mathcal{G}$ of transformations that leave the action $\mathcal{L} = \int_\Omega L(\mathbf{x}, \mathbf{u}, \mathbf{u}^{(n)})d\mathbf{x}$ invariant.[5] (Here $\mathbf{x} = (x^1, ..., x^p)$ stands for the independent variables, $\mathbf{u} = (u^1, ..., u^q)$, are the dependent variables, and the $\mathbf{u}^{(n)}$ are derivatives of the dependent variables upto some finite order n with respect to the x^i.) Every such variational symmetry is a symmetry of the ELEs (i.e. carries solutions to solutions), but in general the converse is not true.

Noether's first theorem concerns the case of an r-parameter Lie group $\mathcal{G}_r$, which I take to be the explication of the (badly chosen) term "global symmetry." The theorem states that the action admits a group $\mathcal{G}_r$ of variational symmetries iff there are r linearly independent combinations of the Euler-Lagrange expressions L_A, $A = 1, 2, ..., q$, which are divergences, i.e. there are r p-tuples $\mathbf{P}_j = (P_j^1, ..., P_j^p)$, $j =$

$1, 2, ..., r$, where the P_j^i are functions of $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{u}^{(n)}$ such that

$$Div(\mathbf{P}_j) = \sum_A c_j^A L_A, \qquad j = 1, 2, ..., r \tag{1}$$

where $Div(\mathbf{P}_j)$ stands for $\sum_{i=1}^{p} D_i P_j^i$ and D_i is the total derivative with respect to x^i. Thus, as a consequence of the ELEs, $L_A = 0$, there are r *conservation laws*

$$Div(\mathbf{P}_j) = 0, \qquad j = 1, 2, ..., r. \tag{2}$$

The $\mathbf{P}_j$ are called the *conserved currents*.[6] It is well to note that there are equations of motion (or field equations) which cannot be derived from an action principle, and in such cases there is no guarantee that a symmetry of the equations of motion (or field equations) will give rise to a corresponding conserved quantity. A concrete application of Noether's first theorem is provided by interacting point masses in Newtonian mechanics, provided that the equations of motion follow from an action principle. The requirement that the inhomogeneous Galilean group is a variational symmetry entails the conservation of energy, angular and linear momentum, and the uniform motion of the center of mass. Conversely, the existence of these conservation laws entails that the action admits a 10-parameter Lie group of variational symmetries.

Noether's second theorem is concerned with the case of an infinite dimensional group $\mathcal{G}_{\infty r}$ depending on r arbitrary functions $h_j(\mathbf{x})$, $j = 1, 2, ..., r$. This I take to be the explication of the (badly chosen) term "local symmetries." The theorem states that the action admits a group $\mathcal{G}_{\infty r}$ of variational symmetries iff there are r dependencies among the ELEs, in the form of linear combinations of derivatives of the L_A, which vanish identically. These dependencies can be interpreted as "strong" (aka "off shell") conservation laws that hold not as consequences of the ELEs but are mathematical identities (aka "generalized Bianchi identities"). Since the ELEs are not independent, we have a case of underdetermination, and as a result the solutions of these equations contain arbitrary functions of the independent variables–an apparent violation of determinism.

Now comes an insight of Noether's that is important enough to be labeled as the third Noether theorem. The action may admit *both* $\mathcal{G}_r$ and $\mathcal{G}_{\infty r}$ as variational symmetries. Noether showed that if $\mathcal{G}_r$ is a rigid subgroup of $\mathcal{G}_{\infty r}$ in the sense that $\mathcal{G}_r$ arises from $\mathcal{G}_{\infty r}$ by fixing $h_j(\mathbf{x}) = \alpha_j = const.$, then the relations (1) are consequences of the Bianchi identities; indeed, each of the relations (1) is a linear combination of these identities. Further each of the conserved currents $\mathbf{P}_j$ can be written as a linear combination of the L_A plus the divergence $\sum_{i=1}^{p} D_i X^{ij}$ of an antisymmetric quantity

$X^{ij}(\mathbf{x}, \mathbf{u}, \mathbf{u}^{(n)})$, $X^{ij} = -X^{ji}$. In this case the conservation laws (2) were dubbed "improper" by Noether. Today mathematicians call such conservation laws "trivial." That $Div(\mathbf{P}) = 0$ is "trivial" means that it is a linear combination of conservation laws $Div(\mathbf{R}_m) = 0$, where each $\mathbf{R}_m$ either vanishes on all solutions of the ELEs or else is such that $Div(\mathbf{R}_m) = 0$ holds for all functions $\mathbf{u}(x)$ regardless of whether the ELEs are satisfied. But note that although Noether's "improper" conservation laws are "trivial," they are not necessarily "strong" or "off shell." Note also that "improper" or "trivial" conservation laws may be far from trivial–for example, in Maxwell's theory of electromagnetism in Minkowski spacetime they imply Gauss' law and the conservation of electric charge![7]

From the perspective of Noether's theorems, the "gauge argument" mentioned in the Introduction amounts to noting that *one* way to move from a group $\mathcal{G}_r$ of variational symmetries (Noether's first theorem) to the bigger group $\mathcal{G}_{\infty r}$ of variational symmetries is to add new terms to the Lagrangian density. To the extent that the move from $\mathcal{G}_r$ to $\mathcal{G}_{\infty r}$ is well-motivated and to the extent that the required additions to the Lagrangian density are unique and/or natural, one can see in the procedure a method of discovery of interactions. One has to examine the details of particular cases to see how well the "to the extent" clauses are fulfilled. And note also that the move from Noether's first to second theorem by embedding $\mathcal{G}_r$ as a rigid subgroup of $\mathcal{G}_{\infty r}$ means that what was initially taken to be a non-"trivial" conservation law is changed into a "trivial" one.

To return to the main theme, the underdetermination encountered in Noether's second theorem is one of the principal roots of the notions of gauge and gauge transformation. In one of its main uses, a "gauge transformation" is supposed to be a transformation that connects what are to be regarded as equivalent descriptions of the same state or history of a physical system. And one key motivation for seeking gauge freedom is to mop up the slack that would otherwise constitute a breakdown of determinism: taken at face value, a theory which admits "local" gauge symmetries is indeterministic because the initial value problem does not have a unique solution; but the apparent breakdown is to be regarded as merely apparent because the allegedly different solutions for the same initial data are to be regarded as merely different ways of describing the same evolution or, in equivalent terms, the evolution of the genuine or gauge invariant quantities is manifestly deterministic.

Now the obvious danger here is that determinism will be trivialized if whenever it is threatened we are willing to sop up the non-uniqueness in temporal evolution with what we regard as gauge freedom to describe the evolution in different ways. Is there then some non-question begging and systematic way to identify gauge freedom and to characterize the genuine observables? The answer is yes, but specifying the details involves a switch from the Lagrangian to the constrained Hamiltonian formalism. To motivate that switch, let me note that, subject to some technical provisos, if one is in the domain of Noether's second theorem (i.e. the action admits "local" symmetries = a group $\mathcal{G}_{\infty r}$ of variational symmetries), as I have been assuming is the case for gauge theories, then the Lagrangian density (more properly, its Hessian) is singular (see Wipf (1994)), and the Legendre transformation which defines the canonical momenta shows that these momenta are not all independent. Hence, one is in the domain of the constrained Hamiltonian theories. To illustrate what this means and to underscore the point that the key ideas about gauge arise in the humblest settings, I will work with examples drawn from classical particle mechanics.

Suppose that we are dealing with a theory whose equations of motion are derivable from the action principle

$$\delta \int L dt = 0, \qquad L = L(q, \dot{q}), \qquad \dot{q} := dq/dt. \tag{3}$$

(Allowance can be made for higher order Lagrangians, e.g. $L = L(q, \dot{q}, \ddot{q})$.) In their familiar form the ELEs are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^n}\right) - \frac{\partial L}{\partial q^n} = 0, \quad n = 1, 2, ..., N. \tag{4}$$

These equations can be rewritten as

$$\ddot{q}^m \left(\frac{\partial^2 L}{\partial \dot{q}^m \partial \dot{q}^n}\right) = \frac{\partial L}{\partial q^n} - \dot{q}^m \left(\frac{\partial^2 L}{\partial q^m \partial \dot{q}^n}\right). \tag{5}$$

If the Hessian matrix $\left(\frac{\partial^2 L}{\partial \dot{q}^m \partial \dot{q}^n}\right)$ is singular, one cannot solve for $\ddot{q}^m$ in terms of the positions and velocities, and determinism (apparently) fails because arbitrary functions of time appear in the solutions. But a way to recoup appears in the Hamiltonian formalism.

When the Hessian matrix $\left(\frac{\partial^2 L}{\partial \dot{q}^m \partial \dot{q}^n}\right)$ is singular, the canonical momenta defined by the Legendre transformation $p_n := \frac{\partial L}{\partial \dot{q}^n}$ are not all independent but must satisfy constraints

$$\phi_m(p, q) = 0, \qquad m = 1, 2, ..., K < N \tag{6}$$

that follow from the definitions of the momenta. These are the *primary constraints*. *Secondary constraints* may also appear when it is demanded that the primary constraints be preserved under the allowed motions. More important for our purposes is a second dichotomy which cuts across the primary vs. secondary dichotomy; namely, a constraint is *first class* if it commutes (i.e. has vanishing Poisson bracket[8]) with all of the constraints, otherwise it is called *second class*. Dirac (1950, 1951, 1964), who along with Peter Bergmann (1949, 1961) was responsible for developing this formalism, proposed that the gauge transformations be identified with the transformations generated by the first class constraints and that the observables be identified with the functions of (p, q) that commute with the first class constraints or, equivalently, are constant along the gauge orbits.[9]

All of this may sound scary, but a feeling for what is going can be gained from the following humble concrete illustration. Those of you who have read Maxwell's (1877) *Matter and Motion* may have been puzzled by his apparently contradictory claim that acceleration is *relative* even though rotation is *absolute* (see sections 32-35 and 104-105). Maxwell is consistent if we take him to be proposing that physics be done in the setting of Maxwellian spacetime. Like Newtonian spacetime Maxwellian spactime has absolute simultaneity, the $\mathbb{E}^3$ structure of the instantaneous space, and a time metric, but it eschews the full inertial structure in favor of a family of relatively non-rotating rigid frames.[10] In terms of coordinate systems adapted to the absolute simultaneity, the $\mathbb{E}^3$ structure, and the privileged non-rotating frames, the symmetry transformations of Maxwellian spacetime are

$$\begin{aligned} \mathbf{x} \quad &\rightarrow \quad \mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{a}(t) \\ t \quad &\rightarrow \quad t' + const. \end{aligned} \tag{Max}$$

where $\mathbf{R}$ is a constant rotation matrix and $\mathbf{a}(t)$ is an arbitrary smooth function of time. In such a setting it seems hopeless to have determinism if, as ordinarily assumed, positions and velocities of particles are regarded as observables. For we can choose $\mathbf{a}(t)$ such that it is zero for all $t \leq 0$ and non-zero for $t > 0$. Since a symmetry of the spacetime should be a symmetry of the equations that specify the permitted particle motions, the application of (Max) to a solution of the equations of motion will produce another solution that agrees on the particle trajectories of the first solution for all past time but disagrees with it in the future–an apparent violation of even the weakest form of determinism.

Now let's see how this example gets reinterpreted when cranked through the Lagrangian/constrained Hamiltonian formalism. An appropriate Lagrangian invariant under (Max) is

$$L = \sum \sum\nolimits_{j<k} \frac{m_j m_k}{2M} (\dot{\mathbf{r}}_j - \dot{\mathbf{r}}_k)^2 - V(|\mathbf{r}_j - \mathbf{r}_k|), \qquad M := \sum_i m_i. \tag{7}$$

The transformations (Max) are global mappings of Maxwellian spacetime onto itself, but they are "local" in that Noether's second theorem applies since (Max) contains an infinite dimensional group $\mathcal{G}_{\infty 3}$. You can verify that the Hessian matrix for (7) is singular. The ELEs are:

$$\frac{d}{dt}\left(m_i(\dot{\mathbf{r}}_i - \frac{1}{M}\sum_k m_k \dot{\mathbf{r}}_k)\right) = \frac{\partial V}{\partial \dot{\mathbf{r}}_i}. \tag{8}$$

These equations do not determine the evolution of the particle positions uniquely: if $\mathbf{r}_i(t)$ is a solution, so is $\mathbf{r}'_i(t) = \mathbf{r}_i(t) + \mathbf{f}(t)$, for arbitrary $\mathbf{f}(t)$, confirming the intuitive argument given above for the apparent breakdown of determinism.

Now let's switch to the Hamiltonian formalism and find the constraints. The canonical momenta are:

$$\mathbf{p}_i := \frac{\partial L}{\partial \dot{\mathbf{r}}_i} = \frac{m_i}{M}\sum_k m_k(\dot{\mathbf{r}}_i - \dot{\mathbf{r}}_k) = m_i \dot{\mathbf{r}}_i - \frac{m_i}{M}\sum_k m_k \dot{\mathbf{r}}_k. \tag{9}$$

These momenta are not independent but must satisfy three primary constraints:

$$\phi_\alpha = \sum_i p_i^\alpha = 0, \quad \alpha = 1, 2, 3. \tag{10}$$

These primary constraints are the only constraints–there are no secondary constraints. They are first class, and they generate (Max). So the genuine or gauge invariant observables of such a theory are the quantities that are invariant under (Max), which include quantities like relative positions, relative velocities, and relative accelerations of the particles. The equations of motion are deterministic with respect to these observables.

In the spirit of *Church's Thesis*[11] I would like to propose what I will call the *Gauge Thesis:* (1) For theories whose equations of motion or field equations are derivable from an action principle, gauge freedom is present when the action admits an infinite dimensional group $\mathcal{G}_{\infty r}$ of variational symmetries. (2) Noether's second theorem implies there are strong conservation laws (Bianchi identities) and that the ELEs are not independent–a case of underdetermination, an apparent breakdown in determinism. The Lagrangian density has a singular Hessian, and therefore the corresponding Hamiltonian formulation contains constraints. (3) The gauge transformations are to be identified with those transformations generated by the first class constraints, and the genuine observables are to be identified with the gauge invariant quantities. Part (1) is widely accepted (among the handful of cognescenti who write on these matters–see, for example, Wipf (1994)). Part (2) is a matter of mathematical proof. Part (3) is controversial for various reasons that will be brought out in sections 3 and 4.

Where do fibre bundles enter this picture? Some theories are constructed in a self-conscious manner so as to emphasize a fibre bundle structure. But for other theories the bundle structure relevant to gauge freedom has to be identified by independent means. In the constrained Hamiltonian formalism the *constraint surface* $\mathcal{C}$ consists of the subspace of the original phase space $\Gamma = \Gamma(p, q)$ where all of the constraints are satisfied. The reduced phase space $\tilde{\Gamma}$ is defined as the quotient of $\mathcal{C}$ by the gauge orbits generated by the first class constraints. One would like it to be the case that the gauge orbits are the fibres of a bundle with base space $\tilde{\Gamma}$. In practice this will not always be the case.[12] If one believes that the fibre bundle apparatus captures an essential feature of nature, then one could posit that the emergence of the appropriate bundle structure is a necessary condition for genuine physical possibility. This is an interesting idea, but obviously it requires critical examination.

3 GTR and gauge

In the late 1940's Peter Bergmann rediscovered a version of Einstein's notorious hole argument (e.g. see Bergmann (1949)).[13] Skipping the technical details, the hole argument can be viewed as a generalization of the quick and dirty argument given above to convince the reader that determinism is, *prima facie*, in trouble in Maxwellian spacetime. In textbook presentations of GTR, a model of the theory takes the form (M, g_{ab}, T^{ab}), where M is a differentiable manifold and g_{ab} and T^{ab} are tensor fields on M which are interpreted respectively as the spacetime metric and the stress-energy tensor. If (M, g_{ab}, T^{ab}) is a solution of Einstein's field equations, then so is $(M, d^{*}g_{ab}, d^{*}T^{ab})$ where $d : M \to M$ is any diffeomorphism of M onto itself. But d can be chosen to be the identity map on and below any chosen timeslice of the spacetime M, g_{ab} but non-identity to the future of this slice. With such a choice the two solutions to Einstein's field equations will agree on the values of g_{ab} and T^{ab} for all points of M on or to the past of the chosen slice but will disagree on the values of these fields to the future of the slice–an apparent violation of determinism. "[I]n a general relativistic theory," Bergmann wrote, "–and this is in sharp contrast to a Lorentz-covariant theory–no initial data enables us to predict the values of local field observables at some later time ..." (1961, 510). Nevertheless, Bergmann held that "such a theory can be completely causal [i.e. deterministic], in a slightly modified sense" (511); namely, the theory enables us to unequivocally predict the values of *observables* from appropriate initial data. Before we can evaluate this claim we need to know what an observable in GTR is. Bergmann's first answer looked alarmingly circular: "We shall call a quantity observable if it can be uniquely predicted from initial data" (511). Fortunately there is a way out of this circle.

Einstein's field equations can be derived from an action principle, the Lagrangian density being the Ricci curvature scalar times $\sqrt{-g}$ (the so-called Einstein-Hilbert action). This action admits the diffeomorphism group as a variational symmetry, so one knows by Noether's second theorem that we are concerned with a case of underdetermination–another way of saying that there is an apparent violation of determinism. When one Legendre transforms to the Hamiltonian form of the theory, one discovers (no surprise) that the Hamiltonian form of Einstein's field equations constitutes a constrained Hamiltonian system. Thus, one can proceed as in the case of the humble example of section 2 and define an observable to be a dynamical quantity that is gauge invariant, i.e. has vanishing Poisson bracket with all of the first class constraints. Since in GTR these constraints generate the diffeomorphism group, determinism as applied to gauge invariant quantities escapes the hole argument. (All of this was discovered simultaneously by Bergmann and Dirac (1950, 1951, 1964), both of whom had been working on a similar program to quantize GTR that came to be known as *canonical quantization*.)

This seems very pleasing. But when one follows through the details of the analysis, pleasure turns to consternation. The first unpleasant surprise is that the Lie algebra of the constraints in GTR is not closed; that is to say, it is not a genuine Lie algebra.[14] Since the closure is a defining feature of Yang-Mills theories, it follows that GTR is not a Yang-Mills theory. Now some writers want to reserve the label "gauge theory" for Yang-Mills theories. This seems to me to be a merely terminological matter–if you do not wish to call GTR a gauge theory because it is not Yang-Mills, that is fine with me; but please be aware that the constrained Hamiltonian formalism provides a perfectly respectable sense in which the standard textbook formulation of GTR using tensor fields on differentiable manifolds does contain gauge freedom. What goes beyond label mongering is the issue of why GTR fails to be a Yang-Mills theory and, more generally, what features separate constrained Hamiltonian theories that are Yang-Mills from those which are not. Some important results of these matters have been obtained by Lee and Wald (1990), but these results are too technical to review here.

The second and even more unpleasant surprise is that taking the Bergmann route leads to a revival of McTaggartism. There are two constraints in the constrained Hamiltonian formulation of GTR: the *momentum constraint* and the *Hamiltonian constraint*.[15] The former generates diffeomorphisms in the initial value surface while the latter generates motion. Since motion is pure gauge, it follows that all of the gauge invariant quantities of the theory are constants of the motion. Leaving aside the questionable business of "A-series" change (change with respect to the monadic properties of presentness, futurity, and pastness), which in any case plays no role in physical theory, the proposed interpretation of GTR seems to imply a static "block universe." Reactions in the physics literature to this result are divided. One side there are those who say that the result is incorrect: obviously (they claim) there is genuine change, which means that the Hamiltonian constraint in GTR cannot be taken to generate gauge (see Kuchař (1993)). On the other side there are those who accept the result and try to explain how the appearance of change arises in a changeless world (see Barbour (1994a, 1994b, 2000) and Rovelli (1991)). This is not the place to thrash out the ins and outs of modern McTaggartism (see Belot and Earman (2000) for details). But even without the details, I think it should be obvious that by concentrating almost exclusively on the *ontological* implications of the hole argument–principally the implications for the status of spacetime points–philosophers of science have missed a number of fascinating issues that flow from the *ideological* implications.[16]

4 Problems and prospects

I have touted the constrained Hamiltonian formalism as a kind of royal road to the understanding of gauge. It is now time to take notice of some potholes and potential dead ends in this road.

The scope of the approach is very broad in the sense that it applies to the vast majority of theories in modern physics. But this is because physicists, for one reason or another, have concentrated on theories where the field equations or equations of motion are derivable from an action principle. There are notable exceptions, such as the Cartan formulation of Newtonian gravitation, a theory which philosophers of science have found attractive and instructive (see Malament (1995) and Norton (1995)). The Cartan theory is a hybrid in which part of the spacetime structure remains absolute (the simultaneity structure, the $\mathbb{E}^3$ structure of the instantaneous space, and the time metric) and part (the affine structure) becomes dynamical. Newtonian gravitation is geometricized in that massive test particles experience no gravitational force but follow the geodesics of the (non-flat) connection. This theory evokes two contrasting attitudes towards the analysis of "gauge." The first is that one wants to use gauge language in describing features of this theory, and since such language isn't sanctioned by the constrained Hamiltonian formalism, there is a need to find a more general characterization of "gauge" which will legitimize such talk. The opposite attitude is that precisely because such talk is not sanctioned by the constrained Hamiltonian formalism, it should be viewed with suspicion. To help settle this dispute, it would be helpful to examine a number of other examples, but that is a task I cannot undertake here.

A more serious challenge is concerned not with the boundaries of the domain of applicability of the constrained Hamiltonian formalism but rather with the fact that within its domain of application it can render different verdicts on what counts as an observable depending on the way in which the theory is formulated. Here are two examples. *Ex. 1* (Janis (1969)). In the conventional formulation of Maxwell's theory of electromagnetism in Minkowski spacetime the electromagnetic field tensor F_{ab} is written in terms of a four-potential A_a, $a = 1, 2, 3, 4$:

$$F_{ab} = 2A_{[a,b]} \tag{11}$$

where the square brackets denote anti-symmetrization. Maxwell's equations in vacuo are

$$\begin{aligned} F^{ab}{}_{,b} &= 0 && (12a) \\ {}^{*}F^{ab}{}_{,b} &= 0 && (12b) \end{aligned}$$

where the $*$ denotes the dual tensor and indices have been raised using the Minkowski metric. The theory can be recast in terms of two four-potentials, A_a and $\bar{A}_a$ by writing

$$F_{ab} = 2A_{[a,b]} + 2\bar{A}_{[a,b]}. \tag{13}$$

With the Lagrangian density $L := \frac{1}{2}F_{ab}F^{ab}$ per usual, this new version of Maxwell's theory also yields the field equations (12). In the first version, the field tensor, but not the potentials, are observables. But in the second version, $\bar{A}^{[a,b]}$ as well as $A^{[a,b]}$ is an observable. *Ex. 2* (Folklore). Start with the standard Hamiltonian formulation of a classical one-dimensional harmonic oscillator. The phase space is $\Gamma = \Gamma(x, p_x)$ and the Hamiltonian is $H = p^2/2 + \omega^2 x^2/2$. There are no constraints and, thus, no gauge freedom, and the position of the oscillator is an observable. Now "parameterize," that is, extend the phase space to $\Gamma' = \Gamma'(x, p_x, t, p_t)$ by considering the time t as a configuration variable, which necessitates also adding the conjugate momentum p_t. This augmented Hamiltonian system has one constraint; namely, the super-Hamiltonian $\mathcal{H} = H + p_t$ vanishes. Since the motion on the augmented phase space is generated by $\mathcal{H}$, the dynamics is pure gauge. The position of the oscillator is no longer an observable (since it does not commute with $\mathcal{H}$), and in parallel with the GTR case all observables are constants of the motion.

My reaction to these and related examples starts from the truism that a physical theory is not simply a set of mathematical equations but the equations together with an intended interpretation. A crucial part of that interpretation is, of course, what is to be counted as an observable–thus, in both of the above examples one should see two different theories rather than two formulations of the same theory. No formal apparatus–the constrained Hamiltonian formalism included–can make the decision of what counts as an observable. What the constrained Hamiltonian formalism does do is to provide a clear and precise list of choices: given any way of presenting the equations of motion (or field equations) in Lagrangian/Hamiltonian form, the apparatus tells you what counts as an observable in that presentation. A combination of experimental and theoretical considerations is required to decide what quantities should be recognized as observables and, thus, which of the theories–formal apparatus plus choice of observables–is appropriate. In the case of electromagnetism we have at present no observational or theoretical reason to doubt that the standard electromagnetic field tensor $A^{[a,b]}$ provides a complete characterization of the electromagnetic field in vacuo. But as Janis (1969) notes, if the existence of magnetic monopoles were verified, the additional field tensor $\bar{A}^{[a,b]}$ would be needed, with magnetic charges and currents serving as its source, in order to fully describe the electromagnetic field. As for the second example, in a pre-GTR setting we would seem justified in dismissing the parameterized formulation on the grounds that obviously there is change in the world–oscillators do oscillate!–and that observables like position of the oscillator are needed to describe change. After the advent of GTR this is no longer so obvious.

Let me now turn to a worry about the status of determinism. I said that determinism becomes a trivial doctrine if whenever cracks appear in the doctrine we are ready to paper over the cracks by seeking gauge freedom. And then I gave the impression that the trivialization is halted by providing a principled way to detect gauge freedom. This impression is badly misleading if the means of detecting gauge freedom is that of Dirac. Start with any theory derivable from an action principle, and suppose that the ELEs fail to determine a unique solution from initial data because arbitrary functions of time appear in the solutions. A cure for this form of indeterminism is always at hand in that in the constrained Hamiltonian formalism the gauge transformations as identified by Dirac's prescription are sufficient to sop up the underdetermination. (Of course, it may happen that the cure takes the drastic form of freezing the dynamics.) To be sure, the assumption that the field equations or equations of motion be derivable from an action principle is a strong restriction, but nevertheless I find it disconcerting that for such cases a cure for underdetermination is always possible.

Finally, I need to say something about Dirac's proposal (or conjecture as it is sometimes called) that *all* first class constraints, secondary, tertiary, etc. generate gauge transformations. There exists an extensive literature consisting of purported counterexamples to and defenses of Dirac's proposal. I cannot even begin to review the literature here, but the issues raised by the purported counterexamples must be treated by any decent account of gauge since they go to the heart of what counts as a gauge transformation and, thus, what counts as an observable. In discussing potential counterexamples, it is useful to start from the conjecture that the total Hamiltonian $H_T := H + \lambda^i \phi_i$ (where H is the canonical Hamiltonian $\sum_i \dot{q}^i p_i + L$, the ϕ_i are the primary first class constraints, and the Lagrange multipliers λ^i are arbitrary functions) and the extended Hamiltonian $H_E := H_T + \beta^j \psi_j$ (where the ψ_j are the secondary, tertiary, etc. first class constraints and again the β^j are arbitrary functions) generate equivalent motions. Counterexamples to this conjecture provide instances where the non-primary first class constraints arguably do not generate gauge. I will illustrate the flavor of the debate by means of an example taken from Henneaux and Teitelboim (1992, 19). The configuration space has two degrees of freedom, x and y, and the Lagrangian is $L = \frac{1}{2}\exp(y)\dot{x}^2$. There is one primary constraint, $p_y = 0$, which turns out to be first class. The canonical Hamiltonian is $H = \frac{1}{2}\exp(-y)p_x^2$ where $p_x = \exp(y)\dot{x}$. This is not the end of the story since we need to assure that the primary constraint is preserved by the equation of motion, i.e.

$$\dot{p}_y = [p_y, H_T] \approx 0 \tag{14}$$

where $H_T = H + \lambda p_y$ and "$X \approx 0$" (read as X is weakly zero) means that X vanishes on the constraint surface. Equation (14) implies that $p_x^2 \approx 0$, which in turn implies that $p_x \approx 0$. This latter secondary constraint is also first class. So if we were to follow Dirac's proposal, the motion is generated by the extended Hamiltonian $H_E = H_T + \beta p_x$. As a consequence the x-degree of freedom is treated as pure gauge. But, arguably, this is wrong because the ELEs imply that $x = const.$ so that this degree of freedom is fully determined. So if an apparent breakdown of determinism is the key motivation for seeking gauge freedom, there is no good reason to treat the x-degree of freedom as gauge.[17]

As far as I am aware all of the examples of this kind have the feature that as one cranks through the Dirac algorithm, one hits an *ineffective constraint* ψ, i.e. $d\psi \approx 0$, which means that the constraint doesn't generate any motion on the constraint surface. (In our toy example $\psi = p_x^2$ is ineffective.) But this ineffective constraint leads, at some stage in the algorithm, to a constraint which is effective. (In our example, the ineffective constraint leads to the effective constraint $\psi' = p_x$.) If this is correct, a technical restriction can be used to fix Dirac's proposal. But the need for a fix suggests that there should be a way to characterize the gauge freedom directly in terms of the Lagrangian formalism. As an example of what I have in mind, suppose that the action $\mathcal{L} = \int L(q, \dot{q})dt$ admits the group $\mathcal{G}_{\infty r}$ of variational symmetries. Define the observables to be functions of the velocity phase space $(q, \dot{q})$ which are invariant under $\mathcal{G}_{\infty r}$. Is this notion of observable equivalent, in some relevant sense, to Dirac's notion? If not, which concept of observable is preferable?

5 Conclusion

"Gauge" is used in so many different ways that there isn't going to be any one univocal concept that corresponds to all the different uses. What I have attempted here is to sketch one approach to capturing what I take to be one of the central features of the gauge notion as it applies to a broad class of physical theories. I welcome the competition from other approaches as a way to sharpen the issues. As for other uses of "gauge" I am willing to ecumenical, but with a caveat. To give an example, some people want to call scaling transformations gauge transformations whereas others don't. My approach explains why the naysayers say what they do. (Typically scaling transformations are symmetries of the equations of motion but are not variational symmetries, and even when they are variational symmetries they are not of the kind that invokes Noether's second theorem.) But I don't see anything against extending what I take to be the core gauge concept to cover such cases. However, it is important to be clear that such an extension involves putting together notions of gauge that have different roots and motivations. The analysis of gauge calls for the kind of multiculturalism that remains sensitive to these differences.

Acknowledgment: Thanks are due to Godon Belot, Chris Martin, and Rob Rynasiewicz for help comments on an earlier draft.

Notes

1. See Auyang (2000) and Teller (2000). See also Morrison (1995, 2000). The members of this symposium do not endorse the approaches or conclusions of any of these authors.

2. Sometimes the notion of gauge is defined in terms of automorphisms of principal fibre bundles, as in Bleeker (1981).

3. Here "normal" doesn't mean problem-free, for as always one can run into operator ordering problems in trying to turn classical observables into self-adjoint operators on a Hilbert space.

4. These theorems were presented in Noether (1918). For relevant historical information, see Kastrup (1987) and Byers (1999). For a modern presentation of the Noether theorems, see Olver (1993).

5. The transformations $\mathcal{G} \ni g : (\mathbf{x}, \mathbf{u}) \to (\mathbf{x}', \mathbf{u}')$ may depend on both the independent and dependent variables. Noether's theorems can be generalized to handle transformations that depend on the $\mathbf{u}^{(n)}$ as well (see Olver (1993)), but these generalized transformations will play no role here. But what is relevant here is the fact that Noether's theorems can be generalized to handle so-called divergence (variational) symmetries that leave the action invariant only to a term of the form $\int_{\partial\Omega} Div(\mathbf{B})$, where the variation of $\mathbf{B}$ vanishes on the boundary $\partial\Omega$. For example, the Galilean velocity boosts do not leave the familiar actions for Newtonian particle mechanics invariant, but these boosts are divergence (variational) symmetries. This is crucial in deriving, via Noether's first theorem, the conservation of center of mass. From here on when I speak of variational symmetries I will mean divergence (variational) symmetries.

6. In some cases this conservation law can be written in the form $D_t T + div(X) = 0$, where div is the spatial divergence. Then if the flux density X vanishes on the spatial boundary of the system, the spatial integral of the density T is a constant of the motion.

7. In Minkowski spacetime, $\mathbf{x} = (x^1, ..., x^4)$, where x^4 is the time coordinate. In this setting, the current $\mathbf{P}$ of an improper or trivial conservation law $Div(\mathbf{P}) = 0$ takes the form $P^i = \sum\limits_{k=1}^{4} D_k X^{ki}$ on a solution to the ELEs. The time component $P^4 = \sum\limits_{k=1}^{4} D_k X^{k4} = \sum\limits_{k=1}^{3} D_k X^{k4}$ since X^{ki} is anti-symmetric. Integrating over a spatial three-volume V_3 gives the conserved charge $Q := \int_{V_3} P^4 d^3x = \int_{V_3} \sum\limits_{k=1}^{3} D_k X^{k4} d^3x =$

$\int_\sigma \sum_{k=1}^{3} n_k X^{k4} d^2\sigma$ where σ is the two-surface bounding the volume V_3 and n_k is the unit normal to the surface. In electromagnetism this relation expresses the charge in a spatial volume as the flux of the electric field through the bounding surface. For an interesting discussion of how the distinction between proper and improper conservation laws illuminates Hermann Weyl's work on gauge theories, see Brading (2000).

8. Strictly speaking, one should say has *weakly* vanishing Poisson bracket with all of the constraints; see section 4 for what this means.

9. Some qualms about this proposal are discussed in section 4.

10. For details, see Chapter 2.3 of my (1989).

11. Church's Thesis, so-called, is best regarded not as a thesis but as a proposal for explicating a vague concept–here, effectively computable function–by means of a precise one–Turing computability. The Gauge Thesis is to be read in a similar spirit.

12. Suppose that the first class constraints form a Lie algebra and that this algebra exponentiates to give a Lie group $\mathcal{G}$ which acts freely on $\mathcal{C}$. Then if the quotient $\mathcal{C}/\mathcal{G}$ is a manifold, it will be the base space of a $\mathcal{G}$-bundle whose fibres are the orbits of the group action. (Private communication from John Baez.) Not all of these conditions are necessary for the gauge orbits generated by the first class constraints to be the fibres of a nice bundle; I don't know what a minimally necessary condition is.

13. For a history of Einstein's struggles with the hole argument, see Norton (1987) and Howard (1999). Bergmann had been an assistant to Einstein at the Institute for Advanced Study. But I know of no evidence that he discussed the hole argument with Einstein.

14. Instead of using the (3+1) Hamiltonian formalism, it is possible to treat GTR as a symplectic system in which the phase space is taken to be not the space of instantaneous states but rather the space of entire histories or solutions to Einstein's field equations (see Ashtekar and Bombelli (1991)). In this treatment the diffeomorphism group acts as a Lie group on the phase space. However, this treatment bypasses rather than resolves the mystery of why the algebra of constraints in the (3 +1) formalism is not a Lie algebra, for in the "covariant formalism" there are no constraints. Rather the phase space is equipped with a "pre-symplectic form" whose null directions correspond to the gauge orbits.

15. Strictly speaking one should refer to momentum and Hamiltonian constraint**s** since there are different constraints for each point of space.

16. I cannot forebear from adding that much of the philosophical literature on the hole argument is wide of the mark–especially that strain that says that the argument shows nothing special about GTR–as opposed to any other spacetime theory–or even about spacetime theories in general. The use of the constrained Hamiltonian formalism shows why this is not so.

17. But note that p_x is pure gauge. If the goal is quantization, it is awkward to say the least to have to deal with a case where the space of physically distinct initial data is one-dimensional. For this reason Henneaux and Teitelboim (1992, 19) take x to be pure gauge.

References

[1] Ashtekar, A. and Bombelli, L. 1991. "The Covariant Phase Space of Asymptotically Flat Gravitational Fields," in M Francaviglia (ed.), *Mechanics, Analysis and Geometry: 200 Years after Lagrange*, 417-450. Amsterdam: Elseiver Science Publishers.

[2] Auyang, S. 2000. "Mathematics and Reality: Two Notions of Spacetime in the Analytic and Constructive Views of Gauge Theory," *PSA 1998*, Vol. 2.

[3] Barbour, J. 1994a. "The Timelessness of Quantum Gravity: I. The Evidence of the Classical Theory," *Classical and Quantum Gravity* **11**: 2853-2873.

[4] Barbour, J. 1994b. "The Timelessness of Quantum Gravity: II. The Appearance of Dynamics in Static Configurations," *Classical and Quantum Gravity* **11**: 2875-2897.

[5] Barbour, J. 2000. *The End of Time*. Oxford: Oxford University Press.

[6] Belot, G. and Earman, J. 2000. "Modernizing McTaggart," pre-print.

[7] Bergmann, P. G. 1949. "Non-Linear Field Theories," *Physical Review* **75**: 680-685.

[8] Bergmann, P. G. 1961. "Observables in General Relativity," *Reviews of Modern Physics* **33**: 510-514

[9] Bleeker, D. 1981. *Gauge Theory and Variational Principles*. Reading, MA: Addison-Wesley.

[10] Brading, K. 2000. "Which Symmetry? Noether, Weyl, and Conservation of Electric Charge," pre-print.

[11] Byers, N. 1999. "E. Noether's Discovery of the Deep Connection Between Symmetries and Conservation Laws," *Israel Mathematical Society Proceedings* **12**:

[12] Dirac, M. 1950. "Generalized Hamiltonian Dynamics," *Canadian Journal of Mathematics* **2**: 129-148.

[13] Dirac, M. 1951. "The Hamiltonian Form of Field Dynamics," *Canadian Journal of Mathematics* **3**: 1-23.

[14] Dirac, M. 1964. *Lectures on Quantum Mechanics*. New York: Belfer Graduate School of Science Monographs Series.

[15] Earman, J. 1989. *World Enough and Spacetime: Absolute vs. Relational Theories of Space and Time*. Cambridge, MA: MIT Press.

[16] Henneaux, M. and Teitelboim, C. 1992. *Quantization of Gauge Systems*. Princeton, NJ: Princeton University Press.

[17] Howard, D. 1999. "Point Coincidences and Pointer Coincidences: Einstein and the Invariant Content of Space-Time Theories." In H. Goenner, J. Renn, J. Ritter, and T. Sauer (eds.), *The Expanding Worlds of General Relativity* (*Einstein Studies*, Vol. 7), 463-500. Boston: Birkhäuser.

[18] Janis, A. I. 1969. "Energy and Observables in Electromagnetic theory," *American Journal of Physics* **38**: 202-206.

[19] Kastrup, H. A. 1987. "The Contributions of Emmy Noether, Felix Klein and Sophus Lie to the Modern Concept of Symmetries in Physical Systems." In M. G. Doncel (ed.), *Symmetries in Physics (1600-1980)*. Barcelona: Universitat Autonoma Barcelona.

[20] Kuchař. K. 1993. "Canonical Quantum Gravity," in R. J. Gleiser, C. N. Kozameh, and O. M. Moreschi (eds.), *General Relativity and Gravitation 1992*. Philadelphia, PA: IOP Publishing.

[21] Lee, J. and Wald, R. M. 1990. "Local Symmetries and Constraints," *Journal of Mathematical Physics* **31**: 725-743.

[22] Malament, D. 1995. "Is Newtonian Cosmology Really Inconsistent?" *Philosophy of Science* **62**: 489-510.

[23] Maxwell, J. C. 1877. *Matter and Motion*. Reprinted by Dover Books (New York, 1951).

[24] Morrison, M. 1995. "The New Aspect: Symmetries as Meta-Laws." In F. Weinert (ed.), *Laws of Nature: Essays on the Philosophical, Scientific and Historical Dimensions*, 157-188. Berlin: Walter de Gruyter.

[25] Morrison, M. 2000. *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*. Cambridge: Cambridge University Press.

[26] Noether, E. 1918. "Invariante Variationsprobleme," *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse.* Heft 2: 235-257.

[27] Norton, J. D. 1987. "Einstein, the Hole Argument, and the Reality of Space." In J. Forge (ed.), *Measurement, Realism and Objectivity*, 153-188. Boston: D. Reidel.

[28] Norton, J. D. 1995. "The Force of Newtonian Cosmology: Acceleration is Relative," *Philosophy of Science* **62**: 511-522.

[29] Olver, P. J. 1993. *Applications of Lie Groups to Differential Equations 2nd ed.* New York: Springer-Verlag.

[30] Rovelli, C. 1991. "Time in Quantum Gravity: An Hypothesis," *Physical Review D* **43**: 442-456.

[31] Teller, P. 2000. "The Gauge Argument," *PSA 1998*, Vol. 2. In press.

[32] Wipf, A. 1994. "Hamilton's Formalism for Systems with Constraints." In Ehlers, J. and Friedrich (eds.), *Canonical Gravity: From Classical to Quantum*, 22-58. New York: Springer-Verlag.